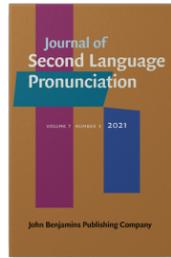


Short title: Customer Service Agents' Performance

To be published in *Journal of Second Language Pronunciation* (John Benjamins)



Listener Perceptions of Customer Service Agents' Performance: Do Comprehensibility, Accentedness, and Fluency Matter?

Aki Tsunemoto, Mark McAndrews, Pavel Trofimovich, & Eric Friginal

This study investigated listener-based assessment of the job performance of second language (L2) speakers employed as customer service agents in outsourced foreign-based call centers, focusing on agents' job performance as a function of the comprehensibility, fluency, and accentedness of their speech. Using Amazon's Mechanical Turk crowdsourcing platform, 116 native English-speaking listeners evaluated two-minute recordings of actual customer service conversations featuring 18 Filipino agents, assessing them for three global speech dimensions (comprehensibility, accentedness, and fluency) and three performance indicators, including agents' confidence, competence, and listeners' interest in future communication with agents (a measure capturing customer patronage). Comprehensibility and fluency consistently predicted how the listeners assessed the agents on all job performance scales, and accentedness was additionally associated with how strongly the listeners wished to communicate with the agents. Findings generally highlight the importance of fluent and comprehensible L2 speech in workplace settings.

Keywords: comprehensibility, accentedness, fluency, perception, customer service, foreign-based call center

CUSTOMER SERVICE AGENTS' PERFORMANCE

Acknowledgments

We would like to thank a US-owned call center company located in Manila, Philippines for providing data for the study. Also, we are grateful to Tamanna Mostafa for her help with material preparation, Thao Nguyen Nina Le for her help with data analysis, and to the Editor and anonymous reviewers for their helpful feedback on earlier drafts of this article.

Funding statement

This research was supported by a Social Sciences and Humanities Research Council of Canada (SSHRC) grants awarded to Pavel Trofimovich (435-2016-1406).

Listener Perceptions of Customer Service Agents' Performance: Do Comprehensibility, Accentedness, and Fluency Matter?

1. Introduction

Since the late 1990s, businesses in English-speaking Western countries have increasingly outsourced their telephone-based customer service hubs to foreign locations, including India and the Philippines. Customer service agents working in these outsourced call centers use English to perform various tasks, including answering product enquiries, resolving administrative issues, and providing technical support for products and services in the retail, financial, and information technology industries (Forey & Lockwood, 2010; Friginal, 2009; Krishnamurthy, 2018). Because the spoken English proficiency of customer service agents is considered the “core commodity for business success” (Lockwood, 2015, p. 1), agents’ performance has been investigated with respect to various characteristics of their speech, such as discourse markers and pragmatic routines (e.g., Friginal & Cullom, 2014; Hood & Forey, 2008), and their job performance, measured through performance reviews (e.g., Friginal, 2007) or breakdowns in communication (e.g., Forey & Lockwood, 2007). However, to the best of our knowledge, no research has to date accounted for the perceptions of customers themselves, examining how customer evaluations of agents’ performance relate to such global dimensions of agents’ speech as comprehensibility, accentedness, and fluency, which was the goal of this study.

2. Background Literature

Much of the previous research into the spoken performance of outsourced customer service agents has sought to describe their use of pragmatic and sociolinguistic resources. Hood and Forey (2008), for example, investigated the linguistic means through which Filipino agents managed the interpersonal intensity of problematic calls, noting agents’ frequent use of

CUSTOMER SERVICE AGENTS' PERFORMANCE

concessive contractors (e.g., *just, already, actually*) to narrow the scope of interactions to specific issues. Hultgren (2011) documented a rapport-building speech style used by call center workers across Denmark, Hong Kong, the Philippines, and the UK. Some notable features of this style, which was used across interactions in four locations and six language varieties, included signals of active listening (e.g., *aha, I see, I know*), signs of understanding (e.g., through paraphrasing and summarizing), and displays of empathy (e.g., by engaging in small talk). Archer and Jagodziński (2015) examined strategies deployed by Polish agents (e.g., expressing empathy, apportioning blame) as they dealt with impoliteness and verbal aggression from customers. Friginal and Cullom (2014) focused on Filipino agents' use of different types of refusals, including indirect refusals through reference to company policies, and their production of pseudo-cleft constructions (e.g., *What I can do is...*) to offer alternative solutions. In a sociolinguistic study, Cowie and Murty (2010) reported that Indian agents were more likely to use the American English postvocalic /r/ (a salient feature of North American English) when speaking with an American interlocutor than in extemporaneous speech among co-workers, using this feature to align their speech to that of their interlocutor.

Despite their general interest in customer service agents' spoken language, researchers have not focused extensively on the relationships between the specific features of agents' speech and their overall job performance. Through an analysis of recorded customer service calls, Forey and Lockwood (2007) investigated the causes of communication breakdowns in interactions between Filipino agents and American customers. While some breakdowns were caused by inaccurate production of intonation, syllable stress, and segmentals, more problematic for overall communicative success were mismatches between the expectations of American customers and the agents' pragmatic and discourse choices (e.g., not acknowledging customers' frustrations). In

CUSTOMER SERVICE AGENTS' PERFORMANCE

another study, Friginal (2007) examined relationships between Filipino agents' scores on an employer-administered English proficiency test and their job performance evaluations. The proficiency test included sub-scores for pronunciation of segmentals and suprasegmentals (including intonation, pitch, volume, and speech rate), vocabulary and grammar, and listening comprehension, while job performance was assessed by company-internal quality monitors, who assigned scores for agents' accuracy of customer support and transaction handling. In a regression model with the overall job performance score as the outcome variable and English proficiency sub-scores as predictors, only the suprasegmental sub-score reached statistical significance, with 16% of variance explained. In follow-up research, Friginal (2009) showed that agents with high performance evaluation scores tended to produce fewer silent and filled pauses than agents with low scores.

Although these previous findings highlight the importance of specific pronunciation features (and especially suprasegmental and fluency phenomena) to evaluations of agents' job performance, little is known about the customer perspective, particularly in relation to such listener-based global dimensions of second language (L2) speech as comprehensibility, accentedness, and fluency. When listeners evaluate comprehensibility, which denotes how easily listeners understand a speaker (Derwing & Munro, 1997), they primarily rely on various linguistic dimensions in L2 speech, including phonology, lexis, grammar, fluency, and discourse (e.g., Isaacs & Trofimovich, 2012). Listener perceptions of accentedness capture how closely a speaker approximates the expected language variety (Munro & Derwing, 1995), which is often spoken by the majority of native speakers in a given context. Compared to comprehensibility, accentedness judgments are narrower in scope and are mostly determined by a speaker's segmental and suprasegmental accuracy (Hayes-Harb & Hacking, 2015; Saito et al., 2017).

CUSTOMER SERVICE AGENTS' PERFORMANCE

Finally, a listener-rated measure of fluency, which typically captures various aspects of utterance flow, can be largely explained through temporal and repair characteristics of speech, including hesitations, dysfluencies, hedging, and articulation speed (Bosker et al., 2013; Kahng, 2018). These global listener-assessed dimensions, which have typically been operationalized in terms of listeners' intuitive judgements through Likert-type scales (e.g., Munro & Derwing, 1995) or continuous sliding scales (e.g., Saito et al., 2017), might capture some key elements of agents' speech that would be of relevance to the evaluations of their performance by customers.¹

Prior research has revealed a great deal of evidence suggesting that global listener-assessed dimensions of L2 speech determine the severity with which listeners evaluate speakers for many personal and professional characteristics. With respect to comprehensibility, for example, when American listeners evaluated L2 English speech recorded by Chinese and Indian speakers, listeners downgraded the speakers whom they perceived as hard to understand, evaluating them as less intelligent and successful and ascribing negative emotions of annoyance and irritation to them (Dragojevic et al., 2017). In an educational setting, after watching an instructional video narrated by an instructor who was rated as hard to understand, university students expressed negative attitudes toward coursework, downgraded the instructor in their assessments, and perceived the video content as more difficult, even though their actual comprehension of the video content was not impaired (Sanchez & Khan, 2016). In a workplace context, when native-speaking employees from two Canadian engineering companies with a large number of workers from various linguistic backgrounds listened to two L2 speakers from their own background, they preferred to interact with the more comprehensible speaker 88% of the time (Derwing & Munro, 2009), implying that workplace communication might be driven by how easy or difficult it was for workers to understand their colleagues. In the customer service

CUSTOMER SERVICE AGENTS' PERFORMANCE

industry, when large-scale surveys of customer experience have been undertaken on behalf of governments and private businesses, such as those conducted by the CFI Group (a founding partner of the American Customer Satisfaction Index), ease of understanding has been consistently cited as a major obstacle to agent–customer communication (e.g., CFI Group, 2020).

When it comes to accentedness, listeners find more accented speakers to be less intelligent, competent, successful, likeable, and even less physically attractive, compared to speakers who are judged to be less accented (Campbell-Kibler, 2007; Hosoda et al., 2007). Accented workers who speak the majority language of the community as their second or additional language are often underemployed (Krahn et al., 2000), discriminated against (Davila et al., 1993), and face isolation in their communities (Gluszek et al., 2011). Accented speech is also a source of bias in the interview process (Almeida et al., 2015), especially in telephone conversations (Cocchiara et al., 2016), and accented speech is often associated with low- rather than high-prestige jobs (Brennan & Brennan, 1981). In the customer service industry, recruiters of call center agents tend to hire employees whose speech is least accented rather than those who might fit job requirements more closely (Lockwood, 2012).

Finally, with respect to fluency, speakers whose utterances include a high proportion of such features as hesitations, dysfluencies, and hedging tend to be downgraded in listener assessments, which include competence, trustworthiness, attractiveness, and intelligence (Hosman, 2015). For instance, when evaluating recordings of applicants for an entry-level banking job, listeners (including human resource professionals) rated the applicants as less competent and employable when their speech contained many hedges (e.g., *kind of, sort of, I think*) and hesitations (e.g., *uh, well, you know*) than when few such phenomena were present (Parton et al., 2002). Similarly, when listeners heard court testimony containing many hedges

CUSTOMER SERVICE AGENTS' PERFORMANCE

and hesitations, they perceived the witness as less attractive and credible and were ultimately less accepting of the testimony, compared to when the same testimony was produced with few hedges and hesitations (Erickson et al., 1978). Listeners also tended to ascribe less knowledge to speakers whose speech contains filled pauses (Brennan & Williams, 1995) and to downgrade speakers in competence and to rate their message as inferior in quality when their speech includes disfluencies (Blankenship & Holtgraves, 2005; Hosman & Siltanen, 2011). Thus, a speaking style that includes disfluencies might be associated with lower listener evaluations of speakers.

3. The Present Study

Across various sectors of the retail, financial, and information technology industries, millions of customers rely on outsourced telephone-based customer service to address their daily personal or professional needs (Friginal, 2009; Krishnamurthy, 2018). Recognizing the importance of communication to customer satisfaction in agent–customer transactions (CFI Group, 2020; Chicu et al., 2019), researchers have investigated the relationships between agents' linguistic performance and their overall job performance, for instance, as measured through company-internal performance reviews or analysis of conversation transcripts (Forey & Lockwood, 2007; Friginal, 2007). However, an important gap in the current knowledge base is the lack of customer- or client-based perspective on the comprehensibility, accentedness, and fluency of call center agents' speech. Considering that listener-based assessments of L2 speakers across various settings, including professional and workplace contexts, are determined by how comprehensible, accented, and fluent they sound to the listener, it is reasonable to assume that potential customers or clients of call center agents might evaluate agents' job performance based on global dimensions of agents' speech.

CUSTOMER SERVICE AGENTS' PERFORMANCE

To address this research gap, we presented excerpts of actual conversations between American customers and Filipino call center agents for assessment by listeners, who were drawn broadly from the same population of speakers as the customers in the audio recordings. The listeners evaluated the agents' speech for comprehensibility, accentedness, and fluency, and then rated their job performance using several scales. From a range of job performance indicators of customer service representatives, which often include measures of knowledge, communication, and personality along with measures of customer satisfaction and patronage (Darian et al., 2001; Forey & Lockwood, 2010), three specific dimensions were selected. The first two dimensions of agents' job performance involved first, their confidence, and second, their competence, which often combines content knowledge and amount of experience (Forey & Lockwood, 2010; Presbitero, 2017) considered to be important components of agents' job performance (Downing, 2011; Lockwood, 2018). Finally, the third dimension captured aspects of customer patronage or loyalty (D'Ausilio, 1998; Whitaker et al., 2020) through listeners' rating of their interest in future communication with agents. A focus on three aspects of agents' job performance—in terms of their confidence, knowledge/experience, and potential to encourage a future interaction with a customer—thus allowed us to provide a fuller picture of agents' job performance.

Because a given customer's contributions to each conversation may have impacted listeners' judgments (e.g., through clear signals of irritation or satisfaction expressed by the customer), the recordings were presented for evaluation in two conditions: with the customer's speech present or replaced by silence. Furthermore, it was possible that various individual differences across listeners, such as those involving their degree of intercultural awareness, personality, or cognitive flexibility, could moderate their judgments of agents' speech and their evaluations of agents' effectiveness. Among these potential variables, some of which are difficult

CUSTOMER SERVICE AGENTS' PERFORMANCE

to capture in a brief listening session, individual differences in listeners' personality were considered particularly important, especially because personality traits have been shown to determine customers' perceptions of service quality (Marbach et al., 2016) and listeners' evaluations of L2 speech (Gaffney & Côté, 2020). Therefore, to control for variance in listeners' personality traits, they were asked to complete a brief personality inventory (Soto & John, 2017) targeting the Big Five traits (extraversion, agreeableness, conscientiousness, negative emotionality, open-mindedness), which is a commonly used personality framework (Goldberg, 1993), and these scores were entered as control covariates in statistical modeling.

Based on prior research about listener-based evaluations of L2 speakers, we predicted that more comprehensible, less accented, and more fluent agents would be generally perceived more positively by the listeners, compared to agents who are judged less comprehensible, more accented, and less fluent. We made no specific prediction as to which of the three dimensions would be more important to listener perceptions, because few studies have examined comprehensibility, accentedness, and fluency in relation to speaker assessments or speaker preferences in a single dataset. However, based on Derwing and Munro (2009), who reported that comprehensibility (along with some aspects of fluency) was a stronger predictor of listener preferences to interact with L2 speakers than accentedness, we anticipated that the agents' accentedness would be the least predictive of their performance ratings. The study was guided by the following question: To what extent do listener perceptions of customer service agents' comprehensibility, accentedness, and fluency predict these agents' job performance, as operationalized through listener-rated measures of agents' confidence, knowledge/experience, and listeners' desire to engage in future communication with the agents?

4. Method

4.1 Audio Recordings

The target customer–agent conversations came from a larger corpus of recorded service center calls provided by a US-owned company employing Filipino agents as telephone support personnel for a variety of businesses (Friginal, 2009). A stratified random sample of 18 conversations was drawn from the smaller subset of the corpus recordings used by the company for training and quality assurance purposes, with the constraint that the selected recordings represented a range of scores (from above to below average) assigned by company-internal quality monitors evaluating the accuracy of agents' customer support and their transaction handling (Friginal, 2007). The 18 conversations represented various types of businesses (e.g., information technology, retail, logistics) and covered different agent task categories (e.g., troubleshooting, product purchase and delivery, product inquiry). Although their background information was confidential, the agents (6 female, 12 male) were all native speakers of Tagalog ($M_{age} = 26.2$ years, $SD = 2.43$) and were employed for at least two years at a call center in the Philippines.

Before presenting the recordings for assessment by listeners, the original audio files, which typically ranged in length from 5 to 25 minutes, were shortened to approximately 2 minutes in length from the beginning of each conversation ($M = 2.02$ minutes, $SD = 0.13$) and were normalized for peak amplitude. All personal information, if any (e.g., customer and company names, addresses), was carefully removed to comply with privacy and confidentiality requirements and replaced by silence equivalent in length to the excised verbal content. Finally, because a customer's contributions to the conversation—for example, the call specifics, the speaking style, or the affect or demeanor displayed during the call (e.g., irritation, relief)—could

CUSTOMER SERVICE AGENTS' PERFORMANCE

impact listener-based assessments of the agent's performance, two versions of each recording (18 per set) were created. The first set contained the original conversations featuring both the agent and the caller (full-audio version) while the second set included only the agent's speech, with all verbal and nonverbal (e.g., filled pauses, backchannels) contributions by the caller replaced by equivalent-length silence (agent-only version). Thus, the same recorded conversation in each set was identical in length but either included or excluded the speech by the customer.

4.2 Listeners

A total of 116 respondents evaluated customer service agents' performance as listeners through Amazon's Mechanical Turk (MT), which is an online crowdsourcing platform where employers can assign specific jobs, such as transcription of audio files or tagging images for content, to registered workers in exchange for compensation. The platform has been increasingly used for research purposes as it allows for reaching a broad base of language speakers in specific geographic locations and is considered a reliable data source in L2 speech research (Nagle, 2019; Nagle & Rehman, 2021). Although job assignments were open to workers from Australia, Canada, New Zealand, the UK, and the US, all listeners came from the US, reflecting the current worker demographics, where 75% of the estimated 100,000-member workforce is US-based (Difallah et al., 2018). The listeners (64 female, 52 male) were on average 43.1 years old ($SD = 11.5$), and all self-identified as native English speakers, with one describing herself as an English-Spanish bilingual and five reporting additional knowledge of Spanish (4) or multiple languages (1). They reported a range of educational backgrounds, including secondary diplomas (33) and undergraduate (74) and graduate (9) degrees. Using a 100-point sliding scale, they also provided a high daily estimate of using English with native English speakers ($M = 95.36\%$, $SD = 11.93$), where 0 indicated "never" and 100 meant "all the time." Given these characteristics, the

CUSTOMER SERVICE AGENTS' PERFORMANCE

listener sample was representative of a typical population of largely monolingual native English speakers residing in an English-dominant context who might contact foreign-based customer service for support with various products or services. As part of a brief questionnaire, the listeners also responded to additional questions targeting their demographic and personality profiles (see Appendix A), so that these responses could be used in subsequent modeling as covariates to control for potential individual differences across the listeners (see below). In terms of their background, they used a 100-point scale to estimate how often they contact customer support for a product or service that they receive in their daily life ($M = 38.49\%$, $SD = 32.10$), where 0 indicated “never” and 100 meant “all the time.” In terms of their personality profile, they completed the Big Five Inventory-2-XS, which is a short form of the Big Five Inventory (Soto & John, 2017), assessing five traits (extraversion, agreeableness, conscientiousness, negative emotionality, open-mindedness) through 5-point scales, where 1 indicated strong disagreement and 5 designated strong agreement.

4.3 Materials and Procedure

The recordings from the full-audio and agent-only versions of the conversations were available to the MT workers for rating in two batches, with the full-audio set rated first. The two batches were separated by two weeks in order to isolate one set of recordings from the other for the workers who have completed the first (full-audio) set and then wished to come back to complete the other (agent-only) set. Each recording was presented as a separate assignment (i.e., a separate rating task), for a total of 18 assignments per batch. Each assignment first provided brief general information about the study, along with the consent form, followed by the background and personality questions, and then one of the 18 recordings to be evaluated. Although the background and personality questions appeared at the beginning of each assignment, those

CUSTOMER SERVICE AGENTS' PERFORMANCE

needed to be completed only once, and the listeners were aware that they could skip these questions if they chose to accept more than one assignment.

The rating interface included an embedded recording with playback controls, followed by several 100-point sliding scales. Each scale contained no numerical markings apart from the anchor descriptors, with the negative descriptor always on the left and the positive descriptor on the right (see Appendix B for a screenshot of the interface). The three speech scales included comprehensibility (*agent is very difficult to understand–agent is very easy to understand*), accentedness (*agent has a strong foreign/non-English sounding accent–agent has no or little accent*),² and fluency (*agent has many hesitations or pauses–agent has few or no hesitations or pauses*). The three job performance scales included agent confidence (*agent is not at all confident–agent is very confident*), knowledge/experience (*agent lacks knowledge and experience–agent is very knowledgeable and experienced*), and the listener's desire to communicate with the same agent in the future (*I would not want to interact with this agent–I would very much like to interact with this agent*). As part of a larger project, other performance indicators targeted agent's personal traits (e.g., trustworthiness, friendliness) and interpersonal skills (e.g., politeness, empathy) and included two open-ended questions eliciting listener comments about the agent's strengths and weaknesses; these data, which targeted a conceptually different research question and whose treatment within the present report would not be possible due to space limitations, are not discussed further. Because the listeners provided multiple evaluations per audio, consistent with prior research (Nagle, 2019; O'Brien, 2016) they were allowed to replay the file as many times as needed; however, tracking the frequency of each audio's playback was not possible.

CUSTOMER SERVICE AGENTS' PERFORMANCE

The rating assignments were available only to experienced MT workers (called Masters on the Amazon platform) who had earned this internally-awarded designation through excellent prior performance in multiple tasks. For each assignment, one of the 18 recordings was presented, and a random set of 40 workers were invited to evaluate each recording. This means that workers could evaluate a minimum of one and a maximum of 18 recordings per batch, with the consequence that some completed all 18 rating assignments per batch (full-audio batch, agent-only batch) while others finished as many or as few as they wished. Because MT workers could accept assignments in any order and could proceed to evaluate as many assignments as they wished, the order in which the recordings were evaluated was effectively random across different workers. Although each assignment was allotted 1 hour for completion, they were completed on average within 20.97 minutes ($SD = 16.70$) for the full-audio recordings and 25.28 minutes ($SD = 15.73$) for the agent-only recordings after they were accepted and initiated. All full-audio and agent-only evaluation batches were completed within 24 hours of their posting. Given the constraints of online data collection and MT workers' expectation that a job task should be straightforward, there was little training provided to the listeners beyond the instructions to listen to the entire recording and to focus only on the speech of the customer service agent. Data collection stopped when each of the 18 recordings across both sets (full-audio, agent-only) was evaluated by 40 listeners.

5. Data Analysis

5.1 Rating Reliability

Of the 116 unique listeners, several individuals evaluated the entire set of full-audio recordings ($n = 13$) or agent-only recordings ($n = 11$). A handful of listeners assessed a minimum of one recording in the full-audio set ($n = 11$) or the agent-only set ($n = 10$). However, the majority of

CUSTOMER SERVICE AGENTS' PERFORMANCE

the listeners completed between 1 and 17 evaluations per set. In the end, each audio was evaluated by a random set of 40 listeners; 69 individuals assessed the full-audio recordings and 81 individuals assessed the agent-only recordings, with 34 listeners common to both sets (i.e., $69 + 81 - 34 = 116$ unique listeners). In total, the listeners contributed 1,440 datapoints for each scalar measure (40 raters \times 18 recordings \times 2 sets). There were no missing data for audio assessments or background and personality variables; however, seven listeners answered background and personality questions more than once, in which case their responses from the first accepted assignment were used. All ratings were transferred into a data matrix as values between 0 and 100, and the scores for the five personality traits (extraversion, agreeableness, conscientiousness, negative emotionality, open-mindedness) were derived per listener by averaging across the three questions targeting each trait, using the test guidelines (Soto & John, 2017).

Because each recording was evaluated by a random set of listeners, where each recording was rated by a nonidentical listener subset, intraclass correlations were a better measure of internal consistency than Cronbach's alpha (Hallgren, 2012; Nagle, 2019; Nagle & Rehman, 2021). Internal consistency was thus checked through one-way random-effects, consistency, average-measure intraclass correlations separately for full-audio and agent-only recording using the psych package (version 2.0.9, Revelle, 2021) in R (version 4.0.2, R Core Team, 2020). As shown in Table 1, the consistency values were high. Because removing the data from the 34 listeners who evaluated both the full-audio and agent-only recordings (separated by two weeks) resulted in no change to the findings and because intraclass correlations yielded high consistency values comparable to those reported previously for speech assessments obtained through the MT platform (Nagle, 2019), all data were retained for statistical analysis.

CUSTOMER SERVICE AGENTS' PERFORMANCE

Table 1. *Intraclass Correlations and Confidence Intervals*

Rated variable	Full audio ($n = 69$)		Agent only ($n = 81$)	
	<i>ICC</i>	95% CI	<i>ICC</i>	95% CI
Comprehensibility	.93	[.89, .97]	.93	[.89, .97]
Accentedness	.97	[.95, .98]	.96	[.94, .98]
Fluency	.90	[.83, .95]	.90	[.83, .95]
Agent's confidence	.94	[.90, .97]	.91	[.85, .95]
Agent's knowledge/experience	.92	[.87, .96]	.89	[.81, .94]
Desire to communicate with agent	.94	[.90, .97]	.91	[.86, .96]

5.2 Statistical Modeling

To address the research question, linear mixed-effects models (Hox et al., 2017; Linck & Cunnings, 2015) were computed in R using the lme4 package (version 1.1-23, Bates et al., 2015). In each set of models, the job performance ratings of confidence, knowledge/experience, and desire to engage in future communication served as the outcome variables whereas condition (full-audio vs. agent-only) and the three speech ratings (accentedness, comprehensibility, fluency) served as fixed-effects predictors, with raters and speakers entered as random-effects predictors. All models also included six fixed effects as control covariates (five personality traits, plus listener-estimated prior contact with customer service), on the assumption that the listeners who had frequent previous contact with customer service and those with specific personality traits (e.g., agreeableness, open-mindedness) might evaluate customer service agents more or less positively, compared to those with less frequent contact with customer service and with different personality profiles. All continuous predictor variables were centered using the grand mean. Although there is no consensus regarding the criteria for considering statistical

CUSTOMER SERVICE AGENTS' PERFORMANCE

significance in mixed-effects modelling, with some scholars using $t > |2.00|$ as the benchmark of significance (Linck & Cunnings, 2015), we obtained p values using MuMIn package in R (version 1.43.17, Bartoń, 2020) but examined 95% confidence intervals (CIs) to check the statistical significance of each parameter (interval does not cross zero). Random slope models were also examined, separately for raters and speakers; however, the inclusion of random slopes did not improve model fit for any outcome variable, so the final models excluded random slopes.

6. Results

The research question asked whether the comprehensibility, accentedness, and fluency of customer service agents' speech contributed to the variance in listener evaluations of the agents' job performance (in terms of their confidence, their knowledge/experience, and listeners' desire to engage in future communication with them). Table 2 summarizes descriptive statistics for the ratings, shown as means and standard deviations across all listeners evaluating the full-audio and agent-only recordings. Generally speaking, the listeners perceived the agents as sounding fairly comprehensible, non-accented, and fluent (rating them at 67–78 on a 100-point scale), and they judged the agents to be reasonably effective, with little difference in ratings between the assessments of the full-audio and agent-only conversations. Nonetheless, the standard deviation values demonstrated considerable variation in listener perceptions, suggesting that the ratings varied across different agents and listeners. Among the three speech ratings, comprehensibility was strongly associated with accentedness ($r = .66, p < .001$) while fluency showed weak-to-moderate associations with comprehensibility ($r = .33, p < .001$) and accentedness ($r = .41, p < .001$), based on field-specific guidelines (Plonsky & Oswald, 2014).

CUSTOMER SERVICE AGENTS' PERFORMANCE

Table 2. Descriptive Statistics (Grand Means and Standard Deviations Across Agents and Listeners) for the Target Ratings (100-Point Scale)

Rated variable	Full audio (<i>n</i> = 69)		Agent only (<i>n</i> = 81)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Comprehensibility	78.20	21.47	73.74	25.47
Accentedness	69.95	26.88	67.14	28.51
Fluency	68.37	26.87	67.52	26.96
Agent's confidence	75.50	22.42	76.31	23.14
Agent's knowledge/experience	77.12	20.67	76.56	22.24
Desire to communicate with agent	71.38	26.46	69.96	27.53

Table 3 summarizes the final mixed-effects model for the agent confidence ratings. There was only a marginally significant difference between the assessments of the full-audio and agent-only conversations ($p = .066$), such that the listeners who had access to the full audio tended to perceive the agent's confidence level lower than those with access only to the agent's speech. In terms of the three speech ratings, the agent's fluency was the strongest predictor of agent's confidence ($Estimate = 0.33, p < .001$), followed by comprehensibility as a predictor ($Estimate = 0.18, p < .001$), such that the agents who were more comprehensible and fluent were perceived as more confident. However, accentedness was not associated with the confidence ratings ($Estimate = -0.02, p = .403$). In terms of the control covariates, two personality traits contributed to listener assessments of confidence, such that the listeners who were more extravert ($p = .009$) and more agreeable ($p = .018$) perceived the agents as more confident. In addition, the listeners with more recent experience with customer service tended to perceive the agents as less confident ($p = .030$).

CUSTOMER SERVICE AGENTS' PERFORMANCE

Table 3. Summary of Final Mixed-Effects Model for Agent Confidence

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	76.84	1.67	[73.45, 80.23]	46.13	< .001
Full audio vs. Agent only	-1.94	1.05	[-4.03, 0.13]	-1.84	.066
Comprehensibility	0.18	0.03	[0.13, 0.24]	6.50	< .001
Accentedness	-0.02	0.03	[-0.07, 0.03]	-0.84	.403
Fluency	0.33	0.02	[0.29, 0.37]	16.10	< .001
Covariates					
Extraversion	2.35	0.89	[0.59, 4.10]	2.64	.009
Agreeableness	2.16	0.90	[0.35, 3.94]	2.39	.018
Conscientiousness	-0.37	9.96	[-2.32, 1.63]	-0.37	.711
Negative emotionality	0.03	0.83	[-1.63, 1.67]	0.03	.973
Open-mindedness	0.01	0.82	[-1.63, 1.62]	0.01	.996
Customer service experience	-0.05	0.02	[-0.10, -0.01]	-2.20	.030
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	26.96	5.19	Log-likelihood	-6134.80	
Agent (intercept)	36.22	6.02	AIC	12297.60	
			BIC	12371.40	
			Marginal R^2	0.31	
			Conditional R^2	0.44	

Note. AIC = Akaike information criterion, BIC = Bayesian information criterion, marginal R^2 = variance explained by fixed factors, conditional R^2 = variance explained by fixed and random factors.

CUSTOMER SERVICE AGENTS' PERFORMANCE

Table 4 summarizes the final mixed-effects model for the agent knowledge/experience ratings. There was no statistically significant difference between the assessments of the full-audio and agent-only conversations ($p = .607$). As with the ratings of confidence, again, agent's fluency was the strongest predictor ($Estimate = 0.31, p < .001$), followed by comprehensibility ($Estimate = 0.16, p < .001$), such that the agents who were more fluent and comprehensible were perceived as more knowledgeable and experienced. Accentedness was again not associated with listener perceptions of knowledge/experience ($Estimate = 0.03, p = .147$). No fixed-effects control covariate contributed statistically significant variance to explaining the knowledge/experience ratings.

CUSTOMER SERVICE AGENTS' PERFORMANCE

Table 4. Summary of Final Mixed-Effects Model for Agent Knowledge/Experience

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	76.93	1.62	[73.65, 80.20]	47.51	< .001
Full audio vs. Agent only	-0.54	1.03	[-2.57, 1.49]	-0.52	.607
Comprehensibility	0.16	0.03	[0.11, 0.21]	6.23	< .001
Accentedness	0.03	0.02	[-0.01, 0.08]	1.45	.147
Fluency	0.31	0.02	[0.27, 0.34]	15.86	< .001
Covariates					
Extraversion	1.42	0.94	[-0.43, 3.30]	1.51	.135
Agreeableness	1.66	0.97	[-0.39, 3.59]	1.70	.092
Conscientiousness	-1.41	1.08	[-3.53, 0.79]	-1.31	.195
Negative emotionality	-0.37	0.90	[-2.15, 1.39]	-0.41	.680
Open-mindedness	0.41	0.89	[-1.38, 2.15]	0.47	.643
Customer service experience	-0.03	0.02	[-0.07, 0.02]	-1.10	.273
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	41.80	6.45	Log-likelihood	-6027.00	
Agent (intercept)	31.64	5.63	AIC	12082.10	
			BIC	12155.90	
			Marginal R^2	0.30	
			Conditional R^2	0.47	

Lastly, Table 5 summarizes the final mixed-effects model for the listeners' desire to engage in future communication with the agent. Again, there was no statistically significant difference between the assessments of the full-audio and agent-only conversations ($p = .534$).

CUSTOMER SERVICE AGENTS' PERFORMANCE

However, all three speech measures contributed to the agent ratings, with agent's fluency being the strongest predictor ($Estimate = 0.24, p < .001$), followed by comprehensibility ($Estimate = 0.17, p < .001$) and finally accentedness ($Estimate = 0.10, p < .001$), such that the agents whose speech was perceived as more fluent, more comprehensible, and less accented were those with whom the listeners wished to engage in future communication. In terms of the control covariates, only one personality trait contributed to the agent ratings, where the listeners who were more agreeable ($p = .023$) indicated greater desire to interact with the agent in the future.

CUSTOMER SERVICE AGENTS' PERFORMANCE

Table 5. Summary of Final Mixed-Effects Model for Listener Desire to Communicate with Agent

Fixed effects	<i>Estimate</i>	<i>SE</i>	95% CI	<i>t</i>	<i>p</i>
(Intercept)	71.13	2.26	[66.57, 75.71]	31.46	< .001
Full audio vs. Agent only	-0.86	1.39	[-3.61, 1.87]	-0.62	.534
Comprehensibility	0.17	0.04	[0.10, 0.23]	4.78	< .001
Accentedness	0.10	0.03	[0.04, 0.16]	3.33	< .001
Fluency	0.24	0.03	[0.19, 0.29]	9.36	< .001
Covariates					
Extraversion	1.13	1.32	[-1.47, 3.73]	0.86	.391
Agreeableness	3.19	1.38	[0.45, 5.91]	2.31	.023
Conscientiousness	-2.09	1.53	[-5.10, 0.98]	-1.36	.176
Negative emotionality	-1.70	1.26	[-4.22, 0.78]	-1.35	.180
Open-mindedness	0.55	1.26	[-1.99, 3.01]	0.44	.664
Customer service experience	0.04	0.03	[-0.03, 0.11]	1.21	.226
Random effects	<i>Variance</i>	<i>SD</i>	Information criteria	<i>Estimate</i>	
Rater (intercept)	93.90	9.69	Log-likelihood	-6423.1	
Agent (intercept)	60.99	7.81	AIC	12874.1	
			BIC	12947.9	
			Marginal R^2	0.20	
			Conditional R^2	0.43	

6. Discussion

Motivated by prior work showing that listener-based assessments are influenced by speaker comprehensibility, accentedness, and fluency (e.g., Campbell-Kibler, 2007; Dragojevic et al.,

CUSTOMER SERVICE AGENTS' PERFORMANCE

2017; Parton et al., 2002), this study examined the relationship between these global dimensions of speech and several listener-assessed indicators of customer service agents' job performance. Comprehensibility and fluency were the two dimensions that consistently predicted how the listeners assessed the agents on all job performance scales, and accentedness was additionally associated with how strongly the listeners wished to communicate with the agents in the future. In all cases, the relationships were positive, in the sense that more comprehensible, more fluent, and less accented L2 speech was linked to higher performance evaluations. These relationships emerged regardless of whether the listeners evaluated only the agent or the entire conversation which included speech by a customer. These relationships also obtained after controlling for individual differences in the listeners' personality variables and their prior contact with telephone-based customer service.

The present assessments of agents' job performance included several measures: confidence (Presbitero, 2017), content knowledge/experience (Forey & Lockwood, 2010), and listeners' desire to engage in future communication with the agent, which encompasses the idea of customer patronage or loyalty (Whitaker et al., 2020). Although these evaluations were associated with both comprehensibility and fluency, fluency emerged as the stronger predictor, as shown through the magnitude of model estimates (cf. Tables 3–5). A key role of fluency in agent evaluations is consistent with prior research which showed various suprasegmental and fluency aspects of agents' speech (e.g., intonation, pitch, volume, speech rate, pausing) to be relevant to their overall job performance as assessed by quality assurance monitors (Friginal, 2007, 2009). In addition, many listeners (including language teachers and laypersons) interpret fluency as a proxy for L2 speaking proficiency (Iwashita et al., 2008; Tavakoli et al., 2020) or language competence (Segalowitz, 2010). Speaking fluently (i.e., with few hesitations and

CUSTOMER SERVICE AGENTS' PERFORMANCE

pauses) is also associated with speaking confidently (Rossiter, 2009; Tavakoli & Hunter, 2017), and listeners attribute judgments of knowledge, credibility, and competence to speaking performances that display few signs of disfluency, such as hesitations and hedging (Blankenship & Holtgraves, 2005; Hosman & Siltanen, 2011; Parton et al., 2002). Against this backdrop, it is perhaps unsurprising that the listeners in this study associated speaker fluency with perceptions of confidence and knowledge, expressing a stronger desire to interact with fluent agents.

Apart from fluency, comprehensibility was the other speech dimension associated consistently with all agent evaluations. Practically speaking, comprehensibility as a predictor of agents' job performance likely reflected the listeners' awareness that the success of customer-agent conversations, at least in part, depends on how easily the customer understands the agent, who is expected to be knowledgeable and helpful in addressing customer concerns. From a theoretical standpoint, the role of comprehensibility in agent evaluations can be explained within the processing fluency perspective (Reber & Greifeneder, 2017), which describes people's subjective experience of ease or difficulty of completing a task, such as solving a math problem or reading a text for comprehension, as critical to their assessments. When listeners struggle to understand a speaker, they may project their feelings of frustration onto the speaker through downgrading the speaker in judgments of success and intelligence, ascribing negative emotions to the speaker, or assessing their own task performance negatively (Dragojevic et al., 2017; Sanchez & Khan, 2016). In fact, the link between people's perceptions of ease or difficulty and their various assessments is well documented in consumer psychology (Schwarz, 2004). For instance, less complex shopping environments, such as neatly arranged product placement, elicit more positive evaluations from consumers than high-complexity retail environments featuring numerous objects placed irregularly (Orth & Wirtz, 2014). More comprehensible speech, similar

CUSTOMER SERVICE AGENTS' PERFORMANCE

to less complex visual scenes, is easier to process, so listeners evaluate their experience positively, assigning higher ratings of confidence and knowledge to, and showing greater interest in communicating with, the agents who require less effort to understand.

Among the three speech dimensions, accentedness was the least consequential for agent evaluations, accounting for listener assessments on only one performance scale. Even though accented speakers generally receive negative evaluations from listeners for many personal and professional attributes, including competence, attractiveness, and integrity (e.g., Campbell-Kibler, 2007; Hosoda et al., 2007), the listeners in this study did not seem to engage in this practice, in the sense that their evaluations of the agent's confidence and knowledge/experience were unrelated to that agent's accent. On the one hand, the listeners may have been aware that a key goal of service transactions is to address a customer's needs, where the agent's professionalism (which includes confidence and knowledge) is more valuable than speaking without a foreign accent. On the other hand, however, the agents included in this study did not have strong accents to begin with (see Table 2), since their speech elicited fairly high ratings from the listeners (67–70 on a 100-point scale, where 100 meant *agent has no or little accent*). Considering that foreign-based call centers tend to hire employees with least accented English (Lockwood, 2012), expecting a strong impact of accentedness on agent assessments may have been unreasonable.

Nevertheless, accentedness was predictive of the listeners' desire to engage in future communication with the agent. This finding is congruent with Derwing and Munro's (2009) observation that, even though the employees in an engineering company generally relied on speaker comprehensibility (along with fluency) to select a potential interlocutor (in 88% of the cases), they also opted for a less accented speaker a significant proportion of the time (65%). A

CUSTOMER SERVICE AGENTS' PERFORMANCE

speaker's accent strength may thus carry real-life behavioral consequences, in that listeners may prefer to communicate with someone who sounds like them as opposed to a perceived outsider, even if there is no expectation that communication (e.g., in terms of understanding the speaker) will be compromised (Bestelmeyer et al., 2015). In this sense, the listeners in this study made a clear distinction between their evaluations of the agent's confidence and knowledge/experience and their own (hypothetical) future behaviors. For example, the listeners in Gathbonton et al. (2005) considered less accented L2 speakers to be less loyal members of their ethnic group yet preferred those same speakers as leaders for potential group projects in the community. In a similar vein, as shown in Tables 3 and 4, the listeners in this study generally valued the agents' fluency and comprehensibility in relation to their professional qualities (confidence, knowledge/experience), irrespective of these speakers' accent, yet the same listeners also preferred to interact with the agents whose accent was less pronounced (see Table 5).

Individual differences in listener background and personality profiles were not the focus of this study and were entered in statistical models as fixed-effects covariates to control for possible influence of these variables on listener assessments. Nevertheless, several findings emerged. Whereas two of the five personality traits predicted agent confidence (extraversion, agreeableness) and listener-rated desire to engage in future communication with the agent (agreeableness), all with a positive association, listener-estimated prior contact with customer service predicted agent confidence, with a negative association. Generally speaking, more extraverted individuals tended to be assertive in their behavior, decisive in their thinking, and positive in their emotion (Wilt & Revelle, 2017), while more agreeable people demonstrated sympathy, generosity, compassion, respect, and cooperation (Graziano & Tobin, 2017). Thus, one possible, though subjective, account of these findings might be that these traits may have

CUSTOMER SERVICE AGENTS' PERFORMANCE

enabled the listeners to provide more generous assessments of the agents on at least some of the performance scales. Needless to say, this speculative interpretation must be revisited in future work. In contrast, the listeners with frequent recent contact with customer service may have been particularly impacted by their (perhaps less than ideal) experiences, leading to harsher agent assessments. It is noteworthy that no control variables predicted listener assessments of the agents' knowledge/experience, implying that perceptions of knowledge or expertise are less prone to the influence of individual listener profiles. Before any definitive conclusions can be drawn, however, these findings must be revisited in focused future investigations into listener personality profiles in relation to various linguistic and professional assessments of speakers.

7. Implications, Limitations, and Future Work

Although this study has limited direct implications for the customer service industry, the findings broadly highlight the importance of fluency and comprehensibility to listener-based assessments of customer service agents' job performance. Practically speaking, customer-agent conversations require time-efficient communication, so that a customer's concerns are addressed thoroughly and expeditiously. Despite the overwhelming belief among many laypersons and professionals that a native accent is critical to the success of the customer service industry (Lockwood, 2012), where recruiters screen applicants based on specific pronunciation errors or accent varieties (Timming, 2017) and training programs focus on pronunciation and grammar drilling (Hamp-Lyons & Lockwood, 2009), the present findings suggest that accent reduction should not be the top priority for agent recruitment and training. Instead, if an agent's fluency and comprehensibility are high, customers may develop a positive view of their experience. However, if an agent pauses or hesitates when explaining procedures and providing solutions or if an agent uses complex terminology or produces long utterances that hinder understanding,

CUSTOMER SERVICE AGENTS' PERFORMANCE

customers might develop negative judgments, perceiving the agent as less confident, knowledgeable, or experienced. Professional training for agents might therefore need to focus on strategies for speaking comprehensibly and fluently while dealing with often unpredictable demands of live conversations.

At a more global level, the present findings highlight the need for researchers and practitioners in various fields of professional communication, including the customer service industry, to problematize and expose various ethical aspects of communication involving non-native speakers in the workplace. In an industry where customer preferences often drive business decisions (Camilleri, 2021), how ethical is it to enforce that non-native speaking employees demonstrate evidence of a natively-like accent, given that the construct of accent is linked to various forms of discriminatory and racist behaviors (Lippi-Green, 2012; Ramjattan, 2019)? Does the burden of ensuring a successful communicative transaction in the workplace lie squarely with the non-native speaking employee (Derwing et al., 2014)? What is the responsibility of a customer in a world supported by transnational corporations, international business, and lingua-franca communication? Is the customer always right (Grandey et al., 2004)? Questions such as these are important to consider if we are to provide a comprehensive account of the role of language in agent–customer communication.

Despite the potential promise of these findings, several key limitations of this study must be acknowledged. First, although the present listener sample was representative of a typical population of customers who might contact foreign-based call centers for support with various products or services, evaluations by external listeners might not be directly comparable to impressions of the actual callers engaged in live conversations with customer service representatives. Real customers may base their assessments on various prior personal

CUSTOMER SERVICE AGENTS' PERFORMANCE

experiences with a product or service (both positive and negative), whereas external listeners, such as those employed in this study, might lack such insight. Future work should compare evaluations of agents' job performance by external listeners and by actual customers, particularly relative to assessments by trained professionals, such as internal or external quality control monitors. Second, the sample of customer service agents' speech in this study was clearly restricted, in that it included only Filipino speakers of English who were relatively fluent, comprehensible, and non-accented. For instance, as shown in Table 2, all agents elicited relatively high scores from the listeners, who rated them at about 67–78 on a 100-point scale across the dimensions of accentedness, comprehensibility, and fluency. Therefore, to provide a comprehensive picture of listener-based assessments of customer service agents, researcher may need to target speaking performances that differ in speakers' accent variety and their comprehensibility, fluency, and accentedness.

Third, several methodological decisions guiding this research must be refined in future work. Although the MT platform ensures worker anonymity, which should contribute to workers' comfort when responding to various questionnaires, similar to any self-report data, some questionnaire items might have been difficult for workers to answer truthfully (e.g., *I am someone who is sometimes rude to others*). Another methodological limitation concerns the joint agent knowledge/experience scale. Although the dimensions of knowledge and experience are often combined in evaluations of agents' performance under the general category of competence (e.g., Forey & Lockwood, 2010; Lockwood, 2018), a dual knowledge/experience scale may have made it difficult for listeners to evaluate two related constructs on a single continuum, which should be corrected in future work. In addition, the scale label for accentedness rating was potentially misleading, as some listeners may have interpreted “non-English sounding accent” to

CUSTOMER SERVICE AGENTS' PERFORMANCE

mean a British variety of English or may have presumed that Philippine English is not a legitimate English variety. Needless to say, it would be important to provide listeners with a clearer label which unambiguously refers to a variety spoken natively by the majority of speakers in a given context. Fourth, in controlling individual differences in listeners' profiles, the present study only captured listeners' personality traits. It is possible that listener-based evaluations of agents might be shaped by other individual listener variables, such as their tolerance for ambiguity, intercultural awareness, or acoustic sensitivity, and the role of these other variables should be explored in future work. Finally, to provide insights into which specific features of customer service agents' speech determine how comprehensible, fluent, or accented they sound, future work should examine these specific features (e.g., segmental accuracy, lexical appropriateness, pragmatic relevance) in relation to listener-based assessments of agents' job performance.

8. Conclusion

Outsourced call centers play an important role in the modern business world, and the language of call center agents has been the focus of a substantial amount of research. The goal of this study was to gain insights into the perspective of customers, by examining relationships between listener-rated dimensions of agents' speech and listener perceptions of agents' job performance. Our findings suggest that agents who were more comprehensible and fluent were perceived as more confident, knowledgeable, and experienced, and that listeners had a greater desire to communicate with those agents in the future. Accentedness was unrelated to perceptions of agents' confidence or knowledge and experience, although listeners had a greater desire to communicate with less accented agents. These findings underscore the importance of comprehensible and fluent speech in customer service contexts, and support a shift in outsourced

CUSTOMER SERVICE AGENTS' PERFORMANCE

call centers' recruitment and training processes, which tend to make accent reduction the top priority.

Notes

1. In keeping with the listener perspective on agents' job performance, we focused only on impressionistic, listener-rated measures of agents' speech, such as accentedness, comprehensibility, and fluency, at the expense of intelligibility, which refers to "the extent to which a speaker's message is actually understood by a listener" (Munro & Derwing, 1995, p. 76). Arguably, intelligibility is a far better indicator of communicative success (Levis, 2018; Nelson, 2012), and its role in agent–customer communication must be explored in greater detail.
2. An anonymous reviewer appropriately pointed out that this scale label is both ambiguous (because it can be perceived to refer to the variety of English spoken in the UK) and inaccurate (because Philippine English is a legitimate, nativized outer-circle variety of English). Although "non-English sounding accent" was meant to provide an intuitive and easily understandable rating anchor point for layperson listeners with little background knowledge in pronunciation research, we acknowledge that it may have created confusion for at least some listeners.

References

- Almeida, S., Fernando, M., Hannif, Z., & Dharmage, S. C. (2015). Fitting the mould: The role of employer perceptions in immigrant recruitment decision-making. *The International Journal of Human Resource Management*, 26, 2811–2832.
<https://doi.org/10.1080/09585192.2014.1003087>
- Archer, D., & Jagodziński, P. (2015). Call centre interaction: A case of sanctioned face attack? *Journal of Pragmatics*, 76, 46–66. <https://doi.org/10.1016/j.pragma.2014.11.009>

CUSTOMER SERVICE AGENTS' PERFORMANCE

- Bartoń, K. (2020). *MuMIn: Multi-Model Inference. R package version 1.43.17*. <https://CRAN.R-project.org/package=MuMIn>
- 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>
- Bestelmeyer, P. E. G., Belin, P., & Ladd, D. R. (2015). A neural marker for social bias toward in- group accents. *Cerebral Cortex*, *25*, 3953–3961. <https://doi.org/10.1093/cercor/bhu282>
- Blankenship, K. L., & Holtgraves, T. (2005). The role of different markers of linguistic powerlessness in persuasion. *Journal of Language and Social Psychology*, *24*, 3–24. <https://doi.org/10.1177/0261927X04273034>
- 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/10.1177/0261927X12455394>
- Brennan, E., & Brennan, J. (1981). Accent scaling and language attitudes: Reactions to Mexican American English speech. *Language and Speech*, *24*, 207–221. <https://doi.org/10.1177/002383098102400301>
- Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, *34*, 383–398. <https://doi.org/10.1006/jmla.1995.1017>
- Camilleri, M. A. (2021). E-commerce websites, consumer order fulfillment and after-sales service satisfaction: The customer is always right, even after the shopping cart check-out. *Journal of Strategy and Management*. Advance online publication. <https://doi.org/10.1108/JSMA-02-2021-0045>
- Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions.

CUSTOMER SERVICE AGENTS' PERFORMANCE

American Speech, 82, 32–64. <https://doi.org/10.1215/00031283-2007-002>

CFI Group (2020). *Contact center satisfaction index: Managing the entire customer experience*.

<https://cfigroup.com/resource-item/contact-center-satisfaction-2020>

Chicu, D., Pàmies, M., Ryan, G., & Cross, C. (2019). Exploring the influence of the human factor on customer satisfaction in call centres. *Business Research Quarterly*, 22, 83–95.

<https://doi.org/10.1016/j.brq.2018.08.004>

Cocchiara, F. K., Bell, M. P., & Casper, W. J. (2016). Sounding “different”: The role of sociolinguistic cues in evaluating job candidates. *Human Resource Management*, 55, 463–477. <https://doi.org/10.1002/hrm.21675>

Cowie, C., & Murty, L. (2010). Researching and understanding accent shifts in Indian call centre agents. In G. Forey & J. Lockwood (Eds.), *Globalization, communication and the workplace: Talking across the world* (pp. 125–144). Continuum.

D'Ausilio, R. (1998). *Wake up your call center: How to be a better call center agent*. Purdue University Press.

Darian, J. C., Tucci, L. A., & Wiman, A. R. (2001). Perceived salesperson service attributes and retail patronage intentions. *International Journal of Retail Distribution Management*, 29, 205–213. <https://doi.org/10.1108/09590550110390986>

Davila, A., Bohara, A., & Saenz, R. (1993). Accent penalties and the earnings of Mexican Americans. *Social Science Quarterly*, 74, 902–916.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19, 1–16.

<https://doi.org/10.1017/S0272263197001010>

CUSTOMER SERVICE AGENTS' PERFORMANCE

- Derwing, T. M., & Munro, M. J. (2009). Comprehensibility as a factor in listener interaction preferences: Implications for the workplace. *Canadian Modern Language Review*, 66, 181–202. <https://doi.org/10.3138/cmlr.66.2.181>
- Derwing, T., Fraser, H., Kang, O., & Thomson, R. (2014). L2 accent and ethics: Issues that merit attention. In A. Mahboob & L. Barratt (Eds.), *Englishes in multilingual contexts* (pp. 63–80). Springer.
- Difallah, D., Filatova, E., & Ipeirotis, P. (2018). Demographics and dynamics of Mechanical Turk workers. *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 135–143. <https://doi.org/10.1145/3159652.3159661>
- Downing, J. R. (2011). Linking communication competence with call center agents' sales effectiveness. *Journal of Business Communication*, 48, 409–425.
<http://doi.org/10.1177/0021943611414539>
- Dragojevic, M., Giles, H., Beck, A.-C., & Tatum, N. T. (2017). The fluency principle: Why foreign accent strength negatively biases language attitudes. *Communication Monographs*, 84, 385–405. <https://doi.org/10.1080/03637751.2017.1322213>
- Erickson, B., Lind, A. E., Johnson, B. C., & O'Barr, W. M. (1978). Speech style and impression formation in a court setting: The effects of “powerful” and “powerless” speech. *Journal of Experimental Social Psychology*, 14, 266–279. [https://doi.org/10.1016/0022-1031\(78\)90015-X](https://doi.org/10.1016/0022-1031(78)90015-X)
- Forey, G., & Lockwood, J. (2007). “I’d love to put someone in jail for this”: An initial investigation of English in the business processing outsourcing (BPO) industry. *English for Specific Purposes*, 26, 308–326. <https://doi.org/10.1016/j.esp.2006.09.005>

CUSTOMER SERVICE AGENTS' PERFORMANCE

- Forey, G., & Lockwood, J. (Eds.) (2010). *Globalization, communication, and the workplace: Talking across the world*. Continuum.
- Friginal, E. (2007). Outsourced call centers and English in the Philippines. *World Englishes*, 26, 331–345. <https://doi.org/10.1111/j.1467-971X.2007.00512.x>
- Friginal, E. (2009). *The language of outsourced call centers: A corpus-based study of cross-cultural interaction*. John Benjamins.
- Friginal, E., & Cullom, M. (2014). Saying 'No' in Philippine-based outsourced call center interactions. *Asian Englishes*, 16, 2–18. <https://doi.org/10.1080/13488678.2013.872362>
- Gaffney, C., & Côté, S. (2020). Does personality influence ratings of foreign accents? *Journal of Second Language Pronunciation*, 6, 68–95. <http://doi.org/10.1075/jslp.18042.gaf>
- Gatbonton, E., Trofimovich, P., & Magid, M. (2005). Learners' ethnic group affiliation and L2 pronunciation accuracy: A sociolinguistic investigation. *TESOL Quarterly*, 39, 489–511. <http://doi.org/10.2307/3588491>
- Gluszek, A., Newheiser, A.-K., & Dovidio, J. F. (2011). Social psychological manipulations and accent strength. *Journal of Language and Social Psychology*, 30, 28–45. <https://doi.org/10.1177/0261927X10387100>
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48, 26–34. <https://doi.org/10.1037/0003-066X.48.1.26>
- Grandey, A. A., Dickter, D. N., & Sin, H.-P. (2004). The customer is *not* always right: Customer aggression and emotion regulation of service employees. *Journal of Organizational Behavior*, 25, 397–418. <https://doi.org/10.1002/job.252>

CUSTOMER SERVICE AGENTS' PERFORMANCE

- Graziano, W. G., & Tobin, R. M. (2017). Agreeableness and the Five Factor Model. In T. A. Widiger (Ed.), *The Oxford handbook of the Five Factor Model* (pp. 105–132). Oxford University Press.
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8, 23–34.
<https://doi.org/10.20982/tqmp.08.1.p023>
- Hamp-Lyons, L., & Lockwood, J. (2009). The workplace, the society and the wider world: The offshoring and outsourcing industry. *Annual Review of Applied Linguistics*, 29, 145–167.
<http://doi.org/10.1017/S0267190509090126>
- Hayes-Harb, R., & Hacking, J. F. (2015). Beyond rating data: What do listeners believe underlies their accentedness judgments? *Journal of Second Language Pronunciation*, 1, 43–64.
<https://doi.org/10.1075/jslp.1.1.02hay>
- Hood, S., & Forey, G. (2008). The interpersonal dynamics of call-centre interactions: Co-constructing the rise and fall of emotion. *Discourse & Communication*, 2, 389–409.
<https://doi.org/10.1177/1750481308095937>
- Hosman, L. (2015). Powerful and powerless speech styles and their relationship to perceived dominance and control. In R. Schulze, & H. Pishwa (Eds.), *The exercise of power in communication* (pp. 221–232). Palgrave Macmillan.
- Hosman, L. A., & Siltanen, S. A. (2011). Hedges, tag questions, message processing, and persuasion. *Journal of Language and Social Psychology*, 30, 341–349.
<https://doi.org/10.1177/0261927X11407169>

CUSTOMER SERVICE AGENTS' PERFORMANCE

Hosoda, M., Stone-Romero, E. F., & Walter, J. N. (2007). Listeners' cognitive and affective reactions to English speakers with standard American English and Asian accents.

Perceptual and Motor Skills, 104, 307–326. <https://doi.org/10.2466/pms.104.1.307-326>

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.

Hultgren, A. K. (2011). “Building rapport” with customers across the world: The global diffusion of a call centre speech style. *Journal of Sociolinguistics*, 15, 36–64.

<https://doi.org/10.1111/j.1467-9841.2010.00466.x>

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, 34, 475–505. <https://doi.org/10.1017/S0272263112000150>

Language Acquisition, 34, 475–505. <https://doi.org/10.1017/S0272263112000150>

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, 29, 24–49.

<http://doi.org/10.1093/applin/amm017>

Kahng, J. (2018). The effect of pause location on perceived fluency. *Applied Psycholinguistics*, 39, 569–591. <https://doi.org/10.1017/S0142716417000534>

Krahn, H., Derwing, T., Mulder, M., & Wilkinson, L. (2000). Educated and underemployed: Refugee integration into the Canadian labour market. *Journal of International Migration and Integration*, 1, 59–84. <https://doi.org/10.1007/s12134-000-1008-2>

Krishnamurthy, M. (2018). *1-800-Worlds: The making of Indian call centre economy*. Oxford University Press.

Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.

CUSTOMER SERVICE AGENTS' PERFORMANCE

- Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, 65, 185–207.
<http://doi.org/10.1111/lang.12117>
- Lippi-Green, R. (2012). *English with an accent: Language, ideology, and discrimination in the United States*. Routledge.
- Lockwood, J. (2012). Are we getting the right people for the job? A study of English language recruitment assessment practices in the business processing outsourcing sector: India and the Philippines. *The Journal of Business Communication*, 49, 107–127.
<https://doi.org/10.1177/0021943612436975>
- Lockwood, J. E. (2015). Language for Specific Purpose (LSP) performance assessment in Asian call centres: Strong and weak definitions. *Language Testing in Asia*, 5, 1–11.
<https://doi.org/10.1186/s40468-014-0009-6>
- Lockwood, J. E. (2018). Theorising communications coaching in Asian contact centres: What is the approach? *Industrial and Commercial Training*, 50, 312–325.
<https://doi.org/10.1108/ICT-05-2018-0045>
- Marbach, J., Lages, C. R., & Nunan, D. (2016). Who are you and what do you value? Investigating the role of personality traits and customer-perceived value in online customer engagement. *Journal of Marketing Management*, 32, 502–525.
<https://doi.org/10.1080/0267257X.2015.1128472>
- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97.
<https://doi.org/10.1111/j.1467-1770.1995.tb00963.x>

CUSTOMER SERVICE AGENTS' PERFORMANCE

- Nagle, C. (2019). Developing and validating a methodology for crowdsourcing L2 speech ratings in Amazon Mechanical Turk. *Journal of Second Language Pronunciation, 5*, 294–323. <https://doi.org/10.1075/jslp.18016.nag>
- Nagle, C., & Rehman, I. (2021). Doing L2 speech research online: Why and how to collect online ratings data. *Studies in Second Language Acquisition, 43*, 916–939. <https://doi.org/10.1017/S0272263121000292>
- Nelson, C. L. (2012). *Intelligibility in World Englishes: Theory and application*. Routledge.
- O'Brien, M. G. (2016). Methodological choices in rating speech samples. *Studies in Second Language Acquisition, 38*, 587–605. <https://doi.org/10.1017/S0272263115000418>
- Orth, U. R., & Wirtz, J. (2014). Consumer processing of interior service environments: The interplay among visual complexity, processing fluency, and attractiveness. *Journal of Service Research, 17*, 296–309. <https://doi.org/10.1177/1094670514529606>
- Parton, S. R., Siltanen, S. A., Hosman, L. A., & Langenderfer, J. (2002). Employment interview outcomes and speech style effects. *Journal of Language and Social Psychology, 21*, 144–161. <https://doi.org/10.1177/02627X02021002003>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning, 64*, 878–912. <https://doi.org/10.1111/lang.12079>
- Presbitero, A. (2017). It's not all about language ability: Motivational cultural intelligence matters in call center performance. *The International Journal of Human Resource Management, 28*, 1547–1562. <https://doi.org/10.1080/09585192.2015.1128464>
- R Core Team (2020). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org>

CUSTOMER SERVICE AGENTS' PERFORMANCE

- Ramjattan, V. A. (2019). Racializing the problem of and solution to foreign accent in business. *Applied Linguistics Review*. Advanced Online Publication.
<https://doi.org/10.1515/applirev-2019-0058>
- Reber, R., & Greifeneder, R. (2017). Processing fluency in education: How metacognitive feelings shape learning, belief formation, and affect. *Educational Psychologist*, 52, 84–103. <http://doi.org/10.1080/00461520.2016.1258173>
- Revelle, W. (2021). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University, Evanston, Illinois. R package version 2.1.3.
<https://CRAN.R-project.org/package=psych>.
- Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *The Canadian Modern Language Review*, 65, 395–412.
<https://doi.org/10.3138/cmlr.65.3.395>
- Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, 38, 439–462. <https://doi.org/10.1093/applin/amv047>
- Sanchez, C. A., & Khan, S. (2016). Instructor accents in online education and their effect on learning and attitudes. *Journal of Computer Assisted Learning*, 32, 494–502.
<https://doi.org/10.1111/jcal.12149>
- Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, 14, 332–348.
https://doi.org/10.1207/s15327663jcp1404_2
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.

CUSTOMER SERVICE AGENTS' PERFORMANCE

- Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory-2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality, 68*, 69–81.
<https://doi.org/10.1016/j.jrp.2017.02.004>
- Tavakoli, P., & Hunter, A.-M. (2017). Is fluency being “neglected” in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teaching Research, 22*, 330–349. <http://doi.org/10.1177/1362168817708462>
- Tavakoli, P., Nakatsuhara, F., & Hunter, A.-M. (2020). Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal, 104*, 169–191.
<https://doi.org/10.1111/modl.12620>
- Timming, A. R. (2017). The effect of foreign accent on employability: A study of the aural dimensions of aesthetic labour in customer-facing and non-customer-facing jobs. *Work, Employment and Society, 31*, 409–428. <https://doi.org/10.1177/0950017016630260>
- Whitaker, J., Krishnan, M. S., Fornell, C., & Morgeson, F. (2020). How does customer service offshoring impact customer satisfaction? *Journal of Computer Information Systems, 60*, 569–582: <https://doi.org/10.1080/08874417.2018.1552091>
- Wilt, J., & Revelle, W. (2017). Extraversion. In T. A. Widiger (Ed.), *The Oxford handbook of the Five Factor Model* (pp. 57–81). Oxford University Press.

CUSTOMER SERVICE AGENTS' PERFORMANCE

Appendix A

Background and Personality Questionnaire

Complete a brief questionnaire about you.

1. Your gender: [options: female, male, other]
2. Your age:
3. Your country:
4. Languages you speak fluently (write down all):
5. Your native (first) language or languages:
6. Your highest level of education: [options: high school or secondary school, college or university bachelor degree, master degree, PhD or other doctoral degree]
7. On a scale of 0–100, indicate your English language ability: *extremely bad* (0) – *extremely good* (100)
8. On a scale of 0–100%, if you use English in your daily life, indicate how often you communicate in English with native English speakers (as opposed to native speakers of other languages): *never* (0%) – *all the time* (100%)
9. On a scale of 0–100%, indicate how often you contact customer service/support for a product or service you receive in your daily life: *never* (0%) – *all the time* (100%)

CUSTOMER SERVICE AGENTS' PERFORMANCE

Select a rating next to each of the following statements to indicate the extent to which you agree or disagree with that statement.

1 Disagree strongly 2 Disagree a little 3 Neutral; no opinion 4 Agree a little 5 Agree strongly

1. I am someone who... tends to be quiet.
2. I am someone who... is compassionate, has a soft heart.
3. I am someone who... tends to be disorganized.
4. I am someone who... worries a lot.
5. I am someone who... is fascinated by art, music, or literature.
6. I am someone who... is dominant, acts as a leader.
7. I am someone who... is sometimes rude to others.
8. I am someone who... has difficulty getting started on tasks.
9. I am someone who... tends to feel depressed, blue.
10. I am someone who... has little interest in abstract ideas.
11. I am someone who... is full of energy.
12. I am someone who... assumes the best about people.
13. I am someone who... is reliable, can always be counted on.
14. I am someone who... is emotionally stable, not easily upset.
15. I am someone who... is original, comes up with new ideas.

CUSTOMER SERVICE AGENTS' PERFORMANCE

Appendix B

Screenshot of the Rating Interface in the Agent-Only Condition

Part B. Rate a customer service representative (agent)

Please listen to a two-minute audio recording of a customer service representative (agent) interacting with a customer. We have removed all confidential information (such as names and places), so you may occasionally hear some unexpected gaps.

Because we wanted you to focus on the speech of the customer service representative (agent), **we have also replaced all speech by the customer (caller) with silence.**

Although the audio will sound a bit strange, pay attention only to the performance of the **customer service representative (agent)** and evaluate the **agent** using the 0-100 rating scales below. Make sure you listen all the way to the end of the file. You can play the file as many times as you wish.



agent is very difficult to understand (0) _____ agent is very easy to understand (100)

agent has a strong foreign (non-English sounding) accent (0) _____ agent has no or little accent (100)

agent has many hesitations or pauses (0) _____ agent has few or no hesitations or pauses (100)