

Linguistic and Conversational Adjustments to Non-Native Speakers

Michael H. Long

University of Hawaii at Manoa

At least 40 studies have been conducted of the linguistic and conversational adjustments made by native speakers of a language using it for communication with non-native speakers. The modifications sometimes result in ungrammatical speech. Generally, however, they serve to provide input that is well-formed, a sort of linguistic and conversational cocoon for the neophyte second language acquirer. Most of the findings hold across age groups, social classes and settings, although some differences, both qualitative and quantitative, have been noted in these areas, too.

In making the adjustments described, native speakers appear to be reacting not to one, but to a combination of factors. These include the linguistic characteristics and comprehensibility of the non-native's interlanguage, but particularly his or her apparent comprehension of what the native speaker is saying. The adjustments appear to be necessary for second language acquisition, in that beginners seem unable to acquire from unmodified native speaker input. There is some doubt as to their sufficiency in this regard.

Introduction

There is a rapidly growing body of work documenting the linguistic and conversational adjustments made by native speakers (NSs) of a language using it for communication with non-native speakers (NNSs). What follows is a survey of findings from these studies with the aim of providing tentative answers to four questions:

How do NSs react to NNSs (a) linguistically, (b) conversationally?

Who reacts like this?

What do they react to?

What difference does it make to second language acquisition by the NNS?

How do native speakers react to non-native speakers?

Linguistic adjustments

Much early work on NS speech to NNSs focussed on what is occasionally one of its most striking characteristics: ungrammaticality. In a much cited paper on the topic, Ferguson (1971) catalogued various features of what he termed 'Foreigner Talk' (FT), a conventionalized

register of simplified speech . . . used by speakers of a language to outsiders who are felt to have a very limited command of the language or no knowledge of it at all (Ferguson 1971:143).

This is the stuff of 'Me Tarzan, you Jane', first encountered (first learned?), in

some Western societies, at least, by children watching certain TV series or reading comic strips.

The ungrammaticality is the product of three main processes, which Ferguson termed 'omission', 'expansion' and 'replacement/rearrangement'. Omission typically involves deletion of articles, copula, prepositions, inflectional morphology, conjunctions and subject pronouns (He live three year Japan). An example of expansion is the addition of unanalyzed (lexical) tags to questions (You have food, yes?/no?/okay?), and insertion of subject pronoun *you* in imperatives (You go). Replacement/rearrangement includes such features as forming negatives with *no* plus the negated item (Me no like), replacing subject with object pronouns (Me go), substituting noun-plus-object pronoun constructions (sister me) for possessive pronoun-plus-noun constructions (my sister), and a preference for uninverted question forms with deletion of the *do* auxiliary (You come?).

Ferguson's data were obtained via an elicitation procedure in which Stanford University students were asked to rewrite 10 sentences in the way they thought they would say them to a group of illiterate non-Europeans who spoke a language other than English. While consisting, therefore, of elicited, written, introspective judgments of what speakers thought they would do in an imaginary situation, Ferguson's findings have subsequently been confirmed not only by several other researchers employing the same or similar elicitation procedures (Andersen 1977; McCurdy 1980; Meisel 1977; Mulhäusler 1981), but also through naturalistic observation of what speakers actually *do* (Clyne 1977, 1978; Dutch Workgroup of Foreign Workers' Language [DWFWL] 1978; Goldberg 1982; Hatch, Shapira and Gough 1975; Heidelberger Forschungsprojekt "Pidgin Deutsch" [HFPD] 1978; Katz 1977, 1981; Ramamurti 1977; Snow, van Eeden and Muysken 1981; Valdman 1976). Further, the same *kinds* of speech modifications identified by Ferguson (Ferguson 1975; Ferguson and DeBose, 1977) have produced the ungrammaticality across the seven languages involved in these studies: English, Dutch, German, French, Portuguese, Finnish and Tâi Bôi.¹

Although several studies have found the kinds of linguistic adjustments which result in *ungrammatical* speech to NNSs, other recent research has shown that ungrammaticality, while striking when it does occur, is by no means the norm. In fact, the studies suggest overall that most speech addressed to second language (SL) acquirers is a modified but *well-formed* version of the target. In other words, the data for second language acquisition (SLA) are no more 'degenerate' than are those for first language acquisition by children. (Snow 1977, 1979, for review of the literature on caretaker speech.)

Among the most commonly observed features of *grammatical* FT are: the use of shorter utterances in T-units (Arthur, Weiner, Culver, Lee and Thomas 1980; Freed 1978; Gaies 1977; Scarcella and Higa 1981), and the use of speech which is syntactically and/or propositionally less complex in various ways, e.g. contains fewer sentence nodes per T-unit (Freed 1978), fewer clauses per T-unit, fewer adjectival, adverbial and noun clauses per T-unit (Gaies 1977), or fewer relative clauses and appositives per T-unit (Scarcella and Higa 1981). The range of vocabulary employed has been found to be more restricted, as measured by type-token

ratio (Arthur *et al.* 1980); impressionistically, idioms and low frequency lexical items seem to be avoided (Henzl 1974, 1975, 1979). Lastly, grammatical FT tends to be a more 'regular' version of the language used in NS-NS conversation. Thus, in a study of English, Japanese and Hindi-Urdu FT (Long, Gambhiar, Gambhiar and Nishimura 1982), speech to NNSs contained higher frequencies (than did speech to other NSs of those languages) of utterances which observed the canonical word order for a given language (SVO for English, SOV for Hindi-Urdu and Japanese), more utterances with optional (S, V or O) constituents retained in surface structure, and more overt marking of grammatical and semantic relations, e.g. Japanese particles indicating topic, comment, subject, object, and various directionals and locatives.

An obvious question arising from these findings concerns the nature of the animal, 'FT'. Why is speech to NNSs sometimes grammatical, sometimes not? Arthur *et al.* (1980) argue that a distinction must be drawn between 'foreigner talk' and 'foreigner register'. The former, they claim, involves rules (adding or deleting elements) which are not found in the standard grammar, whereas foreigner register is defined in terms of the statistical frequency of occurrence or frequency range of various morphological, syntactic, semantic and phonological options present in the standard variety. Consequently, use of FT involves switching; use of 'foreigner register' only involves modifying. The two phenomena are not just points on a continuum, in other words, but qualitatively different.

Whether the ungrammaticality is produced by code-switching or by shifting further down a single linguistic continuum, it is of interest of SLA researchers to be able to identify those variables which predict grammatical, and those which predict ungrammatical, input. For example, it may turn out not to be coincidental that learners who fossilize with a 'pidginized' version of the target are often those who receive restricted input, ungrammatical input, or both (cf., e.g. HFPD 1978; Schumann 1978).

An attempt to account for this aspect of variation (Long 1981a) isolated four factors which appear relevant. Ungrammatical input was found to be more likely when (1) the NNS has zero or very low SL proficiency, (2) the NS either is or perceives him or herself as being of higher social status than the non-native interlocutor,² (3) the NS has prior FT experience, but only with NNSs of low SL proficiency, and (4) the conversation occurs spontaneously, e.g. in task-oriented communication on the factory floor, as opposed to arranged encounters between strangers in the research laboratory (when ungrammatical speech almost never occurs). The review of available studies showed that factors (1), (2) and (4) appear to be *necessary* conditions for ungrammatical FT to occur, but that no single condition alone seems to be *sufficient*. A combination of factors is at work.

Conversational adjustments

In addition to their examination of modified but *grammatical* speech to NNSs, another feature of several recent FT studies is their shift in focus from the *input* (alone) to the structural characteristics of the NS-NNS *conversation* in which FT

occurs, i.e. to the study of 'foreigner talk discourse' (Hatch et al. 1975; Long 1980, 1981b). Some studies have found no statistically significant differences between FT and NS speech to other NSs in the traditional morpho-syntactic areas of analysis, yet have at the same time found the NSs to be doing considerable 'work' at the discourse level. For example, the following two (constructed) sample conversations exhibit identical NS utterance structure, but differ in their *interactional* structure because of the NS's use of exact and semantic repetition:

- (1a) NS-NS conversation
 NS1: Do you like California?
 NS2: I love it
- (1b) NS-NNS conversation
 NS: Do you like California?
 Do you like Los Angeles?
 Do you like California?
 NNS: Huh?
 Uhm-
 Oh!. Yeah I like

To date, some of the ways in which NS-NS and NS-NNS conversations have been found to differ are as follows. Conversational topics are treated simply and briefly in foreigner talk discourse (FTD), as measured by the number of "information bits" supplied by the NS on any topic (Arthur et al. 1980) or by the ratio of topic-initiating to topic-continuing moves (Gaies 1981a; Long 1981b). There is some evidence that the very nature of the topics preferred in FTD differs, too. Scarcella (1981) compared NSs of Spanish (Mexicans) and English (Americans) conversing with each other (informally, the first encounter between strangers), and each of these types of conversation with NS-NNS encounters between members of the two groups. The Spanish NSs in NS-NS conversation spoke freely and early about personal matters (home, age, family, marriage, etc.), whereas English NS dyads tended to discuss only impersonal topics (classes, careers, places of residence, etc.). When the language groups were mixed, in NS-NNS dyads, neither American nor Mexican interlocutors introduced personal topics, perhaps, Scarcella suggests, because the NSs of Spanish had lived in the other culture long enough (a minimum of 12 years) to learn which topics were considered appropriate by the English NSs for discussion with strangers. Possibly arising from the need to negotiate topics across cultural boundaries, the FTD in Scarcella's study was also found to contain more abrupt topic-*shifts* than the NS-NS conversations in either language. It also exhibited a less predictable *sequence* of topics than the English or Spanish baseline data.

FTD has been found to contain at least three additional indications of the topic-negotiation process, but these features seem motivated by the NS's attempt to cater to the NNS's linguistic ability rather than to cultural differences. First, even though the NNS be an adult, with no cognitive limitations, FTD, like caretaker speech (Cross 1978), has been found to be significantly more oriented to the "here and now" than NS-NS conversation, as measured by the relative frequencies of verbs marked temporally for present and non-present (Gaies 1981a; Long 1980, 1981b). This avoidance of reference to displaced time and place is even more apparent in the classroom FTD of ESL instruction (Long and Sato in press). Second, NSs have been found to use significantly more of what Hatch (1978) calls 'or-choice' questions in FTD than with other NSs (Long 1981b). Hatch has pointed

out that such questions allow the NNS to choose from a list of potential topics (or comments), and make his or her participation easier still by containing the "answer" to the question, as illustrated in (2):

(2) NS: Well what are you what are you doing
in the
United States? . . . Are you just
studying?
Or do you have a job?
Or-

NNS: No
I have job

Third, there is a tendency for NNSs to accept unintentional topic-switches by NNSs in FTD when a communication breakdown occurs. The NS may, for example, skillfully treat an inappropriate response as a topic-nomination, thereby repairing the discourse:

(3) NS: Are you going to visit San Francisco?
Or Las Vegas?

NNS: Yes I went to
Disneyland and to
Knotts Berry Farm

Oh yeah?

(Long 1981c, 264)

Several other conversational adjustments have been noted in FTD which concern not the choice of topics, but the way they are introduced by the NS. For example, there is a well-documented preference for questions over statements (Freed 1978; Long 1980; Scarcella and Higa 1981), with questions being especially favored for topic-initiating moves (Gaies 1981a; Long 1981b). The use of or-choice questions described earlier is one manifestation of this. Thus, a NS is more likely to open up a new topic with (4a) than (4b):

(4a) NS: Do you like going to the movies?

(4b) NS: I like going to the movies

and is more likely to use utterances like (4a) with NNSs than with other NSs.

The preference for questions in topic-initiating moves, and in FTD generally, probably has several motivations. First, questions in general are more likely to draw the NNS into the conversation. Second, yes/no questions in particular make the NNS's linguistic task easier by making his or her conversational role easier. They contain a complete proposition, which the NNS need only confirm or deny, whereas *wh* questions contain a missing element, and statements require a complete new proposition from the other speaker. Consider the increasing complexity required of an appropriate response to (5a through c):

(5a) Do you live in Los Angeles?

Yes/No (In Orange County)

(5b) Where do you live?

In Orange County

(5c) I live in Los Angeles

(Do you?) I live in Orange County

Third, questions are useful as comprehension checks (Do you understand?), which help NSs assess whether they are succeeding in communicating with NNSs, and also as clarification requests (What do you mean?) and confirmation checks (The library?), which tell them whether they are understanding what the NNSs are trying to communicate to them. All three functions of questions occur statistically significantly more frequently in FTD than in NS-NS conversation (Long, 1980). This suggests that Varonis and Gass (1982) are correct in claiming that questions occur frequently as a reflection of the NS checking his/her assumptions as to the comprehensibility of the NNS's speech, but not correct when they claim that this is the main or only reason they occur more frequently in FTD (cf. Varonis and Gass 1982: 131-132).

Further devices noted as having more frequent use in FTD than in NS-NS conversation for introducing topics include the use of stress and/or pauses before topic words (Hatch 1978; Long 1980):

(6) NS: Did you ● *like* San Diego?

left-dislocation (Hatch 1978):

(7) NS: Did you ● *like* San Diego? ● San Diego ● did you like it?

question and answer (Long 1980):

(8) NS: Right *When* do you take the break? At ten-thirty?

and "decomposition" (Long 1980):

(9a) NS: When do you go to the uh Santa
Monica?
... You say you go fishing in Santa
Monica, right?

NNS: Yeah

When?

(9b) NS: Uh what does uh what does your
father do in uh you're from Kyoto,
right?

NNS: Yeah

Yeah What does your father do in
Kyoto?

As shown in (9a) and (9b), decomposition starts with a request by the NS for the NNS to comment on a new topic introduced by a *wh* question. This proves too difficult for the NNS. The task is then broken down (decomposed) into two more manageable parts. First, the (sub)topic is established by its repetition in isolation, usually in the form of a yes/no question or uninverted (intonation) question, and often with a tag (right?). When the NNS confirms that the topic has been established, the comment, in the form of a question about the new topic, is restated.

(9a) and (9b), respectively, also show that this device, like most of those described, can serve two functions: to repair the discourse following a breakdown in communication, and to avoid such a breakdown occurring. Elsewhere (Long 1981d), these devices have been classified as *tactics* and *strategies*, respectively, and their realizations quantified and compared. Among the most frequent are (exact or semantic) self- and other-repetitions, expansions, confirmation checks, clarification requests and comprehension checks (Long 1980).

In closing this descriptive section on conversational adjustments to NNSs, one or two methodological points are in order. First, it bears repeating that it is not the *use* of devices like those illustrated which distinguishes FTD from NS-NS conversation, but rather their statistically significantly higher *frequency of use* in FTD. Most, probably all, of them also occur in NS-NS conversation among normal adults, in caretaker-child conversation, and in talk between normal adults and the mentally retarded, although documentation for them is limited with respect to the last two populations. We are dealing, therefore, with quantitative, not qualitative differences between FTD and NS-NS talk.

Second, unlike features of linguistic *input* to NNSs, the analysis of *interactional* characteristics of FTD requires looking at speech by *both* participants in a conversation, as well as at previous speech by each of them. Acts like repetition, expansion and confirmation checks only have life across utterances and speakers in context. Work on FTD, in other words, involves the procedures and problems of discourse analysis, or, more correctly in this case, 'analysis of discourse', since this type of study does not necessarily result in an 'exhaustive' account (Sinclair and Coulthard 1975) of the conversations. While suprasentential units are notoriously difficult to define operationally and to quantify, quantification is as essential here as it is for the grammatical input features, given that it is the relative *frequency of use* (see above) that is at issue. Thus, it is not enough, as some researchers have done, simply to look at FTD (or at FT) and then to assert, e.g. that utterances are short or less complex, or that repetition is used. The question is whether utterances are shorter, syntactically *less* complex, or *more* repetitive *than* those in speech to NSs or in NS-NS conversation in comparable situations.

Last, of the little work done on FTD thus far, most has considered such devices as confirmation checks, clarification requests, repetition and restatement in fairly gross terms. Yet such moves in discourse often have multiple functions, and also multiple realizations, choice among which is not arbitrary. The multifunctionality of such devices as confirmation checks and clarification requests, which may simultaneously serve as corrective feedback, has recently been shown by Chun, Day, Chenoweth and Luppescu (1982). Work by Chaudron (in press) demonstrates the potential of finer-grained analyses of different realizations of the devices, some of which he has shown to facilitate comprehension by the NNS better than others (see section 5, below). This looks a promising area for future research.

Who reacts like this?

In general, the linguistic and conversation adjustments described in section 2 appear to be immune to differences among *groups/types* of speakers. Some of the same linguistic adjustments, for example, have been observed in the speech of (1) children as well as adults (Andersen 1977; Fillmore 1976; Katz 1977, 1981; O'Brate 1980; Peck 1978), (2) upper-middle, middle and working-class adults (Goldberg 1982), (3) individuals with and without prior FT experience (Dahl 1981; Long 1980), and (4) second language teachers, content teachers and non-teachers (Chaudron 1978, 1981 in press; Dahl 1981; Downes 1981; Gaies 1977; Hatch et al. 1975; Henzl 1974; Long and Sato in press; Schinke 1981, in press), although results are eagerly awaited for an explicit comparison of content and second language classrooms (Early in progress). The latter will help sort out just which adjustments are peculiar to FT/FTD, as opposed to teacher talk and/or caretaker speech.

There appears, nonetheless, to be considerable variation at the level of the individual. Thus, while groups of NSs have been found to modify their speech more with decreasing proficiency of the NNS (Chaudron 1978; Dahl 1981; Gaies 1977), *individual* NSs have been found modifying more to more *advanced* speakers (Chaudron 1978; Trager 1978), and even using more 'complex' speech to NNSs than to NS controls on some measures (Dahl 1981; Long 1980). Considerable individual variability has also been noted in the use of ungrammatical versus grammatical speech (Snow et al. 1976), and in the kinds of grammatical modifications and the extent of such modifications (Dahl 1981). Lastly, while most second language teachers studied have been found to use modified but grammatical speech in the classroom (Henzl 1974; Long and Sato in press; Pica and Long 1982), a few teachers have been documented using *ungrammatical* speech (Downes 1981; Hatch et al. 1975).

Whether these and other examples of individual variability are real or 'pseudo' (Dulay, Burt and Krashen 1982) will only be decided by further research. While some apparent differences are probably due to faulty research design, inadequate sample size, subjects' different perceptions of task, etc., other differences may turn out to reflect 'genuine' individual variation in such areas as speech style, empathy/sensitivity to the interlocutor (suggested by Hatch et al. 1975), and the holding of prescriptive norms concerning 'correct' speech (1981 Sato personal communication).

What do they react to?

Most research to date has aimed at *describing* FT/FTD. Very little work has been done to determine what it is that *causes* linguistic/conversational adjustments by NSs. What is it that NSs react to? There are at least five possibilities: the physical appearance (perceived 'foreignness') of the NNS, one or more linguistic features of the NNS's interlanguage, the NS's assessment of the NNS's level of *comprehension* of what the NS is saying, the *comprehensibility* to the NS of what the NNS is saying, and combinations of two or more of these four factors.

The NNS's physical appearance does not, at first sight, seem a very likely candidate. Studies by Abunahleh et al. (in press); Arthur et al. (1980); Chickinsky (1980); and Hatch et al. (1975), have involved FT/FTD corpora obtained from telephone conversations between strangers. The NSs in these studies modified their speech despite the lack of visual contact, and without having met the NNS previously. Arthur et al. also established empirically that their NNS subjects were identifiable as such by their accentedness in English, the language of communication. (The researchers did not claim to have shown that it was accentedness to which their NS subjects in fact reacted, of course.) Second, in a matched guise study, Varonis and Gass (1982) found that one group of 24 NSs modified their speech significantly more to two NSs who were of foreign ethnic descent (i.e. 'looked foreign') and who feigned a heavy non-native accent when they asked for street directions, than did another group of 24 NSs who were asked for the same information by the same two NSs, this time using their normal 'NS' English.

While this evidence would appear to argue against physical appearance as a factor triggering FT/FTD, some anecdotal observations might lead us to hesitate before ruling out physiological characteristics of the NNS. Along with others with extensive experience of talking to Asian (and other) NNSs of English, the writer has occasionally found himself making (grammatical) linguistic and conversational adjustments to Asian-Americans who are NSs of English, not only early, in first encounters, but sometimes even subsequently, temporarily 'forgetting' they are NSs. That these impressionistic observations are real is suggested by the (justifiable) rebukes to which this has led from one of the Asian-Americans who, herself an applied linguist and familiar with the FT literature, recognized what was happening!

Perhaps a distinction needs to be drawn between 'immediate' and sustained reactions of populations with and without prior FT experience. Those with extensive experience may "instinctively" react to perceived 'foreignness', but subsequently readjust to the norms of NS-NS talk. Those with little or no such experience are not so easily misled, as Varonis and Gass (1982) show, and need some other cues before they react. The subsequent readjustment of the experienced Foreigner Talkers, of course, also suggests the longer term importance of one or more variables other than the NNS's physical appearance.

In a series of carefully designed studies, Varonis and Gass (1982) have attempted to tease out the contributions (if any) of two linguistic features of the NNS's interlanguage, accentedness and grammaticality, as triggers of FT/FTD. They began by ruling out grammatical errors alone as the factor by showing (see above) that NSs modified their speech to other NSs talking with a non-native accent, but grammatically correctly. Controlled laboratory studies followed in which sentence pairs with different combinations of accented and ungrammatical speech were presented to linguistically naive NS judges. In general, Varonis and Gass found that the judges rated very good or very poor pronunciation high or low, respectively, regardless of whether the sentences they heard also contained grammatical errors, but rated good pronunciation lower when it coincided with errors of grammar. In other words, grammatical errors led to lower *pronunciation* ratings

for speakers with 'middling' pronunciation, but pronunciation was all that mattered if the NNS had either a very good or a very poor accent. (There was a curvilinear relationship between these two variables.) Next, both pronunciation and grammaticality were found to contribute to the *comprehensibility* of a message, with grammaticality generally having the greater effect. In turn, *comprehensibility* was found to have the strongest effect on the judges' rating of a speaker's pronunciation. If a message was easy to interpret, pronunciation ratings went up; if the message was difficult to interpret, ratings for pronunciation went down. (The importance of grammaticality was stronger than the order in which the sentences were presented.) Varonis and Gass concluded that the *comprehensibility* of a message to a NS, including all the linguistic features contributing to comprehensibility, is the main factor in determining both the NS's evaluation of, and reaction to, non-native speech.

The studies by Varonis and Gass are important and deserve replication by other researchers. Again, however, it may be necessary to distinguish initial and long-term NS reactions. There is, after all, independent evidence suggesting the importance of linguistic features of the NNS's interlanguage in determining both the type and the extent of the adjustments NSs make. First, while a strong version of the 'matching' hypothesis is untenable (cf. Hatch 1979; Long et al. 1982; Meisel 1977), there is some evidence that NSs make finer adjustments to individual NNSs with whose interlanguage they are familiar (e.g. Shapira to Zoila, as described in Hatch 1979). Second, Snow et al. (1981) showed that Dutch municipal workers used more ungrammatical utterances to migrant workers whose Dutch interlanguage contained more grammatical errors. Third, in an eleven-month longitudinal study of the speech addressed to a five-and-a-half-year-old Hebrew-speaking child, Tamar, by an American age-peer, Lisa, Katz (1977, 1981) found that the number of ungrammatical morphosyntactic features of Lisa's speech gradually declined over time, but that the variety of such features increased, as did the variety of communicative functions served by Lisa's FT. While the changes appeared not to be directly related to improvements in Tamar's English, Katz suggests that the persistent interlanguage patterns in Tamar's phonology may have continued to trigger the persistent phonological features of Lisa's FT. Improvements in Tamar's morphosyntax may have promoted the gradual decline in morphosyntactic FT features in Lisa's speech, this decrease mediated by the increase in the range of communicative functions the FT served. Last, studies which have manipulated the linguistic proficiency level of the NNS (Chaudron 1978; Dahl 1981; Downes 1981; Trager 1978, and especially Gaies 1977), have *generally* been supportive of the idea that the level of the NNS's second language proficiency is inversely related to the variety and/or extent of modifications made, although results in all but Gaies' study have been extremely variable at the level of individuals.

Now Varonis and Gass could easily argue that second language proficiency in these studies is simply being confounded with comprehensibility to the NS, such that increased proficiency, including greater grammaticality, means greater comprehensibility. This may turn out to be the case, but there are at least two alternative possibilities. One is that the comprehensibility of the NNS's *initial*

utterance in the street directions study (or of single isolated sentences in the laboratory studies) is a determining factor only in the *initial* evaluation/reaction by the NS. (The studies by Varonis and Gass consider *only* immediate reactions to single NNS utterances.) Another is that comprehensibility may be a *persistent* factor in determining NNS' reactions, but that it is not the only one, and not necessarily even the most important one. All the research cited above, including Varonis and Gass' own studies, would be consistent with either of these interpretations. Even though none of the researchers except Varonis and Gass controlled for comprehensibility, their findings could be showing just what they appear to be showing, namely, that linguistic features of the interlanguage (pronunciation, grammaticality, and/or general proficiency level) determine NS adjustments *beyond* short term reactions to initial utterances.

A more plausible hypothesis than Varonis and Gass' 'comprehensibility' solution, although admittedly a more powerful one, is that while NNSs *can* react to comprehensibility alone, as Varonis and Gass demonstrate, they generally *do* react to a *combination of factors*. It has already been argued that linguistic features of the NNS's interlanguage are among these, but another more important element appears to be the NNS's *comprehension* of what the NS is saying, which the NS assesses from the feedback the NNS provides. Varonis and Gass dismiss this possibility on the basis of their observation that NNSs in their street directions study reacted differentially to NNSs and NNSs in their *first* response (more frequently using repetition of the NNS's question) *before* the NNS had a chance to show lack of comprehension. Note, however, that this observation is entirely consistent with the idea that comprehension (by the NNS) is a, or even *the*, factor determining *subsequent* reactions. One could turn the argument on its head: NNSs *cannot* use their assessment of what the NNS understands to determine how they should adjust their speech until they have a chance to assess that understanding, i.e. until they have engaged him/her in conversation. Thus, they initially *have* to rely on other factors. Beyond their first utterance, however, they begin to receive and to use feedback on the NNS's comprehension, and adjust their linguistic and conversational performance accordingly.

There is a fair amount of evidence consistent with the idea that *comprehension* (by the NNS) is, if not the only, then at least a *crucial* variable in stimulating NNSs' reactions. The evidence consists of the finding that *removal* of feedback on NNS/child comprehension of a NS's/caretaker's message results in speech/conversation which is either hardly adjusted at all or at least not adjusted to the level previously observed in two-way communication between the same parties. Three studies have obtained some version of this result. Snow (1972) found that mothers first observed to modify their speech when talking to their own children, made hardly any adjustments when asked to record a taped message for the same children in the child's absence. Steyaert (1978), in a replication of Gaies (1977), also found few modifications in the speech of NNSs telling stories to classes of ESL students, who were not allowed to respond to the story-tellers in any way. Long (1980) found no statistically significant differences between FTD and NS-NS conversation on various input and interaction measures when the dyads (16 in each condition) were engaged on three tasks which did not require two-way

communication of information (vicarious narrative, giving instructions and giving an opinion as to the purpose of the research study they were participating in). On the two-way tasks, on the other hand, (informal conversation and two cooperative problem-solving games), the same dyads performed significantly differently on the same measures. (The one-way and two-way tasks were carried out in the same *mixed* sequence [211221-way] by all NS-NS and NS-NNS dyads, meaning that the results could not have been due to task order.) In all these studies, when NNS feedback on comprehension was removed, NSs did not adjust, or did so only minimally. Yet the same NSs, or in the study by Steyaert (1978), equivalent NSs in a parallel study (Gaies 1977), *did* adjust in the feedback condition, when they were able to assess NNS comprehension. (cf. Gaies 1981b for further analysis of the role of learner feedback in negotiating comprehensible input.)

In summary, it has been argued that NSs react to a *combination* of factors when they make linguistic/conversational adjustments to NNSs. These include the *comprehensibility* of the NNS's interlanguage, the interlanguage's *linguistic characteristics*, and the NNS's apparent *comprehension* of what the NS is saying. In certain limited circumstances, primarily initial reactions to small interlanguage samples, the NNS's comprehensibility and/or physical appearance may be more influential.³

What difference do the adjustments make to second language acquisition by the NNS?

One of the strongest motivations for the study of linguistic/conversational adjustments to NNSs is the possibility that some or all of them play a role in SLA. They may affect the course or *process* of acquisition. They may increase the *rate* of acquisition. They may affect the learner's level of ultimate *attainment*. They may be *necessary* for acquisition to take place at all. They may even be *sufficient* to guarantee that it occurs. Or they may be altogether *irrelevant*.

One way of investigating these issues would be the use of a true experimental design to test for a direct, casual relationship between some/all of the adjustments and some aspect of acquisition. To my knowledge, this approach has not been tried so far. A difficulty with it in a second (as opposed to foreign) language environment will be that language acquisition takes time, generally too long a time for the researcher to maintain control over the type and amount of input/conversational experience that subjects in a study receive elsewhere, e.g. in street learning or through reading.

Another approach is through *ex post facto* correlational designs in which features of subjects' input/conversational experiences are related to some aspect of acquisition. Snow and Hoefnagel-Höhle (in press⁴) report a study of this type. As part of a larger investigation, samples of the 'comprehensible' speech heard by 13 American children during content instruction in Dutch classrooms was recorded using back-pack microphones attached to the children, one child per classroom. Half-hour segments of the talk were transcribed and speech to the children by their teachers and peers compared with other classroom talk. Speech to the

children was found to contain shorter utterances, and also proportionately more one-word utterances. Three measures of linguistic input were calculated: (1) quantity of speech heard; (2) percent directed to the subject; and (3) quantity directed to the subject. Scores on these three measures when then correlated with scores on seven tests of Dutch proficiency for six of the children 4-5 months later, and (again for these six children) with gain scores from time one to time two. (The children were of differing ages and proficiency levels.) Snow and Hoefnagel-Höhle found almost no statistically significant relationships between *any* of the three input measures and the seven proficiency scores at time two. In fact, relationships were strongest between proficiency scores and the number of utterances heard—not, as might have been predicted, with either measure of the modified speech directed to the learner. The same (null) findings were obtained when gain scores were substituted as the dependent variable. As many *negative* as positive correlations were noted. In other words, results were very variable and did not pattern as predicted.

Snow and Hoefnagel-Höhle are careful to point out several limitations of their study. For example, the n-size (six) was very small, as was the sample of classroom speech (30 minutes). It is also not necessarily justifiable to conclude that because the speech to the children was (a) modified, and (b) comprehensible to the investigator, it was comprehensible to the child NNSs. Also, the study did not control for the amount or type of input subjects obtained outside school during the five-month period. Any or all of these limitations may have effected the outcome. The study should be applauded as a first attempt to address this issue, and its null findings probably need not discourage further research of this type.

Thus far, two approaches to investigating relationships between adjustments and SLA have been outlined. Each has its problems, but each is feasible and will no doubt attract future researchers. Meanwhile, a third, indirect approach to the same issue may also be worth pursuing. This approach breaks the task down into two steps, followed by a deduction. Put simply, if it could be shown that the linguistic/conversational *adjustments* promote *comprehension* of input, and also that *comprehensible input* promotes *acquisition*, then it could safely be deduced that the *adjustments* promote *acquisition*. If A signifies adjustments, B comprehension, and C acquisition, then the argument would simply be:

$$\begin{array}{l} A-B \\ B-C \\ \therefore A-C \end{array}$$

where ‘—’ indicates a causal relationship.

The first relationship, that between some/all *adjustments* and *comprehension*, has attracted only minimal attention from SL researchers and is in urgent need of study. Some work has been done, however, and initial results are encouraging. As part of a larger study, Johnson (1981) found that 46 intermediate/advanced Iranian students of English were able to recall more events from an adapted (linguistically ‘simplified and paraphrased’) version of an American short story they had read than from an unadapted version, although the cultural origin of a text was *generally* more important for recall than was linguistic adaptation.

In another study (Chaudron 1982), this time of aural comprehension, five lecturettes on different topics were scripted such that each lecturette contained a paired set of sub-topics which were mentioned twice and later restated. The (sub)topics were each encoded in one of five different ways:

- (1) repeated noun (The beer . . . the beer tastes terrific)⁵
- (2) simple noun (The beer tastes terrific)
- (3) rhetorical question (What about the beer? It tastes terrific)
- (4) if-clause (If you can afford the beer, it tastes terrific)
- (5) synonym (The brew tastes terrific).

The order of occurrence of the topic-restatement devices and the textual distance between second and restated mentions were control variables. The lecturettes were played to 135 adult ESL students in two conditions: *recognition only* (n = 60), who heard each lecture and then answered 12 recognition questions, and *recall-recognition* (n = 75), who did the same as the recognition-only group, but also completed four short cloze items (two on the restatements and two distractors) during brief interruptions of the lecturettes. Chaudron found that *recognition* scores were significantly higher for repeated noun than for simple noun. *Recall* scores for the recall-recognition group were significantly higher for repeated noun than for if-clause and synonym, and scores for simple noun superior to those for synonyms. (Aggregating scores across the various proficiency levels represented in the sample, the overall pattern for recall scores was [1] through [5] above.)

Chaudron suggests that the *redundancy* of repeated noun probably made it the most successful restatement device overall, at least in its effect on immediate language intake, but notes that different devices proved more useful to learners of different SL proficiency. (cf. Chaudron 1982, in press for full details.) If we assume that 'simple noun' was equivalent to *unadjusted* (NS-NS) speech in his research, then both studies find that at least some kinds of adjustments promote comprehension by NNSs.

With respect to the second relationship, that between *comprehensible input* and *acquisition*, a review of 'natural experiments' (Long 1981c, 1982) found evidence consistent with the idea, and none inconsistent with it. Studies of first and second language acquisition in normal and abnormal populations offer considerable support for three generalizations:

- (1) access to comprehensible input is a characteristic of all cases of successful acquisition, first or second;
- (2) greater quantities of comprehensible input result in better (or at least faster) acquisition;
- (3) lack of access to comprehensible input (as distinct from *incomprehensible*, not any, input) results in little or no acquisition.

Of the three generalizations, the third is clearly the most crucial. A claim that comprehensible input was necessary, let alone 'the single most important' factor (Krashen 1980) in SLA, would collapse if cases could be found of *beginning* learners acquiring a language from linguistically/conversationally unmodified NS-NS talk. No such cases, however, appear to exist.⁶ While not the same as showing a causal relationship between comprehensible input and acquisition, generalizations (1)

and (3), taken together, suggest that comprehensible input is a *necessary* condition for first or second language acquisition and summarize findings consistent with a *hypothesized* causal relationship.

To recapitulate, there appears to be sufficient evidence of (a) an adjustments-comprehension relationship, and (b) a comprehensible input-acquisition relationship to warrant the deduction of (c) an adjustments-acquisition relationship. There is, however, clearly a need for direct tests of (c). Such research addresses what might be termed the 'weak' version of an input hypothesis, namely, that linguistic/conversational adjustments are necessary for SLA.

The 'strong' version of the hypothesis predicts that the adjustments are not only necessary, but sufficient to guarantee acquisition:

'... participation in conversation with NSs, made possible through the modification of interaction, is the necessary and sufficient condition for SLA' (Long 1981c: 275).

This hypothesis was originally based upon the arguments for the 'weak' version, coupled with the observation that near-native fluency in a SL is (supposedly) a common phenomenon among post-critical period starters in many multilingual societies in the Third World in languages for which instruction is unavailable (cf., e.g. Sorenson 1967, and findings summarized in Hill 1970). Findings of a recent study by Schmidt (1981), however, appear to falsify the "strong" version. Schmidt reports on the early fossilization of 'Wes', a middle class, adult Japanese immigrant to the USA (Honolulu). After approximately three years of *extensive* contact with English NSs, Wes' interlanguage is reminiscent of Alberto's (Schumann 1978). Not only is there minimal English morphology, but also no apparent development in this area for about two years. If Wes is more advanced than Alberto, it is chiefly in his greater fluency and higher degree of 'communicative competence', as illustrated by his ability to perform successful and usually appropriate directives in a range of settings. Unlike Alberto, Wes has a nearly optimal acculturation profile, is a well-to-do, highly successful artist (painter), and has a wide circle of friends and admirers, most of whom are NSs of English. His limited progress, in other words, cannot be explained in terms either of high social/psychological distance or as a result of *restricted* input, both of which may have operated against Alberto.

The only way that Wes' lack of development could be compatible with the strong version of the input hypothesis would be if it were shown that, while plentiful, Wes' input was inadequate in other respects. Could it have been undemanding, i.e. pitched too low (at 'i' or 'i-' in Krashen's terms) by people who were very familiar with what Wes knew and adept enough at FT/FTD to eliminate material at 'i+i' in the interest of easy communication? Could some of it have been *too* demanding, i.e. *unadjusted* ('i+50'), and therefore incomprehensible, as when people unfamiliar with Wes and inexperienced at FT/FTD spoke to him? Schmidt (personal communication) reports that both phenomena occur, but that Wes constantly receives large amounts of input that is both comprehensible to him and contains many items not in his interlanguage, yet which he does not proceed to

acquire. Examples (provided by Schmidt) include plural, article, possessive adjectives and pronouns, possessive 's and regular past -ed. While some of these items are 'late acquired', according to the 'natural order' (Krashen 1977), others, such as plural, are not. Yet Wes shows no indication of acquiring these items either, despite adequate opportunity and despite the fact that their absence, Schmidt reports, not infrequently leads to communication breakdowns and, hence, to potentially 'corrective' feedback on communicative efficacy.

While Schmidt's is a case study of a single individual, there are probably a lot of other (non)acquirers like Wes. His apparent uniqueness, in other words, is almost certainly more a reflection of the paucity of studies of this type than of idiosyncratic SLA behavior. As such, Wes provides serious counter-evidence to the *strong* version of the input hypothesis, the status of which, therefore, must currently be considered doubtful.

Conclusion

A considerable research effort in the last few years has served to provide information on the linguistic and conversational adjustments made by NSs communicating with non-natives. The work has also given rise to (sometimes competing) hypotheses as to which kinds of NSs react in different ways, as to what they react to, and as to the effect the reactions have on SLA. In all of these areas, the time is now ripe for research to move from a predominantly descriptive to an experimental stage. The results should prove useful not only to SLA theory construction, but also to applied linguists concerned with such issues as the "simplification" of teaching materials and the design of efficient classroom language acquisition environments.

Notes

1. It will be interesting to see what kinds of changes (if any) result in ungrammaticality in FT in languages with different structural characteristics, e.g. those with little or no inflectional morphology. Mandarin and Japanese have been considered in two studies (Chan and Choy 1980, and Long, Gambhiar, Gambhiar and Nishimura 1982, respectively), but neither found any instances of ungrammatical speech, perhaps because both involved arranged encounters in laboratory settings.
2. A recent study of Portuguese FT in street directions (Goldberg 1982) lends only partial support to this generalization. When addressing a middle class English-speaking tourist, ungrammatical utterances were found in the speech of more upper-middle class (higher status) than middle class (equal status) or working class (lower status) speakers. A few middle and working class speakers did use some ungrammatical utterances, however.
3. The question of whether initial and long term reactions to speech samples differ would also be an important one in research on teacher attitudes to accented speech of various ethnic groups. Most such studies in the U.S. during the last 15 years have reported something like ethnic stereotyping by teachers on the basis of their reactions to small samples of talk by children who were unknown to them. It might be the case that the elimination of other means of evaluating character, attitudes, ability, etc., which is an inevitable by-product of the kind of research design used, virtually forces the teachers to offer a stereotypical reaction, and/or that the same teachers' reactions would have been different after exposure to the same children over time in their own classes.
4. I am grateful to Catherine Snow for allowing me access to the draft version of this report. The final publication will be of a slightly revised version.

5. All examples are from Chaudron (in press).
6. Apparent counter-examples are provided by Ochs (1982) and Schiff (1979). The excerpts of caretaker-child conversation given in Ochs (1982), however, show that while Western Samoan caretakers may not employ a conventionalized variety of Baby Talk, they do make the usual adjustments (repetition, etc.) to the *interactional structure* of conversation. The hearing children of deaf adults, studies by Schiff (1979), who acquired English morphology in the normal sequence, are also not counter-examples since they are reported to have received between 15 and 25 hours per week of normal caretaker-child conversation with relatives and peers outside the home.