Research in Higher Education, Vol. 33, No. 4, 1992

NONLANGUAGE FACTORS AFFECTING UNDERGRADUATES' JUDGMENTS OF NONNATIVE ENGLISH-SPEAKING TEACHING ASSISTANTS

Donald L. Rubin

In response to dramatic changes in the demographics of graduate education, considerable effort is being devoted to training teaching assistants who are nonnative speakers of English (NNSTAs). Three studies extend earlier research that showed the potency of nonlanguage factors such as ethnicity in affecting undergraduates' reactions to NNSTAs. Study 1 examined effects of instructor ethnicity, even when the instructor's language was completely standard. Study 2 identified predictors of teacher ratings and listening comprehension from among several attitudinal and background variables. Study 3 was a pilot intervention effort in which undergraduates served as teaching coaches for NNSTAs. This intervention, however, exerted no detectable effect on undergraduates' attitudes. Taken together, these findings warrant that intercultural sensitization for undergraduates must complement skills training for NNSTAs, but that this sensitization will not accrue from any superficial intervention program.

Observers who have been even moderately sensitive to the voices of North American college students in recent years can appreciate the depth of undergraduates' concerns about the quality of instruction offered at many campuses by nonnative English-speaking teaching assistants (NNSTAs). The NNSTA "crisis" is one of relatively few instructional issues in higher education that has captured the attention of the popular press (Hess, 1987; Verhovek, 1989). In several states, legislatures have directly intervened in policies regarding NNSTAs (e.g., "Instructors' Broken English Prompts Illinois Law," 1987; Thomas and Monoson, 1991).

To the thinking of the press and of the policymakers—and probably also of the undergraduate students—the root of the problem is NNSTAs' poor English-language proficiency (Bailey, 1984). Accordingly, many universities have re-

Donald L. Rubin, The University of Georgia. Address Correspondence to: Speech Communication and Language Education, The University of Georgia, Athens, GA 30602.

0361-0365/92/0800-0511$06.50/0

sponded by erecting procedures for screening NNSTAs' language skills (see discussions of these tests in Abraham and Plakans, 1988; Dunn and Constantinides, 1991; Gallego, Goodwin, and Turner, 1991; Heller, 1985; Yule and Hoffman, 1990). Those deemed deficient are channeled into remediation programs (Turitz, 1984). For the most part, these remedial workshops appear to focus on discrete language behaviors, especially pronunciation and use of idioms (Verhovek, 1989). This focus is not illogical, since many of the current groups of NNSTAs have had little exposure to English as a medium of oral communication. A more recent and educationally progressive approach combines language remediation with training in generic instructional communication skills and in cultural norms associated with American undergraduates (e.g., Althen, 1991; Byrd, Constantinides, and Pennington, 1989; Pica, Barnes, and Finger, 1990).

The common assumption of all these perspectives on NNSTAs is that it is the internationals who bear the onus for poor classroom communication, and therefore it is the internationals who must undergo transformation in the North American mold. An alternative perspective, in contrast, recognizes that communication is a reciprocal process. If communication outcomes are poor in classes taught by NNSTAs, then perhaps the responsibility ought to be shared alike among NNSTAs and the American undergraduate student body.

Scattered evidence warrants the plausibility of this alternative view, the view that the NNSTA "problem" is at least partially a problem of undergraduates' negative stereotyping. For example, Orth (1982) found that undergraduate students' ratings of their NNSTAs' speaking proficiency were only weakly related to expert ratings of the NNSTAs' language proficiency. Instead, undergraduate ratings of NNSTAs were biased by the grades they anticipated receiving from those NNSTAs. For many undergraduates, introductory courses in mathematics and the natural sciences have reputations as extremely inhumane courses designed to winnow out marginal students. And it is well known that NNSTAs are disproportionately assigned responsibility for such high-anxiety classes (Constantinides, 1987). In addition, with an increasing proportion of NNSTAs originating from East Asia (Zikopoulos, 1988), it is likely that powerful ethnic stereotypes also factor into negative perceptions of language proficiency and teaching ability (Brown, 1988).

Rubin and Smith (1990) undertook to explore the role of NNSTAs' ethnicity, the subject matter about which they lectured, and their level of accentedness as determinants of undergraduates' responses. Using a matched guise technique typical of language and attitude research (see review of this body of research in Edwards, 1982), two native speakers of Cantonese each recorded highly accented and moderately accented versions of simulated classroom lectures. One lecture was about a natural science topic (the uses and growing scarcity of helium), and the second lecture was about a humanities topic (the role of the

Mahabarata in Indian society). Undergraduate subjects heard one or the other of the lecture topics delivered with either of the levels of accent, accompanied by a photograph of either a European or an Asian instructor.

Dependent measures included a cloze test of listening comprehension (Bachman, 1985), ratings of teacher quality, and four subscales of a measure of homophily (McCroskey, Richmond, and Daly, 1975). Homophily is a construct that pertains to listeners' perceptions of similarity to message sources. Homophily is an important variable mediating cross-cultural communication (Dodd, 1986). For example, in the diffusion of innovations from one culture to another, people are more likely to accept new ideas from homophilous sources. By the same token, homophily has been found to be an important factor in classroom communication (Elliot, 1979). Students respond more positively to teachers of optimal homophily. In addition to these measures, subjects also replied to questions about their contact with NNSTAs.

Rubin and Smith's (1990) results showed that the effects of accent were not explained by any simple response pattern. On the one hand, manipulated accentedness was not accurately discerned on a questionnaire item designed to check the effectiveness of the experimental manipulation; students couldn't always discriminate between high and moderate levels of accent. On the other hand, manipulated accent did affect undergraduates' perceptions of NNSTAs' ethnicity—but only when other cues were ambiguous, that is, when the photograph was European and the topic pertained to humanities.

Of greater significance, listeners' *perceptions* of the instructors' accent—whether accurate perceptions or not—were the strongest predictors of teacher ratings. When students believed an instructor's accent to be "foreign," they simultaneously perceived him or her to be a poor teacher. Also of particular interest, the best predictor of undergraduates' listening comprehension scores was the number of courses they had taken that had been instructed by nonnative instructors. Those students who had persevered with their nonnative English-speaking instructors had been rewarded by improved skill in listening to accented speech.

On the basis of these findings, Rubin and Smith (1990) concluded that high effort/low results programs to remediate pronunciation ought not constitute the main investment of university resources. Instead, it makes sense to simultaneously educate undergraduates about the nature of NNSTA speech, perhaps to provide instruction in listening to accented speech, and to facilitate the amount of exposure to NNSTAs experienced by undergraduates.

The present article reports three studies that elaborate on and extend the work reported by Rubin and Smith (1990). In that study, speakers altered their degree of accentedness, but they could not completely adopt a native-like Standard American English (SAE) accent. Study 1, therefore, replicates Rubin and Smith, but eliminates accent as a factor. Instead, Study 1 asks undergraduates

to rate SAE-speaking instructors of varying ethnicity. Building on the especially interesting correlational results in the 1990 study, Study 2 utilizes a multiple regression analysis to predict teacher ratings and comprehension scores. Finally, Study 3 is a small-scale attempt to implement a training program for mitigating undergraduates' attitudes toward NNSTAs.

STUDY 1

Pronunciation training is tedious and is rarely 100 percent effective. Even the most motivated and industrious student of a foreign language is likely to retain a vestige of nonnative pronunciation. North American undergraduates apparently are not very discerning in discriminating levels of accentedness. Therefore, it is possible that even those NNSTAs who have excelled in pronunciation drills may still suffer a high level of stigmatization by North American undergraduates.

But what if it were possible to completely supplant a nonnative accent with SAE speech? Would even this unlikely achievement be sufficient to overcome negative ethnic stereotypes and attitudes toward subject matter? The purpose of Study 1 is to ascertain the effects of instructor ethnicity and of lecture subject matter when the instructor's actual language behavior is standard, and therefore not culpable as the source of potentially negative student evaluations.

Methods

Participants

Undergraduates were recruited from basic speech communication classes at a large southeastern university. They participated in this study in lieu of a required out-of-class assignment. Data collected from international students were omitted from analysis. Complete sets of usable data were obtained from 62 North American undergraduates.

Procedures and Stimulus Materials

Each subject listened to a single tape-recorded speech sample lasting about four minutes. The speech was presented as part of a lecture by a university "instructor." As subjects listened to the lecture, a slide photograph representing the instructor was projected on a screen.

Instructor ethnicity was operationalized by projecting a photograph of either a Caucasian or an Asian (Chinese) woman. To avoid confounding ethnicity with physical attractiveness, both models were similarly dressed, were of similar size and hair style, and were photographed in the same setting and pose (standing at a lecturn in front of a chalkboard). No differences between the

photographs in rated physical attractiveness were found in any uses of these photographs (although other factors like speech topic sometimes exerted significant effects on judgments of attractiveness).

Lecture topic—physical science or humanities—was operationalized by using the same two scripts as in Rubin and Smith (1990). Both scripts were adapted from articles originally appearing in the *New York Times*. The topics of helium scarcity and of the Mahabarata were selected to represent science and humanities, respectively, because it was judged that most undergraduates would have relatively little prior knowledge about them. Both passages were edited to approximately 450 words, and both recorded speech samples were approximately four minutes long. They were recorded by a single speaker, a doctoral student in speech communication, a native speaker of English raised in central Ohio, who was well regarded by her own undergraduate students for especially effective and clear classroom delivery.

Subjects were randomly assigned to one of the resulting lecture topics (at two levels) by ethnicity (at two levels) treatment combinations. Complete sets of data were obtained from 13 participants in the Asian instructor/science topic group, 16 in the Caucasian instructor/science topic group, 17 in the Asian instructor/humanities topic treatment, and 16 in the Caucasian instructor/humanities topic treatment.

Measurement and Analysis

Immediately after hearing the lecture, subjects completed a cloze test of listening comprehension. They were presented with a written text of the lecture with every seventh word deleted, save in the first sentence, which was kept intact. Only exact recall was scored as correct (see Bachman, 1985).

Next, subjects completed the McCroskey et al. (1975) homophily instrument. The instrument is comprised of 16 semantic differential items. The items pose polar opposite descriptions at either end of seven equal-appearing interval scales (e.g., "shares my values: : : : : : :doesn't share my values"). The items were arranged on the questionnaire in random order and with randomly alternating polarities. The instrument was developed to yield four subscales, each of which is comprised of four semantic differential items: (1) attitude homophily, (2) background homophily, (3) value homophily, and (4) appearance homophily.

Interspersed among the homophily items were additional semantic differential items checking the ethnicity manipulation ("Caucasian/European ethnicity: : : : : : :Oriental/Asian ethnicity") and measuring impressions of accent ("Speaks with American accent: : : : : : :Speakers with foreign accent") and of teaching competence ("Poor teacher: : : : : : :Effective teacher"), as well as

TABLE 1. Cell Means and Standard Deviations for 8 Dependent Variables (Study I)

Lecture Topic:	Humanities				Science			
Instructor Ethnicity:	Caucasian		Asian		Caucasian		Asian	
N =	16		17		16		13	
Perceived accent	3.44	(2.56)	4.94	(1.98)	2.75	(2.44)	3.77	(1.88)
Perceived ethnicity	2.75	(1.98)	5.53	(1.97)	2.06	(1.39)	6.23	(0.93)
Comprehension	11.94	(4.34)	9.93	(5.70)	12.5	(5.9)	7.31	(4.70)
Teaching qualifications	8.69	(3.61)	8.18	(1.81)	7.25	(2.46)	9.15	(2.73)
Attitude homophily	5.56	(5.51)	6.00	(4.95)	2.38	(4.24)	3.38	(4.37)
Background homophily	12.75	(4.12)	12.18	(4.03)	13.25	(4.48)	10.31	(3.54)
Values homophily	15.88	(4.53)	14.94	(2.88)	14.81	(3.21)	14.46	(3.86)
Appearance homophily	11.13	(5.04)	10.71	(5.77)	10.13	(4.77)	7.69	(3.22)

several unrelated filler scales. In all, the instrument contained 28 semantic differential items.

Dependent variables, then, were cloze test scores, the four homophily subscales, perceptions of accent, and ratings of teaching competence, as well as the manipulation check scale that registered perceptions of ethnicity. The manipulation check was subjected to a 2 × 2 ANOVA with subjects nested in levels of ethnicity (Caucasian, Asian) and lecture topic (science, humanities). The remaining dependent variables were submitted to an analogous 2 × 2 MANOVA. The regression solution, whereby each MANOVA sum of squares is adjusted for each of the others, was used to control potential nonorthogonality due to unequal cell frequencies (Norusis, 1990). Since the MANOVA revealed no interaction effect, the significant main effect was followed up with discriminant analysis to determine which combination of dependent variables best discriminated among groups.

Results

Cell means and standard deviations for the eight dependent variables appear in Table 1. Table 2 shows their intercorrelations. The ANOVA for the manipulation check showed that manipulated ethnicity indeed exerted a main effect on perceived ethnicity ($F_{1,58} = 66.91$, $p < .001$). As expected, the Asian instructor was perceived to be more Oriental/Asian than the Caucasian instructor ($M_{asian} = 5.79$, $M_{cau} = 2.41$). Neither the topic factor nor the interaction between topic and ethnicity significantly affected students' perceptions of instructor ethnicity.

Table 3 summarizes the MANOVA for the seven other dependent variables.

TABLE 2. Correlations Among 8 Dependent Variables (Study 1)

	Ethnic	Comprehend	Teach	Attitude	Background	Value	Appearance
Perceived accent	.448**	.095	.083	.069	−.251	−.123	−.119
Perceived ethnicity	—	−.253	−.003	−.024	−.355	−.038	−.289
Comprehension		—	.096	.135	.150	.192	.097
Teaching qualifications			—	.314*	.023	.283	.056
Attitude homophily				—	.464***	.579**	.618**
Background homophily					—	.492**	.606**
Values homophily						—	.283

*p = .05
**pt = .01

TABLE 3. Multitopic X Ethnicity MANOVA of 7 Dependent Variables (Study 1)

Source of Variation	*df*	Wilk's Lambda	*F*	*D*
Topic	7	0.866	1.15	.35
Ethnicity	7	0.727	2.78	.02
Topic x Ethnicity	7	0.879	1.03	.42
Error	52			

As it indicates, the only statistically significant effect to emerge was a main effect for instructor ethnicity (Wilk's $\text{Lambda}_{7,52} = .727$, equivalent $F = 2.78$, $p < .05$).

To follow up this significant multivariate main effect, a step-wise discriminant analysis, summarized in Table 4, was performed. Comprehension test score was the first variable to enter, based on the highest pooled within-group correlation ($-.63$) with the cannonical discriminant function. Using comprehension test scores alone, 61.2 percent of participants could be correctly classified into groups (in this case the grouping factor being whether they viewed an Asian instructor or a Caucasian instructor). Inspection of Table 1 shows that comprehension cell means were lower for the groups exposed to an Asian visage, and higher for groups that saw a Caucasian instructor.

The final discriminant model also included perceived accent, entered at step 2. Inspection of cell means in Table 1 shows that accent was perceived as more foreign and less standard in the case of the Asian instructor's photograph. An additional 9.8 percent of participants were correctly classified into treatment groups when perceived accent was added to the model.

Discussion

These results provide dramatic evidence that North American undergraduates are reacting to factors extraneous to just language proficiency when they are

TABLE 4. Stepwise Discriminant Analysis of Manipulated Ethnicity by 7 Predictor Variables (Study 1)

Step	Variable Entered	Model Wilk's Lambda	F-to-Enter	Model P-Value	Correlation with Discriminant Function
1	Comprehension	.901	6.61	.01	−.631
2	Perceived accent	.798	7.56	.001	.573

asked to react to NNSTAs. In Study 1, only a single language variety was used: SAE. Still, when they were faced with an ethnically Asian instructor, participants responded in the direction one would expect had they been listening to nonstandard speech. Evidence from the discriminant analysis suggests that participants stereotypically attributed accent differences—differences that did not exist in truth—to the instructors' speech. Yet more serious, listening comprehension appeared to be undermined simply by identifying (visually) the instructor as Asian.

The pessimistic conclusion warranted here is that at least among this particular sample of undergraduates even vigorous pronunciation training for NNSTAs will matter little. Ethnically Asian instructors who speak SAE apparently confront similar dysfunctional attitudes as those who do speak with marked nonnative accents.

STUDY 2

Those sympathetic to the plight of NNSTAs can derive some slight comfort from the results of Study 1: At least ratings of instructional competence were not significantly involved in discriminating between Asian and Caucasian teachers. Indeed, it is uncertain which attitudinal and experiential attributes of undergraduates do affect the way they rate NNSTAs' teaching effectiveness. Rubin and Smith (1990) found that *perceived* accentedness, but not actual accent, was negatively related to teacher ratings. In addition, they found that actual experience attending classes taught by NNSTAs was positively related to accuracy in listening comprehension.

It is reasonable to expect that homophily would also mediate teacher ratings such that students are less antagonistic toward NNSTAs whom they perceive as relatively similar to them. Consistent with the homophily hypothesis, Bailey (1984) reports that students sharing the same major as a NNSTA are likely to rate that NNSTA higher than one from a different academic area.

The purpose of Study 2, accordingly, was to determine the relative contributions of factors that do predict undergraduates' ratings of NNSTAs and their comprehension of NNSTA speech. Based on the earlier studies, it was presumed that it would be fruitful to locate these factors in undergraduates' own attitudinal systems and experience.

Methods

Participants

Participants were drawn from the same undergraduate pool as in Study 1. Again, data from international students were excluded. Usable sets of data were collected from a total of 148 students over the course of four academic terms.

Procedures and Stimulus Materials

To best generalize the results of the regression analysis used in this phase of the research, a wide range of stimulus conditions (i.e., nonfactorial) was desirable. Each subject was nonsystematically assigned to one of the ten stimulus audio tapes developed either in conjunction with Study 1, above, or in conjunction with Rubin and Smith (1990). That is, each subject listened to either high or moderate Chinese accent (produced by one or another of the two native Chinese speakers), or else to SAE accent, in combination with either the science or the humanities tape. About half the participants listened to their simulated lectures in conjunction with a photographic depiction of an Asian teaching assistant. About one-third listened to the lectures while viewing a Caucasian/European face, while one-sixth merely listened to the lectures without any photographic presentation.

Measurement and Analysis

Subjects in this study responded to the same questionnaire prepared for Study 1. This included items comprising the four homophily subscales, and items measuring perceptions of accent, ethnicity, and teaching quality. In addition, a "background questionnaire" asked students to report (among other items not under consideration here) the number of classes in which they had been instructed by NNSTAs and the number of weeks they had traveled outside the United States and attitudes toward nonnative speakers in general. This background questionnaire also included six Likert-type items devised to measure general attitudes toward nonnative English speakers. For each of these Likert-type items, participants read attitude statements and then responded in one of five categories ranging from "strongly agree" to "strongly disagree." Three of these attitude statements pertained to interpersonal relations with international students (e.g., "I would be willing to have a nonnative English speaker as my roommate"). The others pertained to benefits that internationals contribute to the university campus (e.g., "Our campus benefits by having nonnative English speakers attending").

Two step-wise regressions were run. The predictor variables in both analyses were identical: (1) perceived accent, (2) perceived ethnicity, (3) attitude homophily, (4) value homophily, (5) background homophily, (6) appearance homophily, (7) total college credits, (8) number of classes instructed by NNSTAs, (9) a composite scale of attitudes toward interpersonal relations with nonnative English speakers, and (10) a composite scale of attitudes toward the overall contribution of internationals. In one regression, the criterion variable was rated teacher qualification. In the second regression, the criterion variable was listening comprehension (cloze) score. For the purposes of the step-wise procedure, the level of significance for F-to-enter and F-to-exit was set at $P<.05$.

Results

Table 5 contains the zero-order correlation coefficients among the 10 predictor variables and between each predictor and the 2 criterion variables. Table 6 summarizes the final results of the step-wise regressions for judged teacher qualification and for listening comprehension. The total R^2 for the first model indicates that it accounted for 11.6 percent of the variance in teacher ratings (multiple $r = .34$). Inspection of the Beta-weights indicates that perceived accent was inversely related to teacher ratings (i.e., the more foreign the accent was judged, the lower the rating of teacher effectiveness), while attitude homophily was positively related to teacher ratings.

For the multiple regression of listening comprehension scores, only a single variable, number of courses students had taken from NNSTAs, was a statistically significant predictor. It accounted for 9.1 percent of the variance in comprehension ($r = .30$). the positive Beta-weight indicates that students who had taken more courses from NNSTAs generally had higher comprehension scores.

Discussion

Results of Study 2 are consistent with those of Rubin and Smith (1990). The degree to which undergraduates perceive NNSTAs' accents to be markedly foreign (and Study 1 confirms that those perceptions are liable to be inaccurate) undermines those students' evaluations of NNSTAs. On the other hand, when students believe that their instructors' attitudes are similar to their own, they more highly regard their instructors' teaching skills. In this latter finding, the present results confirm those of Elliot (1979), who likewise found an association between students' perceived homophily with their instructors and classroom outcomes.

Actual listening comprehension was not affected by any of the attitudinal variables, but was moderately related to undergraduates' experience sitting in classes taught by NNSTAs. Those students who are willing to subject themselves to NNSTA-instructed classes apparently learn more than just course content; they also learn how to listen more effectively.

Now it is possible that more advanced students simply have better listening skills, and as Table 5 suggests, there is some overlap ($r = .30$) between number of classes taken with NNSTAs and students' overall scholastic experience. To examine this alternative explanation in a post hoc manner, the regression for listening comprehension scores was rerun, this time forcing total number of credits into the equation first. After doing so, number of NNSTA classes remains a significant predictor. The multiple correlation between this two-predictor model and comprehension was .34. This was only a modest gain of .04 over

TABLE 5. Correlations Among 10 Predicator Variables and 2 Criterion Variables (Study 2)

	Accnt	Ethnc	Attde	Backgrnd	Value	Appear	Credit	NNSTA Class	Int'l Relatn	Int'l Cmpus
Cloze test	−.113	−.126	.103	.114	.079	.062	.235*	.301**	−.088	.088
Teaching qualification	−.164	−.052	.304**	.129	.155	.149	.049	.122	.014	.110
Perceived accent	—	.560**	−.040	−.063	.233*	−.191	.178	.018	.298*	.098
Perceived ethnicity		—	.124	.112	−.310**	−.283**	.002	−.003	.065	.029
Attitude homophily			—	.618**	.161	.465**	.016	.004	.090	.188
Background homophily				—	.077	.403*	−.027	−.038	.043	.080
Value homophily					—	.242*	−.083	.063	−.031	.232*
Appearance homophily						—	−.053	−.067	−.035	.202*
Total college credits							—	.305**	.112	−.035
Number NNSTA classes								—	.026	.122
Relationships w/ intern'ls									—	.232*

*$p < .05$
**$p < .01$

TABLE 6. Summary of Stepwise Multivariate Regressions Predicting Teacher Ratings and Listening Comprehension (Study 2)

Criterion	Step	Predictor	R^2 Increment	Beta	F-to-Enter
Teacher ratings	1	Attitude homophily	.09	.30	14.59
	2	Perceived accent	.02	−.15	3.88
Listening comprehension	1	# NNSTA classes	.09	.30	14.58

the single-predictor model presented in Table 6 in which number of classes taught by NNSTAs was the sole significant variable.

STUDY 3

The results of Study 2 suggest that exposure of North American undergraduates to NNSTAs is a good thing. The study, however, offers no clues as to why some undergraduates are willing to stick with a class taught by a NNSTA, while others of their cohorts are among the 40 percent of students that Rubin and Smith (1990) found had avoided a NNSTA-instructed class on at least one occasion. But if students could gain some low-stress, or even some ego-rewarding, exposure to NNSTAs, perhaps they would recognize higher attitude homophily with internationals. Perhaps such exposure might also render perceived differences in accent and in ethnicity less salient.

Study 3, therefore, undertook to pilot an intervention experience in which undergraduates were given structured experience interacting with NNSTAs. The goal of the intervention was to cast the undergraduates in a helping role rather than in an adversarial role vis-à-vis their NNSTA "partners." The literature on cross-cultural training is a rich source of guidance for structuring such experiences. Gudykunst (1977), for example, notes that intercultural contact by itself is no guarantee that attitudes between hostile groups will ameliorate. Instead, successful intercultural contact must be facilitated by a neutral party, must level differences in status between parties, must be informal, and must be gratifying. When peers from diverse cultures work on common academic tasks in small groups, improved ethnic relations are likely to prevail (Sharan, 1980).

Indeed, a recent trend in NNSTA skills improvement programs involves undergraduates in various phases of the process (e.g., Civikly and Muchisky, 1991; Schneider and Stevens, 1991; vom Saal and Sarwak, 1989). Some programs use undergraduates as assessors to help screen NNSTAs for classroom assignment, or to place them in developmental workshops. Other programs

involve North American undergraduates as workshop trainers, or to provide feedback as classroom observers.

The deliberate goal of such programs is to improve NNSTA communication skills. According to anecdotal reports, however, a felicitously incidental outcome pertains to changes in undergraduates' attitudes. Undergraduate participants in NNSTA training programs come to see themselves as allies of the internationals. And as campus opinion leaders, those undergraduates directly interacting with NNSTAs in these helping relationships can propagate their positive attitudes among their peers.

Methods

Participants

Participants for this study were recruited from the same pool as those in Studies 1 and 2. Because this was a pilot effort, only a small number of participants were secured. Complete sets of data were collected from 25 undergraduates, 15 of whom were randomly assigned to the training condition, with the remaining 10 comprising a comparison group.

Procedures and Instruments

The training program was an attempt to model an intervention that could be implemented with large numbers of undergraduates. It was, therefore, admittedly superficial. After pretesting, members of the training group were briefed about the importance of helping NNSTAs by giving them feedback from native English-speaking undergraduates. Throughout the program, those in the training group were encouraged to think of themselves as NNSTAs' allies and benefactors.

Each participant in the training program observed two classes taught by NNSTAs. The NNSTAs were current or former participants in NNSTA instructional development workshops who expressed interest in receiving additional feedback from American students. All the NNSTAs with whom the undergraduates interacted were Asian. The undergraduate subjects recorded their reactions to each class on a structured observation/evaluation sheet that had previously been used in the NNSTA training workshops. Participants and NNSTAs met after each observed class. During these sessions undergraduates explained their recorded observation/evaluations to the NNSTAs in the presence of a neutral facilitator.

The two observation/discussion sessions were approximately one week apart. Posttesting occurred within three days of the final session. No special intervention activities were arranged for comparison group members. The time lag between pre- and posttest for comparison group subjects was identical to the training group's.

All subjects were pretested and posttested with audiotapes developed first in Rubin and Smith (1990). To minimize pretest/posttest variability due to differences in test forms, all testing employed a single language guise, the moderate accent versions, prepared by the same native Cantonese speaker. Accent, therefore, was not a variable in this study. Nor was ethnicity, since subjects were not shown any photograph of the NNSTA whom they were evaluating. Topic was also excluded as a variable in Study 3. About half of the subjects in each of the groups heard the science topic at pretest, and half heard the humanities topic. Topics were flip-flopped at posttest.

Measurement instruments were the same as those used in Studies 1 and 2. These yielded measures of (1) judged teaching quality, (2) attitude homophily, (3) value homophily, (4) background homophily, and (5) appearance homophily, as well as (6) a cloze test of listening comprehension.

Analysis

The six dependent measures were all subjected to a 2 × 2 MANOVA. Participants were nested in treatment groups (experimental or comparison) and crossed with the repeated measure, time of testing (pretest or posttest). A significant interaction between group and time of testing would be necessary in order to probe for effects of training.

Results

Cell means and standard deviations for the six dependent variables appear in Table 7. Table 8 contains the zero-order correlations among the variables. The 2 × 2 MANOVA is summarized in Table 9. No statistically significant main or interaction effect emerged in this analysis.

TABLE 7. Cell Means and Standard Deviations for 6 Dependent Variables (Study 3)

Treatment:	NNSTA "Mentoring"				No NNSTA Intervention			
	(n = 15)				(n = 10)			
Time:	Pretest		Posttest		Pretest		Posttest	
Comprehension	9.73	(3.95)	9.67	(4.89)	10.40	(4.01)	10.70	(4.52)
Teaching qualifications	8.33	(1.84)	7.87	(1.68)	6.22	(3.07)	7.30	(1.57)
Attitude homophily	9.60	(4.26)	10.67	(3.92)	9.56	(3.64)	9.70	(4.32)
Background homophily	12.80	(1.74)	13.47	(2.53)	12.56	(2.35)	14.00	(3.16)
Values homophily	13.80	(3.84)	12.60	(2.85)	12.89	(5.23)	12.90	(4.23)
Appearance homophily	9.00	(4.94)	9.53	(4.81)	6.44	(2.79)	8.70	(3.37)

TABLE 8. Correlations Among 6 Dependent Variables at Pretest and Posttest (Study 3)

	Comprehend		Teach		Attitude		Background		Value		Appearance	
	pre	post	pre	post	pre	post	pre	post	pre	post	pre	post
Comprehension												
pre	—	.66*	.13	.16	−.03	.08	−.25	−.35	−.23	−.16	−.04	.16
post		—	.45	.37	.07	.28	.16	−.21	.16	.05	.14	.06
Teaching qualification												
pre			—	.32	.35	.39	.33	−.09	.40	.14	.54*	.14
post				—	.07	.26	.23	.04	−.03	−.12	.17	.25
Attitude homophily												
pre					—	.69*	.22	.36	.51*	.42	.63*	.52*
post						—	.45	.36	.38	.65*	.61*	.67*
Background homophily												
pre							—	.33	.23	.31	.21	.22
post								—	.04	.24	.09	.16
Value homophily												
pre									—	.47	.54*	.14
post										—	.48*	.52*
Appearance homophily												
pre											—	.47
post												—

*$p < .05$

TABLE 9. Treatment × Time of Testing MANOVA of 6 Dependent Variables (Study 3)

Source of Variation	*df*	Wilk's Lambda	*F*	*p*
Treatment	6	0.722	1.09	.41
Time of testing	6	0.617	1.76	.17
Treatment × Time	6	0.839	0.54	.77
Error	17			

Discussion and Conclusion

Previous studies had demonstrated the potency of undergraduates' own perceptions and stereotypes in determining their response to NNSTAs. Therefore, an intervention program aimed at ameliorating undergraduates' attitudes seems one reasonable approach to reducing the communication interference between North American students and their international instructors. The experimental intervention tested in Study 3, however, was plainly not up to the task. The lack of significant results was not just a matter of low statistical power due to a small sample; cell means showed no consistent movement from pretest to posttest. Indeed, there is some anecdotal evidence (based on undergraduates' feedback to the NNSTAs with whom they were paired) that students in the training group actually became more critical of NNSTAs' communication behaviors during the course of this experiment.

The treatment administered here was motivated by the intent to cast undergraduates into a sort of mentoring role with respect to NNSTAs, as opposed to the antagonistic role they often assume in regular classroom situations. But the stereotypes and prejudices that this experimental program sought to interdict are powerful, indeed. Study 1 showed that the sight of an Asian visage was sufficient to elicit among these students' perceptions of foreign accent, even though the actual speech sample was produced by an expert speaker of SAE. Even more revealing, when students were shown an Asian NNSTA, their listening comprehension scores were negatively impacted. Once again, the speech in each case was SAE.

Most likely the intervention attempted in Study 3 was of too short a duration and too weak an intensity to bring about any profound attitudinal change. Civikly and Muchisky (1991), however, describe a similar project in which students worked with NNSTAs for an entire term under apparently more structured conditions. Their results, however, are likewise no cause for celebration. When asked what they had learned from participating in this project, only 26 percent of those undergraduates offered any comments that indicated increased intercultural sensitivity.

SUMMARY

Brislin (1981) notes that contact alone is surely not a potent enough force to change deeply held prejudices. Instead, changing intergroup attitudes requires an environment rich in informal and pleasant contact, some degree of interpersonal intimacy and equality, and support from the participants' own reference groups. Future programs for improving undergraduate attitudes toward NNSTAs should incorporate these elements. Still, such programs will require labor-intensive and time-intensive efforts, and will not be practical for the sort of large-scale sensitization needed on college campuses. Principles for engineering large-scale attitude change in this particular intercultural context have yet to be articulated, much less tested.

Assuming that positive attitudes toward NNSTAs—or at least open attitudes—will reduce attrition of NNSTA-taught classes, there are very pragmatic reasons to pursue any program that promises to improve those attitudes. According to Rubin and Smith (1990), 42 percent of undergraduates in one sample had on at least one occasion (sometimes more) disenrolled in one way or another from classes they discovered were taught by NNSTAs. Aside from the logistical repercussions and the underutilization of teaching talent this attrition represents, the undergraduates themselves are poorly served by their decisions. Study 2 showed that the best predictor of listening comprehension was the number of NNSTA-taught classes one had attended. Those undergraduates who stuck with their NNSTAs were rewarded by achieving superior communication skills in the form of listening abilities.

Methodologically, each of the three studies reported here follows in the tradition of much language and attitude research by employing the matched guise technique (Lambert et al., 1960) as a measurement device. This has the advantage of controlling for extraneous factors like speakers' vocal characteristics. In other studies, the matched guise technique also permits factorial permutations of language style with nonlanguage factors like ethnicity (Brown, 1988; Rubin and Smith, 1990) or school achievement or social class (Piche et al., 1977). On the other hand, the matched guise technique is vulnerable to the criticism that it requires artificial stimuli (e.g., usually denuded of gestural communication) and testing conditions (Robinson, 1972).

Certainly more studies are warranted in which undergraduates rate NNSTAs with whom they have shared an entire college term of study. In one such study, elements of NNSTAs' interpersonal styles affected judged teaching competence more than did NNSTAs' measured language proficiency (Dalle and Inglis, 1989). The problem inherent in such in situ studies is that—even when ratings of North American teaching assistants can be contrasted with NNSTA ratings (e.g., Briggs and Hofer, 1991)—it is difficult to isolate elements of idiosyncratic teaching behaviors, language proficiency, engendered stereotypes, or course difficulty that might contribute to any differences found.

It would be useful, in addition, to gather data about undergraduates' reactions to non-Asian NNSTAs. The present studies deliberately narrowed their purview to East Asian instructors. This seems a reasonable methodological choice because of the rising influx of university teachers from those particular ethnic groups. Quite possibly, the findings of these studies would not generalize to more Western NNSTAs, say instructors from France or Germany.

Finally, it would be worth replicating these studies with more heterogenous groups of undergraduates. Nonnative English speakers were deliberately excluded from the sample here. It is possible that their reactions to NNSTAs are considerably different from mainstream culture North Americans. Moreover, the geographic region in which this study was conducted has a relatively low proportion of nonnative English speakers and of Asians. These students are exposed to relatively little nonnative accented speech in their daily affairs. Furthermore, few of these students had ever encountered a native English-speaking Asian-American instructor. No doubt stereotypical reactions would be different among students with a greater variety of cross-cultural experience.

Acknowledgments. Kim A. Smith assisted in collecting data for Studies 1 and 2, Kimberly Lichner assisted in collecting data for Study 3 and in data analysis, and Kathleen B. Smith assisted in collecting data for Study 3.

REFERENCES

Abraham, R. G., and Plakans, B. S. (1988). Evaluating a screening/training program for NNS Teaching Assistants. *TESOL Quarterly* 22: 505–508.

Althen, G. (1991). Teaching "culture" to international teaching assistants. In J. Nyquist, R. Abbott, D. Wulff, and J. Sprague (eds.), *Preparing the Professoriate of Tomorrow to Teach* (pp. 350–355). Dubuque, IA: Kendall-Hunt.

Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly* 19: 535–556.

Bailey, K. M. (1984). The "foreign TA problem." In K. Bailey, F. Pialorsi, and J. Zukowski/Faust (eds.), *Foreign Teaching Assistants in U.S. Universities* (pp. 3–15). Washington, DC: National Association for Foreign Student Affairs.

Briggs, S., and Hofer, B. (1991). Undergraduate perceptions of ITA effectiveness. In J. Nyquist, R. Abbott, D. Wulff, and J. Sprague (eds.), *Preparing the Professoriate of Tomorrow to Teach* (pp. 435–445). Dubuque, IA: Kendall-Hunt.

Brislin, R. W. (1981). *Cross-Cultural Encounters*. New York: Pergamon.

Brown, K. (1988). American college student attitudes toward non-native instructors. Unpublished doctoral dissertation, University of Minnesota. (DAI 49: 1710A–1711A).

Byrd, P., Constantinides, J. C., and Pennington, M. (1989). *The Foreign Teaching Assistant's Manual*. New York: Macmillan.

Civikly, J. M., and Muchisky, D. M. (1991). A collaborative approach to ITA training: The ITAs, faculty, TAs, undergraduate interns, and undergraduate students. In J. Nyquist, R. Abbott, D. Wulff, and J. Sprague (eds.), *Preparing the Professoriate of Tomorrow to Teach* (pp. 356–360). Dubuque, IA: Kendall-Hunt.

Constantinides, J. C. (March 1987). The "foreign TA problem"—An update. *NAFSA Newsletter* 38: 3, 5.

Dalle, T., and Inglis, M. (March 1989). What really affects undergraduates' evaluations of nonnative teaching assistants' teaching? Paper presented at the TESOL Annual Convention, San Antonio, Texas.

Dodd, C. (1986). *Dynamics of Intercultural Communication* (2nd ed.). Dubuque, IA: Wm. C. Brown.

Dunn, T. G., and Constantinides, J. C. (1991). Standardized test scores and placement of international teaching assistants. In J. Nyquist, R. Abbott, D. Wulff, and J. Sprague (eds.), *Preparing the Professoriate of Tomorrow to Teach* (pp. 413–419). Dubuque, IA: Kendall-Hunt.

Edwards, J. R. (1982). Language attitudes and their implications among English speakers. In E. B. Ryan and H. Giles (eds.), *Attitudes Towards Language Variation* (pp. 20–33). London: Edward Arnold.

Elliot, S. (1979). Perceived homophily as a predictor of classroom learning. In D. Nimmo (ed.), *Human Communication Yearbook 3* (pp. 585–602). New Brunswick, NJ: Transaction Books.

Gallego, J. C., Goodwin, J., and Turner, J. (1991). ITA Oral Assessment: The examinee's perspective. In J. Nyquist, R. Abbott, D. Wulff, and J. Sprague (eds.), *Preparing the Professoriate of Tomorrow to Teach* (pp. 404–412). Dubuque, IA: Kendall-Hunt.

Gudykunst, W. (1977). Intercultural contact and attitude change: A review of the literature and suggestions for future research. In N. Jain (ed.), *International and Intercultural Communication Annual*, vol. 4 (pp. 1–16). Falls Church, VA: Speech Communication Association.

Heller, S. (September 11, 1985). Colleges try tests and training to make sure foreign T.A.'s can be understood. *Chronicle of Higher Education*, pp. 1, 32, 33.

Hess, J. (November 2, 1987). Instructors' fluency in English debated at University of Florida. *Atlanta Constitution*, pp. 1, 9.

Instructors' broken English prompts Illinois law (October 4, 1987). *Atlanta Journal and Constitution*.

Lambert, W., Hodgson, R., Gardner, R., and Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60: 44–51.

McCroskey, J. C., Richmond, V. P., and Daly, J. A. (1975). The development of a measure of perceived homophily in interpersonal communication. *Human Communication Research* 1: 321–332.

Norusis, M. J. (1990). *SPSS-PC + Advanced Statistics 4.0*. Chicago: SPSS, Inc.

Orth, J. L. (1982). University undergraduate evaluational reactions to the speech of foreign teaching assistants. Unpublished doctoral dissertation, University of Texas. (DAI (1983), 43: 12, 3897-A).

Pica, T., Barnes, G. A., and Finger, A. G. (1990). *Teaching Matters: Skills and Strategies for International Teaching Assistants*. New York: Newbury House.

Piche, G. L., Michlin, M. L., Rubin, D. L., and Sullivan, A. (1977). Effects of dialect-ethnicity, social class, and quality of written composition on teachers' subjective judgments of students. *Communication Monographs* 44: 60–72.

Robinson, W. P. (1972). *Language and Social Behavior*. Harmondsworth, UK: Penguin Books.

Rubin, D. L., and Smith, K. A. (1990). Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of non-native English speaking teaching assistants. *International Journal of Intercultural Relations* 14: 337–353.

Schneider, K., and Stevens, S. G. (1991). American undergraduate students as trainers in an international teaching assistant training program. In J. Nyquist, R. Abbott, D.

Wulff, and J. Sprague (eds.), *Preparing the Professoriate of Tomorrow to Teach* (pp. 361–367). Dubuque, IA: Kendall-Hunt.

Sharan, S. (1980). Cooperative learning in small groups: Recent methods and effects on achievement, attitudes, and ethnic relations. *Review of Educational Research* 50: 241–272.

Thomas, C. F., and Monoson, P. K. (1991). Issues related to state-mandated English language proficiency requirements. In J. Nyquist, R. Abbott, D. Wulff, and J. Sprague (eds.), *Preparing the Professoriate of Tomorrow to Teach* (pp. 382–392). Dubuque, IA: Kendall-Hunt.

Turitz, N. J. (1984). A survey of training programs for foreign teaching assistants in American universities. In K. Bailey, F. Pialorsi, and J. Zukowski/Faust (eds.), *Foreign Teaching Assistants in U.S. Universities* (pp. 43–50). Washington, DC: National Association for Foreign Student Affairs.

Verhovek, S. H. (November 5, 1989). An American accent for teachers from abroad. *New York Times*, Section 4A, pp. 22–23.

vom Saal, D. R., and Sarwak, S. (November 1989). Strengthening the ITA program through the involvement of undergraduates. Paper presented at the Second National Conference on the Training and Employment of Teaching Assistants, Seattle.

Whitla, D. K. (1968). Evaluation of decision making: A study of college admissions. In D. Whitla (ed.), *Handbook of Measurement and Assessment in Behavioral Sciences* (pp. 456–490). Reading, MA; Addison Wesley.

Yule, G., and Hoffman, P. (1990). Predicting success for international teaching assistants in a U.S. university. *TESOL Quarterly* 24: 227–243.

Zikopoulos, M. (ed.) (1988). *Open Doors 1987/88: Report on International Educational Exchange*. New York: Institute of International Education.

Received February 27, 1991.