DENNIS R. PRESTON

Building a New Language:
A variationist account

Dennis R. Preston is University Distinguished Professor of English at Michigan State University, specialising in sociolinguistics and dialectology. A former Vice-president and President of the American Dialect Society, he is a prolific writer and editor in the field of linguistics, with a track record of well over a hundred scholarly articles and twenty-odd books.

1. Introduction

A few years ago I had occasion to take the first semester of a *hard* language. (*Hard* is a euphemism for "non-Indo-European" in the Western language learning and teaching tradition.) Since I was a faculty member, I was on collegial terms with the instructors, and I was even the advisor of some of the teaching assistants who taught the class about half the time. The students in my class consisted of a few speakers of a structurally similar language, one heritage learner, several students who saw business opportunities in the language's homeland or were attracted to its cultural products, some of the latter majors in the university's area studies program connected to this language, and two curious older people (my wife and me).

The teaching assistants were all linguistically sophisticated; in fact, the one who worked with our class was a candidate for an advanced degree in linguistics, and some of the instructors, all trained language teachers, also had a good general linguistics background.

One day one of the assistants, a few other linguists, and I were sitting around discussing life and language, and the assistant observed that she could not understand the strange lack of progress of many in her section. I paraphrase her comment from memory, but I assure you that it was so striking that I do not misrepresent the content in any way:

> Can anybody explain to me why our teaching is so unsuccessful? I spend an entire period explaining a point of grammar; we practice it thoroughly; it's clear the students understand it and can produce it correctly. When they come back for the next class, however, they consistently perform this construction incorrectly. How is it possible for them to learn the correct form, show good evidence, even on tests, that they have learned it, but continue to use it incorrectly?

Researchers in second language acquisition (SLA) and experienced teachers may chuckle at the naivety of this, and the full response to such a mistaken notion of learning would take us deeply into the entire history of the enterprise. It would include at least recent discussions of access (full, partial or none) to universal grammar, of transfer, interlanguage, form-focused teaching and learning, the critical period, memory, and other attempts to account for the character of the learner's states along the path from no knowledge to considerable (if not ultimate) achievement in the second language.

2. Variation or several parallel languages?

I would be equally or more naïve if I pretended that I was about to reveal here a final answer to even part of any of these long standing research initiatives in SLA. I will, however, suggest that one general approach to these mysteries has a greater potential for applied linguistic considerations than has yet been fully exploited in research as well as pedagogical frameworks. I refer to the *variationist* approach to SLA.

Not surprisingly, variationists (sometimes called *sociolinguists*, although that latter term covers a great deal more territory than the quantitative approach I outline here) believe that a speaker's linguistic makeup is variable. In fact, everybody believes this, but variationists take variation to be a serious matter within the linguistic competence of an individual speaker of a single language. Others believe such variation is trivial:

> ...every human being speaks a variety of languages. We sometimes call them different styles or different dialects, but they are really different languages, and somehow we know when to use them, one in one place and another in another place. Now each of these languages involves a different switch setting. In the case of [different languages] it is a rather dramatically different switch setting, more so than in the case of the different styles of [one language].... (Chomsky 1988:188)

For Chomsky, therefore, it would appear that the learning of a second language is not essentially different from the learning of a new style (or dialect) of one's own language. But does the learning of the appropriate switch settings of a dramatically different language differ from the process of learning another style or dialect of the same language? Chomsky does not comment on that possibility. At the least, variationists disagree with his characterization of the cognitive situating of the multiplicity of styles available to every speaker. If the variation evidenced when a speaker shifts from a casual to formal style, for example, requires a shift to a different language (or, as Chomsky would have it, a different grammar), the principle of parsimony seems to me to be seriously violated. In other words, unless a carefully specified set of distinctly different structural features is involved, why would a single speaker need many grammars when one (perhaps a few more for multidialectals) will do?

For example, do I really need several different grammars to cover the dialectal and stylistic fluctuation among the following:

English English Formal: *I have a dollar. Have you a dollar?*
American Formal: *I have a dollar. Do you have a dollar?*
English English and American Informal: *I've got a dollar. Have you got a dollar?*
American Casual: *I got a dollar. (You) got a dollar?*
American nonstandard: *I got a dollar. Do you got a dollar?*

All these varieties share a grammar that already has various means for forming interrogatives: *do*-support ("Do you work every day?"), subject-verb inversion ("Has he worked here all day?"), and intonation only ("You work every day?"). The fact that some dialects assign these strategies to different verbs (e.g., English English allows stative main verb *have* inversion; American English does not) might seem to suggest that different grammars are involved. But in American as well as English English (and within the competence of an individual speaker), it is possible to reanalyze *have got* as a present (meaning "to possess"), and in American English it is possible to even further reanalyze *got* without a perfect marker as a present with the same meaning, and that reanalysis allows only the intonation interrogative (except in some nonstandard varieties). Since these means for interrogative formation are available to all varieties of English and only distributed differently for certain verbs, this seems to variationists to be workable within a single grammar. In short, we find such variation *inherent*. The selection of one form or another is based on linguistically external (topic, interlocutor familiarity, setting) and internal (preceding and following forms, minor shifts in emphasis and meaning) factors. All these factors come into play in determining the choice among these variants in a particular use. Such a speaker has a single grammar with all these variants embedded in it (or characterized by it), and they are available for activation by co-occurring social and linguistic facts.

3. Interlanguages with variable rules

A variationist, then, might offer at least one good explanation for why students who appear to have good knowledge of a rule are nevertheless unable to consistently produce utterances that conform to that rule. Their *interlanguages* are, at that moment, comprised of variable rules, ones that are sensitive to both internal and external pressures on their realization, just as native speakers' linguistic competences are.

Let me be clear by what I mean by *internal* and *external*. By *internal*, I refer to features of the linguistic system itself. By *external*, I refer to features not a part of language structure itself, for example, such *social* factors as gender, ethnicity, status, or social network. There are clearly more or less permanent external factors (e.g., sex) and obviously transient ones (e.g., interlocutor familiarity). Such more or less permanent demographic characteristics have a relatively constant influence on a speaker's choices; transient ones have a much wider range of probabilistic influences since interlocutor familiarity is obviously enormously variable from situation to situation.

Let's pretend, for example, that because of my sex, age, status, regional background, and ethnicity, all more or less permanent aspects of my identity, I am likely to choose the "Do you have…" version of the interrogative described above at an overall rate of, say, 50 percent. But let's say that I run into a good friend of mine; the contribution to informality of meeting my friend will reduce that probability considerably, since I am likely to use the "Do you have…" form only 5% of the time in such casual situations. I will need to go on and calculate the likelihood of this situation contributing to my choice of the remaining "Have you got…," "You got…," and "Got…" forms available to me. Of course, other external and internal factors will need to be calculated to see the probability of my selected form in relation to each of the forms available to me in this particular instance of use. Interestingly, when we have calculated the actual percentage of such alternative forms in authentic discourses, such factors turn out to be very good predictors of actual use.

4. Internal and external pressures on interlanguage

Here, however, I will focus primarily on *internal* factors and will illustrate their operation and type from several previous studies. In a study of Czech/Slovak and Chinese learners of English, Young (1990) looked at the degree to which several internal (*linguistic*) factors influence the likelihood of noun plural marking.

Table 1. The effect of animacy on noun-plural marking in Chinese, Czech, and Slovak learners of English, percent correct (Young 1990)

	Respondents' L1	
	Czech/Slovak	*Chinese*
Animacy		
Animate	68%	34%
Inanimate	84%	59%

Although the Czech and Slovak learners (there was no difference between them) performed better overall than the Chinese learners (and all three groups were at similar levels of proficiency), there is an obvious preference for marking plural inanimates (e.g., *tables*) than plural animates (e.g., *boys*) for both groups. Animacy of nouns is very clearly an internal, linguistic matter, and it also clearly has an effect on plural marking that is similar across very different languages. We might expect, then, that some *universal* features (here, for example, perhaps the marking of animates and inanimates) would have an effect on the developing interlanguages of all learners.

On the other hand, Young also studied the effects of sentence function on plural marking and found the following:

Table 2. The effect of sentence function on noun-plural marking in Chinese and Czech/Slovak learners of English, percent correct (Young 1990)

Respondents' L1

Function	Czech/Slovak	Chinese
Subject	75%	35%
Object	87%	38%
Adverbial	86%	81%
Complement of 'be'	74%	57%

In this case, sentence function promotes greater variability in the Chinese learner group (the high level of correctness for adverbials compared to the much lower level for the other positions), while the same cannot be said for the Czech/Slovak learners. Again, it is an internal or linguistic factor (function of the noun in the sentence) that contributes to the variability, but, since it is different across language boundaries, we might assume that it is not a *universal* feature, as animacy appeared to be, but a *language-specific* one.

Finally, variationists who have studied SLA have also identified internal influences related to neither universal nor language-specific features but to proficiency level.

Table 3. Effect of the preceding segment (of the verb stem) on the past-tense marking of Chinese learners of English, percent correct (Bayley 1994:174)

Respondent proficiency level

Preceding segment	Lower	Higher
Vowel	23%	61%
Liquid	31%	45%
Obstruent	23%	36%

Here there is little influence from the preceding sound on the production of the past-tense marker by the lower proficiency learners, but the advanced learners display a variability that is, interestingly, more like native-speaker variation for this same element. For example, in dialects of English, the past-tense marker of *walked* (with final obstruent /k/) is more likely to be deleted than the past-tense marker of *filled* (with final liquid /l/), which is in turn more likely to be deleted than that of *tried* (with final vowel /aɪ/). Once again, then, an internal factor, the phonetic shape of the preceding

sound, has an influence on interlanguage variability.

Such internal factors may interact with others not themselves a part of the linguistic system, i.e., *external* ones. For example, Tarone (1985) studied the effect of attention to form on third person singular indicative marking on present-tense verbs (3rds) and on article use by Arabic and Japanese learners of English. I report here only the results for Arabic learners, although the two groups were not significantly different overall.

Table 4. The effects of attention to form on 3rds verb marking and article use among Arabic speaking learners of English, percent correct (Tarone 1985)

Attention to Form (high–low)

Feature	Test (high)	Interview (mid)	Narrative (low)
3rds	67%	47%	38%
Article	38%	85%	93%

The grammar test, in which the students identified incorrect usage and supplied the correct form, was assumed to be the setting in which they would pay the most attention to form; the interview, with a native speaker of English, was an unplanned conversation about the student's major field of study and future plans and was assumed to allow less attention to form than the grammar test. The narrative task, in which the student told a story to a fellow non-native speaker from a sequence of pictures presented on video, was assumed to be extremely content-oriented, allowing the least attention to form.

Although attention to form (or *monitoring*) clearly had an influence on interlanguage variability, the internal factor (the form itself) was also extremely influential. In one case, more planning resulted in better performance (the 3rds marker), but in the other (the article), the more monitoring, the worse the performance. The explanation for this odd behavior would seem to lie in the fact that when one monitors for 3rds, one finds an easy, simple-to-apply rule; therefore, the more monitoring, the better. When one monitors for the article, however, as any learner or teacher of English will surely attest, the rule is not so easy, and monitoring is, apparently, more likely to lead to error than the unmonitored dependence on whatever unconscious sensitivity to this feature has been developed by the learner.

Once more, an internal factor (a verbal morphology rule versus an article selection rule) governs interlanguage variability and, in this case, interacts interestingly with attention to form (or *style* as variationists would call this external measure).

It will not do, however, to leave you with the impression that sociolinguists are not interested in the *social*. For example, Young's study of plural marking among Chinese learners of English (1988) shows that learners

with the greatest *convergence* (or *interaction*) with native speakers perform best, and there are many SLA studies that focus on sociocultural, social psychological, and motivational aspects that suggest such findings are not unusual and, therefore, require attention to a wide range of such external factors.

5. Calculating a language learner's expected success rate

What I have hoped to outline briefly here is a model of language acquisition that will assist teacher trainers, textbook writers, test writers, and curriculum designers in preparing teachers, materials, tests, and courses that are sensitive to the fact that a second language grows in the context of a set of influencing factors. At any given moment in the learner's career, those internal and external factors will exert different influences (or *weights* as variationists like to call them) on the probability of the learner's performance of a particular feature.

For example, suppose one of Young's Czech/Slovak speakers is about to say the sentence "Elephants are big." What's the chance of his or her correctly assigning a noun plural marker to *elephant*? Since elephants are animate, the chance is lessened compared to inanimates (68% for animates; 84% for inanimates; Table 1). The chance is also lessened compared to other sentence functions since *elephant* is the subject of this sentence (75% for subjects; 87% for objects; Table 2). Now let's make the phonetically reasonable assumption that Czech/Slovak speakers are sensitive to the same preceding phonetic environment that Chinese speakers are for past-tense marking (Table 3). Since the plural (/s/) to be added to *elephant* is preceded by a /t/ (an obstruent), I will suggest that it is less likely to appear than it would in words that end in vowels (*cow*) or liquids (*bull*). Since Czech/Slovak learners seem to be advanced over Chinese learners, I will suggest a probability of 45% (compared to the Chinese learners' 36% of Table 3). I will also suggest that the appearance of this item (which is always a consonant in regular English plurals) is enhanced by the fact that the following word begins with a vowel (*are*), and I'll make up a percentage of 80%. Let's tally all this:

Internal factor influences:

Animacy	68%
Subject	75%
Preceding obstruent	45%
Following vowel	80%
Total	268
Average	67%

This suggests that on the basis of internal, linguistic factors alone, we would expect our Czech/Slovak speaker to mark this noun (and other nouns

like it in the same environment) about 67% of the time. Just for fun, let's assume that this particular learner interacts very frequently with native English speakers and that that fact contributes a higher likelihood of marking (as Young's 1988 study of Chinese learner's plural marking showed); I'll estimate 90%. Let's also assume, however, that in this particular study the sentence was obtained from a casual conversation, in which the speaker had no opportunity for monitoring. If we assume that noun plural marking is an "easy" rule in English (as we did for 3rds in Tarone's 1985 study), then it's also safe to assume that this inability to monitor would cause noun plural marking to decrease; let's say 60% for speakers like our imaginary Czech/Slovak. We return to our above statistics and add 90% (for convergence with native speakers) and 60% (for an unmonitored or *casual style* situation) and get a new grand total of 418 and an average of 69.33%, our new estimate of the likelihood that our imaginary speaker will pluralize *elephant*.[1]

Once a rule has been taught and successfully practiced, there is, therefore, no guarantee that the variability that arises from such influences as those discussed here will disappear. Some deep-seated influences may last well into the advanced stages of an interlanguage; others may disappear quickly, and still others may interact with such other influences as topic, setting, or the ability to monitor one's performance.

6. Conclusion

The variationist position in SLA suggests that the developing interlanguage of a learner contains a complex bundle of influences best studied by a multivariate approach and best treated as a cognitive competence made up of many competing influences on language use. The more we recognize that learners do not simply switch back and forth between the rules of L1 and L2, but attend to rules whose applications are sensitive to the sorts of factors outlined here, the less likely we will be to complain, as my "hard" language instructor did, that some learners just don't seem to get it right, even after they have been taught.

As one might suspect, there is a large literature on variationist approaches to SLA, including some that touches on the cognitive models such variation implies. A good place to see examples of this work and an outline of the history of it is Bayley and Preston (1996).

---

[1] This is not how such probabilities are calculated since percentages are not reliable indicators. The current practice among variationists is to use a logistic regression format for the calculation of the probability (or *weight*) each of these factors has in its influence on the object of analysis. Several examples of this statistic can be found in Bayley and Preston (1996) along with an explanation of how to use a logistic regression program specifically prepared for such research (VARBRUL or GoldVarb).

References

Bayley, Robert. 1994. Interlanguage variation and the quantitative paradigm: Past-tense marking in Chinese-English. Tarone, Elaine, Susan M. Gass & Andrew Cohen (eds), *Research methodology in second-language acquisition*. Hillsdale, NJ: Lawrence Erlbaum. 157–81.

Bayley, Robert & Dennis R. Preston. 1996. *Second language acquisition and linguistic variation.* Amsterdam: Benjamins.

Chomsky, Noam. 1988. *Language and problems of knowledge: The Managua lectures.* Cambridge, MA: The MIT Press.

Tarone, Elaine. 1985. Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning* 35:373–403.

Young, Richard. 1988. Approaches to variation in interlanguage morphology: Plural marking in the speech of Chinese learners of English. PhD dissertation, University of Pennsylvania (Published as *Variation in interlanguage morphology.* 1991. New York: Peter Lang.)

Young, Richard. 1990. Functional interpretations of variation in interlanguage morphology. A paper presented to the Second Language Research Forum (SLRF), "Variability and second language acquisition," University of Oregon, March.

Two Poems by Dennis Brutus

Buried in Camden

for Walt Whitman

You old greybeard poet
who sang of a free world
where men loved men freely
and all enjoyed earth's bounty
still I salute you
and wish to send words of praise

but if words cannot make something happen
cannot ring out to crack
this obdurate and heartless firmament
let me be dumb
let my lips be silent
let my mouth be stopped
with grime and dust
here where I see decay
where blood runs in the streets
earth shudders in agony
and men gagged by blood clots, plead for pity
let me be courageous
to cry out for a just world
cry 'Justice', 'Justice', 'Justice'
or let me fall silent.

What will it take
to make the heartless humans of the world
listen to our cries?
I will go on my knees
and howl like a dog
a starved chained dog
howling at the moon
and cry 'Justice', 'Justice', 'Justice'
until my corpse falls
worn-out and rigid and silent.

Entering Table Bay

Dance with a skeleton:
macabre jollification:
below saude beat of tam-tams
hear creak of bony joints
shuffle of splayed tarsals:
ambivalent embrace
of bedraggled amour
beauty bedizened for brothels frolics:
rank scent of withering petals
sprawled from dissembling heart
rotted crotch of a decaying rose
where still, absurdly, stamens gleam
with pledges of resilient life