Annual Review of Applied Linguistics (2026), **46**, 199–222
doi:10.1017/S0267190526100257

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Neural mechanisms of attitudinal prosody in a second language: The impact of informal exposure

Victoria-Anne Flood[1], Diego Elisandro Dardon[2], Motoaki Sugiura[3,4] and Hyeonjeong Jeong[1,3]

[1]Graduate School of International Cultural Studies, Tohoku University, Sendai, Japan; [2]Faculty of Education, Miyagi University of Education, Sendai, Japan; [3]Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan and [4]International Research Institute of Disaster Science, Tohoku University, Sendai, Japan
Corresponding author: Hyeonjeong Jeong; Email: hyeonjeong.jeong.b8@tohoku.ac.jp

Abstract

Understanding how attitudinal prosody is processed in a second language (L2) remains an open question, particularly regarding its neural mechanisms and the role of real-world experiences. Using functional magnetic resonance imaging (fMRI), we examined how native Japanese learners processed attitudinal versus linguistic prosody in their L1 (Japanese) and L2 (English) during a forced-choice judgment task. Across languages, attitudinal and linguistic prosody engaged partially dissociable networks: attitudinal prosody recruited socio-cognitive regions involved in inferring speakers' intentions, whereas linguistic prosody engaged phonological-motor regions. Critically, L2 attitudinal prosody elicited distinct frontal modulation, with the left inferior frontal gyrus and middle frontal gyrus showing stronger prosody-type differentiation in English than in Japanese – indicating greater reliance on controlled interpretive and executive processes during non-native attitudinal prosody comprehension. Individual-difference analyses revealed that informal L2 exposure predicted enhanced activation in the thalamus and left hippocampus, as well as better attitudinal prosody identification. These converging neural and behavioral patterns suggest that socially grounded experience plays an important role in developing sensitivity to attitudinal prosody in an L2. Together, these findings provide novel neural evidence for how L2 learners interpret attitudinal prosody and show that L2 exposure is associated with differences in pragmatic prosody at cognitive and neural levels.

Keywords: attitudinal prosody; second language; fMRI; language exposure; prosody perception

Introduction

Prosody – systematic variations in pitch, intensity, voice quality, and duration – shapes how listeners interpret a speaker's message. Beyond its linguistic functions, prosody

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

conveys affective and social information. Attitudinal prosody, in particular, signals interpersonal intentions such as doubt, confidence, or friendliness (Wichmann, 2000) and is essential for understanding meaning beyond the literal words. However, prosodic expressions of social intentions vary across languages, and listeners often misinterpret attitudinal cues in cross-linguistic communication (Chen et al., 2004; Shochi et al., 2006). As multilingual interaction becomes increasingly common, accurately decoding such cues in a second language (L2) has become an important yet challenging skill. The present study, therefore, examines how L2 learners perceive attitudinal prosody and how this process differs from first language (L1) comprehension, adopting an interdisciplinary approach that integrates behavioral, linguistic, and neural evidence. We also consider the potential role of learners' real-world exposure in shaping their ability to interpret attitudinal prosody in the L2.

Attitudinal prosody in the first and second languages

In the L1, attitudinal prosody efficiently conveys speakers' social intentions. Both children and adults reliably categorize attitudes such as doubt, confidence, or enthusiasm based solely on prosodic cues, with clear acoustic correlates in pitch, duration, and intensity (Esteve-Gibert & Prieto, 2013; Hellbernd & Sammler, 2016; Hupp et al., 2021). These findings highlight the intuitive, experience-based nature of attitudinal prosody processing in the L1. In contrast, interpreting attitudinal prosody in an L2 is substantially more difficult. Cross-linguistic studies show that languages differ in how they encode prosodic intentions, leading L2 learners to misinterpret attitudes even when acoustic cues are perceived correctly (Chen et al., 2004; Kang & Kermad, 2019; Rilliard et al., 2014; Shochi et al., 2006). For example, some attitudes, such as authority or interrogation, generalize across languages, whereas others, such as expressions of doubt, are difficult for Japanese learners of English to interpret (Shochi et al., 2009). This difficulty may partly reflect differences between English and Japanese in how prosodic prominence and pitch variation signal both linguistic structure and interpersonal stance. These findings suggest that L2 learners must acquire not only prosodic cues but also the language-specific pragmatic mappings linking prosody to communicative intention.

Attitudinal prosody relies on language-specific pragmatic mappings, making its acquisition in an L2 highly dependent on learners' experience with meaningful exposure. Although higher proficiency reduces reliance on L1-based interpretive strategies (Chen, 2009; Shochi et al., 2016), L2 exposure may be especially critical. Naturalistic exposure allows learners to encounter authentic prosodic realizations of social intentions – patterns rarely taught explicitly in L2 classrooms (Kang et al., 2021) – and provides opportunities to infer how attitudes are encoded in real communicative settings. Behavioral and neurocognitive research shows that immersion supports prosody comprehension (Trofimovich & Baker, 2006) and induces structural and functional adaptations in language-control and socio-cognitive networks (DeLuca et al., 2019; Pliatsikas et al., 2017). Because attitudinal prosody is largely acquired implicitly, the quantity and quality of L2 exposure may be a particularly strong driver of successful interpretation in the L2, relative to learners' proficiency.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Neural mechanisms of prosody processing

Neuroimaging research has shown that prosody processing relies on a distributed network that integrates auditory, linguistic, and socio-cognitive processes. Linguistic prosody, which supports sentence segmentation, focus marking, and syntactic interpretation, typically engages left-lateralized speech-language areas. Core regions include the left inferior frontal gyrus (IFG) and superior and middle temporal gyri, particularly when prosody is essential for sentence comprehension (Friederici, 2002, 2006; van der Burght et al., 2019). When prosodic cues are redundant or nonessential, processing shifts toward right-hemisphere auditory regions. In contrast, affective and attitudinal prosody additionally involve socio-cognitive systems implicated in inferring speakers' emotional states, interpersonal stance, or epistemic commitments. Meta-analytic evidence shows that affective prosody activates audio-vocal and limbic regions (e.g. amygdala, anterior insula), whereas linguistic prosody primarily recruits the speech-language network (Belyk & Brown, 2014; Mitchell & Ross, 2013). Critically, attitudinal prosody, which conveys epistemic stance or interpersonal attitudes, engages theory-of-mind (ToM)–related regions such as the medial prefrontal cortex (mPFC), posterior cingulate cortex, and temporo-parietal junction, supporting the inference of communicative intent (Enrici et al., 2019; Hellbernd & Sammler, 2018; Tettamanti et al., 2017). These findings highlight the dual involvement of auditory–prosodic decoding and higher-order socio-cognitive evaluation in processing attitudinal meaning.

Processing prosody in an L2 imposes additional cognitive and linguistic demands. L2 sentence comprehension generally recruits more extensive frontal and temporal regions than L1, including greater engagement of the IFG and other language-control areas (Jeong et al., 2007; Liu & Cao, 2016; Perani & Abutalebi, 2005). Research on L2 linguistic prosody further shows increased activation in the anterior insula and superior frontal sulcus, likely reflecting heightened attentional and working-memory requirements (Gandour et al., 2007). These findings suggest that the difficulty of L2 prosody processing may arise from both perceptual challenges in decoding unfamiliar acoustic patterns and integrative demands in mapping these cues onto language-specific pragmatic meanings.

Despite the relevance of these systems, the neural mechanisms supporting attitudinal prosody in an L2 remain virtually unexplored. Clarifying these mechanisms is essential for determining whether the additional cognitive load in L2 attitudinal prosody is primarily perceptual (e.g. unfamiliar acoustic contours) or integrative (e.g. inferencing epistemic stance from less familiar pragmatic conventions). This question is central to understanding whether L2 learners recruit the same socio-cognitive and frontal networks as in their L1, or whether they rely on additional executive processes – an issue directly addressed in the present study.

Individual differences in L2 prosody perception

A range of individual learner characteristics may shape how attitudinal prosody is processed in a second language. Among these factors, L2 exposure has emerged as a particularly influential contributor, because it provides learners with opportunities to encounter authentic prosodic realizations of social intentions – patterns rarely taught

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

explicitly in L2 classrooms. Behavioral studies show that greater immersion facilitates more accurate interpretation of prosodic and pragmatic cues (Kang et al., 2021; Trofimovich & Baker, 2006). Neurocognitive research also demonstrates that extensive L2 exposure leads to structural and functional adaptations in multiple learning systems. One line of research emphasizes procedural learning mechanisms, reporting experience-dependent expansions in the caudate nucleus, thalamus, and nucleus accumbens during sustained L2 immersion (DeLuca et al., 2019; Pliatsikas et al., 2017). Another line highlights the role of the hippocampal–parahippocampal system, traditionally linked to episodic and contextual learning, in supporting the integration of rich, naturalistic L2 input. Although these frameworks are often presented as distinct, they converge on the idea that real-world experience reshapes the neural systems that support L2 acquisition. The present study builds on this convergence by testing whether attitudinal prosody, a socially grounded, largely implicitly learned aspect of communication, engages both episodic (hippocampal) and procedural (thalamic–striatal) mechanisms in an L2.

Proficiency has also been shown to influence L2 attitudinal prosody processing. Learners with higher proficiency tend to rely less on L1-based interpretive strategies and demonstrate more accurate mappings between prosodic cues and communicative intentions (Chen, 2009; Shochi et al., 2016). Willingness to Communicate (WTC) may play a role as well, as it shapes how often learners engage in interactional situations in which attitudinal meanings can be perceived (Elahi Shirvan et al., 2019; Yashima, 2002). At the same time, prior work has highlighted naturalistic, informal exposure as a particularly important source of socially meaningful prosodic input, suggesting that exposure may function as a key experiential factor in the development of L2 attitudinal prosody sensitivity.

The present study

Despite growing evidence on the neural bases of prosody and the role of experiential factors in L2 learning, no research has examined the neural mechanisms supporting attitudinal prosody in a second language, nor how learner-related variables, especially naturalistic exposure, modulate this process. Previous work has focused primarily on L1 attitudinal prosody or on L2 linguistic prosody, leaving open the question of how L2 listeners integrate prosodic cues with language-specific pragmatic mappings to infer speakers' social or epistemic intentions. The current fMRI study addresses this gap by comparing attitudinal and linguistic prosody processing in L1 Japanese and L2 English, and by examining how individual differences in L2 experience shape behavioral and neural outcomes. We pursued three research questions:

1. To what extent does L2 attitudinal prosody recruit neural systems that overlap with, or differ from, those engaged in the L1?
2. How does the interaction between prosody type (attitudinal vs. linguistic) and language (L1 vs. L2) modulate neural activation?
3. How do individual differences, particularly naturalistic, informal L2 exposure, shape behavioral and neural responses to attitudinal prosody?

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Although the mapping between specific prosodic cues and attitudes may differ between Japanese and English, interpreting attitudinal prosody in either language requires listeners to infer the speaker's communicative intention. We therefore predicted that L2 attitudinal prosody would recruit intention-inference and socio-cognitive systems that overlap with those engaged in the L1. At the same time, we predicted that L2 processing would place greater demands on executive and controlled interpretive processing regions, particularly the left IFG, because learners must map less familiar prosodic cues onto language-specific pragmatic and linguistic conventions. Finally, we predicted that learners with greater naturalistic L2 exposure would show stronger engagement of neural systems involved in experience-based interpretation of attitudinal prosody.

Methods

Participants

Thirty-six Japanese native speakers participated in this fMRI experiment. Six participants were excluded: four due to excessive head motion (more than 5 mm), and two for accuracy below chance level (50%). The remaining 30 participants (aged 18–30 years, $M = 20.57$, standard deviation [SD] $= 2.22$; 10 females) were included in the analyses. All participants were right-handed (Edinburgh Handedness Inventory), reported normal or corrected-to-normal vision and normal hearing, and had no history of neurological disorder. Participants had intermediate L2 proficiency based on TOEFL ITP scores ($M = 530.27$, $SD = 27.21$), corresponding to the Common European Framework of Reference for Languages (CEFR) B1 level. They began learning English at an average age of 10.5 years ($SD = 2.9$), primarily through classroom instruction. Listening proficiency was assessed via the Oxford Quick Placement test (British English materials; $M = 79.36$, $SD = 6.35$) (Allen, 1992); these scores were used as the proficiency index in the analyses. All participants provided informed consent. The study was approved by the Ethical Committee of Tohoku University.

Individual-difference measures

To examine individual differences in L2 prosody comprehension, participants completed three questionnaires after the experiment.

(1) **L2 exposure and use**

To capture participants' habitual exposure to English, we developed a language background questionnaire assessing both the quantity and context of L2 exposure and use. Participants reported how frequently they encountered English across a range of academic and nonacademic contexts using a 6-point scale (0 = "never," 5 = "every day"). The questionnaire contained 16 items in total, covering listening and speaking across academic (7 items) and nonacademic (9 items) contexts.

- *Academic exposure* included activities such as attending lectures in English, reading academic articles, preparing presentations, and participating in class discussions.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

- *Non-academic exposure* encompassed watching English-language films or YouTube videos, listening to news or podcasts, playing games, and interacting with foreigners in informal settings.

For analysis, listening and speaking exposure items were combined to derive three indices for each participant: overall exposure (mean across all 16 items), academic exposure (mean across the 7 academic items), and nonacademic exposure (mean across the 9 nonacademic items). These indices, therefore, represent mean scores on a 0–5 scale. The overall exposure score was intended to capture learners' cumulative experience with English across contexts and was therefore used as the primary predictor in the main individual-difference analyses. Academic and nonacademic exposure were examined in follow-up exploratory analyses to further characterize the nature of the exposure effect by assessing whether it was more strongly associated with one context type than the other. The full questionnaire is provided in the Supplementary Materials.

Overall exposure was low ($M = 0.70$, $SD = 0.40$), consistent with the limited opportunities for English use in Japanese contexts. Academic exposure was the lowest ($M = 0.48$, $SD = 0.34$), whereas nonacademic exposure was higher and more variable ($M = 0.86$, $SD = 0.64$), suggesting that informal media input accounted for most of the participants' English contact.

(2) **Willingness to Communicate**

The Japanese version of MacIntyre and Clément's WTC scale (Yashima, 2002) was used to assess participants' self-perceived likelihood of initiating communication in English. Using a 0–100 scale, participants rated how likely they would initiate a conversation with friends, acquaintances, and strangers. Given that scores below 52 indicate generally low WTC, the mean WTC score in our sample ($M = 42.04$, $SD = 18.9$) suggests that many participants were reluctant to initiate communication in English, potentially limiting their opportunities for L2 practice.

(3) **Autistic Spectrum Quotient (AQ)**

Because attitudinal prosody is often linked to social–pragmatic processing, we administered the AQ questionnaire (Wakabayashi et al., 2006) to account for individual differences in social-cognitive sensitivity. Higher autism spectrum disorder traits have been shown to negatively impact prosody perception (Sinagra & Wiener, 2022), as well as affect social cognition (Rice & Redcay, 2016). The AQ questionnaire measures the extent of autistic traits, ranging from 0 to 50, with higher scores indicating more autistic traits. In previous research, neurotypical adults typically score around 15–20 (Baron-Cohen et al., 2001), whereas individuals with autism spectrum conditions often score above 33 in the Japanese context, validated by Wakabayashi et al. (2006). AQ scores in the present study ranged from 12 to 35 ($M = 23.27$, $SD = 7.3$). AQ did not correlate significantly with either English ($r = –0.25$) or Japanese prosody comprehension accuracy ($r = –0.23$), and was therefore treated as a control measure rather than a primary predictor.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Stimuli

We created a total of 112 basic English sentences and their Japanese counterparts using high school–level English vocabulary (Yokogawa, 2009), corresponding approximately to the CEFR A2–B1 level. All sentences followed the same syntactic frame – an animate subject, a transitive verb, and an inanimate object – to minimize lexical, syntactic, and semantic variability across conditions. No lexical cues indicating interrogativity or attitude (e.g. question markers, modal verbs, or affective adjectives) were included. Japanese translations were generated via machine translation and subsequently checked and corrected by two Japanese native speakers.

Sentences were recorded by four native speakers of each language (English and Japanese) using a Tascam DR-07C portable audio recorder. Because linguistic prosody (e.g. focus marking and interrogativity) and attitudinal prosody involve different communicative goals, the sentences were recorded separately for the two prosody conditions to ensure that the intended prosodic patterns were produced clearly and naturally. English speakers were native speakers of the North American variety of English, since English education in Japan typically emphasizes the North American variety. Each sentence was produced twice: once with linguistic prosody and once with attitudinal prosody, resulting in a pool of 224 auditory stimuli per language. Using a counterbalanced design, the stimuli were divided into two lists so that each participant heard 112 sentences per language, with each sentence presented in only one prosodic version and no lexical content repeated within a language. For linguistic prosody recordings, speakers produced the sentences with either statement or question intonation and with a prosodic focus on either the subject or object, following the procedure of Gandour et al. (2007). This resulted in four linguistic prosody combinations (statement–subject focus, statement–object focus, question–subject focus, question–object focus), with 14 sentences per combination. For the attitudinal prosody recordings, speakers expressed one of four attitudes while keeping the lexical content unchanged: friendliness, unfriendliness, confidence, or doubt (14 sentences per attitude). Speakers were instructed to convey the intended effect solely through prosody, without altering the segmental content. To ensure the quality of the stimuli, recordings were cross-checked during production: for each language, four native speakers evaluated each other's utterances, and items judged as unclear or ambiguous with respect to the intended attitude were re-recorded. In addition, a perception-based validation was conducted with an independent group of 33 Japanese students, who listened to the English stimuli and identified the conveyed attitudes. The overall mean accuracy was 0.84 ($SD = 0.12$), indicating that the intended attitudes were generally recognizable.

All audio files were edited and amplitude-normalized using Audacity. To preserve the natural prosodic realization of the utterances, no additional temporal normalization was applied; accordingly, sentence duration varied somewhat across conditions (Appendix A).

The final stimuli were divided into four experimental conditions based on the crossing of prosody type (attitudinal vs. linguistic) and language (English vs. Japanese): (1) Attitudinal English (AE), (2) Linguistic English (LE), (3) Attitudinal Japanese (AJ), and (4) Linguistic Japanese (LJ). To characterize the acoustic properties of the stimuli, we analyzed duration, mean intensity, and fundamental frequency for each condition.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Mean intensity and mean F0 did not differ systematically as a function of prosody type or language, whereas F0 range showed a clear effect of prosody type (attitudinal vs. linguistic; $p < .001$), consistent with the expected pitch variation associated with attitudinal prosody. Summary acoustic values for each condition are provided in Appendix A.

Experimental procedures

Inside the scanner, participants performed a two-alternative forced-choice prosody judgment task. On each trial, two response options appeared on the screen, and participants selected the option that best matched the prosodic information in the spoken sentence by pressing a button as accurately and quickly as possible. A one-second fixation cross preceded the next trial. The task was organized into linguistic and attitudinal prosody blocks. In linguistic prosody blocks, participants categorized either sentence type or focus location, depending on the response labels shown on the screen: (i) Sentence type trials: ["Statement"] vs. ["Question"]. (ii) Focus-location trials: ["Subject"] vs. ["Object"] (i.e. which constituent carried prosodic prominence/narrow focus). Participants were instructed to indicate which constituent (subject or object) was prosodically emphasized, rather than selecting a single "most stressed word." In attitudinal prosody blocks, participants categorized the speaker's intended attitude using fixed label pairs: ["Friendly"] vs. ["Unfriendly"] or ["Confident"] vs. ["Doubtful"].

The experiment consisted of four sessions: two in English, followed by two in Japanese. The English and Japanese sessions were split into two because of the session's length. This fixed order was chosen to minimize cross-language interference and to preserve the interpretability of within-language individual differences analyses. Each session contained alternating blocks of linguistic and attitudinal prosody. To counterbalance block order, half of the participants completed linguistic blocks first, and the others completed attitudinal blocks first. As illustrated in Figure 1, each block consisted of four trials, followed by a 15-second rest period, yielding 14 blocks per session (56 trials per session). Trial order within each block was randomized. To counterbalance the stimuli, one group of participants heard half of the sentences in the attitudinal condition and the remaining half in the linguistic condition, whereas the second group heard the opposite assignment. Short rest breaks were provided between sessions. The total task-related scanning time was approximately 44 minutes, after which a T1-weighted structural scan was acquired.

The task was presented using E-prime 3.0 software (Psychology Software Tools, Inc., Pittsburgh, PA, USA). Auditory stimuli were delivered through MRI-compatible active noise–canceling optical headphones (OptoActive; Optoacoustics Ltd., Or Yehuda, Israel), which substantially reduced scanner noise and ensured adequate audibility of the speech stimuli. Prior to entering the scanner, participants completed a practice session of 28 English trials, which did not appear in the main experiment. After completing all fMRI sessions, participants performed the English listening test and completed the three individual-difference questionnaires described in the individual-difference measures section.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

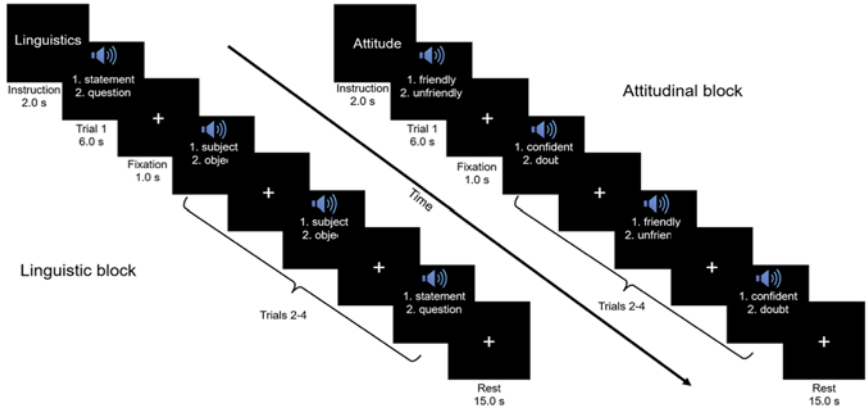


Figure 1. Schematic illustration of the fMRI prosody task. Each block consisted of four trials of either linguistic or attitudinal prosody judgments, followed by a 15-s rest period. Each trial began with a fixation cross (1 s), followed by presentation of an auditory sentence with two response options, to which participants responded via button press. In linguistic blocks, participants categorized sentence type (statement vs. question) or focus location (subject vs. object). In attitudinal blocks, participants categorized speaker attitude (e.g. friendly vs. unfriendly).

Data acquisition and preprocessing

Functional MRI data were collected on a Philips Achieva dStream 3.0T scanner using a gradient-echo EPI sequence (Repetition time (TR) = 2,000 ms, echo time (TE) = 30 ms, flip angle = 80°, 32 axial slices, 3-mm thickness, field of view (FOV) = 192 mm, matrix = 64 × 64). Each session consisted of 333 volumes. A high-resolution T1-weighted structural image (1-mm isotropic) was acquired for anatomical reference. Preprocessing was performed in SPM12 (Wellcome Centre for Human Neuroimaging). Functional images underwent slice-timing correction, realignment, co-registration of the T1 image to the mean functional image, normalization to MNI space, and spatial smoothing with an 8-mm FWHM Gaussian kernel. Motion parameters were inspected for quality assurance, and participants with excessive motion (>5 mm or substantial framewise displacement) were excluded.

Data analysis

Behavior data

Behavioral performance was analyzed to examine how prosody type and language influenced participants' accuracy and reaction time (RT). For both dependent variables, a 2 × 2 repeated-measures analysis of variance (ANOVA) was conducted with Prosody (attitudinal vs. linguistic) and Language (L1-Japanese vs. L2-English) as within-subject factors, resulting in four conditions (AE, LE, AJ, LJ). Accuracy served as the primary dependent variable, as it provides the most reliable measure of prosody comprehension in sentence-level auditory tasks. RT analyses were conducted on correct trials only, with RT defined as the interval between auditory sentence onset and the button press. Significant main effects and interactions were followed up with

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Bonferroni-corrected paired-samples *t* tests. To clarify the Prosody × Language interaction, we compared English and Japanese within each prosody type (AE vs. AJ; LE vs. LJ). The same follow-up procedure was applied to both accuracy and RT analyses. Because sentence duration differed across conditions (Appendix A), RTs are reported as a complementary measure and should not be interpreted as a pure index of task difficulty.

To address the individual differences component of the study, we examined whether participants' behavioral performance – particularly in the attitudinal prosody conditions – was associated with variation in L2 proficiency, WTC, and overall L2 exposure. To further characterize the nature of overall L2 exposure, we then conducted follow-up analyses examining the academic and nonacademic subcomponents of exposure. Pearson correlation analyses were conducted between participants' accuracy scores and each individual-difference measure after inspecting the correlations among the questionnaire variables.

Brain data

Brain data were analyzed using a conventional two-level random-effect approach in SPM12. At the first level, each participant's preprocessed functional images were modeled using a voxel-wise general linear model, with neural activity modeled at the block level. For each condition (AE, LE, AJ, LJ), boxcar regressors were created time-locked to the onset of each block with a duration of 28 s (corresponding to four trials per block) and convolved with the canonical hemodynamic response function. Four regressors corresponding to the experimental conditions (AE, LE, AJ, LJ) were included in the model. Rest periods were not explicitly modeled and served as an implicit baseline. Six movement parameters (three translations, three rotations) were included as nuisance regressors, and a 128-s high-pass filter was applied to remove low-frequency drift. For each participant, contrast images for each condition and for the relevant pairwise comparisons were generated.

At the second level, group-level analyses were performed on these contrast images. To examine the main effect of prosody type, attitudinal conditions (AE + AJ) were contrasted with linguistic conditions (LE + LJ). To examine the effects of language, L2 English (AE + LE) was contrasted with the L1 Japanese (AJ + LJ). The interaction contrast [(AE > LE) > (AJ > LJ)] was used to identify regions showing differential attitudinal–linguistic processing in English relative to Japanese.

The threshold for whole-brain analyses was at $p < .05$, family-wise error (FWE) corrected at the voxel level. For regions with a priori hypotheses, specifically the left IFG and the hippocampus, small-volume correction (SVC) was performed using anatomically defined masks, with significance set at $p < .05$ (voxel-wise, FWE-corrected) following an initial uncorrected threshold of $p < .001$. To visualize condition-specific activation, parameter estimates were extracted from peak voxels within significant clusters.

Finally, a whole-brain voxel-wise multiple regression analysis was performed to examine how individual differences in L2 experience and performance relate to neural activation in the [AE > LE] contrast. The model included four covariates: L2 listening proficiency (Oxford Placement test scores), overall L2 exposure, L2 WTC, and behavioral accuracy for attitudinal prosody. All predictors were entered simultaneously into

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Table 1. Descriptive statistics for the behavioral data

Prosody type	Accuracy (%) English *M* (*SD*)	Accuracy (%) Japanese *M* (*SD*)	RT (s) English *M* (*SD*)	RT (s) Japanese *M* (*SD*)
Attitudinal	75.54 (6.53)	84.88 (7.54)	2.88 (0.52)	2.41 (0.75)
Linguistic	80.18 (7.87)	78.39 (8.77)	2.77 (0.60)	3.14 (0.25)

Note: Values are means, with standard deviations in parentheses. Accuracy is reported as percentages. Reaction times (RTs) are reported in seconds and were calculated from correct trials only.

the SPM12 design matrix, allowing us to assess the unique contribution of each variable while controlling for the others. This analysis used an initial threshold of $p < .001$ uncorrected, with cluster-level correction at $p < .05$.

Results

Behavioral results

A 2 × 2 repeated-measures ANOVA on accuracy revealed a significant Prosody × Language interaction, $F(1, 29) = 27.35$, $p < .001$, $\eta^2 = 0.178$. Bonferroni-corrected pairwise comparisons showed that attitudinal prosody was identified less accurately in L2 English, $M = 75.54$, $SD = 6.53$, than in L1 Japanese, $M = 84.88$, $SD = 7.54$, $t(29) = -5.91$, $p < .001$, $d = -1.21$, whereas linguistic prosody did not differ between languages, $t(29) = 0.99$, $p = 1.00$, $d = 0.23$. These findings indicate a selective L2 disadvantage for attitudinal, but not linguistic prosody, consistent with previous studies showing greater difficulty in interpreting sentence-level attitudes in a second language (Chen et al., 2004; Rilliard et al., 2014; Shochi et al., 2006).

A corresponding 2 × 2 repeated-measures ANOVA on RT revealed a significant main effect of Prosody, $F(1, 29) = 13.53$, $p < .001$, $\eta^2 = 0.318$, and a significant Prosody × Language interaction, $F(1, 29) = 25.98$, $p < .001$, $\eta^2 = 0.473$, but no main effect of Language, $F(1, 29) = 0.44$, $p = .512$, $\eta^2 = 0.015$. Bonferroni-corrected pairwise comparisons showed that participants responded faster to attitudinal prosody in Japanese (AJ: $M = 2.41$ s, $SD = 0.75$) than in English (AE: $M = 2.88$ s, $SD = 0.52$), $t(29) = 3.77$, $p = .001$, $d = 0.69$. Conversely, participants responded more slowly to linguistic prosody in Japanese (LJ: $M = 3.14$ s, $SD = 0.25$) than in English (LE: $M = 2.78$ s, $SD = 0.60$), $t(29) = -3.43$, $p = .004$, $d = -0.63$. Because sentence duration differed across conditions, particularly in the Japanese linguistic condition (Appendix A), RTs are reported as a complementary measure and should not be interpreted as a pure index of task difficulty. Accuracy and RT descriptives are presented in Table 1.

Individual-difference analyses are summarized in Table 2. Pearson correlations revealed that overall L2 exposure was positively associated with accuracy in attitudinal prosody ($r = 0.40$, $p = .031$). To further characterize the nature of this relationship, we conducted follow-up analyses on the academic and nonacademic subcomponents of exposure. This analysis showed that only nonacademic exposure was significantly correlated with accuracy ($r = 0.46$, $p = .011$). No significant correlations were observed for WTC ($r = –0.13$) or L2 listening proficiency ($r = 0.25$).

As a robustness check, we additionally fitted a logistic mixed-effects model for accuracy. This analysis yielded the same significant Language × Prosody interaction

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Table 2. Correlation results for behavioral results

	Accuracy AE	Oxford test	L2 exposure	Academic exposure	Non-academic exposure	WTC
Accuracy AE	–					
Oxford test	0.254 (0.175)	–				
L2 exposure	0.395* (0.031)	0.174 (0.358)	–			
Academic exposure	−0.057 (0.767)	−0.118 (0.534)	0.407* (0.026)	–		
Non-academic exposure	0.455* (0.011)	0.239 (0.204)	0.927** (< 0.001)	0.034 (0.859)	–	
WTC	−0.125 (0.509)	0.121 (0.524)	0.322 (0.083)	0.155 (0.413)	0.288 (0.123)	–

Note: * indicates $p < .05$, ** indicates $p < .01$. Pearson's *r* values are shown. *p*-values are shown in parentheses.
WTC = Willingness to Communicate in the L2 scores.

Table 3. Comparison of prosody types

Brain areas	*x, y, z*	*t*	Cluster size
Attitudinal prosody (AE + AJ) > linguistic prosody (LE + LJ)			
Left anterior temporal lobe	−52, −2, −32	8.81	298
	−48, 6, −32	8.03	
Left lateral orbital gyrus	−40, 22, −16	6.04	167
Medial prefrontal cortex	−4, 52, 32	8.65	735
	−6, 40, 48	6.87	
Left precuneus	−6, −58, 24	7.44	166
Right precuneus	8, −52, 20	5.93	21
Right rolandic operculum	38, −8, 14	7.07	121
Linguistic prosody (LE + LJ) > attitudinal prosody (AE + AJ)			
Left inferior parietal lobule	−44, −44, 40	6.53	1,248
	−44, −48, 54	5.74	
Left precentral gyrus	−44, 4, 28	8.57	662
	−42, 2, 42	8.40	
Right inferior parietal lobule	36, −50, 40	8.35	1,001
	30, −64, 40	6.93	
Right precentral gyrus	30, 4, 50	7.15	303
	42, 8, 32	6.42	35

Note: Coordinates (*x, y, z*) are in MNI space. *t*-values reflect the peak voxel within each cluster. Cluster sizes indicate the number of voxels ($2 \times 2 \times 2$ mm^3). $p < .05$, voxel-wise FWE-corrected.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

pattern as the repeated-measures ANOVA, confirming the main behavioral findings (see Supplementary Materials).

Brain results

Comparison of prosody types (attitudinal vs. linguistic)

Across languages, attitudinal prosody [AE + AJ] elicited significantly greater activation than linguistic prosody [LE + LJ] in the left anterior temporal lobe (ATL), left lateral orbital gyrus (LOG), mPFC, bilateral precuneus, and right Rolandic operculum (Table 3, Figure 2). Conversely, linguistic prosody [LE + LJ] produced significantly greater activation than attitudinal prosody [AE + AJ] in the bilateral inferior parietal lobules and bilateral precentral gyri (Table 3, Figure 3).

Interaction effect between prosody type and language

To identify regions showing a greater attitudinal–linguistic prosody difference in L2 English than in L1 Japanese, we examined the interaction contrast [(AE > LE) > (AJ > LJ)]. This analysis revealed significant activation in the left pars triangularis of IFG and the left middle frontal gyrus (MFG) (Table 4, Figure 4). These effects indicate that the distinction between attitudinal and linguistic prosody engaged left frontal regions more strongly in English than in Japanese. The reverse interaction [(LE > AE) > (LJ > AJ)] yielded no suprathreshold clusters, even at $p < .001$, uncorrected (Table 4).

Individual differences in L2 processing

To examine how learner-related factors influenced neural processing of L2 attitudinal prosody, we conducted a voxel-wise multiple regression analysis using four predictors: overall L2 exposure, L2 proficiency, WTC in L2, and task accuracy. Among these variables, only overall L2 exposure significantly predicted neural activation, revealing two clusters: a thalamic cluster significant at the cluster level (FWE-corrected) and a left hippocampal–parahippocampal cluster significant under SVC (Table 5, Figure 5a and b). This effect remained significant after controlling for proficiency, WTC, and accuracy.

To further characterize the nature of this exposure effect, we conducted follow-up analyses examining academic and nonacademic exposure separately. Scatterplots of beta values extracted from the peak coordinates in each region showed that nonacademic exposure – reflecting naturalistic, outside-of-class contact with English – was significantly correlated with activation in both the thalamus ($r = 0.45, p = .013$) and the left hippocampus ($r = 0.51, p = .004$), whereas academic exposure showed no reliable associations (thalamus: $r = 0.31, p = .092$; hippocampus: $r = 0.15, p = .432$).

Discussion

This study investigated the neural mechanisms underlying attitudinal prosody processing in a second language (L2), revealing three major findings. First, attitudinal and linguistic prosody engaged dissociable neural networks across both L1 and L2. Second, interpreting attitudinal prosody in the L2 elicited greater activation for attitudinal compared to linguistic prosody in the left IFG and MFG, indicating increased

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

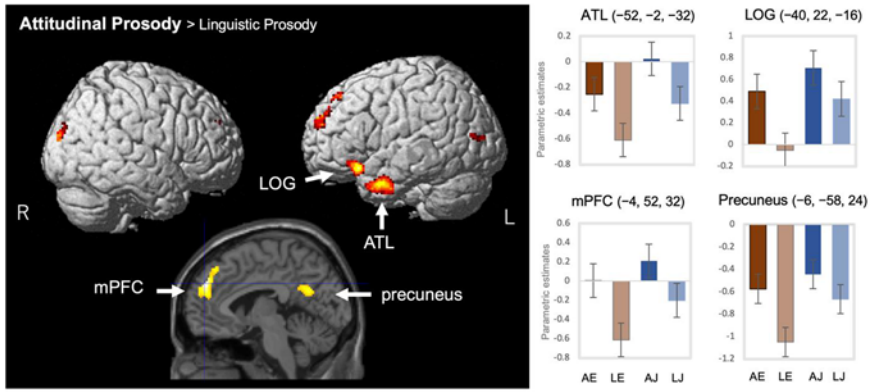


Figure 2. Main effect of attitudinal prosody (attitudinal > linguistic).
Brain regions showing greater activation for attitudinal compared to linguistic prosody across languages, including the left ATL, left LOG, mPFC, and the precuneus. Statistical threshold: $p < .05$, voxel-wise FWE-corrected. Bar plots show mean parameter estimates for each condition (AE, LE, AJ, LJ). Error bars represent the standard error of the mean. R = right; L = left.

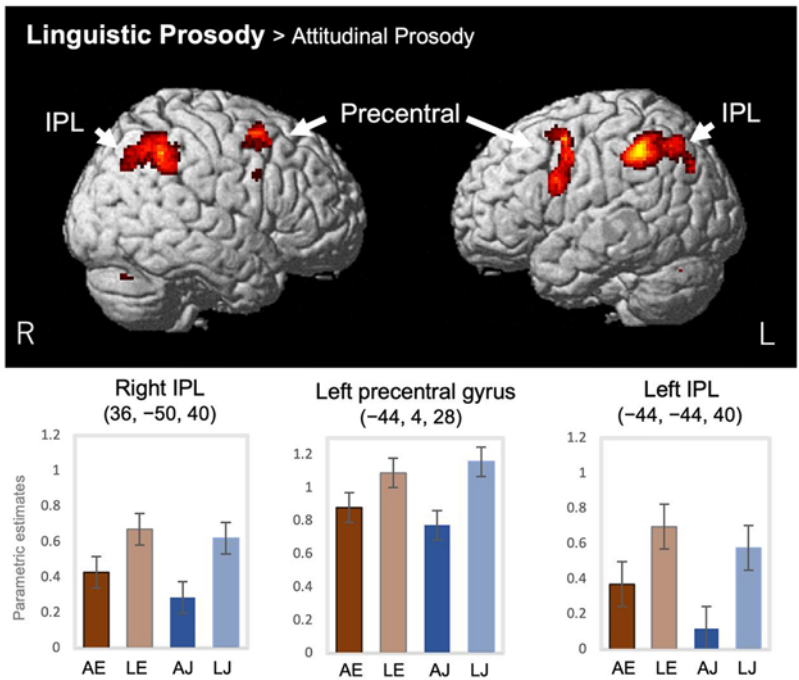


Figure 3. Main effect of linguistic prosody (linguistic > attitudinal).
Brain regions showing greater activation for linguistic compared to attitudinal prosody across languages, including the bilateral precentral gyri and inferior parietal lobules. Statistical threshold: $p < .05$, voxel-wise FWE-corrected. Bar plots show mean parameter estimates for each condition. Error bars represent the standard error of the mean. R = right; L = left.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Table 4. Interaction effect between prosody type and language

Brain areas	*x, y, z*	*t*	Cluster size
[(AE > LE) > (AJ > LJ)]			
Left middle frontal gyrus	−44, 18, 40	4.28	118
Left pars triangularis of IFG	−46, 28, 2	3.74	41
[(LE > AE) > (LJ > AJ)]			
No suprathreshold clusters			

Note: Coordinates are in MNI space. Peak *t*-values and cluster sizes (voxels) are reported. Statistical threshold: SVC at $p < .05$, voxel-wise FWE-corrected (left IFG and left MFG).

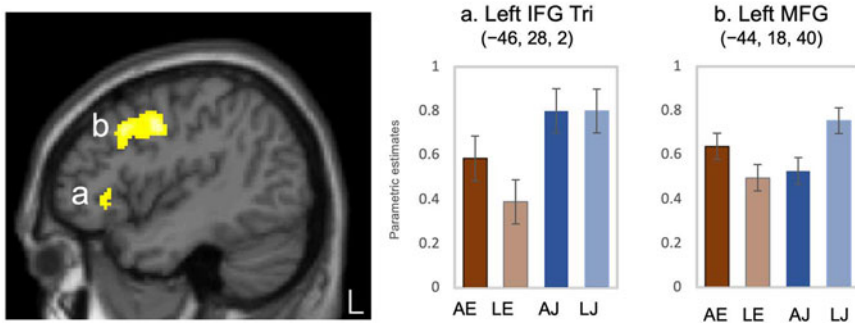


Figure 4. Interaction effect between prosody type and language.
Brain regions showing a stronger attitudinal–linguistic contrast in L2 English than in L1 Japanese, including the left pars triangularis of the IFG and the left MFG. Statistical threshold: SVC at $p < .05$, voxel-wise FWE-corrected. L = left.

Table 5. Language exposure effect on attitudinal prosody processing in L2

Brain areas	*x, y, z*	*t*	Cluster size
[AE > LE] with language exposure scores			
Right thalamus	6, −12, 16	5.76	360
Left hippocampus	−22, −24, −14	4.44	60
Left parahippocampal gyrus	−20, −26, −16	4.02	

Note: Coordinates are in MNI space. Peak *t*-values and cluster sizes (voxels) are reported. The thalamus cluster survived cluster-level FWE correction ($p < .05$; cluster-forming threshold $p < .001$), and the left hippocampal/parahippocampal cluster survived SVC at $p < .05$, voxel-wise FWE-corrected.

cognitive-linguistic demands when processing social intentions in a non-native language. Third, individual differences analyses demonstrated that L2 exposure was the only factor that significantly predicted both behavioral accuracy and neural outcomes in L2 attitudinal prosody processing. Together, these findings illustrate both shared and L2-specific neural mechanisms of attitudinal prosody and highlight the role of language exposure in L2 pragmatic development.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

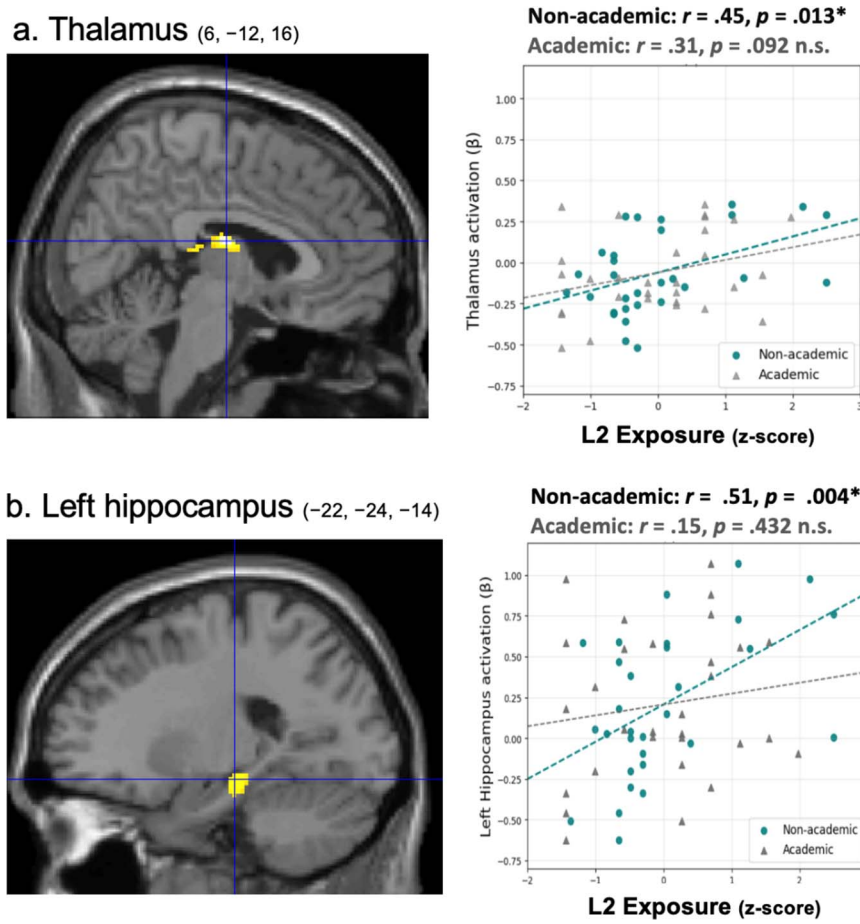


Figure 5. Associations between L2 exposure components and neural activation during L2 attitudinal prosody processing. The brain maps show clusters identified using the overall L2 exposure score. Scatterplots display correlations between neural beta values and nonacademic (teal circles) or academic (gray triangles) exposure. Non-academic exposure showed significant positive correlations with both regions (thalamus: r = 0.45, p = .013; hippocampus: r = 0.51, p = .004), whereas academic exposure showed no reliable relationships.

Attitudinal prosody processing in the L2

A central aim of this study was to determine why attitudinal prosody is more difficult to interpret in a second language and how this difficulty is reflected in behavioral and neural patterns. The accuracy data provided clear evidence of an L2 disadvantage: participants were significantly less accurate at identifying attitudinal prosody in English than in Japanese, whereas accuracy for linguistic prosody did not differ across languages. These behavioral findings suggest that learners experience greater ambiguity and reduced reliability when interpreting attitudinal prosody in their L2, consistent

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

with previous research demonstrating that sentence-level attitudes are more challenging to identify in a second language (Chen et al., 2004; Rilliard et al., 2014; Shochi et al., 2006).

The neuroimaging results provide converging evidence for this behavioral difficulty. The interaction contrast [(AE > LE) > (AJ > LJ)] revealed increased activation in the left IFG (pars triangularis) during L2 attitudinal prosody processing. This region is a core hub for higher-level sentence processing, supporting controlled evaluation, interpretation, and selection of contextually relevant information (Friederici, 2012). Causal neurostimulation work further demonstrates that intonation shapes semantic and syntactic processing within the IFG (van der Burght et al., 2023), suggesting that this region provides a functional interface through which prosodic cues influence sentence-level interpretation. Importantly, the IFG was engaged across both languages and prosody types in the present study, reflecting its general role in supporting the use of prosodic information during sentence processing. What distinguishes the L2, however, is the degree of modulation: activation in the pars triangularis differentiated more sharply between attitudinal and linguistic prosody in L2 English, whereas the difference was minimal in L1 Japanese. This pattern is consistent with Oh et al. (2019), who reported stronger modulation of the pars triangularis in L2 contexts where listeners engage more controlled interpretive processing. Together, these results suggest that L2 listeners rely more heavily on left IFG-mediated controlled processing when interpreting attitudinal prosody, not because the region is uniquely required in the L2, but because its contribution is selectively enhanced for attitudinal, relative to linguistic, prosody in the non-native language.

We also observed increased activity in the left MFG for L2 attitudinal prosody. The MFG is frequently implicated in demanding cognitive tasks involving working memory and executive control (Duncan, 2010; Fedorenko et al., 2013). Its involvement here may reflect the need to maintain task-relevant linguistic and prosodic information while evaluating the speaker's intended social meaning. Notably, the direction of activity in the MFG differed by language: Japanese attitudinal prosody elicited *less* MFG activation than linguistic prosody, a pattern that may be consistent with relatively reduced executive demands in the L1. In contrast, English attitudinal prosody elicited *greater* MFG activation than linguistic prosody, which may reflect greater engagement of domain-general executive resources in the L2.

Taken together, these findings show that attitudinal prosody in the L2 imposes greater cognitive demands than in the L1, requiring greater engagement of controlled interpretive processing (IFG) and executive-control (MFG) mechanisms. Because learners are less familiar with the pragmatic and prosodic conventions of the L2, attitudinal cues become more ambiguous and less reliably mapped onto their intended interpersonal meanings.

Main effects of prosody type: Neural dissociation of attitudinal and linguistic prosody

Across both languages, attitudinal prosody elicited robust activation in ToM–related regions associated with mentalizing and speaker-intention inference (Hellbernd & Sammler, 2018), whereas linguistic prosody recruited dorsal fronto-parietal regions

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

associated with phonological working memory and articulatory planning (Baldo & Dronkers, 2006; Meister et al., 2007). Despite the additional processing demands observed for L2 attitudinal prosody, main effects revealed a stable neural dissociation between attitudinal and linguistic prosody.

Specifically, the main effect of attitudinal prosody, shown in the contrast [AE + AJ] > [LE + LJ], recruited the mPFC, precuneus, left ATL, and left LOG, regions implicated in mentalizing and speaker-intention inference. This pattern aligns with extensive L1 research showing that inferring a speaker's interpersonal stance engages socio-cognitive mechanisms rather than core language regions (Enrici et al., 2019; Hellbernd & Sammler, 2018; Tettamanti et al., 2017). Our findings extend this evidence to the L2, showing that learners rely on the same ToM-related network when interpreting attitudinal meanings conveyed through prosody. These results suggest that listeners engage inferential mechanisms to process attitudinal prosody across both L1 and L2 contexts, highlighting the inherently social nature of attitudinal prosody.

In contrast, the main effect of linguistic prosody, shown in the contrast [LE + LJ] > [AE + AJ], elicited greater activation in the bilateral inferior parietal lobules (IPL) and precentral gyri, regions associated with phonological working memory and articulatory planning (Baldo & Dronkers, 2006; Meister et al., 2007). These findings suggest that sentence-level linguistic prosody depends more heavily on phonological and articulatory planning resources, consistent with its role in conveying structural information such as sentence modality and focus across both L1 and L2.

Taken together, these findings demonstrate that attitudinal and linguistic prosody recruit distinct neural systems that remain consistent across L1 and L2. While L2 learners experience additional cognitive load for attitudinal prosody, as discussed in attitudinal prosody processing in the L2 section, the basic organization of prosody processing appears broadly preserved across languages, with attitudinal prosody relying more on socio-cognitive mechanisms and linguistic prosody relying more on phonological and articulatory planning mechanisms.

Effect of L2 exposure on attitudinal prosody processing

The individual differences analyses revealed a clear and consistent pattern: L2 exposure was the only factor that reliably predicted both behavioral and neural outcomes during L2 attitudinal prosody processing. Learners who reported higher levels of informal English use, particularly listening and speaking outside the classroom, showed higher accuracy in identifying attitudinal prosody and greater activation in the left hippocampus and thalamus during attitudinal prosody processing. These results indicate that naturalistic L2 contact plays a key role in supporting the interpretation of prosodic cues and their mapping onto intended interpersonal meanings.

The regions that correlated with L2 exposure centered on the hippocampus, with the cluster extending into the parahippocampal gyrus. Rather than interpreting these findings in terms of a traditional declarative–procedural distinction, a more contemporary interpretation highlights the role of the hippocampal–parahippocampal system in binding episodic, auditory, and contextual information during naturalistic learning. This system enables learners to integrate past experiences, including

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

prosodic patterns, speaker-specific cues, affective tone, and social–contextual regularities, into predictive representations that facilitate the interpretation of unfolding speech (Davis & Gaskell, 2009; Ranganath & Ritchey, 2012). The extension of the activation into the parahippocampal gyrus further suggests engagement of scene- and context-construction processes that allow learners to situate incoming auditory input within familiar experiential frameworks (Aminoff et al., 2013; Epstein & Kanwisher, 1998; Ranganath & Ritchey, 2012). From this perspective, exposure provides the episodic and contextual grounding necessary for real-time interpretation of pragmatic prosody.

The involvement of the thalamus provides further insight into how exposure is associated with L2 speech processing. The thalamus plays a key role in sensory gating, attentional selection, and the routing of auditory information across cortical networks (Bulut & Hagoort, 2024; Halassa & Kastner, 2017). In the present study, thalamic activation may reflect experience-dependent modulation of auditory attention to socially relevant prosodic cues. This interpretation is consistent with previous findings showing structural changes in the thalamus in individuals with extensive L2 immersion (DeLuca et al., 2019; Pliatsikas et al., 2017), suggesting that repeated L2 use strengthens subcortical mechanisms that support speech processing.

Behaviorally, distinguishing between types of exposure revealed that nonacademic exposure – informal, socially embedded encounters with English – was more strongly associated with attitudinal prosody accuracy than academic exposure. While the lower variance of academic exposure in the present sample may have reduced the likelihood of detecting a significant correlation, this possibility should be considered when interpreting the observed difference between academic and nonacademic exposure, as it may partly reflect differences in variability. At the same time, the observed association between nonacademic exposure and attitudinal prosody performance is consistent with claims that socially grounded L2 use provides richer pragmatic and affective cues than formal instruction alone (Jeong et al., 2021; Li & Jeong, 2020; Liu et al., 2024). Social interactions expose learners to a broader range of communicative intentions, interpersonal stances, and prosodic contours, all of which are encoded as episodic, multimodal experiences that can later be recruited during speech processing. The hippocampal–parahippocampal–thalamic involvement identified here provides a mechanistic account of why such informal experiences matter: they offer episodic and contextual traces that support the extraction and prediction of pragmatic prosodic meaning. Proficiency, WTC, and behavioral accuracy did not show significant associations with neural activation, highlighting that exposure was the primary factor linked to learners' attitudinal prosody development.

Finally, the presence of these exposure effects in the Japanese EFL context is noteworthy. Opportunities for immersive, socially sustained L2 use are limited, and learners typically encounter English primarily in instructed settings. Yet the brain still tracked individual differences in exposure, suggesting that even modest, episodic encounters with English – through media, brief social interactions, or incidental listening – are sufficient to be reflected in the neural systems supporting pragmatic prosody perception. These findings highlight both the sensitivity and adaptability of the L2 learning system and underscore the importance of providing learners with richer opportunities for naturalistic, socially grounded L2 experience.

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Limitations, pedagogical implications, and future research

This study has several limitations that should be considered when interpreting the results. First, the explicit judgment task may have encouraged top-down strategies, potentially altering how learners relate prosodic to semantic and pragmatic interpretations. In addition, prosody type was not independently manipulated from task demands, as linguistic and attitudinal prosody were associated with different categorization requirements and were recorded separately to ensure natural prosodic realization. The present design was intended to compare different task-dependent modes of prosody processing, and the observed differences therefore reflect the combined influence of prosodic features and task-related demands. Although lexical content and syntactic structure were held constant across conditions, the results cannot be attributed to prosodic features alone. More implicit paradigms (e.g. passive listening, mismatch responses, or incidental attitude recognition) would help clarify how attitudinal prosody is processed when attentional demands are minimized. Second, the block design limited item-level analyses, preventing direct comparison of attitudinal subtypes (e.g. confidence, friendliness, doubt) or fine-grained variability across items. Event-related designs would enable more detailed mapping of specific attitudinal categories and their neural correlates. Third, this study focused on a single L1–L2 pairing. Although this provides a clear test case, attitudinal prosody varies cross-linguistically, and future research should examine additional language combinations to assess generalizability.

Despite these considerations, the findings yield several insights into the neural and experiential foundations of L2 attitudinal prosody. First, attitudinal prosody in both L1 and L2 engaged socio-cognitive networks associated with mentalizing, highlighting shared mechanisms for inferring speakers' intentions. Second, L2 attitudinal prosody elicited additional activation in the left IFG and MFG, reflecting increased executive and control-related demands required to integrate unfamiliar pragmatic cues. Third, informal L2 exposure emerged as the strongest predictor of both behavioral accuracy and neural activation, consistent with the idea that socially grounded, real-world encounters provide learners with richer pragmatic, affective, and prosodic information than formal instruction alone. This aligns with emerging perspectives on social learning, which emphasize that episodic, multimodal encounters – in conversation, media, or everyday listening – serve as critical sources of pragmatic input that strengthen experience-based processing mechanisms.

Because attitudinal prosody is rarely addressed explicitly in L2 curricula, incorporating authentic spoken materials – such as films, interviews, podcasts, or conversational clips – may provide learners with more meaningful access to interpersonal intentions. Activities that guide learners to notice and interpret prosodic cues expressing attitudes (e.g. confidence, friendliness, doubt) through perception-based or awareness-raising tasks may support pragmatic development. Interactive communicative activities that require interpreting or expressing attitudes in context may further reinforce sensitivity to prosodic cues. These practices are directly consistent with the present finding that socially embedded, naturalistic experience plays a central role in L2 attitudinal prosody acquisition.

Future research should examine how different forms of L2 exposure – such as conversational practice, multimodal input, or targeted instruction – shape learners'

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

pragmatic prosody skills over time, and how instructional interventions can support the acquisition of socially meaningful prosodic cues. Longitudinal and training studies may clarify how exposure-driven changes in both behavior and neural activation unfold during L2 development, offering new insights into how learners come to interpret interpersonal intentions in multilingual communication.

Conclusion

This study provides the first evidence clarifying the neural mechanisms that support attitudinal prosody in an L2. The findings demonstrate that interpreting attitudinal prosody engages universal socio-cognitive mechanisms as well as additional linguistic and executive processes required when listeners interpret pragmatic cues in a less familiar language. While Theory of Mind–related networks facilitated intention inference across languages, L2 attitudinal prosody was strongly associated with learners' real-world experience, particularly informal, socially grounded exposure. The observed involvement of the hippocampal–parahippocampal–thalamic system further suggests that such exposure supplies the episodic, contextual, and auditory grounding through which learners construct predictions about speakers' attitudes.

By integrating behavioral and neural evidence, this study advances theoretical models of pragmatic prosody acquisition and highlights the central role of naturalistic, socially meaningful input in L2 pragmatic development. These findings underscore the importance of providing learners with rich opportunities to encounter attitudinal cues in authentic communicative contexts and contribute to broader discussions on how experiential factors shape the neurocognitive foundations of pragmatic development in a second language.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0267190526100257.

References

Allen, D. (1992). *The Oxford placement test*. Oxford University Press.

Aminoff, E. M., Kveraga, K., & Bar, M. (2013). The role of the parahippocampal cortex in cognition. *Trends in Cognitive Sciences*, *17*(8), 379–390. https://doi.org/10.1016/j.tics.2013.06.009

Baldo, J. V., & Dronkers, N. F. (2006). The role of inferior parietal and inferior frontal cortex in working memory. *Neuropsychology*, *20*(5), 529–538. https://doi.org/10.1037/0894-4105.20.5.529

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*(1), 5–17. https://doi.org/10.1023/a:1005653411471

Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience*, *9*(9), 1395–1403. https://doi.org/10.1093/scan/nst124

Bulut, T., & Hagoort, P. (2024). Contributions of the left and right thalami to language: A meta-analytic approach. *Brain Structure and Function*, *229*(9), 2149–2166. https://doi.org/10.1007/s00429-024-02795-3

Chen, A. (2009). Perception of paralinguistic intonational meaning in a second language. *Language Learning*, *59*(2), 367–409. https://doi.org/10.1111/j.1467-9922.2009.00510.x

Chen, A., Gussenhoven, C., & Rietveld, T. (2004). Language-specificity in the perception of paralinguistic intonational meaning. *Language and Speech*, *47*(4), 311–349. https://doi.org/10.1177/00238309040470040101

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1536), 3773–3800. https://doi.org/10.1098/rstb.2009.0111

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, *116*(15), 7565–7574. https://doi.org/10.1073/pnas.1811513116

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172–179. https://doi.org/10.1016/j.tics.2010.01.004

Elahi Shirvan, M., Khajavy, G. H., MacIntyre, P. D., & Taherian, T. (2019). A meta-analysis of L2 willingness to communicate and its three high-evidence correlates. *Journal of Psycholinguistic Research*, (48), 1241–1267. https://doi.org/10.1007/s10936-019-09656-9

Enrici, I., Bara, B. G., & Adenzato, M. (2019). Theory of Mind, pragmatics and the brain: Converging evidence for the role of intention processing as a core feature of human communication. *Pragmatics & Cognition*, *26*, 5–38. https://doi.org/10.1075/pc.19010.enr

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601. https://doi.org/10.1038/33402

Esteve-Gibert, N., & Prieto, P. (2013). Prosody signals the emergence of intentional communication in the first year of life: Evidence from Catalan-babbling infants. *Journal of Child Language*, *40*(5), 919–944. https://doi.org/10.1017/S0305000912000359

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, *110*(41), 16616–16621. https://doi.org/10.1073/pnas.1315235110

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84. https://doi.org/10.1016/S1364-6613(00)01839-8

Friederici, A. D. (2006). Broca's area and the ventral premotor cortex in language: Functional differentiation and specificity. *Cortex*, *42*(4), 472–475. https://doi.org/10.1016/s0010-9452(08)70380-0

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–268. https://doi.org/10.1016/j.tics.2012.04.001

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., Li, X., & Lowe, M. (2007). Neural basis of first and second language processing of sentence-level linguistic prosody. *Human Brain Mapping*, *28*(2), 94–108. https://doi.org/10.1002/hbm.20255

Halassa, M. M., & Kastner, S. (2017). Thalamic functions in distributed cognitive control. *Nature Neuroscience*, *20*(12), 1669–1679. https://doi.org/10.1038/s41593-017-0020-1

Hellbernd, N., & Sammler, D. (2016). Prosody conveys speaker's intentions: Acoustic cues for speech act perception. *Journal of Memory and Language*, *88*, 70–86. https://doi.org/10.1016/j.jml.2016.01.001

Hellbernd, N., & Sammler, D. (2018). Neural bases of social communicative intentions in speech. *Social Cognitive and Affective Neuroscience*, *13*(6), 604–615. https://doi.org/10.1093/scan/nsy034

Hupp, J. M., Jungers, M. K., Hinerman, C. M., & Porter, B. L. (2021). Cup! Cup? Cup: Comprehension of intentional prosody in adults and children. *Cognitive Development*, *57*, 100971. https://doi.org/10.1016/j.cogdev.2020.100971

Jeong, H., Li, P., Suzuki, W., Sugiura, M., & Kawashima, R. (2021). Neural mechanisms of language learning from social contexts. *Brain and Language*, *212*, 104874. https://doi.org/10.1016/j.bandl.2020.104874

Jeong, H., Sugiura, M., Sassa, Y., Haji, T., Usui, N., Taira, M., Horie, K., Sato, S., & Kawashima, R. (2007). Effect of syntactic similarity on cortical activation during second language processing: A comparison of English and Japanese among native Korean trilinguals. *Human Brain Mapping*, *28*(3), 194–204. https://doi.org/10.1002/hbm.20269

Kang, O., & Kermad, A. (2019). Prosody in L2 pragmatics research. In N. Taguchi (Ed), *The Routledge handbook of second language acquisition and pragmatics* (pp. 78–92). Routledge. https://doi.org/10.4324/9781351164085-6

Kang, O., Kermad, A., & Taguchi, N. (2021). The interplay of proficiency and study abroad experience on the prosody of L2 speech acts. *Journal of Second Language Pronunciation*, *7*(3), 343–369. https://doi.org/10.1075/jslp.20024.kan

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Li, P., & Jeong, H. (2020). The social brain of language: Grounding second language learning in social interaction. *Npj Science of Learning*, *5*(1), 8. https://doi.org/10.1038/s41539-020-0068-7

Liu, C., Jeong, H., Cui, H., Dewaele, J.-M., Okamoto, K., Suzuki, Y., & Sugiura, M. (2024). Effects of social interactions on the neural representation of emotional words in late bilinguals. *Language, Cognition and Neuroscience*, *39*(3), 383–399. https://doi.org/10.1080/23273798.2024.2307630

Liu, H., & Cao, F. (2016). L1 and L2 processing in the bilingual brain: A meta-analysis of neuroimaging studies. *Brain and Language*, *159*, 60–73. https://doi.org/10.1016/j.bandl.2016.05.013

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, *17*(19), 1692–1696. https://doi.org/10.1016/j.cub.2007.08.064

Mitchell, R. L., & Ross, E. D. (2013). Attitudinal prosody: What we know and directions for future study. *Neuroscience & Biobehavioral Reviews*, *37*(3), 471–479. https://doi.org/10.1016/j.neubiorev.2013.01.027

Oh, T. M., Graham, S., Ng, P., Yeh, I. B., Chan, B. P., & Edwards, A. M. (2019). Age and proficiency in the bilingual brain revisited: Activation patterns across different L2-learner types. *Frontiers in Communication*, *4*, Article 39. https://doi.org/10.3389/fcomm.2019.00039

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*(2), 202–206. https://doi.org/10.1016/j.conb.2005.03.007

Pliatsikas, C., DeLuca, V., Moschopoulou, E., & Saddy, J. D. (2017). Immersive bilingualism reshapes the core of the brain. *Brain Structure and Function*, *222*(4), 1785–1795. https://doi.org/10.1007/s00429-016-1307-9

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, *13*(10), 713–726. https://doi.org/10.1038/nrn3338

Rice, K., & Redcay, E. (2016). Interaction matters: A perceived social partner alters the neural processing of human speech. *NeuroImage*, *129*, 480–488. https://doi.org/10.1016/j.neuroimage.2015.11.041

Rilliard, A., Erickson, D., Shochi, T., & de Moraes, J. A. (2014). US English attitudinal prosody performances in L1 and L2 speakers. *Proceedings of the 7th International Conference on Speech Prosody*. ISCA, 895–899. https://doi.org/10.21437/speechprosody.2014-168

Shochi, T., Aubergé, V., & Rilliard, A. (2006). How prosodic attitudes can be false friends: Japanese vs. French social affects. *Speech Prosody*. ISCA, 692–696. https://doi.org/10.21437/speechprosody.2006-156.

Shochi, T., Brousse, A., Guerry, M., Erickson, D., & Rilliard, A. (2016). Learning effect of social affective prosody in Japanese by French learners. In *Proceedings of Speech Prosody 2016* (pp. 970–973). ISCA. https://doi.org/10.21437/SpeechProsody.2016-199

Shochi, T., Rilliard, A., Aubergé, V., & Erickson, D. (2009). Intercultural perception of English, French and Japanese social affective prosody. In S. Hancil (Ed), *The role of prosody in affective speech* (Vol. 97, pp. 31–60). Peter Lang Publisher.

Sinagra, C., & Wiener, S. (2022). The perception of intonational and emotional speech prosody produced with and without a face mask: An exploratory individual differences study. *Cognitive Research: Principles and Implications*, *7*(1), 89. https://doi.org/10.1186/s41235-022-00439-w

Tettamanti, M., Vaghi, M. M., Bara, B. G., Cappa, S. F., Enrici, I., & Adenzato, M. (2017). Effective connectivity gateways to the Theory of Mind network in processing communicative intention. *NeuroImage*, *155*, 169–176. https://doi.org/10.1016/j.neuroimage.2017.04.050

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, *28*(1), 1–30. https://doi.org/10.1017/S0272263106060013

van der Burght, C. L., Goucha, T., Friederici, A. D., Kreitewolf, J., & Hartwigsen, G. (2019). Intonation guides sentence processing in the left inferior frontal gyrus. *Cortex*, *117*, 122–134. https://doi.org/10.1016/j.cortex.2019.02.011

van der Burght, C. L., Numssen, O., Schlaak, B., Goucha, T., & Hartwigsen, G. (2023). Differential contributions of inferior frontal gyrus subregions to sentence processing guided by intonation. *Human Brain Mapping*, *44*(2), 585–598. https://doi.org/10.1002/hbm.26086

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The Autism-Spectrum Quotient (AQ) in Japan: A cross-cultural comparison. *Journal of Autism and Developmental Disorders*, *36*(2), 263–270. https://doi.org/10.1007/s10803-005-0061-2

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press

Wichmann, A. (2000). The attitudinal effects of prosody and how they relate to emotion. *Proceedings of the ISCA Tutorial and Research Workshop on Speech and Emotion*. ISCA, 143–148.
Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *Modern Language Journal*, *86*(1), 54–66. https://doi.org/10.1111/1540-4781.00136
Yokogawa, H. (2009). *Nihonjin eigogakushūsha no eitango shinmitsudo onseihen* [English word familiarity of Japanese Learners of English, oral edition: Second language database for education and research]. Kurosio Publishers.

Appendix A: Acoustic properties of the stimuli

Condition	Duration (s) *M* (*SD*)	Mean intensity (dB)	Mean F0 (Hz) *M* (*SD*)	F0 range (Hz) *M* (*SD*)
AE	2.41 (0.58)	69.6	177.5 (60.37)	147.7 (97.54)
AJ	2.78 (0.63)	70.6	160.1 (51.25)	117.2 (52.19)
LE	2.47 (0.47)	71.8	179.2 (65.04)	161.1 (106.50)
LJ	3.23 (0.66)	70.6	172.9 (47.88)	162.1 (65.01)

Note: Duration and pitch measures (mean F0 and F0 range) are reported with standard deviations in parentheses. Intensity SD values are not reported because occasional silent intervals produced outlier values, but mean intensity was comparable across conditions.

Cite this article: Flood, V-. A, Dardon, D. E., Sugiura, M., & Jeong, H. (2026). Neural mechanisms of attitudinal prosody in a second language: The impact of informal exposure. *Annual Review of Applied Linguistics*, *46*, 199–222. https://doi.org/10.1017/S0267190526100257

https://doi.org/10.1017/S0267190526100257 Published online by Cambridge University Press