Article

Journal of Cross-Cultural Psychology
2017, Vol. 48(3) 369–383

DOI: 10.1177/0022022116687852
journals.sagepub.com/home/jcc
SAGE

The Social Costs in Communication Hiccups Between Native and Nonnative Speakers

Terry Kit-fong Au[1], Annie Fong-pui Kwok[1], Lester Chun-pong Tong[2], Liao Cheng[3], Hannah Man-yan Tse[1], and Sun-Ah Jun[4]

Abstract

It is well-established that native speakers perceive nonnative speakers with strong foreign accents, compared with those with a more nativelike accent, as less intelligent and competent, less ambitious and dependable as coworkers, and less comfortable around native speakers. But little is known about how nonnative speakers themselves are affected when communication hiccups—often due to incorrect or accented pronunciations—occur in their conversations with native speakers. In this experiment, mispronunciations of an English word were elicited from native Chinese speakers in phone conversations via the Internet with an American English speaker, who then either asked for clarification of the word or showed no confusion about the word but asked about something else. Chinese speakers' reactions were measured using a combination of self-reports, facial affective coding, and skin-conductance responses. When the American asked for clarification—compared with when he did not—Chinese speakers were left feeling more anxious, embarrassed, and unsure of their English abilities, as well as feeling less positive about the American, finding him less attractive socially and their conversation with him less enjoyable.

Keywords

communication, interpersonal relationships, language

With globalization, conversations between native and nonnative speakers have become increasingly commonplace. First impressions are often formed rapidly in such cross-cultural contact, perhaps based on just a few utterances. According to research on such "thin slices of the behavioral stream" (Ambady, Bernieri, & Richeson, 2000; Ambady, Krabbenhoft, & Hogan, 2006), first impressions can be quite predictive or accurate, especially for certain personality traits (Kenny, Albright, Malloy, & Kashy, 1994). Understandably, for better or worse, we come to rely

[1]The University of Hong Kong, China
[2]Stanford University, CA, USA
[3]Harvard University, Cambridge, MA, USA
[4]University of California, Los Angeles, USA

Corresponding Author:
Terry Kit-fong Au, Department of Psychology, The University of Hong Kong, Centennial Campus, Pokfulam Road, Hong Kong, China.
Email: terryau@hku.hk

on them as a heuristic in negotiating our social world. For cross-cultural person perception, it is well-established that speakers with nonstandard accents can suffer serious social costs even when they are able to communicate their main ideas successfully (Gluszek & Dovidio, 2010). For example, native speakers judge them—compared with those with a more standard (i.e., native-like) accent—to be less intelligent and less competent (Bresnahan, Ohashi, Nebashi, Liu, & Shearman, 2002), less ambitious and less dependable as coworkers, and less at ease around native speakers (Cargile & Giles, 1997; Wible & Hui, 1985). As revealed by a meta-analysis on the effects of nonnative speakers' accents (Fuertes, Gottdiener, Martin, Gilbert, & Giles, 2012), speakers with a standard accent are rated much more positively than those with a nonstandard accent on scales of apparent status (how educated, intelligent, or successful they are), dynamism (how active or lively), and solidarity (how similar to the listener, how attractive, kind, or trustworthy). Importantly, negative bias against speakers with a strong foreign accent is often implicit and unconscious (Pantos & Perkins, 2013), rendering it difficult to overcome. As a consequence, native speakers may not always take nonnative speakers seriously (Goffman, 1963; Graham & Requejo, 2009). They may even discriminate against them in the workplace and in courtrooms (Lippi-Green, 1994; Matsuda, 1991).

While much is known about how native speakers' feel about nonnative speakers with a strong foreign accent, much less is known about the reverse. But most of us do not like making mistakes in front of other people, and, in particular, nonnative speakers understandably do not like making mistakes in their interactions with native speakers. And if they are aware that they have made one, they may well have negative feelings about the experience. Indeed, communication apprehension (due to inability to express oneself or to understand another person in a nonnative language) and fear of negative social evaluation of one's nonnative speech have been hypothesized (Horwitz, Horwitz, & Cope, 1986) and shown to be two key components of foreign-language anxiety (MacIntyre & Gardner, 1989). Nonnative speakers' reactions to their interactions with native speakers range from hurt feelings to social anxiety, and from a sense of their own inadequacy to a resentment of native speakers who they may feel are prejudiced against them because of their language deficiencies (Goto, Gee, & Takeuchi, 2002; Lee & Rice, 2007; Stephan & Stephan, 1985). In fact, some have proposed that, to avoid being stigmatized, nonnative speakers may say as little as possible in conversations with native speakers (Gardner, 1979) or even shy away from having such conversations at all (Derwing & Rossiter, 2002).

In this study, we set out to test systematically (a) how nonnative speakers react to interactions with native speakers, and (b) whether the behavior of native speakers can affect these reactions in predictable ways. After all, nonnative speakers may be more or less aware of, and concerned about, their mistakes depending on how native speakers respond to them.

When native speakers are confused by something nonnative speakers say, they often ask for clarification or repetition (Lindemann, 2002). Alternatively, they may feign understanding and wait for additional context to clear things up (Clark & Wilkes-Gibbs, 1986). A native speaker's simple request to a nonnative speaker for clarification (e.g., "Are you talking about a cup or a cub?") could be perceived as too blunt and inconsiderate. A more accommodating strategy might be to feign understanding and wait for more information without calling attention to a likely mispronunciation. According to the Communication Accommodation Theory (Soliz & Giles, 2014), perceived underaccommodation (e.g., asking for clarification bluntly) can alienate the addressee from the speaker psychologically (Gasiorek, 2013; Gasiorek & Giles, 2013). In our study, we compared the impact of these two common strategies—asking for clarification versus feigning understanding—on nonnative speakers' moods and self-perceptions, as well as on their feelings about, and perception of, the native speakers they talk to.

Note that requests for clarification may have social consequences beyond transient negative affect. Bad first impressions of a native speaker as a conversation partner may be particularly difficult to shake (e.g., Delgado, Frank, & Phelps, 2005) and may decrease the likelihood of future interaction and friendship with that conversation partner. These subtle social threats may even sharpen biases against native speakers in general (Riek, Mania, & Gaertner, 2006), which could lead to prejudice and intergroup conflict.

In this study, we asked native Chinese speakers in Hong Kong to talk with an American English speaker in Chicago (a confederate in the experiment) via voice call on Google Chat. The Chinese speakers were asked to role-play telemarketing trainees to promote a new product (i.e., toy bear cubs). During the conversation with the American, mispronunciations of the English word *cub* were elicited from the Chinese speakers. In every conversation, the American then either (a) asked two times for clarification of the word or (b) showed no confusion about the word but interrupted also twice to ask about the product being promoted. Using paper-and-pencil self-reports, we compared how the Chinese speakers rated their mood and their own English proficiency before the conversation and after it. We also evaluated their skin conductance and videotaped facial affect in response to the American's questions. After the conversation, we measured their feelings about their conversation partner.

Hypothesis 1: If a native speaker asks for clarification of mispronounced words, instead of pretending to understand them, nonnative speakers will be more likely to experience negative emotions—as revealed by negative facial affect, psychophysiology such as skin conductance (i.e., sweating), and self-ratings of mood.
Hypothesis 2: They will also make negative evaluations, not only of their own language skills but also of the native speaker (i.e., the conversation partner), and their conversations.

Method

Participants

Ninety adult native speakers of Chinese (Cantonese or Mandarin) participated with written consent (44 men and 46 women), a sample size that should have a power of .8 for detecting medium-size effects. The participants had studied English for more than 12 years, starting around age 6. They were full-time students at the University of Hong Kong, which uses English as the language of instruction and has the highest university-entrance requirements in the city for English language abilities. They received either a half-hour research credit or HK$30 (about US$4) for participating in the experiment. The participants were randomly assigned to either a "Confusion" condition (n = 45; 21 men, 24 women) or a "No Confusion" condition (n = 45; 23 men, 22 women).

Apparatus

Each participant sat in an armchair facing a MacBook Pro computer in a testing room with air-conditioning set at 25.5°C. The experimenter sat nearby facing a Fujitsu laptop computer that was connected via Bluetooth to a biofeedback device—ProComp Infiniti Encoder (SA7500). The MacBook Pro was connected to the Internet and supported Google Chat for a conversation between the participant in Hong Kong and the American confederate in Chicago. A script telling the participant what to say was displayed in PowerPoint on the MacBook, which was also used to audio-record the conversation and video-record the participant's facial expressions with iMovie.

The biofeedback device was used for measuring skin conductance, with electrodes attached to the middle and ring finger of the nondominant hand (Scerbo, Freedman, Raine, Dawson, & Venables, 1992), connected to ProComp5 biofeedback equipment, and recorded with BioGraph Infiniti software (Thought Technology, Montreal, Canada). Skin conductance is conventionally used to assess stress levels by measuring sympathetic nervous system activity (Boucsein, 2012), with larger skin-conductance response (SCR) amplitudes indicating greater stress (Lin, Lin, Lin,

& Huang, 2011). SCRs during three segments of the conversation were identified using the procedure implemented in AcqKnowledge (BIOPAC Systems, Goleta, California).

Each participant's video-recorded facial expressions were coded for displays of anxiety and embarrassment using Ekman and Friesen's (1978) Facial Action Coding System. Anxiety cues include more and faster blinking, widened and tensed eyes, raised and drawn eyebrows, forced smiles, grimacing, and increased self-touching (e.g., Tomkins, 1991), whereas embarrassment cues include averted or shifting gazes, nervous smiles, face touching, speech disturbances, and shifting body posture (Keltner & Buswell, 1997). Overall negative and positive affect were also coded.

Procedure

Briefing. We told participants we were conducting a study on telemarketing conversations, and the experimenter—a young Chinese-speaking woman—asked each one to role-play a telemarketing trainee in Hong Kong talking to an American via Google Chat. A PowerPoint script for the participant to follow was displayed on the MacBook computer: greet the American, read (with enthusiasm) a list of selling points for the product being promoted (namely, toy bear cubs), answer any questions the American might have, and ask the American for his email address so he could be contacted in the future. The participant was told that the conversation would be audio- and video-recorded with the MacBook computer's built-in camera and that sensors would be strapped on two fingertips on the nondominant hand to collect physiological data. Afterward, written consent was obtained.

Preconversation measures. Participants were run one at a time in Hong Kong. They first rated their current mood with respect to four emotional states (interested, satisfied, embarrassed, annoyed) using visual analogue scales (VAS; Ahearn, 1997). Each VAS consisted of a 100 mm horizontal line with a label on each end. For example, the VAS for "embarrassed" had "0% (not embarrassed at all)" on the left end and "100% (embarrassed)" on the right end. Participants marked a point on the line to indicate how embarrassed they felt at that moment. In addition, they rated their overall English abilities compared with other students in their year at the university using a 100 mm VAS ranging from "0% (worst)" to "100% (best)." Next, the experimenter strapped two skin-conductance sensors to the middle and ring fingers of the participant's nondominant hand and began baseline skin-conductance recording. The experimenter then left the testing room to let the participant learn the telemarketing script in private.

Conversation. After 5 min, the experimenter returned, started audio- and video-recording using iMovie, called the American "potential buyer" via Google Chat (audio only), and asked the participant to start talking to the American, who in real life was a Chinese American undergraduate attending the University of Chicago at the time of data collection.

The experimenter then left the testing room again, this time taking along the Fujitsu laptop (with its cable keeping the door open by a tiny crack), and monitored the conversation from outside the room. To synchronize the skin-conductance recording and audio- and video-recording for data analysis, the experimenter time-stamped four events: (a) the first "Hello," (b) the onset of the American's first interruption (with a question about the mispronounced word or something about the product being promoted), (c) the onset of the American's second interruption, and (d) the last "Bye" of the conversation.

To reliably elicit mispronunciations, we capitalized on the tendency for native Chinese speakers to pronounce voiced stop consonants in English (/b,d,g/) as unaspirated voiceless stops [p,t,k] (as in ***spin, stop, skip***, i.e., with no vocal cord vibration during the stop closure and no puff of air at release), regardless of the location of the stop in a word—probably due to interference from

Chinese phonology (Flege & Wang, 1989). By contrast, native English speakers generally produce /b,d,g/ as voiced stops in all locations except the phrase-initial position (Ladefoged, 2005). Moreover, although a word-final voiced stop in English tends to be devoiced during the closure, and is thus similar to a voiceless stop, the vowel is longer before voiced stops than before voiceless stops (Raphael, 1972). As Chinese does not have a voicing contrast in obstruents, all voiced stops are produced as voiceless stops, and the vowel duration before a stop is not manipulated.

To check if participants had indeed mispronounced the target word *cub* as "cup" (or something else), we looked at their first three mentions of *cub* (which occurred before the American confederate asking any questions)—extracting four-word speech segments from their recordings, with one word preceding *cub* and two words following it (e.g., " . . . toy cub is one . . . "). We then used E-Prime 2.0 software to present these 270 segments—three each from the 90 participants—in randomized order to three native speakers of American English in California, who were blind to the group assignment of the participants and were asked to judge independently in each case whether the participant had said "cub," "cup," or something else. Only one participant was judged by all three native English speakers as having pronounced *cub* correctly in all three speech segments; the participant was in the experimental condition that clarification was sought by the American confederate. Because this study focused on the effects of requests for clarification about a mispronounced word, we excluded him from further data analyses. [For the 270 speech tokens, 12% were judged by all three native English speakers as "cub," 32% were judged as "cup," and 1% as something else (e.g., "cop," "cube" but with an unvoiced stop). The native English speakers disagreed on the remaining speech tokens (55%), which were judged to be mispronunciations by at least one of the three native speakers.]

In the telemarketing script, the first two selling points made sense for both toy cubs and toy cups (e.g., "Parents like the realistic design of this toy cub, and children love to play tea party with it"). But after that, the selling points applied to a cub but not a cup (e.g., "The toy cub makes cute sounds if your child touches its head"). The American used a parallel version of the experimental script that specified when to give back-channel responses (e.g., "Uh huh," "Okay" after every two selling points), when to ask questions, and what to ask. Although he knew the hypotheses of the study, we reviewed the audio-recordings and confirmed that he successfully followed the protocol and adhered to the verbal script accurately for all participants.

Participants were randomly assigned to one of two conditions. In the "Confusion" condition, after the third selling point, which was the first time it did not make sense to be talking about a cup ("It's made of 100% organic cotton"), the American asked, "Wait. Are you talking about a cup or a cub?" Three selling points later, he said, "Wait. It sounds like you're saying *cup* not *cub*. Do you mean *cub*?" The American's questions could create further (and genuine) confusion if the participant did not hear the difference between "cub" and "cup" in American English (both ending with an unaspirated bilabial stop but contrasting in voicing if not strongly released—a contrast not found in the phonology of Cantonese, Mandarin, or virtually any other Chinese dialect). In the "No Confusion" condition, the American first asked, "Wait. Is it really safe for young children?" Later he said, "Wait. Do they come in different sizes, and are they easy to carry around?" The conversation ended after the participant answered the American's questions, asked for his email address to use for future promotional contact, and said goodbye.

Postconversation measures. The participant repeated the VAS self-ratings on current mood (interested, satisfied, embarrassed, annoyed) and overall English abilities, using the same procedure as in the preconversation measures. In addition, the participant rated the American using the Social Attraction scale (McCroskey & McCain, 1974), which consists of five positive statements (e.g., "He would be pleasant to be with") and five negative ones (e.g., "He just wouldn't fit into my circle of friends"). We added two statements specifically about the telemarketing conversation (one positive—"I think the experience just now was quite pleasant"—and one negative—"I did not really enjoy the conversation just now"). These 12 statements were rated on Likert scales ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

Results

Preliminary analysis revealed no gender differences in any of the results to be reported here, so gender was not considered in further analysis. As noted earlier, one participant in the "Confusion" condition was excluded from further analysis because, according to all three native-English-speaking coders, he correctly pronounced *cub* all three times prior to the American confederate's questions.

Potential Confound: Confederate's Nonverbal Cues

Recall that our review of the conversation audio-recordings confirmed that the American confederate had faithfully adhered to the verbal script that specified when to give back-channel responses (e.g., "Uh huh," "Okay" after every two selling points), when to ask questions, and what to ask for all participants. But given the confederate's knowledge of the hypotheses, in theory his nonverbal cues such as voice quality and intonation contour could influence the participants' feelings about him, themselves, and the conversations. That is, such nonverbal cues might be less polite in the "Confusion" condition than in the "No Confusion" condition—creating a potential confound of the verbal content (i.e., asking for clarification vs. feigning understanding, respectively).

Voice unpleasantness ratings of the interruptions. To address this concern, we extracted from the conversation audio-recordings the confederate's two interruptions of each participant (1 s to 4 s each) and lowpass-filtered them, using the Praat speech analysis software (Version 5.3; Boersma & Weenink, 2012). The filtering (with frequency from 50 Hz to 400 Hz and smoothing at 100 Hz) removed the phonemic information and rendered the speech unintelligible (see Knoll, Uther, & Costall, 2009, for more information about lowpass-filtering speech). The filtered tokens sounded like muffled speech with the intonation contour and voice quality preserved but the verbal content obscured. We then asked five trained independent raters to judge the unpleasantness of the confederate's voice for each interruption audio clip on a 7-point Likert scale (1 = *not at all unpleasant*; 7 = *extremely unpleasant*), without any knowledge of the experimental design and the source of the audio clips. The interrater reliability was acceptable (intraclass $R = .69$, $p < .001$). We had asked the confederate to deliver the interruption questions with natural intonation. As it turned out, the intuitively blunter questions in the "Confusion" condition (i.e., asking for clarification; $M = 4.4$; $SD = .42$) were delivered with a more unpleasant voice than the questions feigning understanding in the "No Confusion" condition ($M = 3.7$; $SD = .50$); $t(85) = 7.4$, $p < .001$. To assess the effect of the verbal content of the clarification questions, compared with the control questions, we controlled for the confounding nonverbal cues by including the voice unpleasantness rating as a covariate in all the analyses of covariance (ANCOVAs) comparing these two experimental conditions.

Reactions to Confusion in a Conversation

In both conditions, the American interrupted twice: once after the third selling point and again after the sixth selling point. In the "Confusion" condition, the American asked for clarification/repetition of the target word. In the "No Confusion" condition, he did not act concerned about the word and merely asked about the product. To assess how the American's confusion affected non-native speakers, we compared changes in their facial affect and skin conductance right after the American's interruptions occurred.

Facial affect. For each participant, five 10-s clips, without soundtrack, were extracted from the facial-expression videos captured by iMovie with the MacBook computer's built-in camera: (a) after the first "Hello," (b) just before the word "Wait" in the American's first interruption, (c) after the onset of "Wait," (d) just before "Wait" in the American's second interruption, (e) after the onset of the second "Wait." Using E-prime 2.0 software, another five trained independent raters (who were not involved in voice unpleasantness ratings), blind to the group assignment of the participants, viewed all five clips from each participant twice—first to get a sense of the range of facial affect for that participant, and a second time, in the same random sequence and using Ekman and Friesen's (1978) Facial Action Coding System as described earlier, to rate each clip on a 7-point Likert scale from 1 (*not at all*) to 7 (*extremely*) in terms of (a) anxiety, (b) embarrassment, (c) overall negative affect, and (d) overall positive affect. The overall negative and positive affect ratings of a clip were based on a rater's global impression of affect displayed in that clip. The interrater reliability was good (intraclass *R*s: .80, .81, .81, and .85 for anxiety, embarrassment, overall negative affect, and overall positive affect, respectively).

To test whether the "Confusion" condition and the "No Confusion" condition were comparable at baseline, a series of analyses of variance (ANOVAs) was conducted on facial affect ratings for the 10 s after the first "Hello." There were no significant differences between the two conditions during the initial greeting in all four measures (anxiety, embarrassment, overall negative affect, overall positive affect, Fs(1, 87) < 1.1, ps > .3, η_p^2s < .02, for means, see the "Preonset" columns for the two conditions in Table 1). These results are important for two reasons. First, they suggest that the random assignment of participants to experimental conditions was effective. Second, although the experimenter knew the hypotheses of this study, her interaction with the participants (i.e., in the initial briefing and administration of preconversation measures) did not seem to have biased the participants. After she had left the participants alone in the lab to talk to the American via Google Chat, the two conditions still did not differ in facial affect ratings for the 10 s after the first "Hello."

We then compared the effects of the American's interruptions in the "Confusion" and the "No Confusion" conditions with respect to changes in facial affect from the 10 s immediately before to the 10 s immediately after the interruption onsets.

ANCOVAs comparing the two conditions were conducted for both interruptions and for each of the four facial affect measures (anxiety, embarrassment, overall negative affect, overall positive affect), with the affect rating of the 10-s clip after the interruption onset as the dependent measure, and the rating of the 10-s clip before the interruption onset and the voice unpleasantness rating as the covariates.

Table 1 presents the mean facial affect ratings and standard deviations before and after the onset of the American English speaker's first and second interruptions. Taking into account the affect displayed prior to the respective interruption onsets and voice pleasantness, we found that participants in the "Confusion" condition displayed more negative affect—anxiety, embarrassment, and overall negative affect, Fs(1, 83) > 8.1, ps ≤ .005, η_p^2s > .08—and less overall positive affect—Fs(1, 83) > 11.2, ps = .001, η_p^2s ≥ .12—immediately after the interruptions, compared with those in the "No Confusion" condition.

These results suggest that the participants in the "Confusion" condition were distressed by the American's apparent confusion—as revealed by his requests for clarification and repetition, even after controlling for voice unpleasantness—above and beyond the general stress of dealing with questions from an American conversation partner, which was common across both conditions.

SCRs. To assess the stress levels—using skin-conductance data—at baseline and after each of the two interruptions, we defined three 10-s segments starting 4 s after the American first said "Hello," 4 s after the end of his first interruption, and 4 s after the end of his second interruption. These windows were selected to avoid orienting responses, which should not differ as a function

Table 1. Mean Facial Affect Ratings Pre- and Postonset of the Native Speaker's Interruptions (Standard Deviations).

	Confusion condition		No confusion condition		Condition difference in change		
Facial affect ratings	Preonset	Postonset	Preonset	Postonset	$F(1, 83)$	p	η_p^2
Anxiety							
Interruption 1	1.67 (.54)	3.38 (0.80)	1.56 (0.47)	2.60 (0.99)	9.66	.003	.104
Interruption 2	1.80 (.55)	3.54 (0.79)	1.75 (0.57)	2.69 (0.87)	11.36	.001	.120
Embarrassment							
Interruption 1	1.09 (.24)	3.08 (1.01)	1.09 (0.19)	1.93 (0.92)	18.32	<.001	.181
Interruption 2	1.12 (.27)	3.62 (1.19)	1.16 (0.37)	2.07 (1.0)	25.55	<.001	.235
Overall negative affect							
Interruption 1	1.71 (.43)	3.01 (0.94)	1.65 (0.36)	2.27 (0.89)	8.14	.005	.089
Interruption 2	1.69 (.45)	3.21 (1.11)	1.67 (0.41)	2.25 (0.77)	12.10	.001	.127
Overall positive affect							
Interruption 1	2.72 (.86)	2.34 (0.66)	2.77 (1.12)	2.94 (1.0)	11.27	.001	.120
Interruption 2	2.55 (.69)	2.29 (0.42)	2.81 (1.1)	2.77 (0.74)	11.55	.001	.122

Note. ANCOVAs compared the two conditions during the 10 s after each interruption onset, with respective preonset rating and voice unpleasantness rating score as covariates. The effect size measure was η_p^2 (partial eta^2). ANCOVA = analysis of covariance.

of the positive or negative affect induced in the two conditions (e.g., Bradley, 2009). Instead, SCRs during these periods probe affective intensity associated with responding to the American's interruptions. SCR amplitude was calculated by subtracting the onset from the peak value. Within each window, SCR amplitudes were averaged and normalized with a square root transform.

Table 2 presents the means and standard deviations of the participants' SCR amplitudes. An ANOVA revealed no significant difference in SCR amplitude during the initial greeting period between the "Confusion" condition (M = .75) and the "No Confusion" condition (M = .78), $F(1, 67) = .042$, $p = .84$, $\eta_p^2 = .001$, confirming that the random assignment of participants was effective, and that the experimenter's interaction with the participants prior to her leaving the testing room had no apparent bias favoring one condition over the other.

By contrast, with the confederate's voice unpleasantness of the interruptions included as a covariate, an ANCOVA showed that the SCR amplitude was significantly larger in the "Confusion" condition shortly after the first interruption (M = .63) than those in the "No Confusion" condition (M = .49), $F(1, 62) = 4.39$, $p = .04$, $\eta_p^2 = .066$. Although the SCR amplitude after the second interruption did not differ significantly between the two conditions, those in the "Confusion" condition were on average still numerally larger (M = .68) than those in the "No Confusion" condition (M = .56), $F(1, 54) = 2.33$, $p = .132$, $\eta_p^2 = .041$. These psychophysiological results converge with the facial affect results to suggest that, compared with requests for information that seemed like a natural part of successful communication (as in the "No Confusion" condition), the American English speaker's requests for clarification/repetition increased negative affect, reduced positive affect, and caused more stress for nonnative speakers of English.

Self-reports. Questionnaires were administered before and after the conversation. Both consisted of five visual analogue scales: four for affective states (interested, satisfied, embarrassed, annoyed) and one for self-perception of English proficiency. Table 3 presents the means and standard deviations of the self-ratings before and after the conversation with the American.

A series of ANOVAs on the five preconversation measures revealed no significant differences between the two conditions before the conversation in self-reported overall English

Table 2. Means (Standard Deviations) of the Nonnative Speakers' SCR Amplitude.

	Confusion	No confusion	*F* (*df*s)	*p*	η_p^2
Time points					
"Hello" (initial greeting)	.75 (.47)	.78 (.52)	0.042 (1, 67)	.838	.001
Interruption 1	.63 (.19)	.49 (.20)	4.39 (1, 62)	.040	.066
Interruption 2	.68 (.21)	.56 (.24)	2.33 (1, 54)	.132	.041

Note. The degrees of freedom varied across time points because participants who did not exhibit any SCRs were excluded. The proportions of participants thus excluded did not differ by condition (*p*s > .05). For "Hello" (initial greeting), ANOVAs compared the two conditions. For Interruptions 1 and 2, ANCOVAs compared the two conditions, with voice unpleasantness rating score as a covariate. SCR = skin-conductance response; ANCOVA = analysis of covariance.

Table 3. Mean Self-Ratings by Nonnative Speakers Before and After the Conversation (Standard Deviations).

	Confusion condition		No confusion condition		Condition difference in change		
	Before	After	Before	After	*F*(1, 83)	*p*	η_p^2
Current mood							
Embarrassed	30.8 (27.5)	50.8 (28.5)	34.2 (23.5)	37.7 (23.0)	7.26	.009	.080
Annoyed	15.9 (13.6)	21.8 (20.7)	23.3 (19.9)	19.8 (18.1)	4.66	.034	.053
Interested	68.7 (20.2)	60.5 (25.6)	65.7 (21.8)	68.3 (19.6)	1.20	.277	.014
Satisfied	69.1 (17.1)	58.5 (20.9)	62.9 (15.9)	68.0 (16.7)	4.32	.041	.050
English overall abilities							
	60.4 (18.0)	56.9 (21.0)	57.0 (18.0)	59.0 (18.3)	11.18	.001	.119

Note. ANCOVAs compared the two conditions after the conversation, with preconversation score and voice unpleasantness rating score as covariates. ANCOVA = analysis of covariance.

abilities, $F(1, 85) = .81$, $p = .37$, $\eta_p^2 = .009$, and two of the four self-reported mood measures, $Fs(1, 85) < .46$, $ps > .5$, $\eta_p^2 s < .01$, for feeling embarrassed and interested, although participants in the "No Confusion" condition were feeling more annoyed, $F(1, 85) = 4.12$, $p = .046$, $\eta_p^2 = .045$, and those in the "Confusion" condition were feeling marginally more satisfied, $F(1, 85) = 3.15$, $p = .079$, $\eta_p^2 = .035$. Overall, these results suggest that the random assignment was generally effective; there was no evidence that the experimenter's interaction had biased the participants favoring either condition.

ANCOVAs, with respective self-ratings before the conversation and voice unpleasantness rating as covariates, revealed that participants in the "Confusion" condition reported feeling more embarrassed and annoyed, but less satisfied after the conversation than their counterparts in the "No Confusion" condition, $Fs(1, 83) > 4.3$, $ps \leq .041$, $\eta_p^2 s \geq .05$. The two conditions did not have significant difference in interest, $F(1, 83) = 1.2$, $p = .277$, $\eta_p^2 = .014$, although the pattern of means was consistent with the prediction. Moreover, those in the "Confusion" condition lost confidence in their overall English abilities (dropping from 60.4 before the conversation with the American to 56.9 afterward), whereas those in the "No Confusion" condition went up from 57.0 to 59.0, $F(1, 83) = 11.18$, $p = .001$, $\eta_p^2 = .119$.

In addition, participants rated both the American himself—using the 10-item Social Attraction scale (McCroskey & McCain, 1974)—and their conversations with him. Table 4 presents the means and standard deviations of these ratings. ANCOVAs, with voice

Table 4. Means (Standard Deviations) of the Nonnative Speakers' Ratings of the American's Social Attractiveness and of the Conversation.

	Confusion	No confusion	$F(1, 84)$	p	η_p^2
Social Attraction Scale					
Social Attraction Scale Score[a]	3.93 (1.0)	4.54 (0.76)	4.96	.029	.056
Conversation experience					
I did not really enjoy the conversation just now	3.73 (1.5)	2.78 (1.1)	8.87	.004	.096
The experience just now was quite pleasant	4.68 (1.4)	5.20 (.84)	1.92	.169	.022

Note. ANCOVAs compared the two conditions, with voice unpleasantness rating score as a covariate. ANCOVA = analysis of covariance.
[a]With reverse scoring for the negative statements.

unpleasantness rating as covariate, revealed that participants in the "Confusion" condition rated the American as significantly less socially attractive (10-item scale score: $M = 3.93$) than did those in the "No Confusion" condition ($M = 4.54$), $F(1, 84) = 4.96$, $p = .029$, $\eta_p^2 = .056$. Also, participants in the "Confusion" condition strongly indicated that they did not enjoy the conversation compare with those in the "No Confusion" condition ($M = 3.73$ and 2.78 respectively), $F(1, 84) = 8.87$, $p = .004$, $\eta_p^2 = .096$. There was a trend, although nonsignificant statistically, that the participants in the "Confusion" condition found the conversation with the American less pleasant than those in the "No Confusion" condition ($M = 4.68$ and 5.20 respectively), $F(1, 84) = 1.92$, $p = .169$, $\eta_p^2 = .022$.

Notwithstanding some minor exceptions (e.g., a significant difference between the two experimental conditions in preconversation self-reports of "feeling annoyed"), most of the main dependent measures analyzed for our research questions revealed a coherent pattern of findings. Together they suggest that if native speakers want to make a good impression on conversation partners who are not native speakers—and/or if they want to help those partners feel at ease—they should not ask directly for clarification or repetition of any mispronunciations, but instead should feign understanding and wait for additional context to eliminate any confusion they may have.

Discussion

Prior research on the social costs of nonnative accents has focused almost exclusively on the perspective of native speakers, who tend to see nonnative speakers with strong foreign accents—compared with those with a more nativelike accent—as less intelligent and competent, less ambitious and dependable as coworkers, and less at ease around native speakers. But how do nonnative speakers feel?

Our findings suggest that nonnative speakers are quite sensitive about any confusion their language difficulties may seem to cause. If a native speaker asks them directly for clarification or repetition of something they have said, they are likely to feel embarrassed, annoyed, and less satisfied with the conversation, and their faces are likely to show more anxiety, embarrassment, and overall negative affect, but less overall positive affect. Their hands will probably sweat more, and they may well lose confidence in their second-language abilities. The native speakers who ask for clarification/repetition pay a price too. Nonnative speakers are likely to find them less attractive socially than they would otherwise and to consider conversation with them less enjoyable. Such bad first impressions can reduce the likelihood of future interaction and friendship, as was evidenced in the Social Attraction Scale (e.g., "We could never establish a personal friendship with each other.").

While the present results strongly suggest that nonnative speakers are sensitive to how native speakers judge their linguistic abilities, they also afford alternative interpretations. For example, the nonnative speakers' reactions to the native speaker's clarification questions may have little to do with communication hiccups perceived by nonnative speakers to be caused by their own mispronunciations or nonstandard accents. Such reactions could have been general to all speakers—native as well as nonnative—when confronted with a request for them to clarify what they have just said. Further research is needed to evaluate how general the social costs of communication hiccups documented in this study will prove to be, and whether such social costs might be greater for contact between conversation partners from different cultures than from the same culture.

We are also mindful that the nonnative speakers' strongly negative reactions to the native speaker's clarification requests could be attributed to perceived rudeness. In everyday conversations, most people would let others' communication errors go apparently unnoticed if at all possible by adopting the "Let it pass" rule (Garfinkel, 1967). Because this is such a common practice, deviations from it (as in the "Confusion" condition in this study) would tend to be seen as nonaccommodative—and indeed, sometimes even as a sign of relationship breakdown (Brown & Ragan, 1987; Hopper, 1990; Schegloff, 1992). The perceived rudeness of the native speaker can be conveyed via the verbal content of the clarification questions as well as nonverbal cues such as unpleasant intonation and voice quality. As noted earlier, the confederate was instructed to deliver verbal script using a natural voice, and apparently he did. That is, he was less patient and positive when he asked for a pronunciation clarification, as most people would naturally, than when he asked about the product being promoted by the nonnative speakers. This highlights the potential of social faux pas of asking for clarification, namely, we might be inadvertently rude and impatient when we do it. Importantly, even after we have statistically controlled for voice unpleasantness of the confederate's interruptions, our main predictions held up well across a suite of outcome measures. The ANCOVA results, with voice unpleasantness as a covariate, demonstrated that the scripted questions themselves could create negative feelings in those being questioned.

These results also open up interesting vistas for future research. According to the Communication Accommodation Theory (Soliz & Giles, 2014), there are many tools to be more accommodative to our conversation partners. Even if we have to ask for clarification immediately, we can do so more politely (e.g., with apology and/or blaming one's own inattention). Future research can vary the politeness of clarification requests not only to test this theory, but also to find ways to reduce the social costs of clarification requests in communication hiccups.

It also remains to be seen how well these results generalize to other sample and task characteristics. For example, members of collectivistic cultures may be more susceptible to social anxiety, and therefore respond less adaptively to potentially negative social evaluations (Heinrichs et al., 2006). Also, participants in our study may have been especially vulnerable to any negative feedback about their English proficiency, which they had reason to believe was pretty good—good enough, after all, to meet or surpass the rather high university-entrance requirement. Communication hiccups with a native speaker of English may have bothered these participants more than they would bother nonnative speakers with poorer English skills who expected less of themselves in the first place. Perhaps, the higher you rise, the more distressing it is when you fall even a little bit. It remains to be seen how well the present findings generalize to nonnative speakers with lower English proficiency.

Situational factors may also affect the social costs of communication hiccups from the perspective of the nonnative speakers. For example, when the stakes of communicative success are not very high (e.g., in language classes), nonnative speakers may be more likely to respond positively to linguistic challenges. It is also possible that the use of audio chat (as opposed to video chat) exacerbated the observed effects. Phone conversations are generally more challenging because they lack the visual cues that can facilitate communication. Future research can explore whether the phenomenon uncovered here is attenuated in a face-to-face interaction.

Given our findings, what should native speakers (or any speakers) do? If they want to put non-native speakers (or any conversation partners) at ease and/or if they want the latter to like them, they may be wise to feign understanding of confusing language and wait for additional context to provide clarification—as long as such "let it pass" politeness would not lead to terrible consequences (e.g., miscommunication in an emergency or crisis situation). As seen in the "No Confusion" condition, when a native speaker adopts this more tactful tactic, nonnative speakers are apt to be happier (e.g., to feel more positive affect and less anxiety, distress, and doubt about their second-language abilities), to like the native-speaking conversation partner better, and to enjoy the interaction more.

Of course, withholding linguistic feedback may give nonnative speakers the impression that they are expressing themselves clearly, even if they are not, and that could hinder second-language acquisition. Indeed, participants in the "No Confusion" condition became more confident in their second-language ability. Without corrective feedback, nonnative speakers may be less likely, or at least slower, to correct mispronunciations, particularly for contrasts that do not exist in their first language (e.g., Mackey, 2006). Nonnative speakers will do well to see clarification requests from native speakers in a new and more positive light—as learning moments rather than embarrassing moments.

On the other hand, if nonnative speakers get positive vibes from interacting with a native speaker, they may be more willing to interact with other native speakers in the future (Clement, Baker, & MacIntyre, 2003). That is, if they decide that such interactions can be enjoyable, then, in a sort of self-fulfilling prophecy, they may continue to engage in them and find them enjoyable (Snyder, Tanke, & Berscheid, 1977). Through seeking out and enjoying opportunities to talk to native speakers, nonnative speakers may actually become more at ease with their nonnative language and, consequently, enjoy conversing with native speakers even more.

Appendix

Conversation Script

(Participant: in regular font; Confederate: in boldface)

- Hi, I am calling from the Smart Child Company.
- We contacted you earlier.
- Thank you for your interest in the products of our company.
- My name is XX (your name).
- I want to tell you about a wonderful toy cub available only from our company.
- This toy cub is one of "The 10 Best Toys Chosen by Children."
- Parents like the realistic design of this toy cub, and children love to play tea party with it.
- It is made of 100% organic cotton and does not contain any small parts or sharp edges.

Wait, are you talking about a cup or a cub? (Confusion)
Wait, is it really safe for young children? (No Confusion)

- [The Participant reacted with an unscripted answer.]
- Our customer survey found that children love to hug and sleep with their toy cub at home.
- Children can get very creative with the toy cub and toy furniture and clothes.
- The toy cub invites your child to play more by making cute sounds if your child presses its stomach or touches its head.

Wait, it sounds like you're saying cup not cub. Do you mean cub? (Confusion)
Wait, do they come in different sizes? Are they easy to carry around? (No Confusion)

- [The Participant reacted with an unscripted answer.]
- Would you like us to email you a coupon for getting a 50% discount when you place an order?
- If you are interested in the toy cubs, please visit our website www.smartchild.com to place an order. Have a good day/evening. Thank you.

Acknowledgments

We thank the participants and the research funding agencies for their support. We are grateful to our research assistants for data coding and analysis and Karen Ravn for feedback on earlier drafts.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was supported by the Research Grants Council of Hong Kong (HKU740511H) and the University of Hong Kong Seed Funding.

References

Ahearn, E. P. (1997). The use of visual analog scales in mood disorders: A critical review. *Journal of Psychiatric Research*, *31*, 569-579.

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. *Advances in Experimental Social Psychology*, *32*, 201-271.

Ambady, N., Krabbenhoft, M. A., & Hogan, D. (2006). The 30-sec sale: Using thin-slice judgments to evaluate sales effectiveness. *Journal of Consumer Psychology*, *16*, 4-13.

Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer* (Version 5.3). Retrieved from http://www.fon.hum.uva.nl/praat/

Boucsein, W. (2012). *Electrodermal activity* (2nd ed.). New York, NY: Springer.

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*, 1-11.

Bresnahan, M. J., Ohashi, R., Nebashi, R., Liu, W. Y., & Shearman, S. M. (2002). Attitudinal and affective response toward accented English. *Language & Communication*, *22*, 171-185.

Brown, M. H., & Ragan, S. L. (1987). Variations on a theme: An ethnographic case study of family blessings. *Research on Language and Social Interaction*, *21*, 115-141.

Cargile, A. C., & Giles, H. (1997). Understanding language attitudes: Exploring listener affect and identity. *Language & Communication*, *17*, 195-217.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1-39.

Clement, R., Baker, S. C., & MacIntyre, P. D. (2003). Willingness to communicate in a second language: The effects of context, norms, and vitality. *Journal of Language and Social Psychology*, *22*, 190-209.

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, *8*, 1611-1618.

Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System*, *30*, 155-166.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

Flege, J. E., & Wang, C. (1989). Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English /t/-/d/ contrast. *Journal of phonetics*, *17*, 299-315.

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology*, *42*, 120-133.

Gardner, R. C. (1979). Social psychological aspects of second language acquisition. In H. Giles & R. St. Clair (Eds.), *Language and social psychology* (pp. 193-220). Oxford, UK: Basil Blackwell.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Gasiorek, J. (2013). "I was impolite to her because that's how she was to me": Perceptions of motive and young adults' communicative responses to underaccommodation. *Western Journal of Communication*, *77*, 604-624.

Gasiorek, J., & Giles, H. (2013). Accommodating the interactional dynamics of conflict management. *International Journal of Society, Culture & Language*, *1*(1), 10-21.

Gluszek, A., & Dovidio, J. F. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, *14*, 214-237.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Englewood Cliffs, NJ: Prentice Hall.

Goto, S. G., Gee, G. C., & Takeuchi, D. T. (2002). Strangers still? The experience of discrimination among Chinese Americans. *Journal of Community Psychology*, *30*, 211-224.

Graham, J. L., & Requejo, W. H. (2009). Managing Face-To-Face International negotiations. *Organizational Dynamics*, *38*, 167-177.

Heinrichs, N., Rapee, R. M., Alden, L. A., Bögels, S., Hofmann, S. G., Oh, K. J., & Sakano, Y. (2006). Cultural differences in perceived social norms and social anxiety. *Behaviour Research and Therapy*, *44*, 1187-1197.

Hopper, R. (1990). Describing speech phenomena. *Annals of the International Communication Association*, *13*, 245-254.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*, 125-132.

Keltner, D., & Buswell, B. N. (1997). Embarrassment: Its distinct form and appeasement function. *Psychological Bulletin*, *122*, 250-270.

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the big five. *Psychological Bulletin*, *116*, 245-258.

Knoll, M. A., Uther, M., & Costall, A. (2009). Effects of low-pass filtering on the judgment of vocal affect in speech directed to infants, adults and foreigners. *Speech Communication*, *51*, 210-216.

Ladefoged, P. (2005). *A course in phonetics* (5th ed.). Wadsworth: Boston, MA.

Lee, J. J., & Rice, C. (2007). Welcome to America? International student perceptions of discrimination. *Higher Education*, *53*, 381-409.

Lin, H.-P., Lin, H.-Y., Lin, W.-L., & Huang, A. C.-W. (2011). Effects of stress, depression, and their interaction on heart rate, skin conductance, finger temperature, and respiratory rate: Sympathetic-parasympathetic hypothesis of stress and depression. *Journal of Clinical Psychology*, *67*, 1080-1091.

Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society*, *31*, 419-441.

Lippi-Green, R. (1994). Accent, standard language ideology, and discriminatory pretext in the courts. *Language in Society*, *23*, 163-198.

MacIntyre, P. D., & Gardner, R. C. (1989). Anxiety and second-language learning: Toward a theoretical clarification. *Language Learning*, *39*, 251-275.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, *27*, 405-430.

Matsuda, M. J. (1991). Voices of America: Accent, antidiscrimination law, and a jurisprudence for the last reconstruction. *Yale Law Journal*, *100*, 1329-1407.

McCroskey, J. C., & McCain, T. A. (1974). The measurement of interpersonal attraction. *Speech Monographs*, *41*, 261-266.

Pantos, A. J., & Perkins, A. W. (2013). Measuring implicit and explicit attitudes toward foreign accented speech. *Journal of Language and Social Psychology*, *32*, 3-20.

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *The Journal of the Acoustical Society of America*, *51*, 1296-1303.

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, *10*, 336-353.

Scerbo, A. S., Freedman, L. W., Raine, A., Dawson, M. E., & Venables, P. H. (1992). A major effect of recording site on measurement of electrodermal activity. *Psychophysiology*, *29*, 241-246.

Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American journal of sociology*, *97*, 1295-1345.

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, *35*, 656-666.

Soliz, J., & Giles, H. (2014). Relational and identity processes in communication: A contextual and meta-analytical review of communication accommodation theory. *Annals of the International Communication Association*, *38*, 107-144.

Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues*, *41*, 157-175.

Tomkins, S. S. (1991). *Affect imagery consciousness: The negative affects: Anger and anxiety* (Vol. 3). New York, NY: Springer.

Wible, D. S., & Hui, C. H. (1985). Perceived language proficiency and person perception. *Journal of Cross-Cultural Psychology*, *16*, 206-222.