DESERET LANGUAGE
and
LINGUISTIC SOCIETY

Selected Papers from the
Proceedings

Eleventh Annual Symposium

14-15 February 1985
Brigham Young University
Provo, Utah
ACKNOWLEDGMENTS

We gratefully acknowledge the generous support of the Brigham Young University College of Humanities, its language departments, and of CMR Computer Services for the Eleventh Annual Symposium of the Deseret Language and Linguistic Society and the publication of these proceedings.

We are also grateful to those who presented papers at the symposium, to those who chaired sessions or helped with other organizational matters, as well as, of course, to the many who attended and (we hope) benefited from the presentations and from the discussions that followed each.

Special thanks are due our invited speakers, Professor Lily Wong Fillmore (University of California, Berkeley) and Dr. Martin Kay (Xerox, Palo Alto Research Center), for their stimulating presentations. We are happy to be able to include copies of their remarks in these proceedings. As the reader will note, both papers retain the engaging, informal style of our guests' actual presentations. In fact, Dr. Kay's paper was transcribed, with his permission, directly from our video tape recording of his presentation.

Authors and titles of all papers presented at the symposium are listed below, by section, followed by a table of contents which lists the titles of those papers submitted for inclusion in the current proceedings, in alphabetical order by author (after the invited papers). The content and accuracy of the papers contained herein are solely the responsibility of their authors.

Again, thanks to all who contributed to the success of DILS 85, especially to Professor Fillmore and Dr. Kay for their sharing of valuable time and insights with the members of our society.

Robert A. Russell
Thursday, 14 February 1985

Steven Wright, Session Chair
Boyce Fitzgerald  Han-gul and Changes in Korean Phonology

Kwang-Oo Lee and Bill Quinn  Reading Han-gul Made Easy Using a New Simplified Method

Soo-Young Choi  A Paradigm to Solve the Korean Postpositions nun/ka

Harold Madsen, Session Chair
Wilfried M. Voge  Reexamining the Study of Foreign Language as a Universal General Education Requirement

Wilfried Decoo  Looking Back: The Didactics of Semantization in Historical Perspective

Robert W. Blair  Simulation Game as Envelope for Beginning Language Instruction

Gloria Hormazabal, Session Chair
Frank C. Lefrandt, Jr.  Markedness vs. Unmarkedness in the Mayan Verb Paradigm

Rafael Rodriguez  A Transcendental Archaeological Find in Puerto Rico: The Anthropoglyphics of Father Nazario

Invited Speaker
Martin Kay  Unification in Linguistic Computing

Royal Skousen, Session Chair
Frank Otto  Teacher Involvement in Courseware Selection

Randall L. Jones  German CAI with Interactive Videodisc: The Extra Dimension

Alan K. Melby  Using a Text and Translation Data Base

Kenton L. Chauncey  How Useful is Artificial Intelligence in the Study of Language?

Russell Bateman  Linguistic Tools Incidental to Work in Machine Translation
John Robertson, Session Chair

Kathryn Afarian  An Older Learner's Response to a Book of Mormon Burling Reader

Matthew R. Sorenson  A Computerized Diagnostic Test for Dyslexia

Cheryl Brown  Analyzing the Reading Skills of Advanced Language Learners

David Wright, Session Chair

Stephen D. Ricks and John D. Frauzel  Some Observations on Proper Name Formation in the Book of Mormon

Marvin H. Folsom  The Politicization of Repentance

Geraldine Bills and A. Garr Cranney  A Study of LDS Family Scripture Reading Practices

Margaret Baker  Some Effects of the Lay Ministry on Mormon Language

Friday, 15 February 1985

Dilworth Parkinson, Session Chair

Dan Higinbotham  The Banyan: A Partially Ordered Syntactic Structure with Reintersection

J. Halvor Clegg and Mark E. Seely  Set Theory as an Explanation for the Selection between mas que / mas de in Spanish

Matthew R. Sorenson  An Information Theoretic Interpretation of the Variable Rule

Cheryl Brown, Session Chair

Lisa Blair  Performance Differences between PRC and ROC Students on Cloze

Roberta Chock  How Can Proficiency Test Scores be Used to Enhance Classroom Effectiveness?

Harold S. Madsen and Jerry W. Larson  Employing Computerized Adaptive Language Testing Techniques

T. Richard Chi, Session Chair

Marian McMaster  Latest Trends in Curriculum for Translation and Interpretation Programs

Grant Boswell  Imitation to Experience: Changing Assumptions in the Teaching of Writing
<table>
<thead>
<tr>
<th>Name</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sadok Masliyah</td>
<td>Israeli Hebrew—Divergence from Biblical Hebrew and Semitic Features</td>
</tr>
<tr>
<td>Wendy Willes</td>
<td>Eugenics and Changes in the English Lexicon</td>
</tr>
<tr>
<td><strong>Invited Speaker</strong></td>
<td></td>
</tr>
<tr>
<td>Lily Wong Fillmore</td>
<td>Taking Stock: What Do We Know About Second Language Acquisition?</td>
</tr>
<tr>
<td>J. Halvor Clegg, <strong>Session Chair</strong></td>
<td></td>
</tr>
<tr>
<td>J. Halvor Clegg and</td>
<td>Phonological Distribution and Syllable Length in Spanish</td>
</tr>
<tr>
<td>Joseph Brannen</td>
<td></td>
</tr>
<tr>
<td>David K. Hart</td>
<td>Vowel-zero Alternations in Russian Non-derived Words</td>
</tr>
<tr>
<td>Royal Skousen</td>
<td>An Analogical Description of Morphological Variation in Finnish</td>
</tr>
<tr>
<td>Bruce L. Derwing</td>
<td>Phonology or Orthophonology? Orthographic Effects in Phonological Experiments</td>
</tr>
<tr>
<td>Elray Pedersen, <strong>Session Chair</strong></td>
<td></td>
</tr>
<tr>
<td>C. Ray Graham and</td>
<td>Second Language Acquisition as Skill Learning</td>
</tr>
<tr>
<td>Bruce Brown</td>
<td></td>
</tr>
<tr>
<td>Georgia E. Magnera</td>
<td>Why Applied Linguists Need Theories</td>
</tr>
<tr>
<td>Lynne Hansen</td>
<td>First Language Culture in Second Language Learning</td>
</tr>
<tr>
<td>Mark Davies</td>
<td>Evidence for Formulaic Expression and the Monitor Theory Applied to Adults</td>
</tr>
<tr>
<td>Carolyn M. Squires</td>
<td>A Reexamination of some Semantic Feature Hypothesis Predictions about Dimensional Adjective Acquisition</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

## Invited Papers

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fillmore</td>
<td>Taking Stock: What Do We Know About Second Language Acquisition?</td>
<td>1</td>
</tr>
<tr>
<td>Kay</td>
<td>Unification in Linguistic Computing</td>
<td>17</td>
</tr>
</tbody>
</table>

## Contributed Papers

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bateman</td>
<td>Linguistic Tools Incidental to Work in Machine Translation</td>
<td>30</td>
</tr>
<tr>
<td>Bills/Cranney</td>
<td>A Study of LDS Family Scripture Reading Practices</td>
<td>38</td>
</tr>
<tr>
<td>Clegg/Brannen</td>
<td>Phonological Distribution and Syllable Length in Spanish</td>
<td>51</td>
</tr>
<tr>
<td>Clegg/Seely</td>
<td>Set Theory as an Explanation for the Selection between <em>mas que</em> / <em>mas de</em> in Spanish</td>
<td>58</td>
</tr>
<tr>
<td>Decoo</td>
<td>Looking Back: The Didactics of Semantization in Historical Perspective</td>
<td>70</td>
</tr>
<tr>
<td>Folsom</td>
<td>The Politicization of Repentance</td>
<td>80</td>
</tr>
<tr>
<td>Hansen</td>
<td>First Language Culture in Second Language Learning</td>
<td>89</td>
</tr>
<tr>
<td>Hart</td>
<td>Vowel-zero Alternations in Russian Non-derived Words</td>
<td>100</td>
</tr>
<tr>
<td>Higinbotham</td>
<td>The Banyan: A Partially Ordered Syntactic Structure with Reintersection</td>
<td>107</td>
</tr>
<tr>
<td>Madsen/Larson</td>
<td>Employing Computerized Adaptive Language Testing Techniques</td>
<td>117</td>
</tr>
<tr>
<td>Ricks</td>
<td>Some Observations on Proper Name Formation in the Book of Mormon</td>
<td>129</td>
</tr>
<tr>
<td>Skousen</td>
<td>Analogical Descriptions of Variation</td>
<td>132</td>
</tr>
<tr>
<td>Sorenson</td>
<td>A Computerized Diagnostic Test for Dyslexia</td>
<td>146</td>
</tr>
</tbody>
</table>
Sorenson

An Information Theoretic Interpretation of the Variable Rule......................162

Squires

A Reexamination of some Semantic Feature Hypothesis Predictions About Dimensional Adjective Acquisition.....................178
In the past fifteen or so years, there has been a lot of interest in studying the phenomenon of second language learning. It has been studied from many different angles: researchers have investigated the processes and procedures involved in learning, the sequential development of forms and structures, the influence of prior linguistic knowledge, the role of input, the influence of social and affective variables, age differences, individual differences, cultural differences, and a host of other factors that figure in second language learning. I myself have been engaged in studying the social and cognitive processes involved in acquisition, and in identifying some of the factors that produce the variability we find among individuals in the ability to learn second languages. Over the past five or six years, I have conducted several large scale studies in which my colleagues and I have examined the influence of linguistic, instructional, situational and learner variables on second language learning in over two hundred children between ages 5 and 11.

While doing these studies, I focused my attention on the specific variables we were examining, and had little time to think about the larger picture. Now that these studies are largely done, I have begun to take stock. What have we learned? How does it fit with the research findings that have come from the work of others? What do I think I now understand about the process by which people learn second languages? I have not been alone in asking questions like these. A couple of years ago I learned that Merrill Swain, who has been heavily engaged in studies of second language learning in Canada, has been asking the same kinds of questions. Since then, we have had many conversations on these issues with one another, and with other researchers. We didn't have anything very specific in mind when we began our discussions. We simply wanted to compare notes on what we thought we had learned from our own research with what others had learned from theirs. In the end, we arrived at a fairly reasonable formulation of what we think is involved in second language acquisition. I would like to present that formulation for your consideration. The view I present comes from observations of people learning second languages in diverse settings, and are based on findings from many studies--ones that Merrill and I have done, and ones that other second language researchers have done over the past decade and a half.[1]

Actually what I will be describing is a model of sorts. It allows us to identify the critical components and processes in language acquisition; it allows us to account for age differences and individual differences, and it even helps explain why, in some situations people fail to learn second languages.

Let me tell you what appears to be involved in language acquisition by describing its components and processes. The model is a complex one, and it is a little hard to talk about in purely abstract terms. At any rate, it would easier to show how its pieces fit together if they are discussed in relation to situations that are familiar to all of us. Since language
learning is done by people, and it takes place in social settings, let us consider what language learning involves for some learners. For discussion sake, why don't we take as our learners the kind that are the most familiar to us?

Our learners are members of a family—an immigrant family that arrived in Provo in the spring, right after the snow had melted, and the sun was bringing the trees into full blossom. There was Father, Mother, Junior, who was 14 at the time, and 5 year old Sister. They didn't know anyone in Provo, but after a quick look around, they decided to settle there. (It was the sun and the blossoms that convinced them that they should stay. Actually, they had thought about Laie as an alternative, but they did not know what the climate was like in Hawaii, and they were not about to take a chance with anything as important as that. They had had enough of extreme climates back home, what with it being hot one season, cold the next.) Naturally, they were eager to get acquainted with people, and to become a part of their new community. They had a small problem, however—a problem which could hinder them in getting to know people quickly. They didn't speak English, and no one in Provo spoke Patagonian which was their language. If they were to take part in the life of the community, they would have to learn English. It was a necessity for every member of the family. Father needed to learn English because he wanted to get a job selling widgets. He was the top widget salesman in Patagonia—and he believed he could sell his way to the top in Provo too, once he knew enough English to pitch widgets. Mother knew she had to learn English too because she had to conduct the day to day business of the family—she had to deal with shopkeepers, neighbors, doctors, and her children's teachers. Besides, she hoped to get a parttime job to help make ends meet. All of this required a fair level of proficiency in the language. Junior knew he had to learn English—that is, if he was to learn anything at school; most of all, he wanted to get acquainted with his pubescent peers and to take part in the social life of Provo's young set. Sister—well, she could care less about learning English. She just wanted to play with other children. But since all of the likely prospects for playmates that she came across spoke English, she thought she ought to learn it too. In Sister's view, it was really not that big a deal.

So what we see shaping up are ideal conditions for language learning. That is, we have some potential ideal LANGUAGE LEARNERS—the first essential component in our proposed model of language learning. Each had a real need to learn a second language. The members of this family were motivated to learn English because they wanted to become a part of a community that spoke the language. These learners had a lot going for them; they didn't come to the task of learning a new language empty handed. They had a world of prior experience to guide them—they had social knowledge, language knowledge, and a lot of general knowledge about the world around them. They had things to talk about, they knew what kinds of things are called for in different social situations and settings. Because they already knew one language, they knew what people talk about. They had fairly good ideas about what they would have to learn to do and say in the new language. They knew they would have to learn the English equivalents for the many things they wanted to talk about, they knew they were going to have to figure out how to create
sentences, and they knew from personal experience that there are all kinds of things that polite Patagonians say to each other on special occasions, among them phrases such as "hello", "thank you", "you're welcome" and, "may all the warts on your nose be little ones". Naturally, they should look for ways to express the same sentiments in their new language.

And all of this was taking place in a community where the target language was spoken natively by most of the people who lived there. This is the ideal kind of SOCIAL SETTING -- the second essential component in this model of language learning. The specific social settings in which our learners came into contact with the new language were the work place, classroom, neighborhood, and playground. Those were the places in which they came into contact with people who spoke the language well enough to provide them with necessary input to the language -- these SPEAKERS OF THE TARGET LANGUAGE comprise the third essential component. If each of these components is ideal, then language learning is assured. Each of them can vary in a great many ways, however, and some of this variance can crucially affect the processes by which language learning takes place. I will illustrate this shortly, but let me tell you what happened to our language learning family so you won't be kept in suspense.

The members of the family were in an ideal situation to learn a second language. Each of them was motivated, and had a need to learn English. The members of the family were pleased to be where they were, and they were eager to get established socially. They were in an ideal setting for language learning, they were living a community where they were surrounded by speakers of the target language. Indeed, everything would have worked out perfectly if all that was necessary for language learning was motivation, need, and opportunity. But things turned out not to be so easy for the family--at least not for all four members of the family. Two of the members of the family have done very well indeed: they picked up the language in short order--after a couple of years they were speaking English well enough to get by, and now, after 5 years in Provo, they are quite fluent in the language. The other two have not done nearly as well--one of them can barely be understood even after 5 years. The other has found the whole experience to be altogether traumatic, and has had a lot of trouble learning English well enough to get where he hoped he would be by now. Now you must be thinking: I'll bet the parents did not do well, since age makes a big difference. The children no doubt picked up English quickly and easily, since children enjoy a special advantage in learning new languages. Adults have a much more difficult time because they lack the neural flexibility to pick up new languages easily, as we all know. Or you might be thinking: one of the two family members who did not do well was male--of course! Males are said to have a harder time learning new languages than females, so no doubt the two who did poorly were Father and Junior, while Mother and Sister were the ones who did well. Actually, things are not quite as simple as that. The two who did well in this little scenario, were Father and Sister. Mother and Junior did not do well at all.

I will now turn to the processes that are involved in language learning, and will try to show how variation in the various components of the model
affected the outcome of the processes for the four members of the family.

The model, as I said, has three critical components-- LEARNERS, SPEAKERS of the language to be learned, and a SOCIAL SETTING in which language learning takes place. Three types of processes figure in acquisition, each of them intricately connected with the others. The first can be described as SOCIAL, the second LINGUISTIC, and the third, COGNITIVE. By SOCIAL PROCESSES, I have in mind the steps by which the participants in the language learning situation, that is, both the learners and the speakers of the target language, create and shape a social setting in which communication by means of the target language is possible and desired. By LINGUISTIC PROCESSES I have in mind the ways in which assumptions held by the speakers of the target language predispose them to select, modify and support the linguistic data that get produced for the sake of the learner. On the learners' side there are assumptions about the way language works that enable them to interpret the linguistic data they have to work with. And by COGNITIVE PROCESSES, I have in mind the learners use of general cognitive abilities such as perception, memory, association, categorization, inference and the like, along with whatever cognitive abilities humans have that are specialized for learning language. I will characterize each set of processes for you.

Social processes figure in language learning in the following way. In order to acquire any language, learners must be engaged in some sort of social relationship with people who speak the language. Language cannot be learned --at least not by children-- in isolation. It takes at least two persons, someone who speaks the language to be learned, and someone who wants or needs to learn it. More realistically, it takes at least three persons including two who speak the language, since the learner needs evidence on how native speakers talk to each other as well as how they talk to learners. Social contacts between speakers and learners are necessary, since these give the learners opportunities to observe the language as it is used by its speakers in natural communication, these observations providing the learners with the data on which they are to base their learning of the new language. Learners have to participate in these interactions at some level, since the quality of their participation plays a crucial role in getting speakers to use the language in the special ways that make the speech they produce during contacts usable as language learning data. Social contact is necessary for language learning not only because it gives learners the opportunity to hear and use the target language, it also provides them with the need and motivation to learn it.

You will remember that Sister was not so concerned about learning English. She just wanted to be able to play with children her own age. And if she had to learn English to do it, she would do it. Children simply do not learn languages to broaden their minds or to "build character". Their reasons for learning languages are pragmatic ones: they do it so they can interact with people who speak the language, or because they want to understand what people are saying. All of this means that the social settings in which learning is to take place must be ones that allow learners to come into meaningful contact with speakers of the language. Those which promote frequent contacts are the best, especially if the contacts last long enough to give learners
ample opportunity to observe people using the language for a variety of communicative purposes. Those which also permit learners to engage in the frequent use of the language with speakers are even better.

Linguistic processes figure in language acquisition in several crucial ways. The first intersects with the social processes I have just described, and in a sense, involves linguistic processes principally when looked at from the perspective of the speakers of the target language as they interact with learners. I said that certain things have to happen when learners and speakers come into social contact, namely, they have to collaborate in creating situations in which learners have the opportunity to hear and observe the target language in use. Basically, what learners have to get out of these contacts is enough linguistic evidence to allow them to discover how the language works, and how people use it. The end product of the acquisition process is linguistic knowledge—the phonological, lexical, grammatical, pragmatic and sociolinguistic knowledge that eventually allows learners to speak and comprehend the new language in a full range of social and communicative situations. What it takes to acquire this kind of knowledge is exposure to linguistic data in the form of situationally anchored speech produced by speakers of the language in the context of social interaction which involves the learner in one way or another. These linguistic data, together with the supporting social context in which the data are anchored, constitute what researchers studying first and second language acquisition refer to as "input"—the materials on which learners can base their acquisition of the language.

Language produced by speakers in social contacts with learners can serve as input when it has been produced with the learners' special needs in mind. It is not ordinary language, but language which has been selected for content, modified in form, and in presentation. It tends to be structurally simpler, more redundant and repetitive, and as Mike Long (1981) has argued, it is characterized by greater structural regularity than is found in ordinary usage. Linguistic processes figure in the making of these adjustments in that people generally operate with some sort of theory of what their language is like, and have tacit notions of the kind of adjustments they would need to make for the benefit of anyone who doesn't know the language very well—say, when talking to babies, or to foreigners. The evidence which Charles Ferguson found of common baby-talk and foreigner talk features across languages suggests that these features may reflect fairly universal notions of what linguistic novices would find helpful (1977, 1975).

The modifications that speakers make in this kind of language are based partly on notions they have about what people who don't know the language well would find difficult to understand, and what they would find easy. Studies of the phenomenon of "foreign-talk" indicate that modifications made by speakers on the basis of a priori beliefs about the relative difficulty of linguistic forms are not always helpful to learners, and can, in fact mislead them as to what the target forms are like (Chaudron, 1983; Meisel, 1977). More useful accommodations are based on actual feedback provided by the learners as to whether or not they understand what is being said to them (Cross, 1978; Fillmore, 1985; Long, 1983; Long and Sato, 1983; Gass and
Varonis, 1985. When learners appear to understand, speakers can assume that the adjustments they are making are appropriate or even unnecessary. When they appear not to understand or to be having difficulty following what is said, then the speakers make adjustments in the form of what they are saying, or they do something else, verbally or otherwise, to allow the learners to figure out what is being communicated. In an important sense then, it can be seen that learners and speakers collaborate in producing the adjustments which benefit the learners.

It has been argued that the language that learners hear works as input when it is "comprehensible", in Steve Krashen's terminology. According to Krashen (1980, 1981), learners acquire language by having input available to them containing structures he characterizes as being "a bit above the learners' current level of competence"--that is, if the current level of competence is at "stage i", then structures which are "i plus 1". In Krashen's view, what is critical is that the input be comprehensible, that is, more or less transparent in meaning to the learners. Indeed, he argues that learners acquire language, not by focusing on the form of the input, or by analyzing it, but rather by finding or being given access to its meaning.

What Merrill and I have learned in our own observations of children learning second languages is that focus on form is precisely what learners have got to do at some level when they encounter input, and that comprehensibility is important at least in part because it enables learners to make necessary connections between form and function in the new language. In fact, speech which is altogether comprehensible seems to short circuit language learning since when the learners can understand what is being said effortlessly, they have no reason to pay attention to the language itself. Unless learners are actively involved in figuring out what is being said, they do not have any reason to attend to the linguistic forms of the messages addressed to them. Our observations indicate that as Krashen as argued, what seems to work for language learning is indeed speech which is more or less comprehensible by virtue of being used in ways that allow the learner to figure out what is being said. But, the best kind of input is language used in ways that call attention to the form of the message itself. We have found evidence children do pay attention to form, and have found that advances are made in language development precisely when learners appear to take notice of the structural characteristics of the language they are hearing and using.

In short then, we argue that language works as input when it is slightly above the heads of the learners, not structurally, but in meaning, so that learners have to stretch a bit to figure out what people are saying, and when the language is used in ways that call attention to form and structure. In addition, we would argue that learners play a key role in getting this kind of input by their own productive efforts (Swain, 1985). Their efforts at using the language not only gives them an opportunity to test what they think they have learned communicatively, it also tells the people they interact with how much linguistic adjustment they will have to make for their sake.

This leads to the second way in which linguistic processes appear to figure in language acquisition, this one intersecting with cognitive processes.
Looked at from the learners' perspective, the initial problem in language learning is to make sense of what people are saying in the new language. In part, this is achieved by paying close attention to what is going on while people are talking, and in assuming some kind of relationship between language and the events in which it occurs. Basically, the problem for the learner is to guess what people might be saying given the social situation at hand. This might seem like an impossible task, but second language learners have some special resources available to help them in this process. What they have for them is a prior language, thus they have the means to make educated guesses as to what people might say in their L1 under similar situations. Because they already have a language, they know about linguistic categories such as lexical item, clause, and phrase. This awareness of grammatical form and structure will predispose them to look for equivalent properties in the new language data they have available to them.

Similarly, through the experiences they have had in their first language, learners are generally knowledgeable about the speech acts and functions that can be performed linguistically. They know about the uses of declarative and interrogative structures, about affirmation and negation, about expressions of certainty and uncertainty in speech, and the like. They have used, and therefore, are familiar with such forms for making requests, promises, denials, declarations, and questions. They know that one can ask questions, and that questions ordinarily require answers. They know that questions can serve as requests for information, as indirect requests for action, as greetings, and for a host of other communicative functions. This kind of prior linguistic knowledge and experience will lead second language learners to seek and to discover means for accomplishing the same functions in the new language. In other words, they are guided in their language learning efforts by what they know to be possible and useful from their knowledge of the first language. Thus, second language learners start out with a fairly good idea of what to look for in the new language. The assumption that forms will be found in the L2 which are functionally equivalent to L1 forms can lead learners to acquire them more efficiently than they might otherwise, since it's always easier to find things when you know what to look for. At the same time, however, it can also interfere with learning, since this assumption sometimes lead learners to draw largely unwarranted conclusions that L2 forms are functionally and structurally identical to L1 forms and usages. Nevertheless, the net result is positive. By applying the knowledge they have of what people are likely to say in various social situations to what they know are possible forms, patterns, and functions in language, learners are more or less able to give meaningful interpretations to the language they hear, and thus, to discover eventually the principles that govern the structure and use of the language itself.

So now I come to the third type of process in acquisition: those I describe as cognitive processes. In an important sense the cognitive processes in acquisition are the central ones. These involve the analytical procedures and operations that take place in the heads of learners and which ultimately result in the acquisition of the language. Let us be reminded of what the cognitive task involves. The primary linguistic data which learners have available to them as input for their analyses consist of speech samples
produced by speakers of the target language during social contacts in which the learners are themselves participants, as I have argued. Hence what the learners have to work with are observations of the social situations in which the language itself was produced, and streams of vocal sounds produced by human speakers according to complex and abstract systems of grammatical and social rules that systematically and symbolically link up sounds, meaning representations and communicative intentions. What they HAVE to do with these data is discover the system of rules the speakers of the language are following, synthesize this knowledge into a grammar, and then make it their own by internalizing it. That in capsule form is what the cognitive task is for any language learner.

Figuring out how the speakers of the target language are using the noises they produce to represent meaning is the first step. This involves discovering the principles by which segments of the speech produced by target language speakers relate to events, ideas, experiences, objects, and the other things that people are known to talk about. Discovering how the speech serving as input segments in the first place, that is, finding out where one thing begins and another ends, is critical to the procedure. Once the learners know what the pieces are, they can acquire knowledge of how they are used to represent meaning, and eventually, discover how such units can be assembled structurally to communicate more complex ideas and thoughts in the target language. Finally, the cognitive task involves figuring out the principles by which the speakers of the language use it to achieve their communicative goals and intentions: what kinds of things do the speakers of the language talk about, and what can they do with the language they speak? In doing this, learners apply a host of cognitive strategies and skills: they have to make use of associative skills, memory, social knowledge and inferential skills in trying to figure out what people are talking about. They use whatever analytical skills they have to figure out relationships between forms, functions and meanings. They have to make use of memory, pattern recognition, induction, categorization, generalization, inference and the like to figure out the structural principles by which the forms of the language can be combined, and meanings modified by changes and deletions.

The task as outlined here is both complex and enormous. Nevertheless, it's a task that can be handled by any and all ordinary humans, including the members of our immigrant family. According to the prevailing theory of language acquisition, the task is manageable because humans have a special cognitive capacity for learning language—one can even think about this capacity as a kind of cognitive mechanism, which some people refer to as the Language Acquisition Device. This Device, which is affectionately referred to as LAD, operates in a quite different way from ordinary cognitive processes. Its workings can not be observed; they can only be inferred from the fact that all ordinary children learn a first language, and that they appear to do it in ways that can not be explained by ordinary cognitive processes. One of the major arguments for the cognitive processes involved in acquisition being special ones is that many of the features of the grammar that learners eventually acquire can not simply be induced from the linguistic data that are available to them. In fact, the argument goes, it would be impossible to explain how children can arrive at structure as complex and subtle as found
in even the most ordinary run-of-the-mill competence grammars, based on the relatively meager structural evidence they are able to extract from the language spoken around them, unless we assume that a certain amount of that structure is already hardwired into the language learning mechanism children are born with. In this view of acquisition, what I have described as social and linguistic processes are regarded as incidental or peripheral phenomena. If they figure at all, they play only trivial roles; everything that is really important in language learning has to do with the working of LAD. Others have argued that no matter what other kind of information or help is available to learners, the primary data they have to work with are samples of speech consisting of phonological signals which are not cognitively penetrable—they are not tractable to ordinary cognitive manipulations or analytical procedures that are available to children. How, it has been asked, is it possible for children to discover the rules that figure in "parasitic gapping" as they eventually do, based solely on the speech spoken to them? The only explanation to people who hold this view is that such rules are already "known" to the acquisition device in some abstract sense, requiring only exposure to data in which such rules figure, to trigger their discovery.

That may well be the case in first language learning. Nearly everyone does in fact end up learning a first language, despite huge differences in general intellectual endowment and early language experiences, and no matter how difficult or complex the target language is, and indeed, if the language is English they even learn how to deal with sentences with parasitic gaps. No doubt there is more here than meets the eye, and while I am convinced that general cognitive abilities and strategies of the sort I have been talking about also play an important role in first language acquisition, specialized mechanisms are without question much more crucially involved.

What I think happens in second language acquisition, is that the degree of involvement of these two types of mechanisms may be reversed: While specialized language learning processes figure in an important way too, general cognitive processes are much more heavily involved. This, in fact, may be a crucial difference between first and second language learning. There are two kinds of evidence for believing that general cognitive abilities and strategies figure more heavily in the acquisition of languages after the first than do specialized abilities. One consists of observations of strategies that children appear to follow when they tackle a second language. The other relates to observations of individual variation in the learning of second languages. The kind of cognitive strategies and skills that I have been talking about—the ones we find learners applying in getting access to the language, in breaking it down into units, in figuring out its structural properties and in extracting its principles of usage are general cognitive mechanisms rather than specialized ones. The cognitive work learners engage in results in them figuring out and acquiring a lot of rules, principles, and patterns, etc. But such materials do not necessarily add up to a grammar. At some point, the knowledge which has been gained through the workings of general cognitive mechanisms has got to be consolidated,
assembled, in a manner of speaking, into a competence grammar. This, I would like to argue, is where the language specific cognitive mechanisms come into play; through these processes, what the learner has sorted out gets synthesized into a real competence grammar, and perhaps a lot of the details of the grammar get refined here as well; that's the work of LAD or whatever you want to call the innate language learning mechanism people have. This, I admit, is all very speculative: there is no way of proving this last part, or disproving it. We will just have to wait until the day when All is Revealed to know whether or not things really happen that way. Who knows? Maybe it's in this process that all the mysteries of the language get revealed. We know that at some point our learners will be able to handle, more or less painlessly, just about any English sentence anyone cares to spring on them--even those containing parasitic gaps that they might hear without recognizing. I have now described the various processes that I think are involved in language learning. Now let us consider how they work--or don't work in relation to variation in the components I mentioned earlier. Let's return the members of our immigrant family, now more or less settled in their new community. Junior and Sister are in school, and Mother and Father are at work. We will see how they are doing. The children, are in a sense, in a perfect social setting for language learning. Recall that learning is possible when learners come into contact with speakers of the target language in social situations where they can interact in some fashion. They find themselves in their respective classrooms surrounded by English speaking classmates, dealing with a live English speaking teacher.

In such a setting the social conditions for language learning outlined earlier are quite easily met. The learners are in constant social contact with speakers of the target language. The speakers, the teacher especially, but classmates as well, have ample reason to speak to the learners in this setting, and they are generally inclined to do so in a manner that takes into account the fact that the learners don't know the language. What the learners have to do then, is observe carefully what's going on in the classroom, listen to what people say, figure out what they are talking about, and how they are doing it. By doing these things they learn how to do what the speakers can do, so that eventually they can communicate with them in the target language. By making use of the general cognitive strategies and the social and linguistic knowledge that they have, they will be able to figure things out in short order. That's all there is to it.

But as I already told you, this worked for Sister, but not for Junior. To learn why not, we now turn to the question of variation in second language learning, since this is what can convince us that not only are the cognitive aspects of the process handled by the mechanisms just discussed in this paper, it also shows just how crucial are the roles played by the social and linguistic processes described earlier. One of the most striking differences between first and second language learning is in the relative amounts of individual variation that can be found among the two types of learners. Clearly there is variation to be found even among first language learners, and while the differences may not be great, children do vary in how quickly they acquire their primary language, and in how facile they become in their exercise of verbal skills. But the variation we find among first language
learners is still relatively minor compared to those we see in second language acquisition, even among relatively young children. Differences of up to 5 years can be found in the amount of time children take to get a working command of the new language. Learners differ enormously in how easily and completely they master the grammatical details and intricacies of a second language. Some are able to learn it as completely and well as they did their first language; others never totally master the forms or uses of the language.

I believe that a substantial portion of this variation is due precisely to the involvement of the kind of cognitive mechanisms that we have identified as ones that figure most heavily in second language learning. Individuals apparently do not vary in having an innate capacity to learn language; and if this mechanism is as heavily involved in the learning of second languages as it is in first languages, then we would not expect to find much difference in the amount of variation between the two. But, as I have tried to show, the kind of cognitive processes that are most critical in second language learning are the ones which relate to general cognitive abilities. We know that there are enormous differences across individuals in their endowments of this nature. I am not necessarily talking here about differences in those abilities that figure in general intelligence, but about those primarily that we have identified as figuring especially in language learning: verbal memory, auditory perception, pattern recognition, categorization, and so forth. Some of them (generalization and association, for example) are obviously associated with those abilities that are directly related to general intelligence, but most of the ones I mentioned are only incidentally related, I think. The point that I want to make here is that much of the variability found in second language learning can be traced to differences found among learners in the application of these general mechanisms and abilities that figure in language learning. Learners who have poor auditory memory will have a difficult time remembering the things they hear in a new language. If they can't remember what they hear, they will not find it easy to figure things out, or to use them. Those who are poor in auditory perception will have difficulty discriminating between the sounds of the new language, and hence will be poor in learning to make sense of what they hear, and at reproducing anything. Learners who are poor in pattern recognition will have a very difficult time seeing the patterns that they must eventually discover in the new language. But variation in language learning along the cognitive dimension are not just related to differences in learner endowments in cognitive abilities. They are also affected by other learner variables.

Age is an obvious one. Recent research has shown that older learners may be relatively better and quicker at learning certain aspects of second languages than younger learners because they have better developed learning strategies and cognitive abilities. However, this is clearly not true for all older learners. Personality is a type of variable that can interact with age and affect the cognitive processes involved in language learning.

Learning a new language involves the learner in dealing with an enormously complex cognitive task. Handling the various aspects of this task requires no small amount of cognitive flexibility on the part of the learner.
However, as we know, individuals differ considerably in personality related cognitive traits such as such as mental flexibility. Some individuals are quite rigid in their thinking and find it difficult to deal with multiple possibilities, or with things they cannot immediately understand. Unfortunately, there is a lot of that to deal with in the learning of a new language. The unwillingness or inability to accommodate new information or the unknown can make it extremely difficult for learners to handle many of the aspects of the task that we have outlined. Another personality or cognitive style characteristic that can greatly affect language learning has to do with risk taking. Language learning requires learners to apply inferential skills to figuring out what people are saying in the language, and how the language they hear relate to the social situations in which it is used. Some learners find it difficult to take the chances involved in acting upon the things they have learned through guessing, and in fact may be unwilling to risk much guessing at all. They find it hard to try out whatever knowledge they have gained of the new language by their observations because they are afraid of being wrong, or of appearing foolish. Or it might be that they are just unable to take the next step in language learning and draw generalizations from the relationships they do see, and to test them out. Whatever the problem, the cognitive processes that should be operating in acquisition don't function as they should for some learners and so they are not great language learners. At the same time, happy conjunctions of abilities and personality characteristics can result in super language learning in others.

And that may well explain why our brother and sister team differed so much in their language learning ability. If we observed the language learning behavior of these two, we would find, (just as I have for many many language learners I have studied) that rather substantial differences can be found in their cognitive and language learning behavior. Some of this relates to learning style differences, some to personality differences. The net effect, however, is that in some cases, it all adds up to easy language learning, while in others, enormous difficulty. Age might have been an important factor in Junior's case, of course. Adolescence is a time when certain personality characteristics are exacerbated. If an individual is at all inclined to be self-conscious, then it is going to be a major problem during the adolescent years since that is a time when virtually everyone is self-conscious about one thing or another.

Observations of variation in learners provide evidence of the way other types of processes figure in acquisition too. In fact, this is what convinces us that social and linguistic processes are also crucially involved in second language acquisition. Let's consider the way some of the ways in which social variables affect language learning by their influence on the social processes. Consider the effects of differences in the social situations in which learners are to acquire the new language. They can differ enormously in how much contact they provide with speakers, and hence, how much opportunity learners find in them to learn the language. Some settings provide learners with few opportunities to get close enough to speakers of the language to do them any good, or the kinds of contacts they get are inadequate for language learning purposes. They may be too brief, or
too infrequent, or too limiting in the kind of exchanges they allow. Situations can differ enormously in how much of the kind of input needed by learners they provide. If there are few speakers of the target language around, learners will not have adequate opportunities to come into contact with speakers. If the target language speakers are themselves imperfect speakers of the language, the kind of input they provide for learners may not be an adequate representation of what the learners ought to be aiming at. Such input data, which Selinker, Swain and Dumas (1975) have described and which has become known as "junky data", can result in learners acquiring a form of the language which differs substantially from the target. Situations for language learning can also differ considerably in the extent to learners themselves have to play a role in getting the kind of social contact needed for language learning, and in how great a role actual interaction between speakers and learners plays. We have always assumed that direct interaction between learners and speakers is necessary in order for language learning to take place. But in our research in classrooms, Merrill and I have seen how some learners can in fact pick up a language pretty much by observing their teachers and peers, and with very little engagement in direct interactions with them. In other situations we have seen that children who do not get into direct interactions with speakers have an enormous problem acquiring a second language. Is interaction necessary?

What I believe is this: what is necessary is that somehow learners have access to language which is appropriately modified for them, and which is used in ways that allow learners to discover its formal and pragmatic properties. Some settings—for example, classrooms where the language used by teachers meet these criteria provide enough such input so that more direct forms of interaction are not absolutely essential, at least for those learners who are motivated enough to learn the language, and attentive enough to benefit from what they get out of just being in the situation. In others, say in classrooms where the language which is used does not fit the criteria of input-hood that I have outlined, or in the playground, learners play a much greater role in initiating contacts with speakers, and in having sustained interactions with them. In such situations, learner variables such as personality and social skills can play a very substantial role in language learning. Those learners who find it easy or desirable to interact with speakers of the target language will get a lot more of the social contacts needed for language learning, than do those who are not as interested, or motivated, or are less able to manage the kinds of social contacts that are needed for language learning. Variables such as personality, social style, social competence, motivation and attitudes in both learners and speakers of the target language can affect language learning, in fact.

And therein can be found some clues as to why the two adult members of our immigrant family differed in their ability to learn the new language. Father was lucky. He got a job in the widget industry right away—not selling widgets, of course since he did not at that time speak English. Instead, he was placed on a widget repair crew with four of the nicest widget repairmen in the business. These guys took it upon themselves to help the new crew member learn the business right. They were patient, helpful, and wise. They demonstrated things for him, explained, drew diagrams, and they explained
some more. Best of all, they included him in everything. Father joined their bowling team, and even went to baseball games with them and joined them for their weekly pinochle games. And because of all this, and because he basically was a very observant and quick-witted guy—Father picked up English in short order.

But not mother. Although she tried hard to pick up English in short order: she took a parttime job as a short order cook in a cafeteria. Co-workers spoke English to her constantly on the job: "Ham on rye, heavy on the pickles" "Two over easy, side-a-toast" "BLT--hold the mayo!" She did learn some English in this way, but it really wasn't all that useful. And so she signed up for an ESL class at nightschool--twice a week she went, and she diligently practiced the sentences she was taught: "I am Mr. Smith. I am a teacher. I am not a doctor. Are you Mr. Smith? How are you." She met a lot of other ladies there who were also trying to learn English. She visited with them on the nights when there was no class in hopes of practicing English. But as none of them knew much more than she, the practice she got was not much like the English she was hoping to learn. They did talk a lot, however, and while it wasn't very good English, it was good company—and that counted for something.
REFERENCES


---

1. An earlier version of this paper was presented at the TESOL Conference by L. W. Fillmore and M. Swain (1984).
It is really delightful to be here. I have known people at this university for as long as I have known anything, and that’s at least three weeks. And it is delightful finally to be here and see some of them on their turf. You may say that it’s cold outside and it’s warm where I come from. It was actually only 80 degrees yesterday. But actually I feel very warm here. I’m amongst old friends from way back and it’s very nice to be here.

Now I suppose I should worry about the title of my presentation because I realize that in these surroundings 'unification' could mean more than one thing. And all I have to say on that score is that I have checked all this out with the Reverend Moon and it is as it ought to be.

I am a computational linguist and I want to start out by saying just a few things about what I take computational linguistics to mean because that is the backdrop against which all this other stuff will be played out. I take computational linguistics to be a term which has two referents--it can mean two different kinds of things—one of them a big thing and the other one a smaller one included within the big thing. The big thing of course is anything that you care to do with computers that has something to do with language. And a lot of that turns out to be engineering of one sort or another. That is to say, you have something that you want computers to do and you want them to do it with natural language because either the natural language existed before you thought of bringing the computer along or because there are people there that want to deal with the computer and are not prepared to do so in the languages it naturally speaks. You are therefore going to have to persuade them to do it in languages like English, for example. So under this broad heading of computational linguistics come obvious things like machine translation, natural language access to data bases so people can have questions answered about it, and so forth and so on. The use of computers in language instruction, I suppose, would fall under that heading. There is almost an indefinite list of things that you can add to it. Let’s call it 'computational linguistic engineering' for want of a better term. That is all part of computational linguistics in the large.

Then I also like to think of computational linguistics as having a much narrower meaning and it is in this narrower meaning that I pursue it
most of the time. In that case it is not an engineering enterprise so much as a scientific enterprise and its aims are exactly the same as the aims of theoretical linguistics in general: namely, to understand how in the world it can possibly be that by blurting out these noises at one another that we come to have different ideas than we had before the exercise started--how communication, in other words, can be brought about by this really remarkably narrow channel that we manage to establish between one another. Some people are not even amazed by this. But I think if you think about it for a little while and think of the things that you are indeed able to do by blurting noises at other people you will see that it is indeed a scientifically puzzling thing.

Now computational linguists simply bring a new perspective to this same problem that everybody else is trying to answer. They are not trying to do some different thing with it; they are trying to answer the same questions. And in this narrow sense of computational linguistics I take it that they don't actually need to use a computer in principle at all, though they usually in fact do. What they do need to use is their knowledge of how computers work, their knowledge of computer science, their knowledge of symbol processing, their knowledge of certain kinds of logic which turn out to be appropriate inside computers, their knowledge of data structures, and their knowledge of things of that sort. That brings a very different perspective to the kinds of things that you might do as a linguist. This is not surprising if you believe, as I do, that about the only other thing there is around on this planet that performs operations on symbols--properly understood, things which both have a form and a meaning--the only thing that performs operations on symbols in a serious way other than a human being is a computer. Now it may not do it in the same way and it may not be sophisticated in certain ways and it may not have quite that same appreciation for some of the subtleties of poetry that human beings do, but nevertheless it does process symbols. Furthermore, we have some considerable control over the way that it processes symbols. And so it is just possible that from these various ways that are open to a computer scientist we could get a certain amount of inspiration, let us say, about how linguistics might be done. And the people that search for that inspiration in computer science are, to that extent, computational linguists, and some even think they have found it, I among them.

Now there are several questions that you might ask at this point: if there have been a number of computational linguists around doing this sort of thing for a long time, then why haven't they had an impact that I could have noticed on the rest of linguistics, and that is a fair question. And I think there is a good reason why it has taken so long for the impact to be felt but I think it is now being felt. And the reason is that it has taken us this long to understand some of the more fundamental things about computer science in general, and we had to understand those before we could start trying to sell them to somebody else, such as linguistics.

The most interesting thing that I have to say to you today--I hope I can get across the feeling of this even if I cannot give you the details of it--is that noncomputationally-inspired approaches to linguistics in recent years look, on the face of it, more computational than the computationally-inspired ones look. And there is a very good reason for that and it is really quite strange but a fascinating story and it is really the story
that I want to tell. What computationalists have discovered over the years is 'unification' in one form or another and therefore it is going to be up to me to try and tell you what unification is about and why it is such a good thing.

If you look through recent work on linguistics, especially in North America, where formal linguistics is so much in vogue, you will find a lot of procedures involved in doing whatever it is that gets done. If you look in any of the multifarious and extremely challenging work that has been done in the formalist tradition since Chomsky published *Syntactic Structures* in the mid-50's, you will find everything set out with tremendous precision and you will find rules of a more exact sort than was found in linguistics before, and associated with these rules very, very carefully orchestrated sets of procedures for how they are to be applied. So, for example, according to the theory of *Syntactic Structures*, a grammar started out with rewriting rules, famous ones like:

\[ S \rightarrow NP \, VP \]

The way you are supposed to generate a sentence in this is to start by writing down S. Why? Well, because that is how you start. There is no other particular reason, that is just part of this formal system. Then you can take whatever you happen to have written down and you can replace it by whatever the right-hand side of the rule says. And then there are other things that say a noun phrase, for example, can consist of a determiner plus a noun:

\[ NP \rightarrow \text{Det} \, N \]

If you hook these up with lines, which is the traditional thing to do, you get tree-structures out of all this, that you have all seen too many times to care to remember.

So that was the first thing that happened in the generation of a sentence. There was no claim that this went on in people's heads when they were generating sentences; this does not claim to be a psychological model of the steps people go through. This is simply a grammar which is sufficiently precise so that we know for any proposed sentence whether it really is a sentence of the language or not, and secondly what its structure is and therefore how one would go about figuring out what it means.
So you started out by building a tree like this, and then you applied to it a set of transformations, and these transformations take a tree into another tree, and into another tree, and so on, until there are no transformations left on the list, in which case the tree is what is called a 'surface structure'. If you want to know what the sentence really sounds like, then you read off all the words from the bottom of it.

If you are not familiar with all that, it really doesn't matter. Details are not important. The important thing is that you must go through these steps in just the right order. Each of these transformational rules can in general work a massive change on the tree; it may look very different from the one before. The rules are ordered in a list—Transformation 1 through Transformation n—and you must take them exactly in the order given and some of them are marked as obligatory and some of them are marked as optional and by golly you'd better not miss doing any of the obligatory ones and stuff like that, otherwise you'll get the wrong answer.

So there is a very important sequence of steps that gets carried out. It occurs everywhere. It occurs in phonology—let me translate phonology into graphology, which, when you see it, looks a lot simpler. For example, there is a rule in English that says [there is] some funny kind of S—which nobody has ever seen but we can think of as existing in the head and which comes out as the plural on most regular nouns in English and as the 3rd person singular on verbs—that gets rewritten as es in a situation following a number of things, amongst them -s, -sh, -ch, -x, -z, and -j and things like that; otherwise, there is another rule that says S gets rewritten as s:

\[
\begin{align*}
(1) & \quad S \rightarrow es/ \begin{cases} s \\ sh \\ x \\ \text{etc.} \end{cases} \\
(2) & \quad S \rightarrow s
\end{align*}
\]

It is obviously important that the rules should be ordered in this way, because if you were to carry the second rule out first there would be no big Ss left for the first one to work on at all and the words that ended in these magic things would have just a bare s after them and not an es. So you have to get the order of the rules right. You must, in other words, do a certain sequence of steps in a very certain definite order. Transformations were like that; early computational linguistics was like that.

There were types of grammars in early computational linguistics called 'augmented transition networks', and I will give you a sort of comic book view of what those are, the same sort of comic book variety I gave you of the other things. But each one of these is a semester course, you know, and I've got to pretend that I really know what's in it.

In this you imagine machines that can be in a certain state at a given moment and at the next moment the machine can move to another state, and given that it is in a particular state, perhaps of a variety of different states that it could go to, we assume that at a given moment it is sitting
looking at a particular word in the sentence:

And the state it goes to next is determined by the label that we have on the lines, and by that word in the sentence. Some words will allow you to go one way, some words will allow you to go another way, and so forth. So, for example, right at the beginning of the sentence you might do something that says, "look for a noun phrase," and if you find a noun phrase, call it the subject of the sentence.

And then you trundle on through some more of these states here and then over on the right you find a verb which has passive morphology—it is a passive verb. And then you say, "Oh boy, we blew this one didn't we? This wasn't the subject at all. At least not for the answer we would like to get out. We'd really like to fix up passive sentences so that what we learn about them makes them look very like active sentences but we blew it though at the beginning here because we didn't know that the verb had passive morphology." That's all right. Don't worry about a thing. What we will do is take what used to be the subject and put it in the object pigeon-hole and we will clear out the subject pigeon-hole and go on as though nothing happened.

That was wonderful, and they parsed a lot of sentences with this grammar, a lot of sentences about moon rocks and wonderful things like that a long time ago, a part of the folklore of our field, this. Now, the trouble about this is you must start at the beginning and work your way towards the end. Suppose that you were taking the description of a sentence that this gives you, and you were trying to get the sentence from it. Notice that you could not reverse the procedure. This thing will allow you to analyze sentences and it is absolutely hopeless if you try to use it for generating sentences. What would you do? You would say, "Ah, well the first thing I've got to do is to take the subject and put it down here at the beginning of the sentence as a noun phrase." So you do that. There it is, it is down in the sentence and you have moved on. And when you come over to the object, you decide that you need to put out a passive verb and you do that, and furthermore you infer into that that everything you have done up to now was wrong. And there is nothing like the above that you can do that swaps it around because that has already been said. It is all over.
This is standard linguistics. Look, there is nothing wrong with this. The only thing I am trying to point out about it is that it is full of computer procedures, full of programs—that is what is remarkable about it—more full of programs than you will find in any of the things on which it was originally based, and by which it was inspired.

Let me just digress for a little moment about that. Notice how most of the styles of linguistics that we have seen have been inspired by some other scientific enterprise that just happens to be around and popular at the moment. Early linguistics in Germany—the traditional historical linguistics that is still pursued—was a genetic model of language. Let us see what we can do if we treat languages as belonging to genetic families. And we looked into that for a while and we discovered that, by golly, it works. This is a gold mine of interesting stuff.

In American structuralism, we took a chemical model...why not put it that way? What we were trying to do was to find what all the atoms were, find out what the elements were, what are the different classes of things which go in here. So we worried about whether, in 'cranberry', 'cran' is a morpheme or not, because after all, 'blueberry' is a berry and it's blue, but 'cranberry' isn't a berry and it's 'cran'. You worried about whether 'cran' was a morpheme, you worried what that element was. It had to fit into the periodic table somewhere.

Transformational grammar is inspired by logic. Logic tells you how, given one well-formed formula to get another well-formed formula from it by rewriting it. But the operations that you perform in logic are, for the most part, extremely simple and they have the property that if there is a variable, that either that variable has a value, it is bound to a value and it is bound to it for all time and that is the value of that variable, or it is a free variable, and remains unbound. And if you work your way through a proof or a computation or a logical derivation of some sort, you may discover what the value of the variable is, whereas you did not know that before, but you can never change it from what is was to something else. Never in logic do you pull a number like changing the values of variables. You don't say, "Well, I'm sorry. I didn't mean that about 'x'."

Unification, very broadly speaking, is computing with variables whose values you may come to know whereas you did not know them before, but once you know them, they never change. That is what unification is all about. (I checked it with the Reverend Moon, like I said in the brief.) The interesting thing is that you actually can do with systems of this sort everything that you can do the other way, which is perhaps a little bit surprising. It looks as if the other way you had a lot more freedom. You just had to figure out more devious ways of getting away with things and then all was available to you.

But let me give you a feeling for what unification turns out to be like, with what we call 'referentially transparent' variables—referentially transparent because if somebody said "What is the value of this variable?" you don't have to say, "Well, when? What are you talking about? What is it now? What was it yesterday?" No, what is it? Always the same, according to this theory.
It's like this. Suppose that one of your intelligence operatives in a far-off country sent back some photographs of something that you needed information about, but the photographs had been damaged in transport, so that some of the emulsion had come off them so that you can see it all except for one area which has all been sort of scraped off and it is just white and the light passes through it and there is no detail there. Fortunately, you have another slide of exactly the same scene, taken from exactly the same angle. And it has been damaged, too, unfortunately, but the damage is in a different place, in this case. What would you do? Well, what you would do is pretty clear in that you would take one photograph and would superimpose it on the other and if you got the registration just right both of the damaged areas would disappear and you would see the whole thing.

That is what unification is like. Since you can never change the value of a variable, all you can ever do is to take an existing picture, and add more detail to it. There are things that you don't know about it, you can add things to it so that you know more about it, but you can never take something and change it. It has always got to stay. It is only the pieces that you do not know anything about that can be filled in.

So, you can take a string of words and in the course of analyzing it, you can fill in things that you don't know such as that a given unit is a piece of structure, or that another is a phrase, and so on. You can discover what the labels on those phrases are. Now, notice that context-free grammar—that is the grammar that contains only rules like the famous

\[ S \rightarrow NP \ VP \]

rule—that does not violate any of this. That is the most simple kind of unification grammar there is. You can check out whether one of these rules is correctly applied to a particular structure by simply looking at the piece of structure that that rule would have sanctioned. You don't have to say, well, we have to work down from the top, or we have to work up from the bottom, or we have to go left to right; you do not have to do anything. You can just pick some piece of that tree and you can say, "Well, wait a minute, what rule is responsible for this?" And you go look at the rule and if the rule is there, then that's all right and if it's not, it's not. You don't have to go through a certain set sequence of procedures.

The question that arises then is can we do all the things that the people with these more powerful techniques, so-called non-monotonic techniques—never mind about that word, that's the bad guys; the non-monotonic guys are the bad guys, and I'm talking about the monotonic guys and they are the good guys. The question that arises is "Can the good guys do all the things that the bad guys were trying to do?"—the advantages of being able to do it would be tremendous. That would, for example, mean that the theory of language that you wound up with would be the same theory for the speaker of the language as for the understander of the language. And I don't know about you but that sounds to me like rather a good idea. Because without it, you see, in principle, you could have somebody who had learned how to speak English and understand French and could not do the
other thing with either of the languages. It would be possible to communicate with them but it would be sort of funny and roundabout, wouldn't it? But it would be perfectly possible. You would have to account for the fact, if you don't have a theory like this, that if you teach a certain grammatical construction to somebody, if you teach him how to say this, if you teach him how to use this word, or you teach him how to use this tense, or you teach him how to use this construction, and he learns how to use it with great fluency and facility, you would then have to say to him, "Okay, now we can go on to understanding that. Let me teach you how to understand it." Now most people would say, "That's funny. You've just taught me that. I can understand it." It seems that somehow our theories of languages ought to be like that. Monotonic theories of language are like that quite naturally.

So let me write some rules for you. We're going to change gears now and get down to some specifics and I don't know how far I will get with this but I hope that some of these rules will just give you a feel for what you can do. And the rules are going to look like sort of decorated complex tree rules. There is going to be a rule in the simplest possible context-free grammar that you can imagine that says that a verb phrase can consist of a verb and a noun phrase:

\[ VP \rightarrow V \ NP \]

Yes, there are intransitive verbs and there will be other rules for verb phrases and subject-aux inversion and there are lots of other fancy things, but there will be this rule amongst all the rest. Now I am going to write another more fancy version of this rule by replacing each of the three things in it by something else. And at first it won't look as though I am doing very much to these rules except using up more chalk on them.

So, for example, what is a verb phrase? Well, it is something whose category is verb phrase. That, at least, we know about it. And in due course I am going to fill in more of the photograph. I am going to give you more of the emulsion down here. But I am not going to do it just yet. So, that is going to be one description. And the next thing in the rule is going to tell you about something called the category verb, and that will have some information to be added to it. The next thing is something whose category is noun phrase. All right, that much is simply a translation of what we have already.

\[
\begin{align*}
\text{[CAT} = \text{VP}] & \rightarrow \text{[CAT} = \text{V}] \\
\text{[CAT} = \text{NP}] & \end{align*}
\]

Now, different verbs take, let us say, different things. Some of them take a dative object and some of them require certain kinds of complements like 'John is easy to please'--you have probably heard that one. They take to-infinitives. They take other kinds of infinitives. Some of them take certain kinds of prepositions--you know all that sort of stuff. What we would like to have to do is to avoid writing a separate decorated rule for all the different things that verbs might take. We want this object NP to be any of the NPs that it might take next in line. If the verb wants a
dative NP then we would like it to be dative. We would like a lot of things of that sort to fall out. So I am going to assume that there is a thing called a 'frame'—and this is a Fillmorian-like notion—and the frame that this particular verb has, I am going to mention two parts of it: I'm going to mention 's', I'm going to write a dot, and I'm going to write the word 'rest'. And I'm going to put a question mark after each of these things.

\[
\text{CAT} = \text{VP} \\
\text{FRAME} = (s? . \text{rest}) \\
\text{HF} = h? = [o? = \text{obj}\
\]

Now let me explain all that for you. First of all, question mark is simply my way of bringing home to you that it is supposed to be a variable, something that can have any value. Next, this dot is used in the following way: the things in parentheses are lists of things. So whatever is the first thing in this list has got to be the value of the variable 's'. I don't know what it is, but the variable 's' is going to have that value. ('s' stands for subject and 'o' for object.) And the dot means that the next thing that follows here is not the next thing in the list; it is the rest of the list. And I have called that 'rest'. The thing is also going to have an attribute—these things to the left of the equals sign I call 'attributes'—called 'HF' and that stands for 'head features'. And that's for historical reasons. And they are going to do the same as whatever the variable 'h' has, and in particular, there is going to be an attribute and a value pair which is going to be at least part of what that set of head features consists of—and this is the worst piece of complexity; you are about to see the worst of the whole thing—that is going to say 'o? = \text{obj}'.

Now each of these attributes is part of the picture. There is no limit to how complete we can make this picture. We can always add new attributes to any set that already exists. But if we add a new attribute to the picture, we must make sure that its value is consistent with the value we have already given it. We can't have an attribute 'A' with value 'a' and the same attribute 'A' in the same list with the value 'b', because that would be inconsistent. But the head features here could have itself another attribute other than the one that is the value of this variable 'o'. And so long as it is a different attribute it can have whatever value it likes.

Okay, now how are we going to make sense out of all this? Well, the verb itself also has a frame, which will consist of 's?', 'o?', and the 'rest'. And it is going to have some head features (HF) and it is going to have the value 'h':

\[
\text{CAT} = \text{V} \\
\text{FRAME} = (s? \ o? . \text{rest}) \\
\text{HF} = h?
\]
Let me rush through the rest of this now. And this NP (following V) is just going to have some head features. And it is going to have the value 'obj'.

\[
\begin{array}{c}
\text{CAT} = \text{NP} \\
\text{HF} = \text{obj}?
\end{array}
\]

This means an awful lot. This is going to just match the object here. Whatever object is picked for this sentence here is going to have to--and now let me really use the word--it is going to have to 'unify' with the variable 'obj'. That means it can have whatever attributes and values it likes just so long as it has this (NP) value for the category feature. I don't care about the rest. But for the category feature it has this NP value, and for the head feature attribute it must have...anything it likes, so long as it is unifiable with the value of the variable 'obj'.

\[
\begin{array}{c}
\text{CAT} = \text{VP} \\
\text{FRAME} = (s? . \text{rest}?) \\
\text{HF} = h? = [o? = \text{obj}?] \\
\end{array} \rightarrow \begin{array}{c}
\text{CAT} = \text{V} \\
\text{FRAME} = (s? o? . \text{rest}?) \\
\text{HF} = h?
\end{array} \begin{array}{c}
\text{CAT} = \text{NP} \\
\text{HF} = \text{obj}?
\end{array}
\]

Now so far we don't know anything about the variable object and maybe this is all we will ever learn about it. So, just for the time being, let us assume that whatever we find out about this noun phrase, it has these interesting features that we care about and they are simply going to become the value of this variable. The only things that can happen to destroy that idea is if we discover later some further details about the value of this variable. Remember it can never change, we can just learn more about it. That's all we can ever do.

Well, it turns out that that value 'obj' crops up in one other place, namely in the head features of the verb phrase. So, whatever we learn about that noun phrase here--and I am going to assume that the values of head features are the really meaty stuff about it that we really care about--is going to become part of the head features of this noun phrase that we are interested in building; in particular, it is going to become the value of some attribute; unfortunately we don't yet know which attribute. Well, for that we look at the verb frame. Suppose that the frame of the particular verb that we were looking at said that it has an agent and it has a patient, so that it is a standard transitive verb. Well, what that would say is, you would take the second thing in the verb frame--that is the 'o' thing--and that 'o' has to unify with the attribute over here in the VP HF. So that means that this noun phrase is going to become the patient. Notice that it is the verb that decides what it is going to become. Suppose that it was required that it should be in the accusative case. Yet all we would have to do here is to say that the verb has a restriction on it to the effect that the patient must be case=accusative:
And that would mean that the only thing that we would be able to unify with the patient attribute here was something which has this further attribute. So long as that unification could go through, then it would be all right to take that. Notice that we don't have to do anything; we can say anything here we like, about any of the things that the verb picks. You can say they have to be animate or they have to be human or they have to be speckled pink or whatever it happens to be. So that now we know that the head features of the verb phrase have to unify with whatever the head features of the verb happen to be. And in that way this is a feature; in fact, that is why they call it head features, because the verb we regard as being the head of the verb phrase, and the verb phrase in turn would become the head of the sentence. And the head features would be copied over in that case also.

Now let me show you just one other small thing that you buy with this—I've got lots of other examples here that are so meaty and interesting and you won't get to see any of them—how does passive work in a system like this? Remember we have all of these verbs all of which can say whatever they like about the various complements that they take. Well, we regard an English verb as having two kinds of past participles that generally look the same, just the one which is used in perfect tenses and the one that is used for forming passives—let me not pursue that a great deal further for the moment. But there is a rule that tells you how to get from the active form of a verb to its passive form. It says that the verb has a certain frame, and I will say for the moment that that consists of the 'first' thing and the 'rest', so it can be anything, any list of things that contains at least one member, and it has some head features—well I won't spell out the details for the head features, they could be almost anything—that can be rewritten as, can be stated alternatively as, something which is a category verb, which is a frame which is 'rest' and head features which are the same as these (the active verb's) head features except that they have voice=passive:

\[
\begin{align*}
\text{CAT} &= V \\
\text{FRAME} &= (\text{first?} \ . \ \text{rest?}) \\
\text{HF} &= h? \\
\end{align*}
\]

\[
\begin{align*}
\rightarrow
\begin{align*}
\text{CAT} &= V \\
\text{FRAME} &= \text{rest?} \\
\text{HF} &= h? = [\text{voice} = \text{passive}] \\
\end{align*}
\]

Now passive in English is done periphrastically; that is to say it is done by putting certain auxiliary verbs in front of it and it does not get to be passive until that happens. But when you put the passive auxiliary verb in front of it they are also going to say that it has to be voice=passive. And so it is only with this form of the verb that the unification will go through. So we don't have to worry about that aspect of it. The key thing that we have done here is to take the verb in its original active form which consisted of some number of things, the 'first' and the 'rest', and we have simply replaced it by the 'rest'. All right, we have removed the first thing from that list. What that means is that by this rule here
(the unified VP rule above), the next thing, if there were going to be a complement, it would have to be not the second thing anymore, but the third thing, because the thing that used to be the second thing is now the first thing, and it will be the subject. So what used to be the object will now automatically become the subject. This sort of maneuver works for dative movement and translations such as NP complement into VP complements and other technical things of that sort.

I now consider myself to have one minute in which to show you very briefly of unification working in a totally different environment, namely morphology. Consider that you cannot normally reverse rewriting rules. So a rule that says that a y in certain circumstances on the end of an English word was rewritten as ie, for example, in the environment preceding a vowel or the funny S previously spoken about:

$$y \rightarrow ie / \begin{cases} \text{Vowel} \\ S \end{cases}$$

This rule might occur in an ordered list of rules of this sort; in particular, the funny S rule has got to follow it, hasn't it, because it is still there when this rule applies. These rules are not generally reversible. Let me give you a monotonic version of this same rule. What I will give you is a state diagram, which says you can rewrite a y--actually this is going to be part of this rule; it would take another one to finish it--you can rewrite an i into a y, provided that immediately following it you rewrite nothing at all as an e:

So you combine to use both the i and the e, provided that what follows this brings you back to the beginning, either rewrites the funny S as a funny S, or takes some vowel and produces that same vowel as output. If, on the other hand, you were to find on the input a y and were to produce a y on the output so this would be a y that was going to be dealt with not by
this rule but by some subsequent rule, then that's all right, provided that what follows is in fact not the funny S or a vowel, because if you failed to do the correct translation in the case where what follows was one of these things, then you would have done something wrong; that is the case where the rule ought to have applied and you wind up in a state from which there is no way out. If you find anything else, you come back to the beginning. All I have done is to translate that rule into what is called a 'finite-state transducer'. So that now we can take the input to the rewriting rules from which we intend to get some output. And we can put it through a transducer equivalent to the first rule and another one equivalent to the second rule and as we have seen the step here is part of that rule—you haven't actually seen that but I'll tell you it is—you can draw another diagram right here and you keep on like this until you have dealt with all the rules:

Now this is a mathematical object of really rather different character than the rewriting rules, and thanks to that different character there is an algorithm that we can perform on two adjacent members of this sequence here and make one out of them. It means that we no longer have to bother about what would have existed in between them. And we can keep on doing that until there is only one left. We can get one transducer of exactly this form, no more complicated in its general structure, but of course very much larger, and we can be completely indifferent as to whether it takes this string into that string or that string into this string. It is, in other words, a completely monotonic device, and it comes about because this is a unification statement of that rewriting statement (y → i.e...)—notice nothing ever changes in this. One merely chooses the right set of statements to go to. I have now overshot my time by 58 seconds and I invite questions fast.
Linguistic Tools Incidental to Work in Machine Translation

Russell Bateman
Automated Language Processing Systems

During the development of natural language processing software, especially computer translation systems or computed-aids to translation (MT, MAT, etc.), tools are created and research performed which are often wasted when commercial companies die or academic research projects break up. What is the nature of the tools which are developed and just what are we loosing when the projects die with no more legacy than the spreading of good minds and ideas to the four winds?

I should like first to examine the evolution of development environments and the various vehicles which developers have originated to further MT. This is somewhat cursory and amounts to a fairly loose description of what I have encountered in my association with the MT world.

General Purpose Programming Languages

In the early days, the computer languages available to the developer were limited in scope. They had extremely poor string handling capability and were not the least bit suited to expressing linguistic algorithms involving parsing or pattern matching. Among these languages were various machine languages (early SYSTRAN code was written in assembly) and FORTRAN. Severe problems remained even after code was written in that the code was not easily maintained nor was it a simple matter to initiate a newcomer to developing in such an environment.

As early as 1960, formal specifications arose for general purpose programming languages which permitted superior implementation of algorithms through devices such as block structuring and modularity. Throughout the sixties and seventies languages like ALGOL, PL/I, PASCAL and 'C' were born but with the exception of PL/I and 'C', they still lacked formally defined string handling. At Brigham Young University, the Translation Sciences Institute (TSI) as it was to be called, began to tackle machine translation and soon chose PL/I, primarily because they were using IBM hardware but also for its relatively superior algorithmic ability and its string handling. PL/I furnished a built-in data structure for strings but manipulation was still mostly in the hands of the programmer who was required to write libraries of routines for this.

The 'C' programming language has become an excellent system programming language because of its portability and efficiency, but it is nevertheless limited in the same way as PL/I and other 'structured' programming languages in that it has almost no built-in string
manipulation functions at all. Furthermore, these languages do not allow linguistic algorithms, such as parsing or morphological analysis, to be written in concise and intuitively obvious ways. See Figure 1.

Special Purpose Languages

In the early seventies, new languages like LISP and Prolog surfaced with a markedly superior ability in linguistic expression, particularly in expressing and manipulating semantic relationships, which was recognized early on as perhaps the greatest nightmare to overcome in machine translation. The Artificial Intelligence (AI) community has been most active in their use. Few commercial MT concerns have begun to use them. This is, however, understandable for three reasons. First, their initial availability on systems was quite limited and still is in any standard or portable form. Second, few offer native language compilation on any machine thus being difficult to integrate with existing programs which is often a requirement in the commercial setting. Companies like Automated Language Processing Systems (ALPS) rely upon a real-time environment for their translator aids and cannot accept the relative slowness of interpreters, particularly on the existing, general-purpose non-LISP hardware. Last, truly integrated environments where one finds the capability of calling special purpose language functions are scarce or non-existent. Fully compiled versions of these languages cannot, by definition, have all the functionality of the interpreted versions. In any case, LISP and Prolog may be well suited to solving many linguistic or AI problems, but their use as systems programming languages is still not widely accepted nor possible (as least not on a wide enough range of hardware). Realistically, a commercial MT company must provide an entire translator environment, geared to helping the translator translate rather than frustrating him, and in doing such, must integrate a great deal of systems software like word processors and control environments which cannot be done practically in the current array of 'AI' programming languages, because, as we said, of the relative unacceptability of systems programming in these languages. We look forward to the day when LISP machines are widely available and at reasonable prices. See Figure 2.

Comparing general purpose programming languages like ALGOL and 'C' with LISP and Prolog in our context however is quite pointless -- like comparing apples and oranges. To create a good translation environment, one needs document production facilities foremost and that implies systems and general purpose programming languages for the reasons already established. To meet the needs of linguistic expression -- which is worlds apart -- special purpose linguistic languages are necessary, and most companies come to that realization, sooner or later investing a great deal of time and talent in this area.

Linguistic Support Software or 'Lingware'

All projects have used some system supported programming language or another but true innovaters in MT companies, among them SYSTRAN, TSI, TAUM, Weidner and ALPS have designed and implemented what one could call
```c
#include <mtyper.h>
#include <parsingtypes.h>

#define TRUE 1
#define FALSE 0
#define ParseTreeNodeSize sizeof(struct ParseTree)

extern struct ParseTree *BegSeg, *EndSeg;

void build_nph()
{
    char *malloc();
    boolean building_noun_phrase = FALSE,
    first_node = TRUE;
    struct ParseTree *lay, *new_np, *last_brother;

    /*
    * Beginning at right of sentence, parse for simple noun phrases.
    */
    node = EndSeg;
    do
    {
        node = node->left;
        if ((not building_noun_phrase) && (node->cat == NOUN))
        {
            building_noun_phrase = TRUE;
            new_np = (struct ParseTree *)malloc(ParseTreeNodeSize);
            new_np->id = NOUN;
            new_np->son = node;
            node->father = new_np;
            last_brother = node;
        }
        else if (building_noun_phrase)
            switch (node->cat)
            {
                case DET :
                case ADV :
                case ADJ :
                case NOUN :
                    node->father = new_np;
                    node->brother = last_brother;
                    last_brother = node;
                    break;
                default :
                    building_noun_phrase = false;
                    break;
            }
        }while (node != BegSeg);
    }
```

Figure 1: "A Simple Noun Phrase Rule Fragment in 'C'"
sentence( NP, VP) --> noun_phrase(NP), verb_phrase(VP).

noun_phrase( np(Det, Noun, Rel) ) -->
  determiner(Det), noun(Noun), rel_clause(Rel).
noun_phrase( np(Proper_noun) ) --> proper_noun(Proper_noun).

verb_phrase( vp(TV, NP) ) -->
  trans_verb(TV), noun_phrase(NP).
verb_phrase( vp(IV) ) --> intrans_verb(IV).

rel_clause( rel(that, VP) ) --> [that], verb_phrase(VP).
rel_clause( rel(nil) ) --> [].

determiner( a(a) ) --> [a].
determiner( det(every) ) --> [every].

noun( n(man) ) --> [man].
noun( n(woman) ) --> [woman].

proper_noun( name(john) ) --> [john].
proper_noun( name(mary) ) --> [mary].

trans_verb( tv(loves) ) --> [loves].
intrans_verb( iv(lives) ) --> [lives].

Figure 2: "Fragment of a Grammar in Prolog"

Simple Noun Phrase Rules:

\[
\begin{align*}
AJP & \rightarrow \{ADV\} \ ADJ. \\
AJP & \rightarrow \ AJP \ ADJ. \\
NOM & \rightarrow \ NOUN \ NOUN. \\
NOM[1, 2=NOUN] & \rightarrow \ NOUN \ NAJ. \\
NOM & \rightarrow \ NOUN. \\
DET & \rightarrow \ FUNNY WORD(all, both, just, quite, such) DET. \\
NP\text{(features)} & \rightarrow \ {DET} \ {AJP} \ {NAJ} \ NOM\text{(features)}. \\
\end{align*}
\]

Figure 3: "A Rule Fragment in a Lingware Formalism"
'lingware' with the goal of permitting near direct expression of linguistic algorithms: structural analysis, syntactic and semantic transfer and inflection. Some of these new languages received affectionate names from their parents like LINGBOL\(^2\) (which was calqued on SNOBOL and COBOL) and MANIP\(^2\) (for manipulate), others were called simply 'Transfer Language\(^1\)'. They all performed the important role of linguistic support software and had the benefit of being quite maintainable; that is, the algorithms they expressed were easily corrected and improved. Newcomers, even pure linguists, translators and other non-data processing types, learned them easily enough and could express their algorithms adequately in a matter of weeks. I speak from experience here. They also had the advantage of making a particular project come together at an amazingly fast rate -- extremely important in the commercial world -- because the developers could skip the start-up time of performing mundane programming activities like pattern matching and the creation of data structures. In general, this lingware was designed to be used by all the individual language projects and usually succeeded well enough at that. The only drawback was that they were very difficult to enhance, but with experience, the creators began to make succeeding designs more complete and encompassing. Among all the parser tools ever created, few have actually been offered on a practical and commercial basis; one notable exception however is the New York University Linguistic String Parser.\(^3\) See Figure 3.

**Debugging and ISAMs**

Putting together a translation system is a big affair and the software becomes very large and complex. Debugging becomes a concern and is an indispensable tool to the linguist and programmer alike. It is almost always impossible to diagnose a problem in a system based on the translated output. Writing debuggers is something that each project tackles sooner or later whether they think their system-supplied debugger is adequate or not. An important part of creating linguistic support tools like a formalism for syntactic analysis is providing a means of intercepting and debugging their output before it is corrupted by another phase of the translation code.

It is also essential that an MT project have an ISAM or dictionary lookup capability, and if the development system they have chosen does not offer a suitable one, they will be under the obligation of writing one -- a non-trivial endeavour. Their ISAM (or often ISAMs) must provide fast and accurate access to lexicons as well as any tables they may use for grammar, inflection and the like.

**Simple Morphological and Synthesis Tools**

Occasioned by every MT effort are the essential building blocks which must be in place before one can begin the more academically satisfying and more talked-about stage of parsing. By this I am referring to programs that break up and define sentences and words and reduce words morphologically in order to identify and by look-up obtain the necessary syntactic and semantic information used by the parser. In English this
often occurs in the form of a reasonably simple algorithm but in the case of other languages, extensive tables and character matching functions are designed.

On the other hand of course, are the tools with which the target language is produced. We call this synthesis and it entails a host of useful programs to inflect nominal forms, conjugate verbs, determine capitalization, etc. any of which could find their application in assorted CAI (Computer Aided Instruction), writers' workbench packages and the like.

Research Aids, Dictionaries and Grammars

There are differing requirements between the needs of the machine in language processing and the needs of the intelligent human dictionary user. One example of this is the reference work, *L'art de conjuguer, dictionnaire des 12000 verbes*. The 'Bescherelle' as it is commonly called has been the de facto (albeit popular) French verb conjugation bible for a very long time. It treats some eighty conjugation types of the French language in a way that anyone of reasonable intelligence can understand. The problem here is that the computer is not capable of reasonable intelligence, and so each MT project has had to reorganize the tables.

Another example -- because I work most in French -- is the *Larousse Dictionnaire des verbes*, a rich work full of simple, straightforward research on verb valency with examples galore. And yet the work was approached from a traditional grammar standpoint in both the terminology it uses and the categorizations. I have personally adapted much of its codings in my work according to hit-and-miss, practical requirements imposed upon me by the necessity I have to 'get the best out for the current deadline'. There is a dearth of such reference material in languages other than French and German; MT researchers therefore have to research and codify their own. We have linguists at ALPS who have done this sort of thing two and three times as they went from project to project.

Allow me to quickly add to this list of references most often created and exploited by MT researchers, the backward or reversed dictionary and the text corpus. Brown University's corpus has served almost everyone since its tagging was finished in the early seventies, but other languages are not so fortunate. Corpora are difficult to obtain in French and other Romance languages, even though some do exist in academic environments. Many MT companies must resort to compiling their own reverse dictionaries (essential to the establishment of morphological tables) and corpora (used for statistical and contextual analyses of words).

Summary

To recap the products which are created incidental to work on every MT project, allow me to re-enumerate specifically. 1) String handling functions for the programming languages used. 2) Parsers and/or special
purpose programming languages for expressing grammatical formalisms developed by linguists. 3) Other tools often involving compilers and interpreters for performing the steps in the translation process such as morphological reduction, ordering, inflection/conjugation and capitalization or other graphological adjustments of the output. 4) Debugging or diagnostic display packages created for use by the implementors. 5) ISAM capability as a basis for lexicons and tables if none is available or suitable on the development system chosen. 6) Compiled data from grammar research, corpus study, statistical lexicography, semantics research, etc.

Is It Feasible to Sell or Otherwise Distribute these Materials?

Of course that is the question that my superiors and board of directors would be most likely to ask! The problem of feasibility seems to lie in two principal areas: the possibility of producing a workable package to sell and the desire or willingness of the producing company to share its development with potential competitors.

The packaging of such information can take several of many traditional forms, publishing in the case of dictionaries and grammar research or installation in the case of actual software packages like parsers and verb conjugators.

The more obvious impediment to the proliferation of specific linguistic tools in the commercial let alone public domain is the understandable desire of any company which perceives its existence as depending on MT systems sales, to alone reap the benefits of its own R & D lingware and technological edge. To digress I might state that this probably occurs even when the tools they are currently developing and using are academically obsolete when compared with the latest as defined by the participants in conferences like CALICO, COLING and other ACL happenings and the various conferences on AI; for it is doubtful that any of the truly commercial MT companies are now employing any linguistic knowledge or techniques that are not at least five to ten years old. And it is also very doubtful that real AI is even being used in any of the companies with actual products now on the market!

In view of the small number of MT companies in existence, the real market for computational linguistic tools and information might be the academic institution. Much information reaches the public domain through the conferences just mentioned and isn't fully exploited by the MT and university communities. In addition, however, publishers of dictionaries should be interested as the increased computerization of their industry, including their traditionally paper-medium products, will certainly overturn much of what has been compiled over the centuries. This of course applies dramatically to traditional school-boy grammars but can also find its application in age-old authorities such as Bescherelle and Grevisse's Le Bon Usage.

To a large extent, the public domain would be benefitted by the efforts of MT researchers, past and present, particularly in two areas: dictionaries and grammars.
France for example, is currently in a period of *informatisation* or computerization to place a world of information in the form of on-line reference materials, shopping and banking services only as far away from each citizen as his or her telephone and television screen. Prototype systems have already been installed in various French cities. Soon, I believe, the reliance upon manual, intelligent methods will be upstaged by the arrival of automatic, unintelligent ones and the publishers of dictionaries will be urged by software developers of these systems to change their format because the computer cannot operate on their implicit information.

In the area of grammatical theory and research, MT excels as a proving ground. To a great extent, the old school-boy grammars have been shown to be inadequate by attempts made to apply them on the machine. It is true that present machine applications can be unfair, especially in light of difficult semantic considerations, but coding any grammar's rules in a machine can be very instructive as the BYU-TSI project found out during the decade of its work with Junction Grammar.

**Conclusions**

In conclusion, I have shown that a great variety of useful by-products in actual tools and important research are created 'from scratch' each time an MT project is launched. I believe that we in the MT world are doing a disservice to the general advancement of MT by not examining possible and (in the case of private ventures) commercially harmless outlets for the mass of knowledge gained on each project.

**Notes**


2 'LINGBOL' and 'MANIP' are unofficial tradenames belonging to the Weidner Communications Corporation, Northbrook, Illinois. The term 'Transfer Language' was used at the BYU-TSI project.

A Survey of Family Scripture Study in the Latter-Day Saint Home

Gerrie A. Bills
Utah Technical College/Provo-Orem
A. Garry Cramsey
Brigham Young University

In the Latter-Day Saint society, scripture reading, which includes Bible, Book of Mormon, Doctrine and Covenants, and Pearl of Great Price, is given a strong individual and family emphasis. Scripture reading is considered an important instrument by LDS members for shaping of personal values, and much of this reading transpires within the framework of the home and family.

A review of the current literature on personal reading indicates that research has been conducted in areas that concern the influences of reading on the development of personal values (Sabine, G. & Sabine, P., 1983). A substantial literature on Bibliotherapy also exists that is somewhat related to the subject (Salup, A., & Salup, B.J., 1981). However, at this time there is no known professional writing that pertains to the LDS context and this specific topic of study—family scripture reading in the home.

Do Latter-Day Saint families study the scriptures together? If so, how is it accomplished, and to what extent do the family members find it beneficial? For understanding of the LDS point of view, the following are quoted:

...for I know they (the scriptures) shall be of great worth unto them in the last days; for in that day shall they understand them; wherefore, for their good have I written them. (2 Nephi 25:8)

And if you know that they (the scriptures) are true, behold, I give unto you a commandment, that you rely upon the things which are written; for in them are all things written concerning the foundation of my church, my gospel, and my rock. (Doctrine and Covenants 18:3-4)
And the Latter-Day Saints are chastised if they don't read:

Now behold, my brethren, I would ask if ye have read the scriptures? If ye have, how can ye disbelieve on the son of God? (Alma 33:14)

The implication, then, is that LDS individuals must study the scriptures in order to learn, apply, and internalize the precepts contained there. Each individual is charged with that personal responsibility.

This responsibility is taken a step beyond the individual and into the realms of family with the following scriptural admonition:

"Train up a child in the way he should go; and when he is old he will not depart from it." (Isaiah 54:13)

Within the LDS framework, this commandment seems to close the circle of scriptural study endeavors. The individual begins with personal study in an effort to build a knowledge and a conviction of the truthfulness of the scriptural messages. As the adult becomes a parent, he or she feels a strong need to guide and teach the children as they grow in the same cycle. Nephi, a prophet in the Book of Mormon, states the importance of parental teaching of the family so that "our children may know to what source they may look for a remission of their sins" (2 Nephi 25:26).

This brief overview of one element of Latter-Day Saint philosophy underscores the strong feelings of responsibility a parent senses to teach the scriptural principles to family members. As one contemplates venturing into such a reading program, several questions may arise.
1. How does one begin a family scripture study program?
2. How can the sessions be kept interesting?
3. What is the best time for the family members to meet together?
4. Would the individual family members gain personally from such a program?

With these comments in mind, other questions arise.

1. Are some families accomplishing this task?
2. In what ways might family scripture study influence today's families?
3. Are there any ideas for a scripture reading program that could be shared?

The present study will be used to establish a guideline of ideas that may be helpful to family heads and ecclesiastical leaders who wish to encourage and organize family scripture reading programs.

Methodology

Planning

A Provo, Utah ward of the Church of Jesus Christ of Latter-Day Saints was chosen as the survey group. This group consists of 116 families who live in the northwest area of Provo, a high percentage of whom are active in church functions. An average of the past five years' activity level includes:

1. From the ward membership of 523, 72% attended Sunday sacrament meetings.
2. Of the adult men's priesthood quorums, 82% of the 110 members attended weekly quorum meetings.

3. Of the 110 adult priesthood members, 91% completed home teaching assignments (personal visits to homes of members).

Family economic levels in the ward area vary, as do the fathers' choices of employment. Of the 70 members surveyed in the study, 20 are retired wage earners, 18 are professors/staff at Brigham Young University, 19 are professional (engineer, banker, lawyer, physician, artist, pilot, salesman), 14 are in small businesses (owners, managers, employees), 7 are government employees (city, state, federal), 5 are teachers (other than BYU), and 4 are students who work part-time. There are no welfare families within the boundaries of this group.

The bishop over the ward was contacted for approval and input. The questionnaire was formulated and reviewed by two education professors and an educational researcher at Brigham Young University, a religion instructor who works with students from a nearby community college, two lay LDS Church members who are not affiliated with the teaching profession, and the bishop of the chosen survey group. The questionnaire was then revised according to the recommendations of the group. The survey was distributed through the adult priesthood quorums. The bishop stressed that participation was to be voluntary.
The presidents of the adult quorums were contacted and their approval for distribution of the questionnaire received. A time and date were established for the actual survey.

Procedure

Each quorum president introduced the survey (which included an explanatory cover sheet) and allowed ten minutes at the beginning of the Sunday quorum meeting. The participants were asked to complete the questionnaire and leave it at the conclusion of the meeting.

Because there were a number of men whose church assignments required their absence from quorum meeting, the group was smaller than was desired for the survey. Therefore, several questionnaires were personally delivered to the absentees and then collected at a later time, in addition to those given at the meetings.

Within the adult elders quorum, there are 31 members enrolled. Twenty-five of these received questionnaires and 17 members answered and returned them. Several single elders living at home did not participate.

In the high priests quorum there are 80 enrolled members. Sixty received questionnaires and 53 members answered and returned them. One interested visitor asked permission to take his copy to his ward in Idaho for possible use there.

Limitations

The survey was not a total participation for the ward families. Female single-parent families are not included. Illness
and travel caused some interference. Because it was voluntary, several declined to participate. Thirty-six of the 70 total questionnaires were completed in entirety; however, 34 were only partially answered. This was a weakness of the study. The partially completed forms were included in this discussion, however, because several survey questions were open-ended, and this allowed the participants an opportunity to express their opinions.

Some survey items proved to be a problem. No children at home was misinterpreted to mean number of children. Some participants felt there was idea overlap which caused some confusion and repetition of answers.

The findings must be considered highly tentative because of the inability to control the variables involved. Differences in the ward groups such as socioeconomic status, employment, average ages, backgrounds, and education of the adults, as well as personal attitudes and level of activity, restrict the possibility that this study could be widely generalized to other wards of the LDS Church.

Findings

From the results, the survey would seem to indicate that the majority of the families in the Provo, Utah ward are conducting family scripture study programs with varying degrees of success.

From a total of 70 questionnaires, 40 respondents indicated that they do participate in family scripture study, and 30 indicated that they do not. Whether or not the family included children at home did not have as much influence as had been anticipated.
From the elders quorum, 17 (68%) participated in the survey. Seven (41%) indicated that they were participating in a family scripture study program, ten (59%) said they were not. Of the 17 survey participants, seven indicated their family included children at home; four listed husband and wife only. Six did not indicate family number. Of the seven who are holding such programs, four include children; three are husband and wife only.

From the high priests quorum, 53 (88%) participated in the survey. Thirty-three (62%) indicated that they were participating in a family scripture program. Twenty (38%) said they were not. Of the 53 high priest survey participants, 25 indicated the family included children at home; 17 listed husband and wife only. Eleven did not indicate a family number. Of the 33 who are holding the programs, 17 include children at home, 13 are husband and wife only, and three did not indicate family number.

Totals

70 survey participants
40 (57%) hold family scripture study programs
32 (46%) indicate children at home
21 (31%) indicate husband and wife only
17 (24%) did not indicate family number

The majority (30 our of 42 who answered the question or 71%) of those participating in family scripture study find evenings the best time to assemble. A few (9 our of 42 or 21%) prefer early morning or breakfast time. The frequency of the study times varied from once monthly to daily sessions, with the most (18 out of 44 or
41%) choosing to meet once weekly. Sessions, especially with smaller children, are kept short--15 to 20 minutes. Most (14 of 39 or 36%) read about 30 minutes, with some sessions lasting one or two hours. The longer sessions were either by adults only or families with older children at home.

The LDS scriptures--Bible, Book of Mormon, Doctrine and Covenants, Pearl of Great Price--are the basic tools of study. Most of the families (83%) provide individual scripture copies for each member who is old enough to read. Parents usually meet together prior to beginning the program and formulate the basic plans such as time and day of study, and the specific subject. The related materials are gathered and the lesson is planned and prepared according to the ages and needs of the family members. A few (4 out of 39 who answered the question) indicated a preference to meet as a family, rather than just the parents, to make the initial plans.

The following is a list of other supplementary materials referred to in the survey responses as being helpful:

For families with younger children:

1. Children's books (Bible stories, Book of Mormon stories, etc.)
2. Flannel board cut-outs
3. Taped scripture stories
4. Scripture games
For families with wide age differences:

1. Old and new Family Home Evening manuals (published by the LDS Church)
2. Scripture tapes (especially a dramatized version)
3. Pictures
4. LDS Church magazines
5. Marking pens and pencils

For couples or families with older children:

1. LDS Church books on related subjects
2. Published study guides
3. Tapes (usually not a dramatized version)
4. Relief Society and priesthood manuals, both old and new (published by the LDS Church)

One of the main challenges seems to be how best to proceed with the reading-study sessions. Most of those participating read aloud from the scriptures with family members each taking a turn. Other materials mentioned are used as supplementary ideas to keep the sessions interesting and to touch the interests of the various ages. Sixty percent begin their sessions with prayer. Most stress the importance of keeping the reading and discussion within the ability levels of the children. One family solves the age difference problem by dividing the group, with Mother teaching the smaller children and Dad the older ones. Most families agree that the program takes definite planning; it won't just happen. It tends to work better when it is consistent and becomes habitual.
Most families have encountered difficulties in either getting started with a reading program or sticking with it. Persistence seems necessary for success. The greatest deterrents seem to be a lack of commitment and a lack of consistency. Time and opportunity to meet are also pressing problems in today's busy world. Several families indicate that they prefer individual reading-study sessions, and they tend to emphasize this within the family in preference to group study.

Some expressed a desire for a reading-study program but felt they needed help in getting started or encouragement to continue. A desire for a printed guide, outline, or idea list was also expressed. Several stated a wish for more emphasis Church-wide, and particularly for an LDS Church published study guide that would include effective methods of working with families in scripture reading programs.

Other comments seem to indicate that many consider scripture reading to be a vital part of the Church's Family Home Evening program. Many feel Family Home Evening is more important than a separate scripture study time since opportunity to meet together is often limited.

Are there benefits that are worth the effort? All the families who are participating in scripture reading-study programs expressed positive attitudes. All felt that, in differing degrees, it was worth the effort. Many feel their family members are closer, and they are able to resist sibling contentions better than before. The general atmosphere of the home seemed improved and
more peaceful. Parents indicate that their sessions afford an opportunity to handle family or individual member's problems—that answers are found and help is offered within the scriptural framework. Most approved of teaching children to love the stories and identify names and people in the scriptures at a very early age. Parents indicated that they feel the sessions improve their children's overall reading ability since they consistently read (sometimes very difficult material) with the encouragement of the family.

Some indicated a positive influence on their teenagers as the young people were able to find answers to life's problems within the scriptures. Generally, feelings were strong that with persistent efforts, most family members grow to enjoy and look forward to the sessions together.

In summary, the majority (57%) of the surveyed families of the participating groups express positive personal and family benefits that are being acquired from the reading that they indicate could come from no other activity. The families feel they have benefited with a unifying effect, a closeness, that helps bridge daily problems. Positive personal values are taught at a young age and are reinforced throughout one's lifetime. Those who are actively reading scriptures within the family unit overwhelmingly indicated that it is well worth the effort involved to begin and continue a reading-study program based on the scriptures.
The information provided here might also be considered by various ecclesiastical leaders. Scriptural interest is often considered an index to individual or group spirituality. Ecclesiastical leaders might wish to survey their own groups and use the insight gained to guide their planning for membership needs. Parents within the group may appreciate a compilation of ideas gleaned from such a survey.

There are questions that remain unanswered, however. What influence might family scripture study have on individual reading ability, especially among children? Would scripture study influence an adult's or child's choice of leisure reading materials? Does scripture study positively influence a person's self-esteem? What about female single-parent families? Could family scripture study make a positive impact on the lives of a mother and children who face the trauma of a single-parent situation? This study might, then, be of value not only currently to groups, ecclesiastical leaders, families or individuals, but also to those who may conduct future studies of the possible effects of family scripture study on specific family relationships, and the shaping of character strengths and values either within an individual or a family.
References


PHONOLOGICAL DISTRIBUTION AND SYLLABLE LENGTH IN SPANISH

J. Halvor Clegg
Joseph Brannen
Brigham Young University

Much experimental work needs to be done before any understanding of the syllable will be arrived at. This study is part of a larger, long-term study that aims at understanding some of the features of this difficult concept.

It is not the purpose of this study to provide a definition of the syllable. There are many in the different manuals that can be referred to for definitions. We will assume the existence of the syllable and use its classical definition. This means having a nucleus and limits. In our materials the nuclei are always vowels. Our measurements will include these vowel nuclei and their limits. These may be realized as consonants, vowels or modulations in the vowel nucleus itself.

There are many studies on syllable length. These are extremely varied in nature and often conflict. The most common concept that has emerged is that there are languages that are "syllable-timed" and others that are "stress-timed". English is given as an example of a language that is stress-timed. This means that the rhythm of English goes from stress to stress. This results in a stressed syllable that is considerably longer than an unstressed one. For example, in the sentence "This is an incredibly complicated phenomenon", the syllable form would look like this:

-----/ --/ --/ ---/ ---/ ---/ ---/ ---/ ---/ ---/ ---/ ---/
This is an incredibly complicated phenomenon

Note that the long syllables are all stressed and that the short syllables are all unstressed.

The Romance languages have been listed as being syllable-timed. The same equivalent in Spanish (Es un fenómeno increíblemente complicado) (with minor changes for grammatical acceptability) would theoretically look like this:

--/ --/ --/ --/ --/ --/ --/ --/ --/ --/ --/ --/ --/ --/ --/
Es un fenómeno increíblemente complicado
Note that all the syllables are of the same length regardless of stress.

The only exception given among the Romance languages has been Portuguese, but the evidence for Portuguese is still inconclusive. The latest studies give it as being "half and half". Authors such as Peter Ladefoged give French as being an example of syllable-timed but less and less mention is being made of Spanish.

The earliest study done for Spanish was made by Tomás Navarro Tomás. His conclusions were that Spanish syllables greatly varied in length according to whether they were stressed or not, the stressed syllables being much longer. He based his views on Spanish poetry which is interesting but not an example of "real" language. Samuel Gili y Gaya concludes, in his study of read breath groups, that there is a general psychological tendency to syllable-timing, but that there are mitigating physiological factors such as style, number of unstressed syllables between stresses and complexity of the syllable. Pointon (1978) reanalyzed Gili y Gaya's materials and found that the difference between stressed and unstressed syllables amounted to a 50% increase for stressed over unstressed syllables.

Pierre Delattre did a comparative spectrographic analysis of English, Spanish, French and German. His results showed that closed syllables tended to be longer than open ones, that stressed syllables tended to be longer than unstressed ones and that unstressed, open syllables tended to be of the same duration regardless of their position.

Carroll Olsen (1972 and repeated 1984) did a study on a recording of a speech by Octavio Paz. His overall conclusion in both cases was that Spanish is syllable-timed. He recognized that there are differences in length, but the differences are much smaller than in English, for example.

Pointon, in a review of all of the above researchers, found that there are many factors which influence the length of the syllable. These include stressed vs. unstressed syllables (Navarro), speed of movement of the articulators from sound to sound, style, the number of unstressed syllables between stresses and the complexity of the syllable (Gili y Gaya), open versus closed syllables (Delattre), rhythm pattern, sound sequence, structural sequence and length sequence (Olsen). Pointon adds the obvious factors of dialect, elements measured (vowels versus consonants with or without transitions), sense groups versus breath groups versus isolated words, tempo and number of informants. Still other factors, particularly sociolinguistic
ones would add other dimensions. Pointon concluded that Spanish is neither syllable-timed nor stress-timed. He felt that it is "segment-timed" being a mixture of the segments of the syllable combined with stress.

In this study we tried to maintain a uniform sociological level. The informants were all female (6), from the same generation (21-35) and from the same social class (middle class). We felt that we wanted to work with a closed text (which was read by the informants) since earlier studies had been done giving an indication of what basic phenomena we wanted to work with. Besides, Clarke (1975), in his study comparing free text versus read materials found no appreciable difference between the two. We, therefore, created a questionnaire which would examine various aspects of the elements of the syllable.

Álvarez (1981) showed for Spanish what Fails & Clegg (1983) did for Portuguese; that the type of syllable did not affect the vowel quality. Some care had to be taken in the selection of items due to neighboring phonemes since Álvarez showed that some consonants (among them the phonemes /s/ and /t/) can influence the quality of the vowel. For our overall study, position in the phonetic group was very important. It is generally known that acoustic phenomena diminish towards the end of a sentence. We also did our analysis speaker by speaker to mitigate individual differences.

We wanted to compare persons from different dialect areas, hence we selected informants from different general areas in the Americas. The dialects chosen were: Argentine (Buenos Aires), Chilean (Santiago), Columbian (Bogotá), Salvadoran (San Salvador), Puerto Rican (San Juan), and Mexican (Mexico City).

We asked them to read the questionnaire which included 94 words that gave us some 125 tokens. We recorded in an acoustic studio on a Sony cassette recorder. We processed the recordings on a Digital Sona-Graph™ 7800. We only analyzed the first 4,000 cycles, since the elements we wished to study would all show up in this segment. We made contour sonagrams for facility in reading. The sonagrams were measured by a special ruler that provided frequencies. We only used the first two formants for our vowel timbre study. The results show no differences between the timbre of each individual vowel in the different positions (tonic/atonic).

As in the timbre study, we averaged the results for this study, individual by individual and position by position. The individual results were, in general, the same as the average for.
all of the informants, that is, no instance varied appreciably from any other.

We used the exact same data in this study as in the timbre study for purposes of control. The same tokens were measured, only this time for duration. This was done with a ruler calibrated in millimeters, measuring both consonants and vowels that belonged to the syllable in question. This approach was different from previous ones that measured individual sounds, which is difficult at best. This type of measurement eliminated the question of transitions except between syllables. This means that the sonorants that tend to complicate vowel/consonant boundaries were included neatly in their separate syllables.

After measuring the syllables physically, we factored this measurement against actual recording time. There are 5.12 seconds of recording time on a sonagram at 4,000 hertz recording. This recording time converted into a physical measurement of length is equal to 317 millimeters. Converting the time (5.12 seconds) into milliseconds (5120 ms) and dividing by the 317 mm, we derive a factor of 16.1542 ms/mm that we used in calculating the actual duration of each syllable.

The questionnaire gave a list of words that produced each vowel /a,e,i,o,u/ in four different positions: Initial, pretonic, tonic and post-tonic. The words were measured, the measurements then totalled and an average determined. These averages were compared for differences and then averaged among themselves to produce the following results:

### AVERAGES OF SYLLABLE LENGTH IN DIFFERENT PHONOLOGICAL POSITIONS

<table>
<thead>
<tr>
<th></th>
<th>INITIAL</th>
<th>PRETONIC</th>
<th>TONIC</th>
<th>POST-TONIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>8.66 mm</td>
<td>11.31 mm</td>
<td>13.54 mm</td>
<td>9.02 mm</td>
</tr>
<tr>
<td>e</td>
<td>139.87 ms</td>
<td>182.67 ms</td>
<td>218.69 ms</td>
<td>145.69 ms</td>
</tr>
</tbody>
</table>
RESULTS OF THE STUDY

There were some elements that we could not corroborate. For example, we did not study final vowels due to the overall design of the study. Neither did we provide phonic groups nor sense groups. We saw some tendencies in the different speakers, but would hesitate to generalize to the entire dialect due to the limited number of informants even though the research required for six informants is extensive and adequate for our design.

The results of our study do show the following: 1) Tonic syllables are longer than atonic syllables by 3.36 mm and 54.31 ms or 29.18%. 2) The difference among all the atonic syllables only amounted to .372 mm and 6.01 ms indicating great consistency in atonic position. 3) We found a correlation between length and voicing in all syllables where the following consonant was voiced. The more voiced the consonants (sonorants versus unvoiced stops, for example) the longer the preceding syllable. Ladefoged (1975) reports a similar phenomenon for English vowels. 4) We found no definite correlation between open and closed syllables per se. There were cases of very short closed syllables and very long open syllables depending on other factors such as stress and voicing.
Further study will have to be made using minimal pairs that take these factors and others, such as word length into account. These will follow as well as free speech comparisons in a continuation of the broad study.

Is Spanish syllable-timed or stress-timed? Pointon feels that perhaps it's in between. It does not possess the difference in stressed versus unstressed syllable length that English does, nor does it appear to be stress-timed as does English. In fact, it is very close to French in its syllable production and French is considered to be syllable-timed. The only difference between the two lies in where the stress falls. Perhaps a more meaningful question could be asked. Is there such a thing as a syllable timed language? It would have to be one without stress in its system since stress generally gives additional duration to the stressed element. Perhaps a more proper consideration would be to evaluate the amount of additional duration that stress provides and then say, for example, that a syllable with more than a 50% increase, thus having a major impact, leads to calling the language "stress-timed" and less than this makes it syllable-timed. In this sense, Spanish is "syllable-timed".
REFERENCES


Clarke, W. M. 1975. The measurement of the oral and nasal sound pressure levels of speech. JPhon. 3, 257-62.


Gili Gaya, Samuel. 1940. La cantidad silábica en la frase. Castilla (Valladolid). 1, 287-298.


Navarro Tomás, Tomás. 1918. Diferencias de duración entre las consonantes españolas. RFE. 5, 367-393.

Navarro Tomás, Tomás. 1922. La cantidad silábica en unos versos de Rubén Darío. RFE. 9, 1-29.


SET THEORY AS AN EXPLANATION FOR THE SELECTION BETWEEN
MÁS QUE/MÁS DE IN SPANISH

J. Halvor Clegg
Mark Seely
Brigham Young University

Selection between the comparatives más que/más de 'more than' has long been an issue among grammarians of Spanish. Spanish grammars and textbooks contain lengthy explanations and rules with different solutions ranging from a listing of a number of Spanish sentences illustrating their usage to English translation equivalents. Unfortunately, none of these provides a satisfactory indication as to how the distinction is made, and exceptions are frequent.

The use of the notion of 'Set Theory' in grammar is not a new one. It has been written on by 'Cooper 1964' and 'Ortiz & Zierer 1968' and suggested by linguists such as 'Jacobson 1961', 'Bach 1964', and even 'Chomsky 1957'.

In this study we will apply the concept of set theory to the más que/más de distinction in Spanish and through this approach provide a solution that obtains more generally and gives information as to how the distinction functions.

Most Spanish grammars and textbooks list the más que/más de forms under a section titled 'Comparison of Inequality' or 'De and Que as Translations of the English "Than"'. We consulted over twenty Spanish grammars and textbooks as well as several journal articles. A synthesis of the rules which reflects a consensus follows:

COMPARISONS OF INEQUALITY

I más que:

'Pedro es más alto que Juan.'
(Pedro is taller than Juan.)
'Pedro come más que Juan.'
(Pedro eats more than Juan.)

II más de (el, la, lo, los, las) que:

When each member of a comparison contains a different verb (i.e. there are two clauses)

1) If the point of comparison is a noun, use de (el, la, los, las) que:
'Gasta más dinero del que gana'.
(He spends more money than he earns.)

'Ha cometido más crímenes de los que ha confesado.'
(He has committed more crimes than he has confessed.)

2) If the point of comparison is an adjective, adverb, or an entire idea, use de lo que:

'Era más pequeño de lo que parecía sentado.'
(He was smaller than he seemed when he was seated.)

'Hicieron mucho más de lo que era imaginable.'
(They did much more than what was imaginable.)

III In affirmative statements de is placed after más when followed by numerals or a numerical expression:

'Caminó más de cincuenta metros'.
(He walked more than 50 meters.)

'Perdimos más de mil dólares'.
(We lost more than 1,000 dollars.)

IV In negative statements either de or que may be placed after más when followed by numerals or a numerical expression, but with a difference in meaning.

No ... más que ... = 'only'
No ... más de ... = 'not more than'

'Ella no tiene más que veinte años'.
(She is only 20 years old.)

'Ella no tiene más de veinte años'.
(She isn't more than 20 years old.)

V After a superlative, English in is translated by de.

'María es la más guapa de la clase'.
(María is the prettiest girl in the class.)

'Es la montaña más alta de América'.
(It is the highest mountain in America.)

The two forms which seem to cause the most confusion are the no ... más que versus no ... más de followed by a numerical expression and the de (el, la, lo, los, las) que comparative. 'Hayward Keniston 1937:622' notes that '...it is not possible to determine whether a distinction existed in sixteenth-century
Spanish ... but on the whole it seems improbable.' Andrés Bello 1847:331' stated in his grammar that either de or que could be employed in such negative statements, but that all in all the preposition de 'suena mejor' 'sounds better' than the conjunction que. In his revision of Bello's grammar, Rufino Cuervo 'Bello 1847:499' added a note which reads: 'One perceives a difference between No se gastaron más de cien pesos and No se gastaron más que cien pesos. The latter seems to mean that only one hundred pesos were spent and the former that any amount up to and including one hundred pesos was spent'.

'Ramsey 1956:143', first published in 1894, states that either que or de may be used in negative statements, the preference being for que. A footnote explains that theoretically, no ... más que means only and no ... más de means not more than. Finally, the 'Real Academia Española 1973:418' grammar informs us that in modern language the construction with que has gained a lot of ground on the classical construction with de in negative statements.

According to the grammars then, there is not yet a clearcut distinction between que and de in the comparative forms with numerical expressions.

As for the de (el, la, lo, los, las) que form, the 'Real Academia Española 1973:419' grammar says that the preposition de is generally preferred since it has greater traditional force, but que is also permissible. It goes on to mention that perhaps part of the reason for preferring de is for purposes of euphony, so as to avoid a que lo que type of construction by employing de lo que. However, 'Bolinger 1950:35' would disagree with this last statement and insist that the semantic distinction between que and de is respected regardless of euphony.

After reviewing the grammars, we felt it necessary to research current modern educated usage. 'Andrés Bello 1847:27' said, 'The grammar of a language is the art of speaking correctly, that is, according to proper usage, which is that of educated people.' We selected ten novels written by authors representing nine different geographical areas in the Spanish-speaking world. The nine areas were: Argentina (Rio Plate), Chile, Bolivia (Andes), Peru (Coastal), Venezuela, El Salvador (Central America), Mexico, Puerto Rico (Caribbean), and Spain. The novels were all written after 1970. A minimum of fifty pages of each novel was read. This provided just over 1000 pages of text. There were 271 examples of más que/más de found in the sample.

We carefully designed a questionnaire containing 25 items
that would elicit a choice between the two forms. These were presented to ten native speakers of Spanish from the same nine representative geographical areas and then analyzed to produce generalizations which we used to evaluate the rules presented by the grammarians. The speakers were chosen to reflect the same sociological characteristics: female, educated, age 20-30, etc.

We found, as did 'Bolinger 1950' in a similar study that the respondents generally agreed in their usage. They were able to consistently make the distinction between the two forms (147 out of 149 applicable responses). The exceptions to this unanimity occurred in sentences of the type No ... más que/de ... followed by a numerical expression and ... de/que (el, la, lo, los, las) que, etc.

There were five sentences of the type No ... más que/de ... 26 of the applicable responses favored de and 13 favored que, while nine allowed either form and there were two items left unanswered.

There were also five sentences of the type ... de/que (el, la, lo, los, las) que... Both the writers and the grammarians gave four out of the five sentences as evoking the form de. Of 39 responses, 30 preferred de while nine preferred que. In the other sentence, evoking que, five preferred que, two preferred de and one accepted either response. Two of the informants gave no response.

The importance of these results lies in the confirmation of the apparent existence of some underlying generalization that enables the Spanish speaker to make the distinction between the comparatives más que/más de. Their hesitation on the no más que/de followed by a numerical expression and the de/que (el, la, lo, los, las) que forms only points out the fact that the distinction in usage has not been totally refined as of yet.

Nevertheless, among the many solutions brought forward for the distinction between más que/más de, that of set theory appears to be the most plausible.

Set theory is a vast, complicated mathematical concept in its entirety. Its application in this and other linguistic instances can be greatly reduced in scope. For our purposes, it will be treated as the distinction between: 1) one set and 2) more than one set.

In this case, when the comparison is made including elements of one set, the form de is used. If more than one set is used the
form *que* is used. Spaulding comments that the competition between *de* and *que* in comparisons in Spanish has its roots in spoken Latin. 'Saenz 1940:329' suggests that herein lies the 'key to the solution of the problem.' There is a phenomenon in Latin known as the 'genitive of the whole' which is formed with words indicating parts of the whole. This was transmitted into the Romance languages in the form of the preposition *de*. This idea 'relation of a part to the whole to which it pertains' seems to be an integral part of the application of set theory to this distinction. Several of the grammarians such as Bolinger, Gili y Gaya, Bull, Keniston, and Anderson included this perception in their presentations on *más que/más de* 'De Mello 1977:510-11' came closest when he said that '...*que* is used in comparisons involving two different entities, while *de* is required when reference is to a single entity.' While he understands the concept, he fails to recognize that the theoretical basis is set theory.

'Solé 1982' wrote an excellent survey article showing all the syntactical possibilities for *más que/más de*. De Mello wrote a review pointing out that all of her cases could be answered by his approach. They never came to a meeting of the minds, since Solé responded citing the same examples and saying that De Mello didn't understand what she was stating.

The essence of the matter is that both are right. Solé provides the syntactical structures but in the end her conclusions only say that there are two types of modification; one that is phrasal (*de*) in nature and one that is clausal (*que*) in nature. She still doesn't provide a basis for selection between the two forms.

The solution to the selection lies in set theory. We have taught this as being theoretically feasible since 1974. The research to verify it wasn't carried out until 1984. In the review of the literature we found that mention was made of this application by 'Ortiz & Zierer 1968'. Their observations appeared tucked away at the end of a general article on set theory and linguistics. They only explored the possibility and provided general suggestions.

Our research suggests that the underlying distinction here is that the Spanish speaker perceives each comparison in his mind as being one of two types:

1) An extension of one entity, set, or thing, which would require the use of *de*.
2) A comparison of two different entities, sets, or things, which would require the use of *que*.
The following examples with illustrations will demonstrate how the distinction works.

1. Pedro es más alto que Juan.

2. Gastas más dinero del que ganas.

3. Hicieron mucho más de lo que era imaginable.

4. Caminó más de cincuenta metros.
5. No tiene más de 20 años.

\[
\begin{array}{c|c|c|c}
\text{años a la derecha} & \text{años a la izquierda} & \text{21 no tiene} \\
19 & 18 & \\
18 & 17 & \\
17 & 16 & \\
16 & 15 & \\
15 & 14 & \\
14 & 13 & \\
13 & 12 & \\
12 & 11 & \\
11 & 10 & \\
10 & 9 & \\
9 & 8 & \\
8 & 7 & \\
7 & 6 & \\
6 & 5 & \\
5 & 4 & \\
4 & 3 & \\
3 & 2 & \\
2 & 1 & \\
1 & 0 & \\
0 & &
\end{array}
\]

6. No tiene más que 20 años.

\[
\begin{array}{c|c|c|c|c}
\text{tiene 20} & \text{20} & \text{no tiene} & \text{años/ X} & \text{X}
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\text{años} & \text{X} & \text{no tiene} & \text{un car} & \\
15 & 10 & 5 & \\
\end{array}
\]

7. Tiene más que un carro.

\[
\begin{array}{c|c|c|c}
\text{carro} & \text{casa, amigos} & \text{esposa, motocicleta, etc.} & \text{años/ X}
\end{array}
\]

8. María es la más guapa de la clase.

\[
\begin{array}{c|c|c|c}
\text{guapas} & \text{María} & \text{las otras guapas} & \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\text{guapas} & \text{María} & \text{las otras guapas} & \\
\end{array}
\]
Our approach provides a solution that applies more generally and provides information on how the distinction is made by Spanish speakers between the forms más que/más de in the comparative. This illustration of how the distinction is made through the application of set theory is an improvement over the somewhat lengthy and involved traditional grammar explanations. It goes to the real root of the distinction and handles all of the más que/de comparatives encountered in the Spanish language in a neat and concise way.
APPENDIX

QUESTIONNAIRE WITH RESULTS SUPERIMPOSED

FAVOR DE DAR LA FORMA CORRECTA DE MÁS DE LO QUE, MÁS QUE LO QUE, ETC. PARA LAS ORACIONES SIGUIENTES. SI HAY MÁS DE UNA POSIBILIDAD, Dé LAS ALTERNATIVAS CON UNA EXPLICACIÓN DEL SIGNIFICADO DE CADA UNA.

<table>
<thead>
<tr>
<th>DE QUE</th>
<th>EITHER</th>
<th>ORACIÓN</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>1</td>
<td>1. Gastas más dinero ___________ ganas.</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>2. Me pareció más serio ___________ hacían creer.</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>3. Se quedó en Madrid más tiempo ___________ pensaba quedarse.</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>4. Hoy nace mucha más gente ___________ se muere.</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>1  5. Un episodio de la Conquista - más fiel a la realidad, en cierto modo, ___________ hasta ahora se ha visto aquí.</td>
</tr>
</tbody>
</table>

FAVOR DE DAR LA FORMA CORRECTA DE MÁS QUE, MÁS DE PARA LAS ORACIONES SIGUIENTES. SI HAY MÁS DE UNA POSIBILIDAD, Dé LAS ALTERNATIVAS CON UNA EXPLICACIÓN DEL SIGNIFICADO DE CADA UNA.

<table>
<thead>
<tr>
<th>DE QUE</th>
<th>EITHER</th>
<th>ORACIÓN</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2</td>
<td>6. No (traigas) más _______ una docena.</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>7. Tiene más vidas _______ los gatos.</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>8. Tenían una carta más _______ los demás.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>9. (Querían) ser algo más _______ ricos.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>10. Es la montaña más alta _______ América.</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>11. No tienen que hacer más _______ sentarse en sus oficinas.</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>12. (Estaba) sin más adorno _______ una cruz.</td>
</tr>
</tbody>
</table>
13. En la vida, más vale una vez colorada ________ mil descoloridas.

14. ...porque son más tristes ________ entierro a pie.

15. No la tomamos en serio más ________ cuatro docenas de mujeres.

16. No era más ________ una niña asustada.

17. Jaime, que entonces no tenía más ________ siete años.

18. Ángela Vicario era la más bella ________ las cuatro.

19. No se demoró más ________ diez minutos.

20. Estaba más bella ________ nunca.

21. No duraría más ________ un minuto.

22. Perro más inteligente ________ los hombres.

23. Esta llamada capital no era más ________ un poblacho.

24. La moral de hoy, en más ________ un sentido, había cambiado mucho.

25. Vengo también de un ambiente más ________ acomodado.
REFERENCES

BIBLIOGRAPHY OF WORKS CITED


De Mello, George. 1977. Que and de as translations of 'than'. Hispania 60.510-11.


Spaulding, Robert K. 1962. De (i, la, lo, los, las) que vs. (el, la, lo, los, las) que or the force of tradition. Hispania 45.309-14.
SOURCE BIBLIOGRAPHY OF NOVELS EXAMINED


Looking back:
The didactics of semantization in historical perspective
Wilfried Decoo
University of Antwerp (UIA)
Belgium

0. Introduction

0.1. The history of science and language learning

In most branches of science there is a keen awareness of the many contributions that have built up, through the centuries, the total spectrum of the present knowledge and upon which further research continues to build. Especially in physics, in biology, in chemistry, in medicine, do we witness this historical perspective in the many names tied to scientific definitions, which point back to discoverers of decades and even centuries ago.

Not so in language learning: the names of past researchers in this field and the work they have done are unknown to practically all in the profession. One could even say that quite often present-day researchers present their work (usually a new methodology) as the first serious answer to successful language learning. No reference to previous similar answers is mentioned, no reflection on past experience with those answers is made.

We feel it important to look back, not only to give credit to those who have preceded us, but especially to learn from them, whether from their achievements or from their mistakes.

0.2. The question: the transfer of meaning in FL-education

The central questions of this presentation have to do with the transfer of meaning in foreign-language teaching: In what ways can a teacher make students understand what a new word or structure means in the foreign language? What are some of the problems connected to this understanding? What kind of strategies does a student use to grasp and to retain these meanings? We will approach this problem of semantization in didactic terms, namely in immediate relation to efficient classroom-instruction. And we will do this in historical perspective: the many ways in which teachers and researchers have experienced with the question of semantization are rich and revealing.
0.3. Methodology and didactics

I would like to stress the difference between methodology and didactics. Foreign-language teaching is mostly controlled by so-called methods, namely a number of precise instructions to transfer subject-matter. Very often a particular method is born in specific circumstances where it works successfully. It is then being commercialized, usually through a textbook to be applied by teachers who have not been involved in the growth and the specificity of the method. This situation leads to a number of tensions between a methodology and the teachers or between teachers themselves using different methods.

The study of these tensions is revealing. It helps us to understand the interaction and the problems between teacher-trainers, student-teachers, and teacher-mentors. A method can become an immutable anchor for some, and, by the same token, a target for critique for others. Very often a methodology helps to escape complex questioning, because dogma's are easier to handle.

Methodology implies a unilateral, normative approach to education. Didactics, on the other hand, is interested in the many possible teaching and learning processes that could make a specific situation more effective. In many cases didactics will gnaw at methodologies, not to reject them, but to question their validity in giving circumstances.

What we will discuss, therefore, is not the methodology of semantization, namely what is right or wrong as a dogma, but the variety of approaches.

1. From antiquity to the 19th century

1.1. Antiquity and Middle Ages

As far as we can tell from the sources available, the basis for the transfer of meaning in foreign language learning has always been translation. However, care was taken to avoid the word for word equalizations: dialogues and prose texts were taught in global sentences, showing a healthy balance: the naturalness of complete utterances was saved, while semantization was realized as efficiently as possible.

The use of illustrations to convey meaning remained limited, although there are some attempts to use the pictorial ornaments in manuscripts for didactic purposes, as for example in bestiaries.

Another way to transfer meaning is by using the target language itself, mostly by giving definitions of new words in the target language. In the Middle Ages this became a widely used procedure, especially to explain hairsplitting differences in Latin, as part of theological and scholastic training.
1.2. The Renaissance and the 17th century

The Renaissance, although considered a turbulent time of renewal, did not do much to revolutionize foreign language learning. The obvious and traditional techniques of centuries past were used; namely a fundamentally bilingual approach, using translation as the foremost means to convey meaning; the importance of the global sentence to ensure naturalness and fluency; a slow systematization of the pictorial support, providing charts and detailed drawings to name objects within a semantic field; and a continued use of definitions in the target language. After the first global language learning, wordlists would expand the subject-matter for direct use in the target language.

One could summarize the approach as follows:

- Phase I = practical dialogues
  - global bilingual first encounter
  - eventual pictorial support
  - followed by synthetic > analytic integration

- Phase II = wordlists (glossaries and lexicons)
  - analytic > synthetic integration
  - global monolingual application

1.3. Comenius: more bilingual than iconic

We should of course draw the attention to the foremost representative of foreign language learning of the period, namely Comenius, who has discussed in some detail the problem of semantization in his Didacta Magna (1650). Most people know Comenius from his Orbis sensualium pictus, a broad pictorial presentation of all aspects of life, which makes some believe that Comenius stressed pictures as the ideal way to transfer meaning. This is not so: in his Didacta Magna, he presents the use of translation as the best way for semantization, coupled, however, with the synthetic approach to full sentences. He also pleads for the use of grammar and dictionary on a contrastive basis with the mother tongue.

1.4. 18th and 19th centuries

The 18th and 19th centuries see the development towards more and more theoretical language instruction, whereby the study and translation of texts becomes an end in itself. Pictures to transfer meaning tend to disappear from textbooks. Dictionaries have become the most important tool, often less as an aid than as a method in their own right.

However, one should note that also a much more lively tradition of direct language learning continues to flourish, namely through the use of private foreign tutors in the higher social classes. Semantization, in such cases, made much more use of gestures, mime, pictures and contextual learning, instead of translation.
2. The Reform-movements: 1880-1900

The first in-depth research and experimentation with semantization happened in the last twenty years of the 19th century, known as the Reform-movement, putting all the elements in place which we still discuss today.

The battle of the Reform-movement is basically one between two extremes:

grammatical theorization \(\leftrightarrow\) practical proficiency

But because grammatical theorization grew out of approaches which were fundamentally bilingual, the strategy of semantization through translation is pushed into the grammatical camp and the opposite strategy is added to the Reform-movement, namely the elimination of the mother-tongue, leading to the following polarization:

grammatical theorization \(\leftrightarrow\) practical proficiency  

bilingual semantization \(\leftrightarrow\) elimination of mother tongue  

direct semantization

This is the way the Reform-movement is often represented, as if all Reformers defended this strategy as a main characteristic of their approach. However, when studying the articles and books of these linguists and learning psychologists, it becomes clear that their standpoints were not so simplistic and categorical.

Semantization was indeed one of the area's that stirred up a lot of study and controversy, leading to various approaches and many helpful insights. To name just a few of these:

- the function of direct procedures to transfer meaning:
  - objects and pictures: relation studies between word and concept
  - situational and intuitive identification: ways and degrees
  - activity integration; the power of mime

- nominal versus verbal basis semantization
- subjective versus objective language use for semantization
- presentation versus integration phases
- complete versus partial translation

The war over these aspects was thus more internal than against the past. Most Reformers advocated clearly that the use of the mother tongue was a simple necessity in the semantization process (Sallwürk, 1898, Sweet, 1899, Viëtor, 1882, Wendt, 1898). History shows that within the Reform Movement only a small group adopted extreme standpoints, one of which was the principle of absolute monolingualism: it is forbidden to use the mother tongue in a foreign language classroom.
This principle did not last long in the reality of the classrooms. Teachers, even those adopting such an extreme direct or natural method, quickly felt the limitations of monolingualism. Many articles of the time show the practical reactions against this strategy.

3. The eclectic period: 1900-1940

3.1. The International Conference of 1909

The reactions culminated in the international language teacher's conference of 1909, where a vast majority voted against the continued implementation of too strict procedures, among others monolingualism in the classroom. But research would continue, trying to combine the best of all possible methods and bringing us into a rich eclectic period.

3.2. H. Büttner

In one of the first major studies on semantization, Büttner (1910) adopts the principle of natural language command, namely that it is possible to have a direct tie between the foreign word and the concept, excluding the mother tongue medium; however, this tie can only come through integration: when presenting the foreign word it is not possible to eliminate the mother tongue reaction.

Büttner makes a distinction between "das Verstehen" (understanding) and "die Aneignung" (integration). He made a thorough study of the problems of unsatisfactory semantizations when applying extreme direct approaches. He rejects the argument that translations do not always cover concepts precisely, by studying these differences and concluding that the fundamental meanings for the fundamental vocabulary are indeed very much identical. Semantic nuances always come at a later stage, when the context will allow the student to grasp these nuances without reference to the mother tongue. The main advantage for a semantization through the mother tongue is the gain of time in the learning process.

For the second phase, namely "die Aneignung", Büttner develops a number of didactic strategies to develop and strengthen the direct tie between the foreign word and the concept, and to lessen, at the same time, the translation reflex.

The work of Büttner is a remarkable and quite thorough approach of the problems of semantization, providing at the same time very concrete suggestions for effective language learning.

3.3. Ernst Otto

Through a number of thorough experiments, Otto (1925) comes to the same conclusion as Büttner, namely that semantization through the mother tongue remains the most efficient and most precise way to transfer meaning. But he seems to go beyond Büttner by advocating the use
of bilingual exercises, also in the integration phase. One of his strategies is a remarkable combination of translation and automating drill, precisely meant to eliminate the word for word translation. In fact Otto defends the same principles as Büttner, namely to integrate the direct ties between foreign word and concept, but he does this more systematically over a longer period in the learning process.

3.4. Other important researchers of this period

Among those who continued to study semantization, we should mention people like Flagstad (1913) Aronstein (1926), Palmer (1917), Glauning (1910), Kirsten (1920). All these continue to recommend a bilingual approach for semantization, with a combination of various strategies to ensure the transfer of meaning. They warn against unilateral approaches.

4. Behaviorism and the audio-revolution: 1940-1965

In spite of all the research done and the quite effective strategies developed in the first 40 years of this century, all this is forgotten in a repetition of the Reform Movement, but much more extreme and much more arrogant in its statements of absolute originality.

However, it would be a mistake to think that the audio-movement of the 40's started out as a complete and clear-cut revolution. Historical research shows that the 40's stressed the importance of the behavioristic drill, but that the strategies were not that extreme at first. For example, in connection with semantization, translation was used without constraint in the presentation phase. The pattern-drill books contained the equivalent of the drills in the mother tongue, so that a student could check the meaning of what he was practising.

The extreme positions develop later, comparable to what happened within the Reform movement of the 19th century.

At the end of the fifties a few specialists in foreign language learning, but not in its history, decide to launch a so-called new strategy: monolingualism. In 1959 Nelson Brooks publishes his Language and Language Learning in which he posits without nuance for the English-speaking student who will learn a foreign language:

What the learner must not do:  
- he must not speak English  
- he must not learn lists of English-foreign language equivalents  
- he must not translate from the foreign language into English

Brooks bases his ideas on the linguistic theories of Erwin and Osgood who studied the lingual systems of bilinguals and who made the famous distinction between compound bilinguals, using one conceptual system, and coordinate bilinguals, using two conceptual systems. For Brooks, language learning means to become a coordinate bilingual and thus the
conceptual system of the mother tongue must be avoided by teaching directly and only in the target language. This reasoning is not only simplistic, but it passes over all the fundamental research and insights of the Reform Movement and the eclectic period.

Brooks' so-called discovery is repeated in a few other places, among others in France, where the principles of the audio-visual methodology also contain a categorical rejection of any semantization with the help of the mother-tongue. For the French, however, it was more than a pedagogical issue. Their political action to preserve and strengthen the French language in the many newly independent countries in Africa entailed the "de-conditioning" of the mother tongue.

In the sixties monolingualism thus becomes again a battering-ram, taught as a fundamental principle in the training of foreign language teachers: as a teacher you are not supposed to use one word of the mother tongue in classroom. Semantization will be done through visuals and through contexts.

5. Since 1965: the cycle continues
5.1. The cognitive reaction: probing deeply into semantization

The reaction against extreme monolingualism came slowly, because the audio-proponents defended their principles with a lot of pseudo-scientific terminology and vague references to "recent discoveries in psychology and linguistics". The reaction came in the first place from academicians, like Wilga Rivers (1964), J.B. Carroll (1966) and D. Ausubel (1964). In his article, Ausubel stated that the "avoidance of the mediational role of the native language" was "psychologically incompatible with effective learning processes in adolescents".

The controversy resulting from these opposing standpoints led to scores of experimental projects, illustrating among other things that monolingualism certainly was not the final answer to semantization (Chastain & Woerdehoff, 1968; Smith, 1969; Von Elek and Oskarsson, 1972; Janssens, 1974).

But just like in the first half of this century, studies continued, repeating the research done and coming up with insights that often did not match the depth and the nuances of work done decades ago.

However, four researchers deserve special mention for their specific and thorough work on the problem of monolingual versus bilingual semantization, namely C.J. Dodson (1967), W. Butzkamm (1973), T.T. Meijer (1974) and S. Olsson (1973). Their conclusions are similar:

- monolingualism is a strategy without any scientific foundation
- monolingualism leads to more mother tongue interferences than bilingual semantization
- a bilingual approach leads to more correct and stronger integration of new material
- a bilingual approach does not lead to translation habits and does not hamper automatization of lingual reflexes.

Other researchers who have studied the same problem and who have come up with the same conclusions and with various didactic recommendations, include Hammerly (1982), Lübke (1971), Rattunde (1971), Dietrich (1973), Koster (1975), Hüllen (1971), Lim (1970).

5.2. New methodologies and old dogma's

But the cycle continues. New methodologies that have appeared in the eighties, although some did not posit monolingualism at first, tend to become entrenched in extremes by well-meaning revolutionaries. The principle of monolingualism is alive and well. The didactics of semantization continue to go through the same cycle. In classrooms all across the world many teachers are convinced they should never speak one word of the mother tongue to get the meaning of foreign words or sentences across. The arguments they use are those that were used by some extremists from the Reform Movement and from the audio-revolution. The careful refutations, the experimentations and insights of all who responded to these extremes are unknown to these teachers and new methodologists.

This is why it is helpful to teach them to look back. Only then can language teaching mature to a balanced and conscious science.

References

Aronstein, Ph., Methodik des neusprachlichen Unterrichts, I/II (Leipzig, 1926).


Büttner, H., Die Muttersprache im neusprachlichen Unterricht (Marburg, 1910).


Flagstad, Chr. B., Psychologie der Sprachpädagogik (Leipzig, 1913).

Glauning, F., Didaktik und Methodik des englischen Unterrichts (München, 1910).


Hüllen, W., Linguistik und Englischunterricht (Heidelberg, 1971).

Janssens, Walter, Vreemde taal op de basisschool: Een didactische benadering van de problematiek, proefschrift (Gent: Rijksuniversiteit, 1974).

Kar, R., So lernt man Sprachen (Meisenheim, 1959).


Lübke, D, "Die Rolle der Muttersprache beim Vokabellernen", Die neueren Sprachen, XX (1971), 169-177.
Meijer, Tjeerd T., De globaal-bilinguale en de visualiserende procedure voor de betekenisoverdracht: Een vergelijkend methodologisch onderzoek op het gebied van het aanvangsonderwijs Frans (proefschrift, Vrije Universiteit Amsterdam, 1974).


Olsson, S., A learning experiment in English with monolingual and bilingual word-lists, CUME-project (Gothenburg, 1973).

Oskarsson, M. "Monolingual and bilingual vocabulary learning", paper (Budapest, 1974).

Otto, E., Methodik und Didaktik des neusprachlichen Unterrichts (Bielefeld, 1925).


Saltwürk, Ernst V., "Leitmotiv des Unterrichts der neueren Fremdsprachen", in Lehrproben und Lehrgänge, Heft 19 (Halle/S, 1898).


The Politicization of Repentance

Marvin H. Folsom
Brigham Young University

I would like to begin by relating an incident recorded in one of the letters of Augustine. It occurred in the North African branch of Oea (now Tripoli), when the bishop introduced the new edition of the bible (by Jerome) for use in the church. As a verse in Jona was read aloud, in which Jerome deviated from the traditional (Old Latin? or Greek Septuaqint) text, a great tumult arose among the audience. The Greeks said that the reading was incorrect. The bishop asked the opinion of some Jewish townspeople, presumably priests. Either out of ignorance or spite, they said that the Hebrew text was best rendered by the older version. For the sake of peace in the church, the bishop felt it necessary to go back to using the old edition. This example shows how sensitive the ordinary, faithful reader of the scriptures is to changes in the traditional text. I don't know how extensive the demonstration among the audience was but it forced the bishop to change his decision and reinstate the traditional text. (Reported in Ulrich Köpf, "Hieronymus als Bibelübersetzer," Eine Bibel—viele Übersetzungen, ed. Siegfried Meurer, Stuttgart, 1978, p. 76.) We see also how the reactions of the congregation can have an influence on the nature of a religious text.

Now to an example closer to our own experience in English. In 1525, in the first New Testament in English to be translated from the original Greek texts, Tyndale used the words repentance and repent where earlier translations had used penance and do penance. (Others of the controversial words which triggered the heated discussion were congregation for older church, love for older charity, elder for older priest and favour for older grace.)

Tyndale was a superior scholar of Greek but he had controversial theological opinions. He published his views in polemical tracts and also in tendentious glosses and notes. Sir Thomas More criticized Tyndale for this and many other "errors" in a dialogue which Tyndale answered as follows:

And in like manner, by this word penance they make the people understand holy deeds of their enjoining; with which they must make satisfaction unto God-ward for their sins: when all the scripture preacheth that Christ hath made full satisfaction for our sins to God-ward; and we must now be thankful to God again, and kill the lusts of our flesh with holy works of God's enjoining. And I am bound to take patiently all that God layeth on my back; and, if I have hurt my neighbor, to shrive 'confess' myself unto him, and to make him amends, if I have wherewith; or if not, then to ask him forgiveness, and he is bound to forgive me. And as for their penance, the scripture knoweth not of (it). The Greek hath Metanoia and Metanoite, repentance and repent; or forethinking and forethink.
As we say in English, 'It forethinketh me, or I forethink;' and 'I repent, or it repenteth me;' and 'I am sorry that I did it.' So now the scripture saith, 'Repent, or let it forethink you; and come and believe in the gospel, or glad tidings, that is brought to you in Christ, and so shall all be forgiven you; and henceforth live a new life.' And it shall follow, if I repent in the heart, that I shall do no more so, willingly and of purpose. And if I believed the gospel, what God had done for me in Christ, I should surely love him again, and of love prepare myself unto his commandments. These things to be even so, M. More knoweth well enough: for he understandeth the Greek, and he knew them long ere I. But so blind is covetousness and drunken desire of honour. (Tyndale's Answer to Sir Thomas More's Dialog, Cambridge, 1850, pp. 22-23)

Tyndale's attacks on the church and the clergy were among the reasons Henry VIII condemned his New Testament (1530) and banned it in England. It had been printed in Germany but had secretly found its way to England. Tyndale fled to the continent, but was finally arrested, imprisoned, condemned to die, strangled as a heretic and his body was consumed to ashes in September 1536.

Permit me to backtrack just slightly to following the development in German. When Luther first translated the New Testament 1522 (hatth 3:2, 4:17, Mark 1:15), he used bessert euch 'better yourselves, improve yourselves, amend your lives,' but five years later in the edition of 1527 he had already decided to change it tut Buße. Zwingli (1523) and the Züricher Bible (1531ff) both use the Alemannic form of the same word Luther used in the beginning: Besserend euch. The majority of the translations in German from then until now have for the most part followed Luther's choice of Buße but there have been several exceptions. Early translators such as Felbinger 1560, Zinzendorf 1744, Bahrdt 1777 used alternatives for Buße tin such as 'change your ways, change your way of thinking, turn around, turn away from, begin anew,' etc. (ändert euch, ändert euren Sins, kehrt um, stellt euch um, ihr müßt völlig neu anfangen), but I am not acquainted with the political ramifications, although I suspect there were some. I believe the dominant position of the Luther text precluded any large scale or organized effort for revision. The greatest number and variety of these alternatives begin to appear about 1930.

In his 95 theses, Luther quotes Matthew 4:17 Tut Buße, das Himmelreich ist nahe herbeigekommen (KJ Repent for the kingdom of heaven is at hand) He might just as well have chosen the parallel verse from Mark 1:15 Die Zeit ist erfüllt, und das Reich Gottes ist herbeigekommen. (KJ The time is fulfilled, and the kingdom of God is at hand.) These verses were interpreted both by those in the centuries right after Christ as well as by those at the time of Luther to mean that the second coming with its last judgement was very close. "Judgement is coming. Don't close your eyes to this fact. Live as though the second coming were tomorrow. Or more ominously: "Judgement is near, repent of your sins and do better."
Let us now take a brief look at some of the changes that have been associated with the development of the theological meaning of repentance. According to the Theologische Realenzyklopädie the central concept of Christian theology and piety is Buße 'repentance,' yet, inspite of its great importance, its history has been one of distortion. (Vol. VII, p. 473)

Latin, which replaced Greek as the language of the (Roman) church had paenitentiam agite 'show repentance, penitence.' Here the notion in the Greek of 'change your thinking' has been left behind and a new meaning arises to take its place. In colloquial Latin, there was an inaccurately pronounced form poenitentia along side the etymologically correct form paenitentia. In the minds and language of the linguistically unsophisticated, the careless form with pe was easily associated with the phonetically similar word poena 'punishment' (in our penitentiary). According to Franz josef Michels (Martin Luther, Inter Nationes: Bonn, 1983, p. 20ff), the unsophisticated Germanic tribes, were familiar with punishments but not with penances or repentance. If one had killed a man for example, he was guilty of murder but was not required to pay with his own life, instead he could pay Wergeld 'manbote, man money, compensation to an overlord for the loss a man.' (The German Buße and English to boot 'to the good' and poet of bale 'relief of bale' are both related to OE manbote.) In Germanic law, the emphasis was on restitution and not necessarily in a literal sense, that is, money could be substituted for the life of a man. A Frank of the 6th or 7th century might well have interpreted the call to repentance as: Judgement day is soon. Examine your life and make restitution for whatever things you have done wrong toward man and God. By the 10th century, the substitution had become quantified. In 970, for example, King Edgar of England decided that an earl who had been condemned to seven years of austere fasting could reduce the sanction to three days if he could persuade 12 and then 120 of his clansmen to fast three days for him. This contains in a nutshell the basic model for indulgences. A punishment of the church can be transmuted just like any other. All that remains is the assignement of amounts to each violation and we have created the Bußgeldkatalog 'fixed penalty code of traffic infractions.'

Indulgences as we understand them were granted for the first time around the middle of the 11th century in southern France, which according to Michels was more than just coincidental. The early Christian church anticipated an imminent second coming. Since it had not happened, several generations had lost faith in the prophecies about the last days. Not until the beginning of the 10th century did they begin to take them more seriously, based in part on the interpretation of a verse from which the end of the world could be calculated to be the year 1000 or 1001. In anticipation of this judgment a first great revival calling the believers to repentance ensued and it produced monastaries, pilgrimages, good works, etc. Delay forced a reinterpretation of the year of the advent as 1300 (1000 + 300) and produced a jubilee year and another wave of intensified religious feeling.

As time went on a distinction gradually came to be made between
'first repentance' as a one-time conversion and rebirth and a 'second repentance' as a lifelong process of daily repentance, a distinction which the reformers later did away with.

One other piece of information I would like to relate in connection with this discussion is found in the introduction to the 1891 edition of the Bible published by Brockhaus (Elberfeld), which includes some excerpts from the preface to the earlier edition of 1855. It informs the reader as to the reasons for certain choices of words in the translation, for instance, why der Christus (with the article) was used to designate the office of the 'Anointed One,' whereas Christus (without the article) was used for the personal name. In the discussion of the word chosen for 'repentance,' the editor expresses the futility of searching for a more appropriate equivalent for Greek metanoia (KJ repentance).

Even though we have used the word Buße, we are not satisfied with it because it has more of an external character and signifies outward works (Werketun). The word Eekehrung 'conversion' was suggested as an appropriate alternative; yet, though many translators have used it, we have not, because it does not convey the meaning of metanoia. Metanoia is the moral judgment of the soul concerning everything in one's past as well as everything the soul is in the flesh before God. Others have preferred Sinnesänderung 'change of heart or mind' and really have come a little closer to the true meaning. However, because the judgement of the soul with respect to the past was lacking in this word, we felt it necessary to retain the word Buße.

Here we see the conscientious translator struggling with a term which in the course of time has assumed inappropriate connotations, yet he does not feel the price of selecting some other term is worth it and does not want to make the break with tradition, since the new word itself would likewise not be adequate to express such an important theological concept. We do not expect the single word baptism to carry all the meaning. When necessary we modify a concept by adding a qualification. We use the same word for baptism as most other Christian churches, yet our concept is very different from many of them. When a distinction is necessary, we say 'baptism by immersion' to avoid any misunderstanding. We could make a similar qualification for repentance where necessary without replacing the basic terminology.

A radical change in terminology was instituted by the Swiss Reformed theologian Karl Barth (1886-1968), opponent of the National-Socialist regime, father of the Bekennende Kirche who pondered the question of how people become Christians. He replaced the traditional term Buße with Umkehr 'turn around' in order to avoid the notion that we are dealing with a single or several individual events in the sense of the Catholic sacrament of penance or of the pietistic experience of conversion and penitent struggles. The meaning most often associated with words with the root umkehr- is 'reverse, invert or make opposite.' Although it is supposed to effect the individual in all
aspects of life, the word itself focuses on the single event of turning around and reduces the emphasis on constant repentance throughout one's entire life, that is it also does away with any distinction between the first or great repentance and the second or small repentance.

As an example of the usual meaning of the verb umkehren, let me refer to a 19th century pamphlet on language teaching by the noted linguist Wilhelm Vietor. He entitled it Der Sprachunterricht muß umkehren, not 'language instruction must repent' but with the older meaning 'language instruction must reverse its course, do something entirely different'.

This radical change in terminology is reflected in the modernizing translation known as the Einheitsübersetzung (vorläufige Endfassung 1971, Endfassung 1980) inaugurated by the catholic bishops of Germany, Austria, Switzerland, Luxemburg, Lüttich, Bozen-Brixen and ultimately approved by the Council of the Protestant Church in Germany and the Protestant Bible Association of the Federal Republic of Germany. This edition systematically substitutes Umkehr for Buße to refer to the central Christian concept of repentance. In the wake of this sweeping change, the Luther translation of 1975 was also modernized in an attempt to make it more understandable for the modern reader. Besides these and other changes in the vocabulary of this latest revision, it removed the most striking stylistic features of Luther's Bible (the Saxon genitive 'in my father's house' (in meines Vaters Haus, bis an der Welt Ends) and the positioning of the verb in mid-sentence instead of at the end (Matth 25:31): when the Son of man shall come in his glory, and all the holy angels with him, then shall he sit upon the throne of his glory (Wenn aber der Menschensohn kommen wird in seiner Herrlichkeit und alle Engel mit ihm, dann wird er sitzen auf dem Thron der Herrlichkeit). This 5th revision in a century set off a wave of controversy in the popular and scholarly press, especially among those who were familiar with Luther's style and did not want to give up there cherished religious vocabulary. The publication of this latest revision was discussed in an article by Renate Schostack in the Frankfurter Allgemeine Zeitung of 31 October 1984). Its German title has echoes of the present discussion: "Rückkehr zum Reformator" 'Return to the Reformer.' Its English title ignores the essence of the article "New Bible gets unheralded welcome."

As a result of all the furor, the Council of Protestant Churches in (the Federal Republic of) Germany commissioned a revision of the 1975 version of the Luther text, which in essence turned out to be a work of restoration, a return to the reformer's original translation. They changed Luther's text only where they felt it could not be understood by modern readers. In the summer of 1984, after three years' work by German theologians and language experts, it came off the presses.

One of the words that was replaced in the 1984 revision was the controversial word Umkehr. Luther's choice (Buße) was restored except in two passages which have seem to have more general meaning
One of the central considerations mentioned in the report was the fact that Protestant hymns and the Book of Common Prayer used Luther's wording almost exclusively and a radical departure would create a chasm between the two segments of religious life. The Synod of the Protestant Church in the Federal Republic of Germany to be held in Lübeck will likely recommend the use of the restored text of 1984. Church officials in the German Democratic Republic have already done so. The prospect is that the restored text of 1984 will be approved for official use by all German speaking protestants.

An inquiry to a member of the commission on revision into possible reasons for the return to Luther's for 'repentence.' brought the following response:

With regard to the change of Umkehr to Buße, we decided to restore Luther's concept in its biblical meaning even if or precisely because of the fact that for the most part it is used only in compounds like Bußgeld in modern German. Letter from Dieter Gutzen, 21 January 1985)

I believe he is saying that because the word Buße is used in the negative religious sense of penance only in compound words, it is free to be used in its positive biblical and religious sense as envisioned by Luther. He goes on to point out their general attitude of making sure that the revision maintained the style of Luther as much as possible. The restoration of the term Buße definitely belongs in this category. We can conclude that the strongest motive in the restoration of the term, is to be found in the dissatisfaction with the modernizing approach of the last revision and in an attempt to rehabilitate the image and style of Luther.

Let me turn now to developments outside the realm of the text itself. Beginning in 1968, progressive elements in the Catholic Church have repeatedly tried to bring about some reforms. In 1960, several groups decided to organize and hold their own Katholikentag at the same time as the regularly established general meeting of the Catholic Church in West Germany. Among them were three well-known personalities: 1) Hans Küng, reform minded theologian from Tübingen who was refused permission to teach because of his doubts about the infallibility of the pope, 2) Johann Baptist Metz, professor of theology at Münster who was refused the chair of ecumenical theology in Munich because of his leftist political leanings, 3) Norbert Greinacher, professor of pastoral theology in Tübingen who formed the committee for Christian rights in the church after the Küng affair. These reformers felt that the church was just going around in circles and not addressing the needs of the young people, especially with respect to giving the sacrament to homosexuals and divorced people, the marriage of priests, and the peace movement. In a demonstraton at the meeting, the protesters had on their placards the picture of a nun in religious habit striking a rocket with her umbrella. The motto around the edge reads: Kehrt um. Entrüstet euch. 'Turn around, be indignant,' or 'Change your course. Be full of indignation.' Those of you who know German realize, of course, that their is a double meaning
for each part of the slogan. The new, religious meaning for the first part is 'repent,' conjuring up a revivalist atmosphere and the new meaning for the second part is 'disarm yourselves.' Actually, it is a very clever, double pun for the peace movement against the nuclear arms race: "Repent, Disarm." A second version of the motto in another protest movement of the grass roots church reads: Kehrt um--widerstellt den Herren der Welt. 'turn around/repent, resist the lords of this world.'

The motto for the peace weak in Heidelberg in October 1984 was:
Noch ist Zeit zur Umkehr: Weg mit Rüstungswahn und Militarisierung! Nein zu Pershing II und Cruise Missle! 'There is still time to turn back/repent. Down with the insanity of armaments and militarization. No to Pershing II and Cruise Missiles.' Once again the atmosphere of final judgement is conjured up from the religious heritage to frighten people, in this case, to oppose military and political activity.

The most recent issue of Der Spiegel (4 February 1985, Nr. 6/1985) contains part three of an article by Oskar Fontaine ("Der andere Fortschritt"), in which he discusses the protest vote and the prospects of the Social Democrats and the Greens ('Environmentalists') forming a working coalition. In the article there is a picture of a person at a demonstration holding up a neckerchief. There is the outline of a church and a hand. It is indistinct and I can't tell what the relationship between the church and the hand is. Across the top of the neckerchief is the motto: Umkehr zum Leben, 'Return to life' or 'repent and turn to a better life.' Underneath the church is the phrase: Die Zeit ist da für ein Nein ohne jedes Ja zu Massenvernichtungswaffen. 'It's time to say no without any yes (at all) to weapons of mass destruction.' The caption over the picture reads: Das Kirchenvolk ist dabei, die Bibel beim Wort zu nehmen. 'Lay members are taking the bible literally. Beneath the caption we read: Der Glaube an den Schöpfer verbietet es, die Natur zu plündern und auszunutzen. 'Faith in a creator forbids the plundering and exploitation of nature.'

Clearly kehrt um has become associated in the press and in demonstrations within the Catholic Church and elsewhere with reform, resistance and rebellion. In the popular press, it no longer has only the older theological meaning 'repent' but often it has in addition the older meaning 'turn around, turn back' and the new meaning 'rebell.' It now is most likely to mean 'Change your political course,' and not 'sorrow and change because of moral shortcomings.'

In our own sphere, the EU came into the hands of the church translators at a time when the standard works had just been retranslated (not revised) from the English and was at the press being typeset. Thus the new translations of the standard works in German (1980) in a manner similar to the revision of the Luther text swept aside 130 years of tradition and replaced all examples of Buße with Umkehr, even going so far as creating the unique phrase Umkehr üben on the model of the older Buße tun instead of using the corresponding verb umkehren. At the same time, the Luther translation of the bible
was abandoned as the official bible of the church and the Einheitsübersetzung was required for all official church uses. There were a few voices that objected to the wholesale discarding of the language of Luther and wrote letters and articles but for the most part these radical changes were accepted quietly. The ones most effected by the decision were those who were very close to the traditional text of Luther and were prevented from using it in Sunday school classes and other official church contexts.

In a letter of 23 January 1981, leaders of the church in German speaking areas were enjoined to use in all official contexts the new Einheitsübersetzung and the new standard works which had adopted the vocabulary of the EU. They were also asked to urge members to do the same. Leaders and members interpreted this so narrowly that in some cases members were not allowed to quote the scriptures in Sunday school according to the traditional Luther text. Taken literally, this single directive would have eradicated from the religious vocabulary of German speaking saints traditional terms used by Protestants since Luther and in the LDS church since the first translation of the Book of Mormon into German in 1852, words like Buße 'repentence,' Heiland 'savior,' Nichtjuden 'gentiles' and many others. Over the next few years, there was enough resistance to the exclusive use of the new text and the stigmatizing of the traditional text that a new directive (13 January 1984) was issued to clarify the use of the approved scriptures. It explains that the official church Bible (EU) will be used for bible quotations in all publications. Members are free, however, to use any Bible version they choose. It notes that it may even be helpful to use several versions at once so as to arrive at a better understanding of a given passage. There is also a caution about the fact that the introductions and notes do not necessarily reflect church doctrine. The fact is that the introductions in the Einheitsübersetzung deny the unity of Isaiah (it has three subheadings: Protojesaja 1:1-39:8, Deuterojesaja 40:1-55:13 and Tritojesaja 56:1-66:24), dispute the Pauline authorship of certain letters and otherwise include the results of modern textual and higher criticism, including the interspersing of the books of the Apocrypha throughout the Old Testament. A letter by the translation department in Frankfurt to the editor of Dialogue (Vol. 17, No. 4, Winter 1984) about the new German standard works reports that efforts to have an edition published without the Apocrypha or with the Apocrypha in a separate section will depend on ecumenical recognition—surely a lengthy and uncertain process. In the meantime, members are expected to rely on the statement in the 91st Section of the D & C. It does concede that it is unfortunate that the apocryphal books are difficult to identify as such.

When I first heard of the new directive, I thought it would reflect a policy similar to that expressed by the Protestant Church in Germany who have in essence reversed themselves and returned to a slightly modified version of the traditional text of Luther. In my naivete, I thought it would at least allow members to use either the EU or the revised Luther text (of 1984) for general use in the church. When I had a chance to examine the letter in greater detail, it became clear that they had not changed their position at all with
respect to the use of the official text but had merely allowed members to use any other version they might choose as an aid to interpret or clarify the "official" text. Since the basic concepts of the gospel are now clothed in the language of the EU (Umkehr for Buße, etc.) and the new standard works (Erretter for Heiland, die Andern for die Nichtjuden, etc.), a linguistic dichotomy has been created. On the one hand, we have the Catholic Church and the official part of the LDS Church in the West. On the other hand, we have the Protestant Church of the Federal Republic of Germany and of the German Democratic Republic, and the unofficial portion of the members of the LDS Church in the West and the LDS Church in general in the East, all of whom recommend and use the traditional text of Luther. (A colleague who recently returned from Leipzig reports that the sacramental prayers are given there in the traditional form and not in the form in the new standard works.)

The adoption of the EU and wholesale modernization in the standard works gives rise to a paradox. The church has now approved for German what it has refused for decades to do for English. On the one hand, we cling in English to an archaic text not based on the best manuscripts and depol the critical inroads in more modern texts, but on the other hand, we abandon a similar Luther tradition based on the same manuscripts and embrace for German the very things we stubbornly resist in English. I believe this is a real problem raised by the new translation of the standard works in German and the controversy related to it. We must make up our mind. Should we pursue one course in English and a very different course in German? If we should choose to follow for German the course presently followed in English, we would return to the Luther text for official use. If we should choose to follow for English the course recently adopted for German, we would discontinue using the King James version and approve a more modern version or make our own translation.

In my opinion, further revision of the EU will be minimal or limited to ecumenical compromise in the Old Testament and the Apocrypha, and that would mean that the vocabulary in the New Testament that has been adopted in the standard works in German will remain unchanged. The question in the Church for the future revolves around whether the projected revision of the standard works in German, in addition to incorporating the changes of the revised text of the English edition, will revise the basic theological concepts ('repentance' and 'savior') to correspond to the traditional Mormon and Protestant terminology. The alternative is to allow the present dichotomy to continue until time eventually erodes and overcomes it. The conservative stance of the LDS Church in the German Democratic Republic and the recent return to the language of Luther by the Protestant Church in both the Federal Republic of Germany and the German Democratic Republic are strong arguments for us to do the same, that is, to admit that it was premature and precipitous to alter the traditional text so radically and to restore the traditional terminology of Luther in the next revision of the standard works.
FIRST LANGUAGE CULTURE IN SECOND LANGUAGE LEARNING AND TESTING: EVIDENCE ON FIELD SENSITIVITY

Lynne Hansen
Brigham Young University--Hawaii Campus

In comparing first and second language acquisition, one obvious difference lies in the amount of individual variation in language learning success. Despite considerable variation across cultures in the structuring of first language learning experiences (Ochs 1982), children around the world achieve quite comparable levels of control over their first language in about the same amount of time. This relative uniformity in rate of acquisition has been used as evidence for the argument that first language acquisition is under the control of universal learning mechanisms which are impervious to cultural influences.

In second language acquisition, on the other hand, we find substantial amounts of variation in how quickly and how well a language is learned, even among young children (Fillmore 1976). Clearly individuals vary greatly in their ability to learn additional languages after the first one. In the development of these individual abilities, however, the influences exerted by the cultural milieu are not well understood. Is it the case that superimposed over the individual differences that are found in second language learning strategies and achievement there are cultural group differences as well?

Among language educators in multicultural settings, strong impressions of group differences in the ways in which students go about learning second languages and in their levels of success in accomplishing the task are frequently reported (James 1983). Upon beginning teaching duties in an English as a second language program in Hawaii, for example, the writer was told by teachers there that the Polynesian students would be eager to speak up in class discussions but would not benefit much from the explicit teaching of grammar while the Asians, on the other hand, would be reticent about participating in class but would excel in reading and the learning of grammar rules.
How much research evidence is there for the existence of such group differences in second language acquisition? Hardly any. There have, in fact, been very few investigations that have made explicit cultural comparisons such as we are attempting to do, in a three year longitudinal study of ESL acquisition in the English Language Institute at BYU--Hawaii Campus. In this study we are looking at a number of variables, cognitive, social and affective, which we think might vary between cultures and that may affect language learning outcomes. We shall examine in this paper one of these cognitive variables, field sensitivity, and look at our findings with regard to three questions: 1) Are there significant differences between the cultural groups in the English Language Institute in their levels of field sensitivity? 2) What is the relationship between field sensitivity and second language achievement? Does field sensitivity affect performance on language tests which require inferencing, such as the cloze test?

Field dependence/independence (FD/I) is defined as "the extent to which a person perceives part of the field as discrete from the surrounding field as a whole, rather than embedded, or...the extent to which a person perceives analytically" (Witkin et al. 1977).

A FI person may approach problem solving situations analytically while the FD person may approach them in a more global way. In the area of intellectual problem solving it is claimed that a highly FI person is able to detect patterns and sub-patterns, while a FD person tends to get lost in the totality of the stimulus. Further, it is claimed that FI persons are better at problem solving, restructuring data and critically evaluating data. A FD person, in contrast, is capable of seeing the total picture in a given situation.

In terms of learning strategies, FI persons use their own abilities to structure information while FD learners rely on cues from others, and allow others to structure information for them. FD students have been found to learn and remember material with a social content better than material without social content, and appear more able to pick up clues from teachers than FI students are (Witkin et al. 1977).

In the affective domain, a FD person tends to be more empathetic, more socially attuned and more sensitive to subtle social clues than FI persons. In contrast, a highly FI person may be seen as impersonal, individualistic, and distant. Because FD persons rely more on external referents, they are more likely to be influenced by the opinions of others than are FI persons (Witkin and Goodenough 1981).
There is a gradual increase in field independence through childhood, but from the mid-teens through adulthood an individual's field sensitivity is relatively stable. The extent of development with maturity of a field independent cognitive style is related to the type of home and society in which the child is reared. For example, Mexican-American children from traditional Mexican homes have been found to be more FD than Mexican-Americans from non-traditional homes (Ramirez 1973; Ramirez, Castaneda and Herold 1974). It has also been found that high school students in Hawaii are significantly more FI than students in Fiji, Samoa, Tahiti and Tonga (Hansen 1984). From these and other comparisons of field sensitivity in different cultures (reported in Witkin and Berry 1975) it appears that cultures with more elaborate social structures, and pressures to conform tend to have children who are more FD; while democratic industrialized societies with more relaxed rearing practices tend to produce more FI.

FI has been found to be positively related to second language acquisition as measured by a number of different tests: a general French achievement test (Tucker, Hamayan and Genesee 1976), a French imitation test (Naiman, Frolich and Stern 1975), an ESL dictation, a grammar test, and the TOEFL (Chapelle and Roberts 1984), a sentence disambiguation test (Seliger 1977), the Michigan Test of English Language Proficiency and English cloze and 'noise' tests (Hansen 1984), overall "better performance in a Spanish course" as measured by five test types (Hansen and Stansfield 1981).

The nature of the measures on which a FI student performs significantly better than his FD counterpart is of interest. By definition, a FI person should be good at analytic language skills such as those that must be used for good performance on many language tests. The FI person's ability to approach problem solving analytically, i.e., find patterns, organize information to make generalizations and impose structure on a situation should facilitate performance in many of the tasks that students are asked to perform on language tests.

Although FI may be a trait that is beneficial to students for performance on many different kinds of L2 measures, it has recently been proposed that performance on the cloze test is affected to a greater extent by a test-takers field sensitivity than is performance on other L2 tests. In a study of first-year Spanish students at the University of Colorado, it was found that the significance of correlations between FI and cloze test scores...
was stronger than those between FI and other language measures. After verbal aptitude was partialed out, in fact, the relationship between FI and the cloze test remained significant while the relationship between FI and other language measures fell below the level of significance prompting the researchers (Stansfield and Hansen 1983) to conclude that the cognitive restructuring abilities of FI individuals are conducive to success on a cloze reconstruction task.

Another study which examined the relationships between field sensitivity and language measures included students from several cultural backgrounds (Hansen 1984). The findings largely validate the Stansfield and Hansen (1983) hypothesis of a minor cognitive style bias in the cloze test, but sizeable differences between academic ability groups and cultural groups in the relationships found between FD/I and the language measures led the researcher to suggest a cautionary approach in generalizing findings from the investigation of any single ethnic or ability group.

METHOD

Subjects

Subjects for the study are the foreign students who have come from Asian and Pacific cultures to study English in the English Language Institute at Brigham Young University--Hawaii Campus during the past two years. Their cultural backgrounds and group N sizes are as follows: 20 Samoan, 28 Tongan, 15 Micronesian, 53 Hong Kong Chinese, 26 Japanese, 28 Korean, 16 Filipino, 23 other Asian (including PRC Chinese, ROC Chinese, Thai, Vietnamese and Indonesian).

Testing

Upon entering the English Language Institute students are given three tests: the Michigan Test of English Language Proficiency, the Michigan Test of Aural Comprehension and an essay which is graded by two readers on a ten point scale. After each semester in the ELI program these measures are administered again, together with three additional tests: a dictation, a speaking test and a cloze passage. The dictation is read three times, the
second with pauses for writing. The speaking test consists of subtests for conversational ability, two-and-a-half minute prepared speech and reading pronunciation. These three subtests are administered on three consecutive days and each rated by a different pair of examiners. The cloze passage has every seventh word deleted and is scored by the acceptable synonym method.

The FD/I measure was the Group Embedded Figures Test (GEFT) (Witkin et al. 1971). This is a standard measure of field sensitivity that has been used in the studies of L2 acquisition reported above. The GEFT requires the subject to outline a simple geometric shape embedded within a complex design. One must separate the relevant information from the contextual visual field in order to find the correct shape.

RESULTS AND DISCUSSION

The mean scores on the GEFT for the groups tested are given in Table 1. Notice that the Asians (with the exceptions of the Filipinos) score higher on this test (indicating a higher level of FI) than do the South Pacific students. Table 2 which separates the male and female scores for each group reveals interesting sex differences as well with the males in most of the cultures showing a higher level of FI than the females. In order to determine whether the these cultural and sex differences in performance on

<table>
<thead>
<tr>
<th>Culture</th>
<th>Mean</th>
<th>S.D.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tongan</td>
<td>7.6</td>
<td>5.0</td>
<td>28</td>
</tr>
<tr>
<td>Samoan</td>
<td>7.5</td>
<td>5.8</td>
<td>20</td>
</tr>
<tr>
<td>Micronesian</td>
<td>8.5</td>
<td>4.4</td>
<td>15</td>
</tr>
<tr>
<td>Filipino</td>
<td>5.5</td>
<td>4.8</td>
<td>16</td>
</tr>
<tr>
<td>Other Asian</td>
<td>10.6</td>
<td>4.4</td>
<td>23</td>
</tr>
<tr>
<td>Korean</td>
<td>10.8</td>
<td>5.2</td>
<td>28</td>
</tr>
<tr>
<td>Hong Kong Chinese</td>
<td>12.2</td>
<td>4.4</td>
<td>53</td>
</tr>
<tr>
<td>Japanese</td>
<td>14.0</td>
<td>4.3</td>
<td>26</td>
</tr>
</tbody>
</table>
the GEFT are statistically significant, a two-way ANOVA for South Pacific and Asian groups was run. The results, presented in Table 3, show that the F value for culture is significant at the .000 level and the F value for sex at the .01 level, indicating highly significant differences in FD/I between cultural groups and between males and females. The Asians tend to be more FI than the South Pacific islanders, and the males more than the females.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>558.8</td>
<td>2</td>
<td>279.4</td>
<td>10.7**</td>
</tr>
<tr>
<td>Sex</td>
<td>244.0</td>
<td>1</td>
<td>244.0</td>
<td>9.4*</td>
</tr>
<tr>
<td>CS</td>
<td>72.2</td>
<td>2</td>
<td>36.1</td>
<td>1.4</td>
</tr>
</tbody>
</table>

*p < .01
**p < .001
The second research question in the present study concerns the relationship between field sensitivity and the second language achievement of the ESL learners at the English Language Institute. To examine this question, correlations are made between the students' ESL test scores and their scores on the measure of FI, the GEFT. In table 4 we see the resulting correlation coefficients from the application of the Pearson product moment procedure to our data. Here the GEFT is related to three tests which were administered three times to the subjects in the present study: first at the beginning of their first semester at the English Language Institute, second after the first semester, and third after the second semester of enrollment in the ELI program. Like previous researchers who have correlated GEFT scores with second language measures we find a pattern of low positive correlations. Since a higher score on the GEFT indicates a higher degree of FI, these results indicate that greater field sensitivity is associated with a higher level of second language achievement as measured by all of the language tests.

Table 4
GEFT Correlations with Language Measures

<table>
<thead>
<tr>
<th>MTELP 1</th>
<th>MTELP 2</th>
<th>MTELP 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>.12*</td>
<td>.15*</td>
<td>.16*</td>
</tr>
<tr>
<td>N-201</td>
<td>N-199</td>
<td>N-129</td>
</tr>
<tr>
<td>MTAC 1</td>
<td>MTAC 2</td>
<td>MTAC 3</td>
</tr>
<tr>
<td>.08</td>
<td>.18*</td>
<td>.11</td>
</tr>
<tr>
<td>N-190</td>
<td>N-139</td>
<td>N-54</td>
</tr>
<tr>
<td>Essay 1</td>
<td>Essay 2</td>
<td>Essay 3</td>
</tr>
<tr>
<td>.19*</td>
<td>.21*</td>
<td>.30*</td>
</tr>
<tr>
<td>N-181</td>
<td>N-184</td>
<td>N-121</td>
</tr>
</tbody>
</table>

*p< .05

Notice here the variation in the strength of the relationships across the three administrations of the same tests to the same subjects. While the MTAC relationship with the GEFT is not statistically significant at testing periods 1 and 3, at period 2 it is significant. The relationship of the Essay
with the GEFT varies from .19 at the first testing period to .30 at the third. In light of the importance attributed in recent studies to variations of comparable magnitude across different language tests administered to the same subjects (Stansfield and Hansen 1983), these data suggest caution in the interpretation of such variation. Not only is a cautious approach called for in the generalization of findings from a single ethnic or ability group, as was urged in a previous research report which demonstrated sizable cultural and ability group differences in the relationship between language test scores and GEFT score (Hansen 1984), but also in the making of generalizations on the basis of a single administration of language tests.

<table>
<thead>
<tr>
<th>Time</th>
<th>MTELP</th>
<th>MTAC</th>
<th>Essay</th>
<th>Dictation</th>
<th>Speaking</th>
<th>Cloze</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.01</td>
<td>.00</td>
<td>.03</td>
<td>.07</td>
<td>16</td>
<td>.29</td>
</tr>
<tr>
<td>N=51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.26</td>
<td>.44</td>
<td>18</td>
<td>.06</td>
<td>16</td>
<td>.52</td>
</tr>
<tr>
<td>N=47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.19</td>
<td>-.20</td>
<td>.15</td>
<td>.02</td>
<td>-.15</td>
<td>.53</td>
</tr>
<tr>
<td>N=29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.30</td>
<td>-.26</td>
<td>.23</td>
<td>-.07</td>
<td>.05</td>
<td>.04</td>
</tr>
<tr>
<td>N=28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.36</td>
<td>.23</td>
<td>-.39</td>
<td>-.05</td>
<td>.33</td>
<td>.23</td>
</tr>
<tr>
<td>N=31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5 shows Pearson product moment correlations of the GEFT with the six English language measures which are administered at the end of each semester in the ELI program as part of the final examination battery. The subjects for each test time are those ELI students for whom all of the language test scores and the GEFT score are available. At each testing period the same students took the six language tests, but different people are involved across testing periods though each generally represents the cultural composition of the English Language Institute.
In these data we see a pattern of low, primarily positive correlations between language test scores and GEFT score. Again the variation is considerable, however, in the relative correlation strengths for the tests across the five examination periods, although generally a higher level of FI is associated with higher performance levels on the English tests. It may be that the inconsistency in the results over testing periods is due, at least in part, to the diminished validity of small group correlations.

Evidence on our question concerning the relationship between field sensitivity and cloze test performance may be found through a comparison of cloze test correlations with those for the other ESL measures in Table 5. Overall we see that the cloze correlation coefficients do tend to be larger than those for the other tests (see Times 1, 2, and 3, but notice also Times 4 and 5 where a stronger correlation for the cloze test does not obtain), thus lending support to the Stansfield and Hansen (1983) hypothesis of a cognitive style bias in the cloze procedure which gives FI people an advantage unrelated to their language achievement. Such a minor bias does not show up consistently, however, at least in small group data, for all culture and ability groups (Hansen 1984), nor even for the same population when different cloze passages are used over testing periods.

Clearly the relationship between field sensitivity and language test performance is not a simple one, and these results suggest that probably no more than a superficial understanding of its complexity can be obtained through the analysis of group test data with a simple correlational research design. The group mean scores may well obscure significant individual differences in the relationship between FD/I and performance on language tests. Future research must explore the possibility that the usefulness of a FI cognitive style in finding test solutions may be greatly enhanced or mitigated by the presence or absence of other variables which may interact with FD/I differently in different testing situations.

In sum, the results of this study do indicate a slight advantage for FI people in the classroom learning of a second language as measured by a variety of language achievement tests. Among these tests, the cloze tends to correlate most highly with FI (though not consistently across all examination periods), supporting the findings of previous research of an apparent minor cognitive style bias in this testing procedure in favor of FI people. The finding, in addition, of highly significant differences in field sensitivity across the cultural (and sex) groups examined, provides evidence on one way in which culture affects the learning of second languages.
A cautionary approach is urged, however, in the application of these findings to educational practice. Although group differences in field sensitivity (or other learner variables) in multicultural classroom settings may be substantial, they are never absolute but are, rather, expressed as group tendencies towards one end of a continuum or the other. Individual differences within a given culture in a characteristic or behavior are likely to span the entire continuum, as is the case with the FD/I of the subjects in the present study. No matter how much culture may be found to influence language learning, these influences will be expressed in unique ways in individual learners.

REFERENCES


Vowel-zero Alternations in Russian Non-derived Words

David K. Hart, Assistant Professor of Russian
Brigham Young University

One problem facing linguists dealing with Russian morphology is how to treat the vowel-zero (V-Ø) alternations found in various case forms of non-derived words and in various partners of derivationally related words. These alternations have been traced to historical sound changes, but due to subsequent sound changes the situation in modern Russian has become fairly complex. Researchers attempting to provide a synchronic account of word-formation have had to resort to abstract underlying representations when dealing with words that exhibit the V-Ø alternation. In this paper I briefly review the historical events that resulted in the V-Ø alternations. I point out weaknesses in the solutions that have been proposed to account for these alternations and propose a different solution, one which makes use of the synchronic manifestations of the vowels in question.

Words that exhibit identical lexical meaning, which differ only in grammatical endings are normally considered to be grammatical variants of a single word, or more precisely, variants of a single stem. Due to historical sound changes, the form of a stem may differ from one grammatical variant to another. For example in the English word hoof, the voiceless continuant [f] alternates with its voiced counterpart [v] in the plural form hooves. It has been shown that given adequate phonological rules these words can be derived from a single underlying stem. This approach accounts for the identity in meaning shared by the parts of these words not involved in inflection, since both words are derived from a single stem.

In Russian historical sound changes have produced many such alternations within stems. One such sound change that occurred in Russian was the development of lax vowels. Old Russian contained two lax vowels, called jers, a high front vowel and a high back vowel, transcribed here as [i] and [u]. These two vowels were in complementary distribution: back jers occurred only after nonpalatalized consonants and front jers only after palatalized consonants. In time the jers that were stressed became fully vocalized, i.e. i and u lowered to e and o, respectively. Jers that were before a syllable containing another jer were also vocalized. Most other jers were dropped (Kantor:396-397). These historical events can be schematized:

(1) Old Russian    sūnu    sūna    d'in'i    d'in'a
    u,i→o,e    sonu    --    d'en'i    --
    u,i→j    sna    d'en'a    d'n'a
    'sleep'    'sleep'    'day'    'day'
    Nom/Sg    Gen/Sg    Nom/Sg    Gen/Sg

It can be seen from this example that one phonological alternation in Russian that resulted from the development of the jers is that of a vowel (either o or e) with zero. Only those stems that once had a jer exhibit this alternation: Nom/Sg STOL 'table' - Gen/Sg STOLA. Here o does not alternate with zero; historically it was an o, not a jer.
A subsequent sound change in Russian was the shift of Old Russian e to o when stressed and before a nonpalatalized consonant: OR [m'eda] 'of honey' - Modern Russian [m'oda]. Since this shift occurred after the change of i to e (vocalization of jers) it affected the new e also:

\[(\text{2})\]  
<table>
<thead>
<tr>
<th>Old Russian</th>
<th>zig</th>
<th>zigla</th>
<th>zigt'i</th>
</tr>
</thead>
<tbody>
<tr>
<td>palatalization</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>ū, ţ → o,e</td>
<td>zig</td>
<td>zigla</td>
<td>zigt'i</td>
</tr>
<tr>
<td>ū, ţ → ø</td>
<td>zigla</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>e → o</td>
<td>źog</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

'he burned' 'she burned' 'to burn'  

Due to these sound changes, modern Russian exhibits the following alternations:

\[(\text{3})\]  
| o (derived from ū) | -- | Ø | as in SON SNA |
| e (derived from ţ) | -- | Ø | as in DEN' DN'A |
| o (derived from ţ) | -- | Ø | as in źOG źGLA |
| e (derived from ţ) | -- | o | as in źOG źEC |

These alternations cannot be accounted for in the modern language by means of a simple phonological rule of the type o→Ø when followed by a vowel, since words with original o do not truncate this vowel. Similarly, the rule e→o before a hard consonant is inadequate because there are other sources for e in Russian which did not undergo this change (cf. SNEG 'snow' from OR SNÉG).

In order to account for the vowel-zero and e-o alternations in Russian Worth (1968:110) proposed a 'complex set of morphophonemic entities'. For vowels alternating with Ø, Worth proposed the existence of a morphophoneme which he transcribed as ź. Thus, for Worth, d'n' is the stem for all inflected forms of den'. He proposed other abstract symbols, namely ź and Ø, to account for similar alternations in derivationally related words; and he uses the symbol ź in derivation as well. Worth sees the underlying stem of okno 'window' as ok#n-. A 'flexional rule' generates phonetic zero from this morphophoneme to produce Nom/Sg okno. Another rule realizes ź as o in the genitive plural.

Let us examine more closely Worth's proposed underlying segment ź for use in non-derived words, or using Worth's terminology, flexional stems. Worth (1968:114) says, 'the choice of full vowel or zero is determined partially phonetically, partially by paradigm class or stylistically in endings, but is conditioned almost exclusively by phonetic environment in the case of stems.' He then proposes the rules:

\[(\text{4})\]  
\[# \rightarrow Ø / _C_oV\]
\[# \rightarrow \text{vowel} / \text{elsewhere, i.e. } _C_C\]
\[(_C_c#)\]
\[(_C_c)\]

Since ź has traditionally been used in underlying representations for phonetic [Ø] (i.e. word boundaries) and since Worth does not
differentiate this symbol from the one he proposes, Worth's novel (and ingenious) use of it to account for the V-Ø alternations in Russian runs into trouble. First, this symbol differs from word boundary in that the former can be realized as a phonetic entity, namely a vowel. Word boundary, on the other hand, never is. Second, Worth's symbol must be differentiated from word boundary in some way, otherwise we would get non-occurring forms such as okono from underlying okon#bylo#, where # represents the third part of Worth's vocalization of rule. Third, Worth's rules do not take into account one of the most striking regularities of the V-Ø alternation in non-derived words, namely that the full vowel (ö or ø) is always realized under stress. Finally, the use of an abstract symbol violates the naturalness condition of underlying representations, for which Kiparsky (19) has convincingly argued.

Theodore Lightner has proposed another solution to the V-Ø problem. He suggests that the underlying system of Russian contains two lax vowels, namely a lax high back vowel ū and a lax high front vowel ɨ. He calls these vowels jers. In order to account for the V-Ø alternation in non-derived forms, Lightner suggests the following ordered rules:

(5) 1. ū, ɨ → Ø except / ē ū, ɨ 
2. ū, ɨ → ø, e 

Here are some sample derivations.

<table>
<thead>
<tr>
<th>Underlying</th>
<th>ñu+ū</th>
<th>ñu+a</th>
<th>oku+n</th>
<th>oku+n+ū</th>
</tr>
</thead>
<tbody>
<tr>
<td>rule 1)</td>
<td>sun</td>
<td>sna</td>
<td>okno</td>
<td>okun</td>
</tr>
<tr>
<td>rule 2)</td>
<td>son</td>
<td>--</td>
<td>--</td>
<td>okon</td>
</tr>
</tbody>
</table>

These rules work only if there exists a universal convention which states that stressed vowels do not truncate: cf. dialectal rúta 'mouth' which according to Lightner is derived from rúta; there are many more examples from derived words, cf. rútık 'small mouth' derived from rútik+u. (examples are Lightner's).

Lightner's approach to the V-Ø problem works well, given the universal convention that stressed vowels are never truncated, which he proposes. The approach has several drawbacks, however. First, it requires the presence of the underlying desinences ū and ɨ in the Nom/Sg and Gen/Pl of many nouns, where it is just in these cases in Russian that there are no endings. This is necessary to retain and eventually vocalize the jer in the stem. It may be argued that a 'trace' of the ending is found in the vocalized jer—which otherwise would be truncated. Lightner postulates these desinences solely for this reason. The endings themselves are never vocalized; they are always dropped. In short this desinence is really an abstract, albeit historically accurate, symbol whose only function is to allow stem-bound lax vowels to become vocalized according to rule 1 above. Without independent motivation for these desinences, it is hard to accept this proposal.

The second problem with Lightner's approach is that, implicit in his analysis the segment ū somehow underlies o and its phonetic manifestations, [o] and [ø]. In fact, o never alternates with ū.
This is further evidence that \( \tilde{u} \) is an abstract symbol, void of any psychological reality, much like Worth's \( \tilde{o} \). There seems to be no valid phonetic evidence to the statement that in Russian some \( o \)'s are derived from underlying \( \tilde{u} \).

Finally, while Lightner's analysis recognizes that stress plays an important role in determining which member of the alternation is realized on the surface, it requires an uneven distribution of stress. For example, the word stol 'table' is end-stressed; the final vowel of each form in the paradigm is stressed. In Lightner's approach, the Nom/Sg desinence, however, cannot be stressed or else it would be vocalized to produce the non-occurring \( \tilde{stol} \).

The analyses of Worth and Lightner represent the current approaches to the V-\( \tilde{O} \) problem. Lunt (1978:230) states 'analysts nearly always end up positing some sort of underlying unit to take care of the vowel-zero alternation.' He concludes that the loss of jers 'has not yet taken place in Russian' and adopts a position similar to that of Lightner.

I will now present an alternate approach to the V-\( \tilde{O} \) problem. This approach holds strictly to the condition that phonological segments that underly surface forms must occur in the surface forms. Thus, a segment such as \( \tilde{u} \) is not considered a viable possibility for an underlying segment for modern Russian. The segment \( \tilde{o} \) will be considered a word boundary which is never phonetically realized except as a pause in deliberate speech. For ease of reference, here again are the alternations involved:

\[
\begin{array}{ccc}
\text{\( o \)} & \text{\( \tilde{O} \)} & \text{\( \tilde{\tilde{O}} \)} \\
\text{\( \text{son} \)} & \text{\( \text{sna} \)} & \text{\( \text{sn\text{\-}} \)} \\
\text{\( \text{okon} \)} & \text{\( \text{okno} \)} & \text{\( \text{okn\text{\-}} \)} \\
\text{\( \text{den'} \)} & \text{\( \text{dn'\text{\-}} \)} & \text{\( \text{dn\text{\-}} \)} \\
\text{\( \text{sosen} \)} & \text{\( \text{sosna} \)} & \text{\( \text{sosn\text{\-}} \)} \\
\text{\( \text{zec} \)} & \text{\( \text{zog} \)} & \text{\( \text{zog\text{\-}} \)}
\end{array}
\]

Word pairs such as spisok - kiosk show that epenthesis can be ruled out as the source of the alternating vowels. Russian \( o \) is lax and unrounded [\( \text{\( o \)} \)] and [\( \text{\( A \)} \)] when not stressed: \( \text{stol} [\( o \)] \) - \( \text{stol\text{\-}} [\( A \)] \), 'm\text{\-}олод' 'young' [\( o \)] - \( \text{molod\text{\-}} [\( A \)] \). Note, however, that this \( o \) differs from the one which alternates with \( \tilde{O} \). The difference lies in its surface distribution. When \( o \) from historical jer is stressed, two consonants or a consonant and a word boundary follow. When normal \( o \) is stressed, it may occur at word-final position (alternating \( o \) never does), before a consonant followed by a word-final vowel or anywhere else in the word. Alternating \( o \) does not occur in these positions. In short, while both these \( o \)'s alternate with reduced vowels [\( \text{\( e \)} \)] and [\( \text{\( A \)} \), only one alternates with \( \tilde{O} \), the one which is restricted in distribution. In order to account for the vowel-zero alternations discussed above, and in order to account for the dichotomy in distribution just discussed, I suggest that there are at least two sources for Russian phonetic [\( o \)], namely underlying \( o \) and underlying \( \tilde{O} \). The latter underlying segment is different from Lightner's \( \tilde{u} \) in two ways. First, it is a segment which occurs in speech, i.e. its use as an underlying segment does not violate the naturalness condition. Second, it occurs in underlying forms only where it occurs on the surface. When no phonological rules alter
it, it is realized as surface [ə]. Its presence is not required to account for other ə's becoming vocalized in non-derived words. A third source of [o] is underlying ɨ, which is phonetically similar to Lightner's ɨ only ɨ, too, is restricted to environments where it occurs in speech. This segment also alternates with e and o.

We will discuss below how these segments are related by phonological rules. First it is worthwhile to point out the theoretical reasons for suggesting underlying ə and ɨ are the source of the V-Ø alternations. Jakobson (1949:156) discusses the underlying form of the root for the word smotrét' 'to look'. He compares the 1st Sg and 2nd Sg forms which are stressed differently: smotr'ú - smotiš, which he renders phonetically as [smatr'ú] and [s'sm'òr'tís]. Jakobson takes 'as basic the alternant which appears in a position where the other alternant too would be admissible' and concludes that the underlying form of the stem must be smotr'ı because given the alternation of the stem vowels o-a, the o occurs under stress. Since ə is permitted under stress too, it must be secondary in this stem. While Jakobson's conclusion is certainly correct, it does not follow from his example. The phonetic rendering of the 1st Sg form is [smatr'ú] and [A] never occurs under stress. In order to determine which is underlying here, o or A, we must return to Bloomfield's statement, as Jakobson (p. 166) does in relation to a question regarding morphology: 'When forms are partially similar, there may be a question as to which one we had better take as the underlying form... The structure of the language may decide this question for us, since taking it one way we get an unduly complicated description, and taking it the other way, a relatively simple one.' Bloomfield's statement can be applied to underlying segments as well as to underlying morphemes. Thus, while [o] and [A] both occur in the stem of the verb 'to look' in Russian, it is easy to settle on o as the underlying segment since A, when stressed, can be either a or o (cf. mol[A]ko 'milk' - mol[o]ny 'milky'; sekret[A]n' 'of a secretary' - sekret[a]l' 'secretary'. In short, it is impossible to predict phonetically whether in a given word [A] represents underlying o or a. Phonetic [A] is ambiguous in Russian. For smotret', o alone can be considered basic. The same reasoning suggests that in the case of the V-Ø alternations, it is impossible to predict which segment of ə underlying o represents: the one which never is realized as zero or the one which sometimes is. In this instance the o is ambiguous. Consequently ə, which alternates with zero must be taken as basic; the ə that does not alternate with zero must be taken as derived. The relationship between ə and o can be described by phonological rules involving stress and consonant distribution. That ɨ, which alternates with zero, underlies both o and ə can be established by similar reasoning.

We now proceed to the phonological rules which relate underlying ə to o and ə and underlying ɨ to o, e, and Ø.

(7) 1. ə -→ o / (+stress)
2. ɨ -→ o / (+stress) -→ C", C" (-delayed release)
3. ɨ -→ e / (+stress)
4. ə, ɨ -→ Ø / — CV
Here are some sample derivations. The term 'other' refers to other well-known phonological rules.

| Underlying | sən | san+ə | žIg+1 | žIg+1ə | žIg+č' | okən+o | əkən
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Other</td>
<td>--</td>
<td>--</td>
<td>žIg</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>rule 1)</td>
<td>sən</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>rule 2)</td>
<td>--</td>
<td>--</td>
<td>žog</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>rule 3)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>žč'</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>rule 4)</td>
<td>--</td>
<td>sna</td>
<td>--</td>
<td>žgla</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

| Underlying | dIn' | dIn'a | səsIn+a | səsIn
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>rule 1)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>rule 2)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>rule 3)</td>
<td>den'</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>rule 4)</td>
<td>--</td>
<td>dən'a</td>
<td>səsna</td>
<td>--</td>
</tr>
</tbody>
</table>

Phonological rules accounting for akan'e and ikan'e will be ordered after the set of rules proposed above.

This paper has been limited to proposing a solution to the V-Ø alternations in non-derived words. The phonological rules proposed above should also apply to words formed by productive derivational processes, such as knižka 'small book', which they do. More perplexing problems arise, however, when the derivation involves multiple lax vowels such as in the pair bůlka 'bun' - bůlčka 'small bun'. In the latter, the o never alternates with zero. According to what has been presented above, this must represent underlying o. Yet, in the source word for this diminutive, the o does alternate with zero bůlka - bůlčok (Gen/Pl). According to what I have suggested above, this o is underlying ź. Which is it, o or ź? This question has been the subject of much investigation and has helped lead to the kind of abstract formulations already discussed. It seems possible that a minor change in one of the proposed phonological rules would produce the correct results in the majority of cases. This change mirrors the historical sound change involving vocalization of jers before a syllable containing another jer. Rule 4 can be rewritten:

4. ź, I, -→ Ø / CV (+tense) or (+stress). Thus, in the underlying form bůlčk+t+k+a, rule 4 would not delete ź, which is realized phonetically [bůlčke]. It appears that the approach suggested in this paper for the V-Ø alternations in non-derived words holds promise for accounting for similar processes in derivationally related words.
REFERENCES


THE BANYAN:
A PARTIALLY ORDERED SYNTACTIC STRUCTURE
WITH REINTERSECTION

Dan W. Higinbotham
The University of Texas at Austin

Context free phrase structure rules have long been used as a basis for parsing natural language sentences and describing their structure. The so-called non-configurational languages, those with relatively free word order, pose a problem for phrase structure rules because each of the possible orders requires a separate rule. Elements of a sentence which have more than one function also pose a problem, because a constituent's functional relationship to the remainder of the sentence is normally related to what it is immediately connected to above, and in a tree, there can be only one such connection.

Recent theories continue to have trouble with relatively free word order and elements with multiple functions. Lectures on Government and Binding, for example, insists that Japanese has a verb phrase constituent, and that the grammatical function subject is the noun phrase dominated by the sentence node, and the grammatical function object is the noun phrase dominated by the verb phrase node. The configurations that determine these grammatical functions, however, are not represented in the syntax in the X-bar system in D- or S- structures in Japanese.' (Chomsky 1982, p. 129). The problem stems from the fact that subject and object may appear in either order preceding the verb; according to the theory, the one case involves a discontinuous verb phrase constituent. But phrase structure trees simply cannot show discontinuous constituents. Government and Binding theory also has problems with elements with multiple functions. In the Japanese causative, the indirect object of the main clause is always the same as the subject of the subordinate clause. The Government and Binding analysis predicts that this element will be marked with the subject marking postposition ga, but it in fact appears with the indirect object marker ni. This incorrect prediction results from the faulty assumption that each element can have only one 6-role, an assumption necessitated by the fact that in the D-structure tree, an element can occupy no more than one position (Chomsky 1982, p. 131).

The inherent properties of trees themselves
appear to be at the root of these problems but nature itself has given us a cue to look beyond traditional restrictions. The banyan is a type of tree which grows to a certain size and then lets down vines from its branches. Some of these vines reach the ground and take root. The result, when a new root has grown, is a single branch connecting two roots. Banyans have been known to have forty or fifty trunks each over a foot in diameter, all interconnected with common branches. The banyan tree in nature is an organic acyclic connected directed graph which may have multiple roots.

Just as the banyan in nature is a slight generalization of what is normally expected for a tree, the banyan structure in linguistics is a slight generalization of the traditional sentence tree. The banyan is an acyclic connected directed graph. Just as each trunk of a natural banyan may be the endpoint or origin of a number of branches, each node of a linguistic banyan may be the endpoint of several arcs, and the origin of several others.

Just as sentence trees are generated by phrase structure rules, linguistic banyans are generated using an enhanced version of the constituent rules, or c-rules, of Lexical Functional Grammar (Bresnan 1982).

Lexical Functional Grammar. The structures generated by the c-rules of Lexical Functional Grammar are the c-tree, or constituent tree, and the f-structure, or functional structure. The following pair of structures are taken from the book The Mental Representation of Grammatical Relations, edited by Bresnan (p. 221, 236, 238):

(1a)

```
The girl wondered who the boy persuaded the girl to see the boy persuaded the girl to see.
```
"The girl wondered who the baby persuaded the boy to see."

Although a full explanation of these structures is well beyond the scope of this paper, a few notes are in order. The c-tree of (1a) is a traditional phrase structure tree. In (1b), the material between each pair of brackets comprises one f-structure, and each f-structure corresponds to one or more nodes of the c-tree of (1a). The dashed line connecting the f-structure labelled OBJ in the SCOMP f-structure, to the SUBJ of VCOMP of SCOMP, means that the boy has two functions in the sentence; it is the object of persuade and also the subject of see. The dashed line connecting who to the OBJ of VCOMP of SCOMP means that who has two functions as well; it is the complementizer (or Q-FOCUS) of the relative clause, and also the object of see.

Some of the c-rules used to generate (1) are as follows:

\[
\begin{align*}
(2a) \quad S & \rightarrow \ NP \quad VP \\
& \quad (\uparrow{\text{SUBJ}})=\downarrow \quad \uparrow\downarrow \\
(b) \quad VP & \rightarrow \ V \quad NP \\
& \quad \uparrow\downarrow \quad (\uparrow{\text{OBJ}})=\downarrow \\
(c) \quad VP & \rightarrow \ V \quad NP \quad VP' \\
& \quad \uparrow=\downarrow \quad (\uparrow{\text{OBJ}})=\downarrow \quad (\uparrow{\text{VCOMP}})=\downarrow
\end{align*}
\]
The top line of each of these rules appears like a traditional phrase structure rule, and is basically used as such to build the c-tree. The other lines refer to the associated f-structures. The notation \((\uparrow_{\text{VP'}})=\downarrow\) underneath NP in (2a) means that the NP node in the c-tree corresponds to an f-structure \((\downarrow)\) which has the label SUBJ within the next higher f-structure \((\uparrow)\), namely, the largest f-structure, or the one corresponding to the S node of the c-tree. The notation \((\uparrow_{\text{VP}})=\downarrow\) under the VP node means that the VP node corresponds to an f-structure \((\downarrow)\) which is the same as the next higher f-structure \((\uparrow)\); in other words, the S and VP nodes of the c-tree correspond to the same f-structure. Rules (2a) and (2b) generate SVO sentences; more complex verbs like persuade require rules such as (2c) and (2d). Although the to element of rule (2d) is optional, the lexical entry for persuade specifies that its VCOMP must have a feature TO with the value +. The notation \((\uparrow_{\text{INF}})=\name{c}+\) means that the next higher f-structure \((\uparrow)\) is required to have a feature INF with the value +; since the f-structure of the VP of VP' is the same as that of the next higher f-structure, this is a restriction that the verb of the VP' be in an infinitive form. The notation \(=\name{c}\) is used for constraint equations, which restrict both the construction of the c-tree and the forms of the related f-structures. A similar notation will be used in the rules generating banyans in this paper.

The main clause of Japanese. The arguments of Japanese verbs are marked with postpositions, which are similar in many ways to case markings. Since the postpositions marking subject and object clearly identify their grammatical function, the ordering of these elements is not nearly as fixed as in English. In fact, the only real ordering constraint dealing with the verb and its arguments, is that the verb must be last.

In mathematics, a set may be unordered, linearly ordered, or partially ordered. The situation of the main clause in Japanese seems to fall in the last category, since it is neither totally free, nor totally fixed. The mathematical definition of partial ordering is as follows:
A set $Z$ is partially ordered with respect to a relation $R$ if and only if

$R$ is a subset of $\{ (x, y) \mid x, y \in Z \}$

and $(x, y) \in R$ implies $x \neq y$

and $(y, x) \notin R$

and if $(y, z) \in R$ then $(x, z) \not\in R$.

Given any two elements $x$ and $y$ of the set $Z$, either $x$ precedes $y$ ($(x, y) \in R$), or $y$ precedes $x$ ($(y, x) \notin R$), or $x$ and $y$ have no necessary ordering relationship.

This is precisely the case in the Japanese main clause. The subject and object must precede the verb, but the subject and object have no necessary ordering with respect to each other. This situation could be reflected by the following rule:

(4) $S \rightarrow NP \rightarrow V$

($\uparrow$SUBJ)$\downarrow$ ($\uparrow$OBJ)$\downarrow$ ($\uparrow$HEAD)$\downarrow$

($\uparrow$CASE)$\rightarrow$ GA ($\uparrow$CASE)$\rightarrow$ O

$\downarrow \leq_c$ ($\uparrow$HEAD) $\downarrow \leq_c$ ($\uparrow$HEAD)

The subject is required to be marked with the postposition $ga$, and the object with the postposition $o$. The ordering restrictions are given by the notation $\downarrow \leq_c$ ($\uparrow$HEAD). The head element is the one marked by $\uparrow=\downarrow$. Both subject and object are constrained to appear before the verb, but no ordering constraint is placed on the relative order of the subject and object; consequently, they may appear in either order. The actual order of elements in the rule reflects only the default order, or the order of highest frequency. The notation $\succ_c$ can be used in a similar way.

Since the ordering restriction is really a restriction on the placement of the verb, (4) could also be written as follows:

(5) $S \rightarrow NP \rightarrow NP \rightarrow V$

($\uparrow$SUBJ)$\downarrow$ ($\uparrow$OBJ)$\downarrow$ ($\uparrow$HEAD)$\downarrow$

($\downarrow$CASE)$\rightarrow$ GA ($\downarrow$CASE)$\rightarrow$ O $\downarrow \succ_c$ @

The constraint $\downarrow \succ_c$ @ means that the verb must appear after all of the other elements in its domain.

The topic in Javanese. In active sentences in Javanese, locative and temporal phrases can topicalize freely by being placed sentence initially; these same phrases, however, cannot be topicalized in the passive.
The subject of a passive is also required to be definite, leading to the conclusion that the subject of a passive in Javanese is also a topic. When speakers of Javanese want to topicalize a first or second object of the verb, they do so by making the sentence passive; in other words, topicalization triggers passivization.

(6a) ibu nambil katerku
mother patch pants-my
Mother patched my pants.

(b) katerku ditambah ibu
pants-my patched mother
My pants were patched by mother.

(c) \( S \rightarrow NP \quad V \quad NP \)
\( (\uparrow \text{SUBJ})=\downarrow \quad \uparrow=\uparrow \quad (\uparrow \text{OBJ})=\downarrow \)
\( \downarrow \downarrow \quad \downarrow \downarrow \quad \text{ACTIVE} \quad \downarrow \downarrow \quad \text{(HEAD)} \)

(d) \( S \rightarrow NP \quad V \quad NP \)
\( (\uparrow \text{SUBJ})=\downarrow \quad \uparrow=\downarrow \quad (\uparrow \text{OBJ-AG})=\downarrow \)
\( \downarrow \downarrow \quad \downarrow \downarrow \quad \text{PASSIVE} \quad \downarrow \downarrow \quad \text{(HEAD)} \)

Rule (6c) is the rule for active sentences with one object, and (6d) is the corresponding passive. The element preceding the verb as SUBJ in (6c) appears in (6d) after the verb as the oblique agent OBL-AG, and the element that was OBJ in (6c) appears as the SUBJ in (6d).

The following rule can describe topicalization in nonstative sentences of Javanese:

(7) \( S' \rightarrow NP \quad S \)
\( (\uparrow \text{TOPIC})=\downarrow \quad \uparrow=\downarrow \)
\( (\uparrow \text{SUBJ})=\downarrow \quad \downarrow \downarrow \quad \text{\( \emptyset \)} \)

This requires that the topic be first in its sentence, and must be the same element as the subject. This gives us the following banyan structure:
Notice that the additional ordering constraint of (8) makes (6d) the only choice if *katɔ'ku* 'my pants' is to be the topic. Also notice that the topic does not appear before all of the elements of the sentence, since it is one of them itself. The ordering constraints of (6d) and (7) are satisfied, however, if they are construed to apply to the arcs leaving each node of the banyan, rather than the destination nodes of these arcs. The TOPIC arc precedes the HEAD arc, from the S' node, and the SUBJ arc precedes the HEAD arc, from the S node.

The Japanese causative. In the Japanese causative construction, as discussed above, the indirect object of the main clause is also the subject of the subordinate clause.

(9) *Takahashi ga Michiko ni ringo o tabesaset*.

Takahashi SUBJ Michiko OBJ apple OBJ eat-cause-past

Takahashi made Michiko eat the apple.

The fact that one noun phrase functions as both the indirect object of the main clause and the subject of the embedded clause, can be captured by the following rule:

(10) $S \rightarrow NP$

$$
\begin{align*}
& (\uparrow \text{SUBJ}) = \downarrow \\
& (\uparrow \text{OBJ}) = \downarrow \\
& (\uparrow \text{SCOMP}) = \downarrow \\
& (\downarrow \text{CASE}) = c \text{GA}
\end{align*}
\begin{align*}
& (\uparrow \text{SCOMP SUBJ}) = \downarrow \\
& (\downarrow \text{CASE}) = c \text{NI}
\end{align*}
\begin{align*}
& (\uparrow \text{SCOMP HEAD}) = \downarrow \\
& (\uparrow \text{CAUSE}) = c^*
\end{align*}
\begin{align*}
& \downarrow > c @
\end{align*}
$$
The noun which is the indirect object (IOBJ) is also the SUBJ of SCOMP. The notation (↑SCOMP HEAD)=↓ requires the subordinate and main verbs to be the same element, with the added restriction, (↑CAUSE)=c⁺, that the verb be in the causative form. These rules lead to the following banyan:

(11)

The ordering in (10) is the only ordering possible. A look at (11) makes it clear why this is so. Only elements on the periphery of the subordinate clause can be shared with the main clause, or else lines will cross. The verb of the subordinate clause is the same as that of the main clause, and that must be last. Since the subject of the subordinate clause is the indirect object of the main clause, it must be on the periphery of the subordinate clause. It must therefore be first in the subordinate clause, since the verb is last. The subordinate object must therefore be between them and the subject of the main clause has nowhere to go but to be first in the sentence.

Notice that according to (5), the postposition of the subordinate subject should be ga. It must be ni according to (10), however. Since a noun phrase in Japanese cannot be marked by both ga and ni, these constraints conflict with each other. The actual winner, as indicated in (9), is ni, and this is indicated in (10) with the (↓CASE)=ni notation, which indicates that the postposition is not only required to be ni, but this constraint overrides others specified by c⁺.
The relative clause of English. Above, we saw a case where the value of a feature on a reintersection node was required to have two conflicting values, and one constraint took precedence. The same can occur with partial ordering features.

A relative pronoun in English has two functions; it is the complementizer of the relative clause, and also one of the arguments of the verb. When the relativized element is the subject, no ordering constraints conflict. When the relativized element is the object, however, there are conflicting constraints. Complementizers in English must occur clause initially, but objects follow verbs.

In this structure, the subordinate OBJ arc is ordered after the subordinate SUBJ and HEAD arcs, but the position of the terminal element, the relative pronoun, is determined not by the $(\uparrow \text{OBJ}) \triangleright_o (\uparrow \text{HEAD})$ restriction, but by the overriding $(\uparrow \text{HEAD}) \triangleleft_o (\hat{\otimes})$ restriction of the R(elative) level.

Treating ordering with partial ordering features allows us to view the unusual features of reintersection nodes in a uniform way. The reintersection node in the Japanese causative was constrained to have two
conflicting postpositions, but only one prevailed. The reintersection node in the English relative was required to have two conflicting positions, but only one prevailed. The notations $\langle c \rangle$, $\langle c \rangle = 0$, $\langle 0 \rangle$, and $\rangle 0$ allow us to capture these generalizations in a rather simple way.

Conclusion. This paper has introduced partially ordered phrase structure rules and the banyan in order to simplify the analysis of relatively free word order languages and cases where a single element has more than one function. This has simplified the analysis of cases where a multi-functioned element appears with an unexpected feature, because there have been two conflicting constraints, and one has taken precedence over the other.

Among the mechanisms used by language to convey structural information are inflection, order, and prepositions or postpositions. Partially ordered phrase structure rules and the banyan allow all of these to be described in an analogous way. This unification of perspective should prove to provide many new insights.

REFERENCES


Computer applications in the field of testing are numerous and tantalizing. But in most educational settings in the United States and abroad, the use of computers as a test delivery system is simply not yet feasible. The purpose of this presentation is to discuss ways in which the computer can assist teachers in carrying out more effective traditional ESL or foreign language testing—by aiding in their preparation rather than in their classroom administration. We will begin by placing the discussion in perspective by contrasting computer assisted and computerized adaptive procedures. Then after a brief consideration of a few undergirding principles, under the rubric of latent trait analysis, we will cover computer applications in the preparation of standard language test items.

Computer Assisted and Computerized Adaptive Language Testing

Just as educational technology has advanced in recent years so has the use of this technology increased not only in the field of language instruction but also in language testing. Considerable effort is being expended in the development of computer assisted instructional materials for use in second- and foreign-language programs. In addition to the creation of computerized learning materials, computer-delivered tests are also being developed. As with the instructional packages, the testing programs vary in quality and usefulness. And, though some claim the computer is the answer to several of the testing problems that have plagued us for decades, we must still be cautious about how and when it is used for testing. Certain drawbacks should be considered when deciding whether to implement a computer assisted testing (CAT) program. The most obvious disadvantages include the lack of availability of the necessary hardware and software to implement an effective CAT program, due to limited budgets; an unfamiliarity with computers, which, as Cohen (1984) claims, can seriously
affect test outcome; and, the fact that it is difficult to test productive language skills and evaluate them holistically using CAT.

Unquestionably, however, some applications of the computer in testing are clearly more effective and efficient than conventional paper-and-pencil-type tests. While the aforementioned limitations must be taken into account when considering the use of computers for language test delivery, the advantages must also be weighed. These include such things as flexible scheduling of tests, self-paced testing, immediate feedback and record keeping. (See Larson and Madsen 1985 for a more complete discussion of advantages and disadvantages of computer assisted testing.)

One of the benefits of using the computer in test delivery is the ability to control which items are administered to which students. This capability of CAT leads directly to the development of the still more sophisticated computerized adaptive language test. This type of test differs from the computer assisted tests referred to above in that it goes beyond the scope of routine item delivery. While CAT tests may have the ability to present items on a predetermined basis, a CALT test administers items based on the individual examinee's performance on the test being taken. This means that after the examinee answers a given item, his or her response is evaluated and the next item presented is based upon that evaluation. For example, if the student answers the question correctly, a more difficult question is then given. If the answer is incorrect, an easier item is presented. Thus, the test adapts to the level of ability of the examinee, ultimately providing a more precise—but shorter—measure of that student's real ability than conventional tests are presently able to do.

CALT and latent trait analysis. Key to the development of effective computerized adaptive language tests is the ability to produce test items that do indeed measure differing degrees of student latent ability. The idea of adjusting test questions to match the ability of the examinee is not a new one. Psychometricians have known for many years that the most efficient tests are those whose items are centered closely around the actual performance level of the person being tested. For example, if we wished to measure precisely how high a person is able to highjump, it would be senseless to set the crossbar only at increments of one foot or more rather than in inches or centimeters. Furthermore, as expressed by Thurstone (1928), it is important that the measuring instrument not be seriously affected by the
object of measurement. In other words, the test should be able to measure everyone independently and not function as a 'rubber yardstick' (Wright 1977).

In order to have a CALT test that meets the accuracy requirements stated above, it must have items that are precisely calibrated along the trait or ability continuum from below the lowest student's ability range to above the highest student's range. Several latent-trait statistical procedures are available to perform item calibrations of this type. They vary considerably in their sophistication and practicality. The most appropriate procedure, however, for small-scale test analysis is the Rasch one-parameter model (Wright 1977; Stocking 1985).

Microcomputer Rasch Analysis and Test Item Calibration

In addition to using the computer for test delivery, it can also be used for other conventional testing purposes. While access to main frame computers to run item analyses is out of reach for most language teachers, it is beneficial to note that the Rasch latent-trait analysis can be run on an IBM-PC or other compatible microcomputers. The authors have used Microscale Plus™, a Rasch-analysis software package from Mediax Interactive Technologies, Inc.¹, to calibrate various sets of test items. This analysis program can be run on either dichotomous response (right/wrong) or equal-interval scaled data. Data input is done simply by recording 0 (incorrect) or 1 (correct) in each cell of the program's spreadsheet. After the responses have been entered, two convergence routines are performed on the data, comparing individual student ability values with individual item difficulty values.

Conclusions of the Microscale Rasch analysis are presented via several tables, maps, and charts. Results tables are given for both students and items. Table 1 shows a results table for an ESL reading test of 60 items administered to 183 students at Brigham Young University. (A similar table based on student performance is also available.)

The analysis program also generates a map that compares student ability to item difficulty on the same scale. (See Illustration 1.) The distribution of the students is represented above the horizontal line, while the distribution of items is shown below the line. This visual comparison allows for quick and easy determination of the relative difficulty of the test for this group of students.
<table>
<thead>
<tr>
<th>ITEMSCOR</th>
<th>LOGIT</th>
<th>ERROR</th>
<th>INFIT</th>
<th>OUTFIT</th>
<th>COUNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>107</td>
<td>.01</td>
<td>.16</td>
<td>-.11</td>
<td>.04</td>
</tr>
<tr>
<td>2</td>
<td>82</td>
<td>.61</td>
<td>.16</td>
<td>-.05</td>
<td>.05</td>
</tr>
<tr>
<td>3</td>
<td>47</td>
<td>1.52</td>
<td>.18</td>
<td>.01</td>
<td>.06</td>
</tr>
<tr>
<td>4</td>
<td>52</td>
<td>1.36</td>
<td>.17</td>
<td>.14</td>
<td>.18</td>
</tr>
<tr>
<td>5</td>
<td>112</td>
<td>-.18</td>
<td>.16</td>
<td>-.01</td>
<td>.18</td>
</tr>
<tr>
<td>6</td>
<td>58</td>
<td>1.18</td>
<td>.17</td>
<td>.03</td>
<td>.09</td>
</tr>
<tr>
<td>7</td>
<td>116</td>
<td>-.27</td>
<td>.17</td>
<td>-.12</td>
<td>.10</td>
</tr>
<tr>
<td>8</td>
<td>91</td>
<td>.34</td>
<td>.16</td>
<td>.08</td>
<td>.21</td>
</tr>
<tr>
<td>9</td>
<td>119</td>
<td>-.48</td>
<td>.17</td>
<td>-.14</td>
<td>.06</td>
</tr>
<tr>
<td>10</td>
<td>104</td>
<td>-.02</td>
<td>.16</td>
<td>-.15</td>
<td>.06</td>
</tr>
<tr>
<td>11</td>
<td>83</td>
<td>.50</td>
<td>.16</td>
<td>.08</td>
<td>.22</td>
</tr>
<tr>
<td>12</td>
<td>56</td>
<td>1.26</td>
<td>.17</td>
<td>-.15</td>
<td>-.08</td>
</tr>
<tr>
<td>13</td>
<td>51</td>
<td>1.44</td>
<td>.17</td>
<td>-.03</td>
<td>.07</td>
</tr>
<tr>
<td>14</td>
<td>56</td>
<td>1.28</td>
<td>.17</td>
<td>-.21</td>
<td>-.19</td>
</tr>
<tr>
<td>15</td>
<td>42</td>
<td>1.69</td>
<td>.18</td>
<td>.04</td>
<td>.15</td>
</tr>
<tr>
<td>16</td>
<td>59</td>
<td>1.17</td>
<td>.17</td>
<td>-.13</td>
<td>.16</td>
</tr>
<tr>
<td>17</td>
<td>33</td>
<td>1.94</td>
<td>.20</td>
<td>-.07</td>
<td>.02</td>
</tr>
<tr>
<td>18</td>
<td>100</td>
<td>-.10</td>
<td>.17</td>
<td>-.03</td>
<td>.07</td>
</tr>
<tr>
<td>19</td>
<td>67</td>
<td>.80</td>
<td>.17</td>
<td>-.06</td>
<td>.01</td>
</tr>
<tr>
<td>20</td>
<td>47</td>
<td>1.37</td>
<td>.18</td>
<td>.02</td>
<td>.15</td>
</tr>
<tr>
<td>21</td>
<td>101</td>
<td>-.25</td>
<td>.18</td>
<td>-.03</td>
<td>.10</td>
</tr>
<tr>
<td>22</td>
<td>54</td>
<td>1.19</td>
<td>.18</td>
<td>-.13</td>
<td>-.09</td>
</tr>
<tr>
<td>23</td>
<td>45</td>
<td>1.41</td>
<td>.18</td>
<td>-.10</td>
<td>.08</td>
</tr>
<tr>
<td>24</td>
<td>81</td>
<td>.28</td>
<td>.17</td>
<td>-.08</td>
<td>-.01</td>
</tr>
<tr>
<td>25</td>
<td>54</td>
<td>1.00</td>
<td>.18</td>
<td>.03</td>
<td>.23</td>
</tr>
<tr>
<td>26</td>
<td>82</td>
<td>.02</td>
<td>.18</td>
<td>-.14</td>
<td>-.10</td>
</tr>
<tr>
<td>27</td>
<td>39</td>
<td>1.46</td>
<td>.20</td>
<td>-.02</td>
<td>.19</td>
</tr>
<tr>
<td>28</td>
<td>54</td>
<td>.92</td>
<td>.19</td>
<td>.19</td>
<td>.38</td>
</tr>
<tr>
<td>29</td>
<td>161</td>
<td>-1.74</td>
<td>.23</td>
<td>-.18</td>
<td>-.31</td>
</tr>
<tr>
<td>30</td>
<td>158</td>
<td>-1.58</td>
<td>.22</td>
<td>-.18</td>
<td>-.14</td>
</tr>
<tr>
<td>31</td>
<td>176</td>
<td>-3.10</td>
<td>.41</td>
<td>-.05</td>
<td>.08</td>
</tr>
<tr>
<td>32</td>
<td>160</td>
<td>-1.68</td>
<td>.23</td>
<td>-.18</td>
<td>-.32</td>
</tr>
<tr>
<td>33</td>
<td>160</td>
<td>-1.63</td>
<td>.23</td>
<td>-.18</td>
<td>-.32</td>
</tr>
<tr>
<td>34</td>
<td>175</td>
<td>-2.94</td>
<td>.38</td>
<td>-.06</td>
<td>.03</td>
</tr>
<tr>
<td>35</td>
<td>150</td>
<td>-1.59</td>
<td>.21</td>
<td>-.14</td>
<td>-.19</td>
</tr>
<tr>
<td>36</td>
<td>136</td>
<td>-.76</td>
<td>.18</td>
<td>-.05</td>
<td>.01</td>
</tr>
<tr>
<td>37</td>
<td>140</td>
<td>-1.01</td>
<td>.19</td>
<td>-.17</td>
<td>-.18</td>
</tr>
<tr>
<td>38</td>
<td>171</td>
<td>-2.56</td>
<td>.32</td>
<td>-.18</td>
<td>-.69</td>
</tr>
<tr>
<td>39</td>
<td>147</td>
<td>-1.16</td>
<td>.20</td>
<td>-.05</td>
<td>.12</td>
</tr>
<tr>
<td>40</td>
<td>141</td>
<td>-.95</td>
<td>.19</td>
<td>-.27</td>
<td>-.35</td>
</tr>
<tr>
<td>41</td>
<td>80</td>
<td>.73</td>
<td>.15</td>
<td>-.04</td>
<td>.04</td>
</tr>
<tr>
<td>42</td>
<td>135</td>
<td>-.79</td>
<td>.18</td>
<td>-.11</td>
<td>-.11</td>
</tr>
<tr>
<td>43</td>
<td>134</td>
<td>-.67</td>
<td>.17</td>
<td>-.13</td>
<td>-.13</td>
</tr>
<tr>
<td>44</td>
<td>122</td>
<td>-.33</td>
<td>.16</td>
<td>-.24</td>
<td>-.18</td>
</tr>
<tr>
<td>45</td>
<td>89</td>
<td>.50</td>
<td>.15</td>
<td>-.12</td>
<td>-.03</td>
</tr>
<tr>
<td>46</td>
<td>31</td>
<td>2.18</td>
<td>.20</td>
<td>.12</td>
<td>.68</td>
</tr>
<tr>
<td>47</td>
<td>120</td>
<td>-.31</td>
<td>.16</td>
<td>-.08</td>
<td>-.01</td>
</tr>
<tr>
<td>48</td>
<td>173</td>
<td>-2.93</td>
<td>.38</td>
<td>-.12</td>
<td>-.25</td>
</tr>
<tr>
<td>49</td>
<td>94</td>
<td>.38</td>
<td>.15</td>
<td>-.02</td>
<td>.07</td>
</tr>
<tr>
<td>50</td>
<td>129</td>
<td>-.52</td>
<td>.17</td>
<td>-.08</td>
<td>-.01</td>
</tr>
<tr>
<td>51</td>
<td>95</td>
<td>.34</td>
<td>.15</td>
<td>-.12</td>
<td>-.06</td>
</tr>
<tr>
<td>52</td>
<td>148</td>
<td>-1.14</td>
<td>.19</td>
<td>-.03</td>
<td>.07</td>
</tr>
<tr>
<td>53</td>
<td>151</td>
<td>-1.29</td>
<td>.20</td>
<td>-.05</td>
<td>-.01</td>
</tr>
<tr>
<td>54</td>
<td>157</td>
<td>-1.53</td>
<td>.22</td>
<td>-.17</td>
<td>-.25</td>
</tr>
<tr>
<td>55</td>
<td>90</td>
<td>.46</td>
<td>.15</td>
<td>-.04</td>
<td>.07</td>
</tr>
<tr>
<td>56</td>
<td>55</td>
<td>1.39</td>
<td>.17</td>
<td>.03</td>
<td>.21</td>
</tr>
<tr>
<td>57</td>
<td>59</td>
<td>1.27</td>
<td>.16</td>
<td>.01</td>
<td>.19</td>
</tr>
<tr>
<td>58</td>
<td>101</td>
<td>.20</td>
<td>.15</td>
<td>-.25</td>
<td>-.21</td>
</tr>
</tbody>
</table>

Table 1

ESL Reading Test: Item Results Table
Illustration 1

ESL Reading Test: Map of Student Ability Compared to Item Difficulty
Illustration 2

ESL Reading Test Items: Outfit by Logit
Illustration 3

ESL Reading Test Items: Infit by Logit
Other charts that provide valuable information at a glance are the outfit and infit plots. (See Illustrations 2 and 3.) The outfit plot is useful in spotting 'outliers,' students or items that do not 'fit' the model for one reason or another. The infit plot provides, among other things, information about item interdependency and precision. Further discussion of the usefulness of these charts and tables is presented in the section that follows.

Computer Applications in Standard Testing Situations

Utilizing a single microcomputer at one's school, or collaborating with a colleague who has access to one, a teacher can now analyze, refine, and develop language tests in a way never before possible.

At the same time, it should be recognized that there are widely used evaluation measures the computer cannot yet help us with, such as essays, dictation, translation, precis, and oral interviews.1 The test form most amenable to the computer routines discussed below is multiple choice or right-wrong items such as cloze or completion-type questions. We need to recognize this limitation at the outset and avoid letting enthusiasm for computer applications deter us from employing legitimate evaluation measures that do not lend themselves to computer applications.

The range of options in making computer applications. Before employing the computer in test making, it is advisable to identify the options available to us. First of all, the Rasch analysis discussed earlier can enable us to improve an existing test: It can help us identify redundant items. It can assist us in identifying biased items as well as questions unsuited for our student group--either because of item difficulty level or inappropriateness of content. And it can point out bad items that are cued by earlier questions on the test.

Besides helping us refine an existing test, the computer (in this case the Microscale Plus routine) can help us with exam security by facilitating the development of alternate test forms without the need to try out both forms on the same group of students. It provides us with the option of test item banking, and it can help us develop tests tailored to a specific level; and such
tests can be considerably shorter than usual and still provide us reliable results.

And after our new test has been administered, a computerized Rasch analysis can help us identify persons who have cheated as well as individuals for whom the exam is inappropriate.

Applying the computer. To use the computer as an aid to developing a language test, one begins typically with an existing exam that has already been administered and scored. Every item from each test paper is entered onto the Microscale Plus spreadsheet, as explained earlier, and then Rasch analyzed. With the printouts generated by this routine, we are now able to eliminate inappropriate items and select those questions that are the most appropriate for the test we are planning to construct.

First of all, we can identify redundant items by referring to the item bar graphs at the bottom of the map (Illustration 1). Noting the longest bars and the logit value (difficulty index) of these bars, we can eliminate excess questions with these logit scores. Such items are not contributing any useful information to us.

We have written elsewhere about item bias (1985). One approach to identifying bias is to look at "outliers" on the item outfit chart (Illustration 2). Outliers are those items (or students, if examining a student outfit chart) that do not cluster with the bulk of the questions—notably those in the two upper quadrants. For example, a vocabulary test administered to a class of students with mixed language backgrounds, may result in several outlying items. Upon examining these, we may find that each of the outliers inadvertently consisted of a Spanish cognate. The student outfit chart would likely corroborate our findings: Low ability Spanish speakers getting these difficult cognate items right would appear as outliers: persons performing unexpectedly on the test.

Referring to the map once again (Illustration 1), we can quickly identify items that are much too easy or much too challenging: those that fall far to the left or right of the student group charted above the horizontal line. These we would eliminate from our test, assuming our students were about the same in ability as those pictured in our analysis. Such items might include a simple grammatical construction long since mastered by all of the students.

Or these outlying questions might be cloze items from an inappropriate (say, scientific) prose passage not readily comprehended by your arts-stream students.

Moreover, our computer printouts can also identify test questions that interact inappropriately with other
items on the exam. For instance, if item 12 on a reading passage were "Why did the little old man in the green VW shout at the policeman?" and item 15 were, "What color VW was the little old man driving?" many students could be expected to rely on the former question when answering the latter. Outlying items in the lower quadrants of the infit scale (Illustration 3) identify problems such as this and once again enable us to eliminate or reword such questions.

Finally, the map of students and items shows us gaps where more items are needed on a particular level of difficulty.

Besides enabling us to effect test improvement or the editing of an older test in constructing a new one, the computer can help provide increased flexibility in evaluating students. For one thing, computerized Rasch analysis enables us to calibrate the difficulty level of items, with great precision. By selecting representative items from one test (say items 10 or 15), we can add these to another exam. After the second exam has been administered, the computer enables us to calibrate very precisely on a single scale all of the items from both tests, even though different groups of students took the two exams.

Such procedures enable us to develop banks of items that can be used in developing subsequent tests. And these carefully calibrated item banks make it simple to create alternate test forms that match each other very closely as to difficulty level.

Tailored testing is likewise facilitated by the Microscale computer routine. While truly individualized tests require a computer delivery, an approach to tailored testing is possible using a Rasch analysis of conventional tests. One possible approach is to administer a screening exam first. Depending on the nature of one's program, a branching test could be administered: the first part might be anything from an oral interview to a dictation or essay. Students could then be grouped into three general proficiency levels, and their evaluation fine-tuned by providing them with a Rasch-calibrated test tailored to their particular level. Items of overlapping difficulty could be included at the upper ends of tests at the first two levels and the bottom ends of the second two levels to accommodate persons not accurately identified by the screening exam:

<table>
<thead>
<tr>
<th>screening test</th>
</tr>
</thead>
<tbody>
<tr>
<td>low test</td>
</tr>
<tr>
<td>mid test</td>
</tr>
<tr>
<td>high test</td>
</tr>
</tbody>
</table>
Tailored tests such as these make it possible to reduce the length of the instrument rather substantially without sacrificing accurate measurement—perhaps up to 50 percent. (Computerized adaptive tests presented to students on a personal computer have been successfully reduced by up to 80 percent.)

Finally, the computerized Rasch analysis described in this article makes it possible to spot abnormalities in student performance once the exam has been prepared and administered to your students. For example, checking the student outfit chart can help one identify medium or low proficiency students who are cheating on occasional items. The student infit chart helps pinpoint students for whom the exam may be inappropriate, such as native Americans taking an ESL test, or capable foreign speakers who miss achievement-type items simply because they had been absent when a segment of material was presented that appears on the test.

In summary, then, while computerized Rasch analysis has its limitations in language test applications, it can provide an array of benefits to ESL and foreign language teachers in the area of multiple-choice exams, right-wrong items such as found on cloze or completion tests, and exams with equal-interval scaled scoring.
Footnotes

1 Further information about Microscale Plus™ can be obtained by writing to Medix Interactive Technologies, Inc., 21 Charles Street, Westport, Connecticut 06880-5889.

2 Actually, guided interviews with set questions can be Rasch analyzed and calibrated; therefore, this type of interview can be refined by computer analysis.

References


Some Observations on Proper Name Formation in the Book of Mormon
Stephen D. Ricks

The Book of Mormon contains several hundred names that offer an opportunity for examination and study. In the following sections are some observations on a few of them.

The Proper Name Anti-Nephi-Lehi

The proper name Anti-Nephi-Lehi (the name given to the Lamanite converts of Ammon who returned with him to Nephite territory) is striking for a number of reasons: (1) the division of the morphemes by hyphens; and (2) the familiarity of the latter two elements of the name and the apparent familiarity of the first. While it is undeniably true that the current edition of the Book of Mormon places hyphens between the elements of the name, this was not the case in the manuscript versions. This leads me to observe a peril associated with determining morphemic elements of a name and of proposing etymologies—variant spellings. Of the 188 Nephite names found in the Book of Mormon (Tvedtnes 1984) there are variant spellings in the manuscripts and in the various editions in over fifty of them. Although most of these differences are in the vowels (e.g., Onti/Anti, Morianton/Morionton), occasionally there are differences in the consonants as well (Paanchi/Paachi, Ammon/Amon, Ammonihah/Amonihah, although any of these may, of course, be the result of a scribal error), which could lead to a very different proposed etymology.

While I do not disagree with the division of the name Anti-Nephi-Lehi into the discrete elements suggested by the hyphens, the hyphens cause us to dissociate the "Anti" in this name from other "Anti" elements in proper names, such as those occurring in Antion, Antionah, Antionum, Antiparah, Antipas, Antipus, Antum, and possibly also Onti. While these may or may not contain the same element, if they do, it would be just as appropriate to write, e.g., Anti-On as to write it in the manner which is currently found in the Book of Mormon (I suspect that it does not contain the same element, since the separate elements in the name Antion are not as obvious as they are in Anti-Nephi-Lehi). Finally, I would like to mention a possible etymology which has been proposed for the "Anti" element in Anti-Nephi-Lehi (Tvedtnes 1984: N 29). On the fairly likely assumption that the language of the Nephites (and, consequently, of the plates) is Afro-Asiatic, the Egyptian ntr, "the one who is of, belongs to" renders the sense of the whole admirably. In any event, this gives a more meaningful sense for the name than any provided by the English—or Greek—“anti,” although here too, it ought to be noted that occasionally the meanings of names are translated in the Book of Mormon (e.g., Bountiful and Desolation).

Proper Names Common to the Nephites and Jaredites or Jaredites and the Bible

As noted above, in the Book of Mormon there are 188 Nephite names attested, not including names with variant spellings in different editions or sometimes within the same edition (e.g., Amalekites/Amilikites or Ammaron/Ammoron/Amoron), those which have what are clearly or apparently English suffix morphemes (Amlici/Amlicites), and those biblical names which are mentioned outside of a biblical citation but not specifically applied to a Nephite or Lamanite. Similarly, there are sixty-two Jaredite and 148 biblical names mentioned in the Book of Mormon (Welch). Of these, nine names are common to the Jaredites and Nephites, sixteen names are common to the Bible and the Jaredites, and thirty-seven names are common to the Nephites and the Bible. The names common to the Bible and the Jaredites are Aaron, Ephraim, Esrom, Ethem, Ether, Gilead, Gilgal, Heth, Jared, Kish, Levi, Nimrah, Nimrod, Noah, Ramah, and Seth. The names found among both the Jaredites and the Nephites are Aaron, Coriantumr, Gilgal, Moriancumr,
Morianton (Morionton), Nehor, Noah, Shiblon (Shiblom), and Shim. Given the origin of the Jaredites, it is not surprising to find certain names common to the Jaredites and to the Bible in the patriarchal period predating the Tower of Babel, such as Jared, Nimrod, Noah, and Seth. On the other hand, the fact that some of the biblical names found in the Book of Ether are post-Tower (i.e., Aaron, Ephraim, Esrom, Ethem, Ether, Gilead, Gilgal, Heth, Kish, Levi, Nimrah, and Ramah) requires a somewhat different explanation. Among these possible explanations are (1) linguistic coincidence among two otherwise unrelated languages (an explanation which I find rather unlikely in the light of the sheer number of these coincidences: sixteen of a total of sixty-two names); (2) proper names shared by both of the languages, either as borrowings from other language families (Heth, Kish, and Levi may fall into this category) or because the languages are related to each other (it is clear that Hebrew itself has borrowed a number of proper names--e.g., Pinchas, which is a Nubian name--from other languages. It is possible, then, that Hebrew--or proto-Hebrew--borrowed from the Jaredite language or vice versa, resulting in the number of shared names, or that they both derived them from a common source); and (3) the similarities in the names are the result of levelling that took place during the translation process. As the record itself shows, it was Mosiah II who translated the Jaredite record. Although it is quite apparent that Moroni edited the Book of Mormon in its present form (consider the number of times--eleven--that Moroni includes editorial asides in the course of the book), it is not equally clear that Moroni translated, or even knew how to translate, the record of the Jaredites, and it is possible that he was forced to rely solely on the work done by Mosiah II. Although there is no reason to suppose that the translation work of Mosiah was on the whole untrustworthy, it may be that his rendering of certain of the Jaredite names was influenced by his knowledge of biblical names that sounded similar to those found in the Jaredite record. Thus, there may have been a name in the Book of Ether that sounded like Ephraim, and was consequently regularized to a name that was already familiar to Mosiah from his knowledge of the brass plates, even if it was current as a personal name among the Nephites (cf. the practice of finding Hebrew names which sound somewhat like the original name, e.g., Reuben for Robert, Simcha for Stephen, etc.).

What has been said of the biblical names that are also found among the Jaredites holds true for the commonly shared Nephite and Jaredite names. However, it is also possible that certain of these shared names that are not also found in the Bible (Coriantumr, Morionicum, Morianton, Nehor, Shiblom, Shiblon, Shlom, Shim) may have been borrowed from the Jaredites following the Mulekite discovery of Coriantumr, the last of the Jaredites. This supposition is, I believe, strengthened by the observation that all of these names occur in the Nephite record only after the Mulekite contact with Jaredite civilization. I think that Coriantumr in particular may have come into use among the Nephites in this manner, since the finding of the Jaredite Coriantumr must have created something of a sensation among the Mulekites and Nephites.

Glosses in the Book of Mormon

Five words in the Book of Mormon are glossed: deseret (a word, apparently of Jaredite origin, meaning "honey bee," Ether 2:3), Irreantum (the name given to the sea in Bountiful by Lehi and his party, 1 Nephi 17:5), Liahona ("a compass," Alma 37:38), rabbah (powerful or great king, Alma 18:13), and Rameumptom (the "holy stand," Alma 31:21). What do all of these glossed words have in common? They all seem to be in dialects or languages foreign to the language of the editor or the translator within the Book of Mormon itself (there is no strong reason to suppose that the glosses are placed in the Book of Mormon by Joseph Smith himself. Rather, he simply translated what was already there). They seem to be placed there for their linguistic interest, or as antiquarian curiosities. This contrasts with the glossed words in the New Testament, for instance. There, the words are given because they were deemed particularly holy or powerful. I
think that the Book of Mormon use of glossed words may be indirect evidence for the role played by the editor in its compilation.

Conclusion

Of the rather considerable amount of literature devoted to the study of the Book of Mormon, comparatively little, even of a preliminary nature, has addressed the topic of personal names. This paper represents a small contribution in this direction. Clearly, much more can--and should--be done.

References


The purpose of this paper is to compare two different approaches to describing linguistic variation. The first of these approaches, commonly called structuralist, is the traditional method for describing behavior. Its methods are found in many diverse fields -- from biological taxonomy to literary criticism. A structuralist description can be broadly characterized as a classificatory system (normally defined as a system of rules). The fundamental question that a structuralist description attempts to answer is how a general contextual space should be split up. Structuralists have implicitly assumed that descriptions of behavior should not only be correct, but should also minimize the number of rules and permit only the simplest possible contextual specifications. It turns out that these intuitive notions can actually be derived from more fundamental statements about the uncertainty of rule systems.

Within a structuralist framework, we can identify three fundamental types of description according to how the contextual space is split up: (1) exceptional versus regular behavior; (2) categorical behavior; and (3) idiosyncratic behavior:

![Diagrams showing exceptional versus regular, categorical, and idiosyncratic behavior]

Problems arise when we try to use rules to predict behavior. A system of rules partitions the contextual space. This naturally implies a rule of usage which permits one and only one rule to apply to any given fully-specified context. When we use this rule of usage, the contextual partitioning forces an exact demarcation in predicted behavior as we move across the contextual space from one rule context to another. Consider the predicted behavior for our three fundamental types of description; in each case we get abrupt, distinct shifts in behavior:

![Diagrams showing exceptional versus regular, categorical, and idiosyncratic behavior]
Yet in all three of these cases evidence from language behavior indicates that transitions across the contextual space are gradual and probabilistic:

As an example, consider the well-known results of perceptual tests between voiced and voiceless stops which show a gradual shifting towards the voiced stops as voicing onset time is increased:[2]

Or consider Labov's semantic experiment in which he found that as the relative width of a cup-like object is increased, the greater the chances speakers will identify the cup-like object as a bowl:[3]

Similarly, evidence from linguistic behavior shows that forms close to exceptional items may occasionally behave like those exceptions. Consider, for instance, repeated misspellings such as GREAT for grade,[4] which is apparently based on the exceptional spelling of the word great. A similar example is the spelling INCHOIR for inquir[e], obviously based on the uniquely-spelled word choir.[5] Johnson and Venezky have provided
examples of synthetic English words taking on exceptional pronunciations when these made-up words are close to certain exceptionally-spelled words. For example, PLOOD was frequently pronounced with the /ʌ/ vowel (as in the exceptionally spelled but very similar words blood and flood). Similarly, THEAT was frequently pronounced with the /ɛ/ vowel (as in the word threat, a very similar form).[6] The same effect has been observed in studies of word formation. For instance, I have recorded incorrect plural forms such as axen (for axes), which is obviously based on the exceptional plural exen.

Another difficulty with rule approaches is that without additional interpretations of how to use the rules, we cannot make predictions about behavior when a given context is deviant in some way or when crucial contextual information is missing. Yet there is abundant evidence from language usage that we can interpret improperly-formed contexts, such as slips of the tongue and ungrammatical sentences, including most of the starred sentences constructed by linguists. In addition, we can usually understand non-native and dialectal speech, proving these speech types are reasonably close to our own.

We can also use redundant information to predict outcomes when normally expected information is missing. Consider the following partially-obscured word from McClelland and Rumelhart:[7]

\[
\text{WORK}
\]

Even though the final letter is partially covered, it is obvious from what is given that the word must be work, not word nor the impossible word. Similarly, we can usually delete all the vowels from sentences of English without totally impairing our ability to understand what is intended, as in the following passage from Chomsky:[8]

The problem for the linguist, as well as for the child learning the language, is to determine from the deterministic system of rules that has been mastered by the speaker-hearer and that he puts to use in actual performance.

Traditionally, linguistic analyses have been based on the idea that language is a system of rules. Saussure, of course, is well-known as an early proponent of linguistic structuralism. Yet linguistic structuralism did not originate with Saussure -- nor did it end with "American structuralism". The neogrammarian approach to historical change is clearly structuralist. And it must be recognized that Chomsky himself is a structuralist par excellence. His attack against the American structuralists was not an attack against structuralism per se, but instead was an attack against the methodological assumptions that these linguists had espoused. For Chomsky (and virtually all other linguists today) there is no doubt that language is rule-governed
and that language behavior must be accounted for in terms of explicit rules. As a corollary, language acquisition is viewed as learning rules and language change as a change in the rules.

Within the structuralist framework, linguists have usually operated under the hypothesis that language rules are deterministic. Cases of variation have long been recognized, but until fairly recently these cases were vaguely identified as "free variation", a term which essentially meant that the behavior was non-deterministic. Labov and his co-workers have provided many examples of language variation which cannot be reduced to deterministic explanation. The question no longer is whether there is probabilistic behavior in language. Instead the question is: How do we account for this behavior? Following the traditional assumption that language is rule-governed, Labov has proposed that variable rules be used to account for probabilistic behavior.[9] Yet there are some serious conceptual difficulties with variable rules.

As I see it, there are two specific problems that have arisen in the study of language variation. One is the assumption that multiple-factor effects are simple mathematical functions of single factors.[10] This is, of course, a very specific assumption and one wonders whether there is much empirical evidence from language learning for such an underlying separation of variables. This assumption also explains the lack of reference to standard statistical approaches (such as discrete multivariate analysis) which directly consider the possibility of multiple effects.

The second specific problem follows from the first; namely, the question of which model (additive, multiplicative non-application, or multiplicative application) fits the data best -- or, equivalently, which parameters should be assigned to a logistic model.[11] There appears to be little evidence for a principled basis on which to choose the appropriate model.

But there are also two general problems that have arisen in variation theory. One is the status of probabilities in the model -- do they actually exist? and if so, how are they learned from the statistics and then used to predict behavior? Many have argued that the probabilities do not in fact exist, but this still leaves us with the question of what does account for the probabilistic behavior.[12]

But probably the most serious problem is the seemingly unlimited number of variables (linguistic as well as social) that affect any given non-deterministic phenomenon in language. In other words, there seems to be no end to the variation. This reminds one of the well-known suspicion in statistics that completely independent variables are rare, that given enough data almost any two variables can be shown to be dependent to some degree.
As a specific example of this problem, consider Guy's discussion of final /t,d/ deletion in English. Guy first lists Labov's three main linguistic variables that affect the probability of final /t,d/ deletion. Originally, Labov assigned two variants to two of these linguistic variables and ignored the lesser effect of the third:

(1) **grammatical conditioning**: Is the word monomorphemic or bimorphemic?
(2) **conditioning by following segment**: Is the segment a vowel or a non-vowel?
(3) **conditioning by preceding segment** (not specified in Labov's original study)

Over the years additional studies have shown that each of these variables should be assigned more variants — that additional distinctions are necessary if we want to account for empirically significant effects on final /t,d/ deletion. Guy partitions these three variables more finely:

(1) **grammatical conditioning**: Is the word monomorphemic, the past tense of an ambiguous verb, or the past tense of a regular weak verb?
(2) **conditioning by following segment**: Is the segment a consonant, a glide, a liquid, a vowel, or a pause?
(3) **conditioning by preceding segment**: Is the segment a sibilant, a non-sibilant fricative, a nasal, a stop, or a lateral?

So from an original 4-way distinction we now have a potential distribution of 75 possibilities (3×5×5 = 75).

But this is not all. Other variables have been discovered that affect the probability of final /t,d/ deletion: lexical stress, rate of speech, length of consonant cluster, articulatory complexity of clusters, speech style, and social factors (such as age, sex, social class, race, geographic background, and so on).

Guy rightly observes that this rule of final /t,d/ deletion shows that "variation is inherent, and cannot be scrubbed out of our linguistic description by ever-finer subdivisions of the data." In other words, final /t,d/ deletion cannot be reduced to deterministic phenomena. But the history of this example also implies that final /t,d/ deletion cannot be correctly described either! There doesn't seem to be any limit to the number of variants that affect final /t,d/ deletion. As more data is collected, more distinctions are discovered. The effect of these additional variants is less important, but they are
still statistically significant. It appears as if no simple correct description of final /t,d/ deletion is forthcoming. In attempting to describe non-deterministic phenomena optimally, we will have to sacrifice minimality in order to achieve correctness.

This example seems to point to the following conclusion — that ultimately the correct description may have a separate rule for every different set of conditions. Taken to its logical conclusion, this would mean that each rule would represent a single occurrence since no two occurrences are completely identical. In other words, instead of representing types of occurrence, rules would represent tokens of occurrence.

In the second part of this paper this idea will be developed as an alternative to structuralist descriptions. In fact, the notion of rule will be abandoned in favor of an analogical approach that avoids the conceptual difficulties of rule approaches. Instead of trying to predict behavior by using a system of rules (probabilistic or otherwise), an analogical description predicts behavior by means of a collection of examples called the analogical set. Given a context x, we construct the analogical set for x by looking through the data for

1. classes of examples that are most similar to x, and
2. more general classes of examples which behave like those examples most similar to x.

In order to show how to construct the analogical set for a given context, let us consider final stop deletion once more. Suppose we are interested in predicting final stop deletion when we are given the context vws — that is, when the final stop is not followed by a vowel (v), is not word-final (w), and is preceded by a sonorant (s). For this given context, we construct a hierarchy of supracontexts by systematically eliminating the factors v, w, and s. For each of these supracontexts we also determine the number of times the final stop is deleted and the number of times it is retained:

\[
\begin{align*}
&\text{vws (21,4)} \\
&\text{vw- (34,6)} \quad \text{v-s (58,14)} \quad \text{-ws (28,9)} \\
&\text{v-- (117,37)} \quad \text{-w- (55,16)} \quad \text{--s (80,67)} \\
&\text{--- (172,198)}
\end{align*}
\]
We next test each of these supracontexts for homogeneity. A supracontext is homogeneous providing all its subcontexts (as defined by the given context) behave alike. Using a statistical procedure that I will explain momentarily, we obtain the following results for this hierarchy of supracontexts:

(1) The given context (vws) is by definition homogeneous since it has no subcontexts.

(2) There are two supracontexts that are statistically homogeneous: vws and v-s.

(3) Two supracontexts are statistically heterogeneous: -ws and v-. These supracontexts have been circled.

(4) If a supracontext is statistically heterogeneous, then all supracontexts containing that supracontext are automatically considered heterogeneous. In our hierarchy there are three such inclusively heterogeneous supracontexts: -w-, -s, and -. These supracontexts are enclosed by boxes.

Now the analogical set is formed by all the occurrences from each homogeneous supracontext. In our example we obtain 137 occurrences, of which 113 involve final stop deletion:

\[
\begin{array}{ccc}
\text{vws} & 21 & 4 \\
vw- & 34 & 6 \\
v-s & 58 & 14 \\
\hline
113 & 24
\end{array}
\]

In order to predict behavior, we posit a rule of usage called random selection:

Randomly select one occurrence from the analogical set and use it as a model for predicting behavior.

With this rule of usage, we obtain the probability that the final stop will be deleted when the given context is vws:

\[
P(\text{deletion}|\text{vws}) = \frac{113}{137} \approx 0.825
\]

But the question still remains: How do we statistically determine whether a supracontext and its subcontexts behave alike. Traditionally, in using a statistical test, we must know either the underlying probability distribution for the test or a distribution that approximates the actual distribution. In this paper, I use a natural statistic called the rate of agreement.
that avoids this problem. If we have a context with $n_i$ occurrences of each outcome $\omega_i$, with $n = \sum n_i$ occurrences in all, then the rate of agreement is defined as

$$f = \frac{1}{n-1} \sum n_i(n_i-1)/2$$

Using this statistic, our decision procedure for testing homogeneity of behavior turns out to be very simple:

Always try to increase the rate of agreement.

To show how this works, consider the following array of data derived from our hierarchy of supracontexts:

<table>
<thead>
<tr>
<th>$\bar{\bar{\omega}}$</th>
<th>21</th>
<th>4</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>$\bar{\bar{\omega}}$</td>
<td>28</td>
<td>9</td>
<td>37</td>
</tr>
</tbody>
</table>

For each context we calculate the rate of agreement:

$$\bar{\bar{\omega}} \quad (1/24)[21 \cdot 20 + 4 \cdot 3]/2 = 9$$
$$\bar{\omega} \quad (1/11)[7 \cdot 6 + 5 \cdot 4]/2 = 2.818$$
$$\bar{\bar{\omega}} \quad (1/36)[28 \cdot 27 + 9 \cdot 8]/2 = 11.5$$

Since the two subcontexts ($\bar{\bar{\omega}}$ and $\bar{\omega}$) form a partition on the supracontext ($\bar{\bar{\omega}}$), we sum up the rates of agreement for the two subcontexts and compare that sum with the rate of agreement for the supracontext:

$$\Delta f = f(\bar{\bar{\omega}}) + f(\bar{\omega}) - f(-\bar{\omega})$$
$$\approx 9 + 2.818 - 11.5$$
$$\approx 0.318$$

Since we always try to increase the rate of agreement, we have the following decision procedure:

If $\Delta f > 0$, then the array is statistically heterogeneous.

If $\Delta f < 0$, then the array is statistically homogeneous.

Therefore, our sample array is statistically heterogeneous.

This statistical procedure is very simple. It never requires us to calculate probabilities or use an approximate distribution to estimate those probabilities. Another advantage
is that it is parameter-free. Its level of significance is asymptotically less than one half, but nonetheless fairly close to one half. This of course means that this test is very powerful. From a decision point of view this procedure can be defended in that it equally favors heterogeneity and homogeneity -- unlike traditional statistical procedures which strongly favor homogeneity. (It also turns out that we can redefine this statistical test so that decisions are made at smaller levels of significance.)

This procedure is also biased towards deterministic predictions of behavior, especially under certain well-defined conditions; namely,

(1) when the number of occurrences is low; or

(2) when imperfect memory reduces the number of occurrences.

This biasedness towards deterministic predictions helps explain several well-known observations about language variation; for example,

(1) the historical tendency to replace synchronic language variation by deterministic behavior; and

(2) the historical tendency to split up a fairly frequent non-deterministic context into a class of deterministic contexts.

In addition, deterministic behavior is favored whenever there is a need to maximize utility (that is, maximize rewards or minimize losses). Under such conditions a different rule of usage called selection by plurality is used:

Select the outcome which occurs most frequently in the analogical set and predict that the given context will take that outcome.

In order to compare this analogical approach to a rule approach, let us consider an example of morphological variation from Finnish. In Finnish certain bisyllabic verb stems ending in

\[
\text{[non-obstruent]} \text{ [dental stop]} \text{ [low vowel]}
\]

take imperfect forms ending in ti or si. Some of these verbs take only ti, some only si, and others can take either ending but with different degrees of acceptability (depending on the particular verb). The contextual space for this class of verbs can be split up into four distinct categories according to the length of the vocalic portion (either short V or long VV) and whether or not the syllable ends in a consonant C:
The last category, VVC, is statistically deterministic since it only takes the șl outcome (in the standard language). An optimal structuralist description of this class of verbs would at least combine all the verbs in the VVC category into a single rule. Such a rule would predict that a verb like viertaa, would take the ending șl in the imperfect. In contrast to the rule approach, consider the analogical set for viertaa (based on statistics from Tuomi's statistical analysis of standard Finnish):[16]

<table>
<thead>
<tr>
<th>distance from verb</th>
<th>verb frequency</th>
<th>number of homogeneous supracontexts</th>
<th>frequencies in analogical set</th>
<th>probability of random selection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>viertaa</td>
<td>ti si</td>
<td>ti si</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>viertaa</td>
<td>0 0</td>
<td>38</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>kiertaa</td>
<td>0 16</td>
<td>22</td>
<td>0 352</td>
</tr>
<tr>
<td>2</td>
<td>pitaa</td>
<td>0 24</td>
<td>14</td>
<td>0 336</td>
</tr>
<tr>
<td></td>
<td>siirtaa</td>
<td>0 8</td>
<td>14</td>
<td>0 112</td>
</tr>
<tr>
<td></td>
<td>kielitaa</td>
<td>0 22</td>
<td>10</td>
<td>0 220</td>
</tr>
<tr>
<td></td>
<td>riintaa</td>
<td>0 8</td>
<td>10</td>
<td>0 80</td>
</tr>
<tr>
<td></td>
<td>viinta</td>
<td>0 1</td>
<td>12</td>
<td>0 12</td>
</tr>
<tr>
<td>3</td>
<td>pyörtaa</td>
<td>0 4</td>
<td>6</td>
<td>0 24</td>
</tr>
<tr>
<td></td>
<td>kiintaa</td>
<td>0 1</td>
<td>6</td>
<td>0 6</td>
</tr>
<tr>
<td></td>
<td>vääntaa</td>
<td>0 1</td>
<td>4</td>
<td>0 4</td>
</tr>
<tr>
<td>4</td>
<td>murtaa</td>
<td>0 6</td>
<td>2</td>
<td>0 12</td>
</tr>
<tr>
<td></td>
<td>sortaa</td>
<td>1 0</td>
<td>2</td>
<td>0 2</td>
</tr>
<tr>
<td></td>
<td>huoitaa</td>
<td>0 1</td>
<td>2</td>
<td>0 2</td>
</tr>
<tr>
<td></td>
<td>juontaa</td>
<td>0 1</td>
<td>2</td>
<td>0 2</td>
</tr>
<tr>
<td></td>
<td>kuulta</td>
<td>0 1</td>
<td>2</td>
<td>0 2</td>
</tr>
<tr>
<td></td>
<td>kääntaa</td>
<td>0 26</td>
<td>2</td>
<td>0 52</td>
</tr>
<tr>
<td></td>
<td>muuntaa</td>
<td>0 1</td>
<td>2</td>
<td>0 2</td>
</tr>
<tr>
<td></td>
<td>myöntaa</td>
<td>0 30</td>
<td>2</td>
<td>0 60</td>
</tr>
<tr>
<td></td>
<td>puolitaa</td>
<td>0 1</td>
<td>2</td>
<td>0 2</td>
</tr>
<tr>
<td></td>
<td>työntaa</td>
<td>0 10</td>
<td>2</td>
<td>0 20</td>
</tr>
<tr>
<td>5</td>
<td>sääntaa</td>
<td>0 1</td>
<td>1</td>
<td>0 1</td>
</tr>
</tbody>
</table>

We first observe that the predicted behavior for viertaa looks rule-governed. In the analogical approach the predicted probability for the șl outcome is very close to one, the value that an optimal rule description of the data would predict.

But we also note that the analogical approach is, in a sense, messy: it permits the verbs murtaa and sortaa (which do not have the same syllabic structure as viertaa) to affect the predicted outcome. In contrast to an optimal rule approach, the
analogue description allows a slight "leakage" in favor of the random outcome.

The analytical approach also assigns probabilities to particular proportional analogies. The preferred analogy is, in the case of viertää, a rhyming analogy:

\[
\text{kierää : kiersi :: viertää : viersi}
\]

This analogy occurs 27% of the time. Rhyming analogies tend to be fairly significant, especially when the given context is non-occurring. But we must keep in mind that the analytical approach does not necessarily prefer rhyming analogies: the occurrence of such analogies depends on whether rhyming contexts are homogeneous in behavior.

This example also shows that the effect of a particular verb depends upon three factors:

(1) the amount of similarity between the verb and the given context;

(2) the frequency of the verb; and

(3) the number of homogeneous supracontexts that the verb occurs in (or, equivalently, the extensiveness of the homogeneity).

These same multiple analogical effects were noticed by Johnson and Venezky in their study of speakers' pronunciation of unfamiliar words:[17]

A model that might provide a higher degree of predictability ... is a final consonant model [in essence, a rhyming model] based on token counts rather than on type counts. This model would be especially effective if the final consonant influence derives from analogy with a few high frequency words rather than from a generalization based on all real words that contain a particular spelling.

This model is, of course, compatible with an analytical approach based on frequency of occurrence.

In many cases (like this one) the predicted behavior is nearly the same no matter which approach is used, but conceptually the two approaches are quite different:
RULE APPROACH

a system of rules
based on types of behavior
contextual space is partitioned
global, macroscopic
must learn rules from the data
static, rigid
usage: find the applicable rule
must specify how rules interact
well-defined boundaries
sharp, precise transitions
rule-governed
predictions made by rules alone
explicit, direct

equivalent to the original data
based on tokens of behavior
contextual space remains atomistic
local, microscopic
need for large memory capacity
dynamic, flexible
usage: find an appropriate example
must be able to access data quickly
no boundaries directly defined
gradual, fuzzy transitions
appears to be rule-governed
predictions only for given contexts
implicit, indirect

Many of these same distinctions are found in Winograd's terms "declarative" versus "procedural".[18] This distinction seems to be particularly relevant in explaining language performance, as has been pointed out by Rumelhart:[19]

Perhaps the classical case of using knowledge how (procedural knowledge) to produce knowledge that (factual knowledge) occurs in the domain of grammatical judgements. The knowledge that we have about language seems to be largely embedded in the procedures involved in the production and comprehension of linguistic utterances. This is evidenced by the relative ease with which we perform these tasks when compared with our ability to explicate the knowledge involved in them. Semantic knowledge would appear to be the same. Whereas we can quickly interpret sentences, it is only with the most painstaking effort that we can produce definitions of terms with any generality.

Despite these arguments, both empirical and conceptual, in favor of an analogical approach to the description of language (as well as other forms of behavior), there is a place for structuralism too. Structuralist descriptions are properly used to describe actually-occurring behavior. An optimal structur­alist description serves as a kind of meta-language that effi­ciently describes past behavior and allows us to talk about that behavior. Whenever we attempt to summarize behavior or to discover relationships in data, our viewpoint is structuralist. But if we wish to predict behavior rather than just describe it, it may be necessary to abandon structuralist approaches in favor of an atomistic one.
NOTES

1 This paper is based on my forthcoming book *Analogy and Structure*.

2 Abramson and Lisker 1972:19.


4 Gates 1937:72.

5 I owe this example to material provided by Thomas D. Horn.

6 Johnson and Venezky 1976:262.


9 Labov 1969:737-739.

10 Cedergren and Sankoff 1974:335-336.


14 Guy 1980:11.


16 Tuomo Tuomi, statistical analysis of approximately 600,000 words of text from 1975-76 issues of *Suomen Kuvalehti* (available on microfiche).

17 Johnson and Venezky 1976:266.


REFERENCES


McClelland, James L. and David E. Rumelhart (1980). "An Interactive Activation Model of the Effect of Context in Perception." CHIP 91 (part I) and CHIP 95 (part II) (University of California at San Diego: La Jolla, California).


A Computerized Diagnostic Test For Dyslexia

Matthew R. Sorenson
WICAT Education Institute

The Department of Health, Education and Welfare has estimated (HEW 1969) that 15% of the nation's schoolchildren suffer from specific reading disability, or dyslexia. Most of them remain undiagnosed during the early years when the foundation skills and emotional outlook towards learning are established. Because reading is important for studying all subjects, dyslexia can be a particularly unfortunate learning disability.

There has been some disagreement on how to define dyslexia (Benton and Pearl 1978, Boder and Jarrico 1983, Kinsbourne 1982, 1983, Liberman 1983). If the definition is too broad, dyslexics may not receive the specialized attention they need. Yet if the definition is too narrow, many dyslexic kids will remain undetected and untreated.

The lack of agreement on even the basic concepts has been characteristic of dyslexia research. The field has suffered from poor tests, poor research design, poor models, overspecialization by related disciplines, and a general lack of linguistic sophistication.

Problems with reading tests

Educators and researchers have been using standard reading tests for evaluating dyslexics (appendix A). These tests are not diagnostic, since they don't indicate specific problem areas. They are normative and therefore uninterpretable for non-normal readers such as dyslexics. They do not address the necessary linguistic issues involved in a language deficit like dyslexia. And finally they do not yield the fine-grained information needed to set up a useful remediation program.

Problems in reading research

Vellutino's (1979) review of dyslexia research pointed out the inadequate and inconsistent methods that have been used in sampling, experimental controls, data analysis, and interpretation. The result is a muddle of basic concepts and findings.

'Although competent researchers have recently become more actively invested in the study of dyslexia in young children, most descriptions of the disorder are based either upon clinical studies and informal observations or upon loosely designed experimental contrasts that have typically yielded equivocal and conflicting results.' (Vellutino 1979:3)

Models of reading

One reason for and one result of this lack of productive research is the lack of any useful models of either reading or reading disability (Doehring et al 1981:26, Calfee 1982, Kinsbourne 1982:209). Without a theoretical model, experiments cannot be fruitfully constructed, nor experimental results fruitfully compared.
Too many specialists

There are many disciplines interested in dyslexia and each one has their own focus to pursue. In neuropsychology questions have arisen about the role of hemispheric dominance patterns (Satz 1976, Aaron 1982), localization of functions (Benson 1983), high level integration of information from the various association areas (Geschwind 1979), and the cytoarchitectural pathology of the cortex areas involved in reading (Galaburda 1983).

Vellutino (1979) and Liberman (1978, 1983) have argued convincingly that the behavioral problems observed among dyslexics can be best explained in terms of a verbal/decoding deficit.

That approach was taken a step further by those who would explain low level symptoms by positing higher level cognitive and metalinguistic dysfunction (Orton Society Bulletin 1980, Pirozzolo and Wittrock 1981, Tunmer and Bowery 1984).

Studies on monozygotic and dizygotic twins, as well as family histories, have shown a significant genetic factor in the distribution of dyslexia (DeFries and Baker 1983, Decker and DeFries 1981, Finucci 1978).

Also, delays in the basic maturational processes have been blamed for dyslexia (Fletcher and Satz 1980, Elkind 1976, DeHirsch 1984). Although the longitudinal studies for investigating the role of maturation are difficult to pursue, they are necessary; both for a clear understanding of the disability itself, and for creating and evaluating remediation programs.

A narrow view of language

A general problem in this multidisciplinary landscape is that the redundant, recursive and abstract nature of language is unappreciated. The linguistic aspects of reading are usually taken to amount to a linear coding between sound and print. Even those who consider dyslexia to be primarily a language problem have been reluctant to look beyond word-level decoding. With very few exceptions (Vogel 1975, Kean 1984), the extent to which dyslexics have problems in morphology, syntax, lexical structures, and discourse processing have remained unexamined.

The sorts of cognitive tasks that are required for dealing with language have also been slighted in favor of memory and sensory processing. For example, the skills we have included under the label classification need more attention. The ability to class elements into abstract sets must be involved in establishing a phonology, in making sense of morphological alternations, in relating words into semantic and syntactic word classes, in paraphrasing sentences, and in associating language tasks with the appropriate situations.

Yet the dyslexia literature and the tests used by dyslexia researchers have not been sensitive to these issues.
Figure 1. WRAP's three dimensional measurement space.
It appears to the WICAT Education Institute that there is a clear need for a fine grained and comprehensive diagnostic test for dyslexia. We have undertaken the development of a computerized test to fill this need. For the moment, we will call it the WICAT Reading Abilities Profile (WRAP).

In establishing a conceptual basis for our computerized diagnostic test, we have tried to be as eclectic and empirical as possible. A framework that addresses the important issues and variables in the field is schematized in figure 1. Here we have a matrix defined by three dimensions: the level of language organization, the complexity of the cognitive task, and the input/output sensory modalities.

**Language level**

The language level that has commanded attention in past research has been the word. The literature has dealt with word-attack skills, word-level phonics, blending of phonemes into words, sight word recognition, and vocabulary. The instruments used have primarily been structured word lists. There was some early attention to letter reversals but it was found that the phenomenon was constrained by context in the word. The use of Berko's (1958) closure test of inflectional morphology has been increasing lately, but only in a few circles. Concern with higher levels of language organization has been largely limited to very gross measures of oral reading of passages.

Yet we know that reading involves a delicate coordination of all the various language levels, from segments, to affixes, to words, to phrases, to sentences and into discourse structures (cf. the variety of views in Spiro, Bruce and Brewer 1980). The size of the chunk that can be processed may be different for kids with different kinds of deficits.

**Cognitive tasks**

Along the second axis are the operations that must be applied to the different sized chunks of language.

The first thing a reader must do is direct attention. The units must be identified, then they can be discriminated from each other. Similarities and differences are compared so as to classify the units into functionally equivalent sets. They each must be remembered, and then put into or pulled out of strings, which demands ordering and analysis. And we need to find out how readers learn to manipulate and create these meaningful strings.

All of these cognitive tasks have been implicated as contributing to reading and must be involved in reading deficits. In dyslexia research, three have been emphasized: identification (recognition), memory, and analysis (decoding).
<table>
<thead>
<tr>
<th>discourses</th>
<th>texts</th>
<th>contexts</th>
<th>sentences</th>
<th>phrases</th>
<th>syntactic</th>
<th>affixes</th>
<th>lexical</th>
<th>words</th>
<th>letters</th>
<th>sounds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>match pictures to words</td>
<td>recognize items embedded in a matrix</td>
<td>are two things the same or different?</td>
<td>match words to pictures</td>
<td>identify the part that changed</td>
<td>which of four is different from others?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>use two or more screens</td>
<td>closure / fill in the blank</td>
<td>grammaticality judgement</td>
<td>divide up a string into components</td>
<td>comprehension</td>
<td>completing analogies or patterns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>cumulative subtasks</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>building whole from parts</td>
<td>free description of pictures</td>
<td>attention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>identification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>discrimination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>classification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>memory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>learning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>creating</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2.** Types of items to test cognitive tasks.
Figure 3. Traditional categories of reading subskills.
Item types

Of course, the problem in constructing a test is to isolate the variables to be measured. This can be done by using certain presentation formats for eliciting a particular cognitive task. Figure 2 shows the types of items that address the respective cognitive tasks. Then various sized language units can be inserted into that invariant presentation format. Using this 'slot-and-filler' approach will allow us to isolate the variables and will also make the test more simple and compact.

Both of these primary dimensions are hierarchical. The simpler components contribute to the more complex. Given this relation between variables, we would expect that the effects of a deficit in one area will spread to others.

If the spread is from complex to simple, then the disrupted processes are most likely of the top-down sort. A case of this sort arise if a kid cannot interpret words without knowing the meaning of the sentence that contains them. On the other hand, a spread from the simple into the complex would indicate problems in bottom-up processing (cf. Haber and Haber 1981, Rummelhart 1977).

Sensory modality

The third dimension of our conceptual framework would allow us to investigate which sensory modalities were either a problem or a crutch for a dyslexic. I will defer discussion on this point because the actual test will not implement this dimension right away. The kinds of technology and interface devices that will allow manipulation of how subjects receive the items and how they input responses to the computer are going to be emerging in the near future.

Within this framework we can identify where traditional categories of reading subskills would lie, as in figure 3. For example, the problem of perceptual reversals would be tested by items that fall in the slot defined at the letter level of language organization and involving the identification level of cognitive operations. Word attack skills would be located by crossing the word level with analysis. Storytelling would show up in the slot where texts and creating intersect, etc.

Items and subtests relevant to these areas are being collected from the best of the existing tests. However, these tests address only small parts of the model. So where good items do not exist, they are being written. The range of variables examined by existing tests is shown in figure 4.
Figure 4. Existing tests used in dyslexia research.
WRAP logic

Once we have collected or created items that can test the combinations of variables we want, they will be computerized. Item sequencing will be governed by intelligent and computer-adaptive diagnostic logic. The test will proceed to search for areas in the measurement space that seem to present problems to the student. Problems are indicated by slow response latency, inaccuracy, and inconsistency. All of these can be monitored and manipulated by the computerized test.

Computer-adaptive tests, like the WRAP, can adapt to the subjects' responses, administering only those items that are appropriate in terms of content and difficulty. If they miss an item, an easier one follows; if they pass an item, a harder one is used. This process continues until the correct level of ability is estimated. Besides being more efficient and accurate, computer-adaptive testing avoids the problems of giving every item to every student.

There is more to a test items than just right or wrong. By using informative distractors, we can find out not only where the dyslexic has problems, but what compensatory strategies are being used to deal with them. That is, if a kid consistently keys in on sound patterns when confused (or on word shape, or on semantic similarities) that information can help to establish a subtype.

Where there is evidence of a problem, the focus of the test would shift to that particular subspace. Related items would be chosen and a search would begin for the boundaries of the problem. The test would end at the point where the problem area, or areas, were well defined.

Every time the test is administered, we could add the result to a cumulative data base, which would tell us about which configurations of problems can be expected, which variables are correlated, which items are the most predictive, and which diagnostic strategies are the most productive. As a result, then, of its experience over time the test could provide the basis for its own refinement. The refinements themselves would take the form of a structured body of inferences, called an 'inference engine', which would take past results and present subject responses to make judgements about what items to present next.

Such a combination of tests, diagnostic logic, user-adaptivity, inference engine, and growing knowledge base constitutes an expert system (Hayes-Roth et al 1983, Winston and Pendergast 1984). Another essential feature of expert systems that we need to include in the WRAP is a flexible output interface. Different people are going to require different explanations for what has been determined by the test.

The results could be expressed in the language of reading pedagogy for teachers, in psycholinguistic jargon for researchers, or in very general terms for the student or parent. In any of these cases, the test should also be able to respond to user queries about procedures that were used and conclusions that were reached.

An overview of the structure of the WRAP is shown in figure 5.
Figure 5. Overview of WRAP
Subtypes

There is a consensus in the field that there are several subtypes of dyslexia that have specific deficit patterns and demand specific kinds of remediation. However, there is no consensus on what the subtypes are, and classification systems abound (Boder and Jarrico 1983, Mattis et al 1978, Lyon 1983, Malatesha and Dougan 1982, Rosenthal et al 1982, Doehring et al 1981, Satz and Morris 1981).

Because of the multiplicity of interests and biases, these typologies are built upon different sets of assumptions and findings. It is difficult to compare or synthesize them. What is needed to make sense out of the many schemes is consistent test results within a comprehensive model. This is the contribution that the WRAP can make.

Using the conceptual framework of the WRAP, some comparison can be made of several of the most prominent subtype classifications, as in figure 6.

Although admittedly there is some oversimplification in this comparison, we can see that in spite of the plethora of labels, there is some overlap. The first group of subtypes concerns problems of visually processing shapes of words. The second grouping clusters around tasks of analyzing phonological sequences. The third cluster seems to involve problems in meaning, naming, and interpreting sentences. A fourth grouping, which should be plotted on the third dimension of our measurement space, implicates problems of the hand and mouth, in the output process.

These four subtypes of reading deficit are very idealized. Most kids have been found to suffer from a mixture of problems, and there are likely more subtypes than we have been able to measure.

A complicating factor in subtyping is the issue of compensation. It has been noted that kids with sight problems tend to rely on phonics, and vice versa. This factor can only be controlled for if we have a sensitive enough test, apply it early in the acquisition process, and follow a stable population longitudinally with consistent measurements.

Progress

In pursuing the goal of constructing the WRAP test we have reviewed the research literature and consulted with the top experts in the country. The result of this groundwork is the conceptual framework we have described. The meat on this skeleton is the pool of subtests and items that we have either created or culled from existing tests. We are now on the verge of creating a prototype that will be used for trial tests in the following year, with subjects, both dyslexic and normal, provided by local school districts.
Figure 6. Comparison of several subtyping schemes.
Initially the WRAP test will be implemented on WICAT computers, using our established test development software. Later, we will make updated versions of the test available on IBM and Apple microcomputers.

In the future, there are several enhancements and extensions that we would like to make. First, using the knowledge of what dyslexic kids do wrong and what compensatory strategies they use, we can construct an effective computerized remediation program (cf. Clancey 1984). Second, we can exploit new developments in voice recognition and synthesis, new manual interface devices, and graphics and animation to present items in a variety of formats that test the participation of each sensory mode in the reading process. Third, the diagnostic and remediation programs need to be integrated into the normal reading curriculum. The materials used in classes can be parsed and coded to allow for automatic construction of new test items that will indicate problems and progress in that particular reading program.
References


Neuropsychology of learning disorders: Theoretical approaches, ed. by 
Knights and Bakker. Baltimore: University Park Press.

Neuropsychological and cognitive processes in reading, ed. by 

Spiro, Bruce and Brewer. 1980. Theoretical issues in reading 

Turmer, W. E. and J. A. Bowey. 1984. Metalinguistic awareness and 
reading acquisition. Metalinguistic awareness in children, ed. by 

MA: MIT Press.

Vogel, Susan Ann. 1975. Syntactic abilities in normal and dyslexic 

Winston, Patrick H. and K. A. Pendergast. 1984. The AI business: 
An Information Theoretic Interpretation of the Variable Rule

Matthew R. Sorenson
WICAT Education Institute

'Information' has been used informally and as a technical term in several disciplines, but has not been well-developed in the social sciences. This paper will look at information theory and its relevance to linguistics, trying in particular to show how it illuminates the nature of the variable rule.

Claude Shannon, John Von Neuman, and Norbert Weiner were the major figures in the postwar interdisciplinary science of systems. Von Neuman contributed game theory, dealing with strategic decisions. Weiner founded cybernetic control theory and focused on how systems regulate their own behavior to meet goals. And Shannon established a mathematical definition of information that encompassed both communication and thermodynamics.

Although Shannon is the recognized father of information theory, it is the main contention of this paper that the ideas contributed by Weiner and Von Neuman are essential in a correct application of it. It is because decision and control have been neglected by linguists that information theory has not been found fruitful.

By 1953, when Charles Hockett published a review of Shannon and Weaver's Mathematical Theory of Communication, there had already been some interest by linguists in information theory. Hockett lists three references and adds that

'It is not certain that all these references are based on adequate understanding of the theory.' (Hockett 1953: footnote 2)

Hockett's (1953) review of Shannon and Weaver was tentative and ambiguous. He approached the topic with the attitude that it must be valuable to linguists, but he failed to show how.

The review's first section describes the essential concepts of the theory. The following are paraphrases of Hockett's main points.

Signals are chosen out of a set of possible alternatives. Information is transmitted when the variety in the alternatives is narrowed down by selecting one of the possibilities. If there is no choice, there is no information transmission.

If the variety in the set of alternative signals increases, then the system has a greater informational capacity. A system with 'yes', 'no', and 'maybe' can transmit more information than a system with only 'yes' and 'no'.

The basic terms apply to the capacity of an information source, but not to single instances. Information theory is concerned with averages.
The source, channel, and receiver of information are distinguished, but the signal at any point in this sequence may be transformed (or transduced) into various material forms. Information is independent of mechanism.

The code, or the syntax and semantics of the signal set, may be manipulated for the sake of efficiency or redundancy, depending on the demands on the system. The two primary means of manipulation are adjustments on the frequency and on the interdependence of signals.

Shannon's first theorem says that codes can be made maximally efficient, if desired, so that channels can always be used at their maximum capacity. Shannon's second theorem is a generalization to cases involving a noisy channel, and says that any degree of noise can be offset by the use of redundancy.

Having established these basic notions, Hockett begins a discussion of how they may apply to linguistics. His first point is that the continuous nature of speech and the discrete structures of language are not inconsistent and may be mathematically transformed into each other within information theory. Speakers (and field linguists) learn through experience to act as transducers to interpret continuous signals into discrete language units.

Then Hockett addresses the extreme redundancy in language. Stress, for example, is redundant because it can be transmitted simultaneously with other signals. Writing is estimated to be about 50% redundant since we can read text in which half the letters have been removed.

Problems of indeterminacy and transformation in morphology, and encoding through writing are muddled through in information theoretical terms, but as Hockett admits,

'Since there is currently no way in which all this can be disproved, it does not qualify as a scientific hypothesis; it is merely a terminology.' (p. 42)

Hockett's final section outlines 'general implications' and touches on issues that would arouse discussion for years to come: the confusion of information and meaning, the tempting parallel between information and energy, and the idea that structuring of energy processes is accomplished by informational 'triggers'.

The understanding of information theory in this review is unfortunately structural instead of regulative. The new ideas are seen as ways of describing the way language is, instead of the way it works or how it is used. For example, Hockett raises the idea that social information can be transmitted by variation in language, but immediately dismisses the possibility as uninteresting (Hockett 1953: 39).
After the Hockett review, information theory was not absorbed by linguistics. Bar Hillel and Carnap (1952) made a valiant but uninterpretable effort at extending information theory into semantics. Greenberg (1956) made a well-conceived but superficial application in historical linguistics. Hockett continued to be interested and his 1977 volume contains early articles on several related topics. Zelig Harris tried early to examine syntactic structure in terms of statistical distribution of elements, and apparently is still at it (Harris 1987). Jakobson edited a volume on mathematical linguistics containing a piece by him on information and communication (Jakobson 1961). The same volume also included a chapter by Rulon Wells referring to Church and Wittgenstein as the conceptual forebears of information theory.

Although cluttered and unsystematic, On Human Communication by Colin Cherry (1957) was for years the best effort at applying information theory to language. It suffered, however, from too much engineering jargon and not enough appreciation of the human dynamics of communication.

In 1966 Alfred Smith edited a volume on Communication and Culture. Chapter two brought together articles which outlined the basic concepts. The section on social interaction included Goffman. Chapter six discussed networks, and seven dealt with noise. The sections on feedback and control and on redundancy and equilibrium included some rich ideas by great thinkers like William Powers, Ross Ashby and Karl Deutsch. Sections on semantics and pragmatics pushed into frontier areas.

However, the bulk of the work in Communication and Culture was either suggestive or peripheral. The book was designed with a clear idea of the interdisciplinary theory that could be applied, but the content never did rise to the concept. Particularly disappointing was the section on linguistic theory (chapter four). Greenberg, Gleason and Pike, pillars of the old guard, presented a very static and structural view. They did not understand the new ideas, and linguistics was not ready to deal with problems of decision and control.

Also included was the classic Chomsky piece called Three Models for the Description of Language. We must remember that inasmuch as revolutions impose order, they destroy variety. The unflowered bud of new scientific ideas are uprooted along with lots of dead wood when a new order is ushered in. Perhaps the most compelling reason for the failure of information theory to take hold in linguistics was that it was associated historically with the structural and quantitative world view that was swept away in the rationalist fervor of the sixties.

Reminiscent of the early interdisciplinary days, a recent symposium discussed the development of information theory throughout the thirty years since Shannon and Weaver (Machlup and Mansfield 1983). The physicists at the symposium were well ahead of the pack. The linguists still failed to understand the nature of information theory and how it could be applied to human communication.
Recently some efforts have been made to take the rigorous notions of information and entropy from physics and apply them to language (Campbell 1982 and Eigen and Winkler 1981). Unfortunately, these writers assumed that all there was to linguistics was Chomsky, and their presentation suffered.

Presently there are several areas on the 'fringes' of linguistics where information theory is being either implicitly or explicitly applied, including semiotics, formal semantics, discourse theory, and variable rule analysis. We will focus on the variable rule.

The variable rule

As variation involves the probabilities of selecting signals out of sets of alternatives, it can be described in the language of information theory. Linguistic variables that can assume a minimum range of values (say two) have the capacity of transmitting the minimum quantity of information (one bit). Variables that accomplish a greater degree of variety reduction have a greater informational capacity.

A serious question that must be addressed at this point is whether information theory can apply at all to language or social interaction. It has been argued (Waddington 1977:140) that if we don't know (ie. can't list) the set of signal possibilities, then we can't quantify how signals are selected, and therefore information theory is inapplicable in any interesting human domain.

A reasonable response is given by Dretske, who suggests that it is valuable to use information theory comparatively rather than absolutely.

'Such comparisons can be made without ever determining absolute values for either magnitude. That is, one can use these formulas in the way one would use a piece of string that is not marked off in inches and feet. One can use the string to determine whether A is longer than B without determining the length of either A or B.' (Dretske 1981:54)

Another qualitative approach to using the originally quantitative notions of information theory is proposed by Ashby (1972). He suggests using set theory to describe the relationships between ensembles of alternatives, replacing probability with membership as the central construct.

Anyway, to the extent that we can specify the range of values that can be taken on by a linguistic variable, we can use absolute informational measurements. In other cases, there are other means.

The variable rule, as developed by Labov and others is a generalized statement of the distribution of the probabilities of alternative patterns, as taken from a measured corpus. It is therefore a statement of informational capacity of that variable and its tokens as sociolinguistic signals.
It is clear that the frequency of occurrence of a variable (or the frequency of application of a variable rule) is both monitored and meaningful among speakers of a community. The signals are transmitted, received and have interactional consequences.

Since the variable can occur within many different kinds of words, or sentences, or situations, there is not a single probability for its occurrence. The constraints on the probability of application of a rule (called 'factors' in the variable rule literature) may be due to phonological, grammatical, or interactional environment. This requires us to use conditional probabilities in calculating the informational capacity (Cherry 1957:182ff. and Lerner 1972:63ff.). This does not affect the applicability of the informational approach, but is just a means of relativizing for context.

Factored into the machinery of the variable rule is an 'input probability', which is a sort of baseline probability, independent of all environmental constraints. The other factors are adjustments on this baseline.

By using measurements of either conditional probability or input probability, the informational capacity of all the signals in the set of alternative tokens can be established. The tokens with little probability of occurring will have the greatest capacity. Highly probable tokens will have less capacity.

Linguistic variables contribute to style (Labov 1972: chapter 3). As style changes, by cooperative shifting to a new set of 'contextual cues' (Gumperz 1982: chapter 6), the frequencies and expectations for signals are 'recalibrated'. The new becomes the norm. The variables that define a style become expected and can carry lower amounts of information. There is a sort of self-organized 'bootstrapping' going on between speakers when signals structure style, and style in turn structures the informational capacity of those very same signals.

Guy (1980) discusses whether the variable rule should apply to individuals or groups. Guy concedes to the Bickerton-Bailey axis that the variable rule applies where variation is ubiquitous and homogeneous. In communities where individual patterns are stable, but there is stratified variation between speakers, the averaging effect of the variable rule analysis obscures the facts. The important point here is that variation is meaningful only where there is a conventional code for its use. Variation, because of its informational nature, tells us about interaction, not about the gross structure of the speech community: about style, not stratification.

Guy (1980) also discusses the problem of how big the corpus should be for a good variable rule analysis. The method of the Labovians has been to think up strategies for enriching the quality of the interview. If there were a greater appreciation for the role of variation in regulating interaction, we would focus on more natural conversation. If the data is artificial and static, lots of it is going to be needed.
So just as the proper appreciation of information theory requires the consideration of decision and control, so the proper appreciation of the variable rule requires that we apply it to natural conversation where speakers are deciding, and controlling the speech situation.

Frake showed how asking for a drink in Subanun required competence far beyond the grammatical (Frake 1964).

'If messages were perfectly predictable from a knowledge of the culture, there would be little point in saying anything. But when a person selects a message, he does so from a set of appropriate alternatives. The task of an ethnographer of speaking is to specify what the appropriate alternatives are in a given situation and what the consequences are of selecting one alternative over another.'

The selections made in generating messages reduce possibilities, and thus transmit information. The content of the information is the concern of semantics and ethnography. The form and quantity of the information is characterized in the variable rule.

The selections may be grammatical/semantic (which modifier? which quantifier?) or may be less referential (which form of address? which dialect?). Some sets of alternatives are more easily specified than others. A choice between tu and vous (Brown and Gilman 1960) is more discrete than a choice about how to formulate place (Schegloff 1972). But in any case, there are choices made by speakers with consequences for the subsequent discourse.

The variable rule has too often been used in a way that describes only the static structure in a speech community. Language variables have been portrayed as passively responding to external aspects of the situation (see the various applications in Labov 1980). Bell reports dissatisfaction with this view. He discusses both 'responsive' and 'initiative' style shifts and concludes that

'Language is to be seen not merely as a dependent variable, manipulated by non-linguistic factors. It may be an independent variable which itself influences the situation.' (Bell 1984:183)

Brown and Levinson also ascribe to this view, in which variables are seen as manipulated by speakers for social-interactional purposes.

'... language usages are tied to strategies rather than directly to relationships, although relationships will be characterized by the continued use of certain strategies' (Brown and Levinson 1978:286)

Accommodation and control

One of the reasons that speech communities are coherant systems is because the members actually do 'cohere'. Language behavior can contradict the natural tendency toward entropy as speakers put more and more constraints on their selections.
'Variation on the style dimension within the speech of a single speaker derives from and echoes the variation which exists between speakers on the social dimension. This cause and effect relationship holds on three levels. First it operates synchronically for an individual speaker who, in specific situations, shifts style to sound like another speaker. Second, it operates diachronically for individual speakers who, over time, shift their general speech patterns to sound like other speakers. Third, it operates diachronically for an entire group of speakers which, over time, shifts its speech to sound like another group.' (Bell 1984:151)

Here Bell has suggested accommodation as the basic mechanism of language dynamics for both individuals and groups. Accommodating speakers who move into a new mode of interaction have changed selection patterns, but not grammars. This is indicated by the possibility of reverting back to older patterns, or employing 'inappropriate' forms for the sake of changing frame, testing reactions, or making jokes.

The 'work' that is constantly being done in speech communities is this negotiated convergence to specific sets of conventions. When interaction is weak, the conventions decay and expectations are fuzzy. But the highly predictable patterns of usage are reestablished as speakers accommodate. This is not to say that there are no conventions for interacting with strangers. But the more intimate or specialized the relationship is between speakers, the more discriminations must be made, which involve greater variety reduction, and greater information flow.

Information serves control, the manipulation of response toward some established goal. Inasmuch as speakers have goals and process information, they are able to control the speech situation. According to Ashby (1964) the basic principle of control in information systems is the need for requisite variety. In order to control the environment a system must have as many effective options to respond as there are effective constraints imposed on it.

If speakers share goals, their accommodation reduces the variety in the set of alternatives. They require less variety to control their shared environment, and so the code is restructured accordingly.

If speakers do not share goals, then they have different ideal states toward which they are trying to steer the situation. In this case, the requisite variety for maximizing control must be generated by increasing the range of alternative responses. Variation will increase and speakers will tend to diverge from shared conventions. Thus, Roloff suggests this general principle.

'As a relationship escalates from non-interpersonal to interpersonal, the variety of communication strategies will decrease; as a relationship de-escalates from interpersonal to non-interpersonal, the variety of communication strategies will increase.' Roloff (1976:189)
Speakers must have a soft touch on the brakes, however, and they generally negotiate their position between the extremes, like Schopenhaur's porcipines (Eigen and Winkler p. 153), which must stay close enough for warmth and far away enough for comfort. The signals are neither categorical nor discrete. Instead they are like the impulses that build up in a neuron until a firing threshold is reached.

In living systems, small changes can have large effects, and individuals can structure groups. In a speech community, accumulated fluctuations can result in shifts from one macroscopic state to another. In his review of Shannon, Hockett describes the effect of such a cumulated informational input as 'trigger action' (Hockett p. 47).

Variable rule analysis can tell us about the informational inputs to such triggers. But the variable rule must be a part of a larger view of interaction and social structure, including considerations of decision and control. Helpful for this larger view is an understanding of entropy and nonequilibrium.

**Entropy**

'... entropy is a macroscopic property (i.e. one which does not apply to individual atoms or molecules but only to appreciable aggregates of them) and is thus comparable to other macroscopic properties such as temperature and pressure.'

(Denbigh 1975:67)

Contrary to the accounts of some writers (Waddington 1977:143), information is not the negative of entropy (negentropy), and it should not be said that information decreases as entropy increases. It may be true that when entropy increases, an observer's ability to predict its behavior is decreased. But that is not to be confused with the capacity of signals (events) in the system to carry information. Entropy is a structural property of a system whereas information is a relation between elements or processes within a system.

This confusion is handled by Gatlin (1972:48ff) in her distinction between 'potential information' (which increases with the entropy or potential variety of a system) and 'stored information' (which increases with the degree of variety reduction of a signal). The latter is what Shannon calls 'redundancy'.

Gatlin (1972:chapter 2) also makes a distinction between two different types of stored information, or redundancy. The first is what she calls 'divergence from equiprobability'. She defines and contrasts a higher level of redundancy as 'divergence from independence'. Both can be measured in DNA and can vary independently of each other. Her measurements suggests that lower organisms attain informational complexity by increasing divergence from equiprobability. Vertebrates, however, have evolved to the highest levels of complexity by increasing divergence from independence, i.e. by imposing structure not only on the probability of DNA substructures, but of ordered sequences of these substructures (Gatlin 1972:80).
Non-equilibrium

'At equilibrium molecules behave as essentially independent entities; they ignore one another. We would like to call them 'hypnons,' 'sleepwalkers.' Though each of them may be as complex as we like, they ignore one another. However, nonequilibrium wakes them up and introduces a coherence quite foreign to equilibrium.' (Prigogine and Stenger 1984:181)

Equilibrium is the extreme independence and equiprobability of microstates. Strictly speaking, it is impossible to find in social systems. But we can discuss different social systems, or speech communities, in terms of how far from equilibrium they appear to be functioning. This degree is a matter of interaction between parts of the system itself, and interaction between the system and its environment.

'So far we have discussed isolated systems... However it is possible to extend Boltzmann's explanation to open systems that interact with their environment... Equilibrium is the result of competition between energy and entropy. Temperature is what determines the relative weight of the two factors. At low temperatures, energy prevails... At high temperatures, however, entropy is dominant and so is molecular disorder.' (Prigogine and Stengers 1984:125-6)

At high temperatures, interaction between elements is disrupted. The regularity or predictability of the ongoing relationship between any two given parts of the system is lost. You can't tell who you are going to bounce up against next, or what the nature of the interaction will be. Urban anonymity may be a social reflection of this high temperature disruption.

In these circumstances, it takes more communicating to maintain the same degree of structure. Thus it is the informationally more complex systems that can regulate their behavior and protect their integrity at the higher levels of energy flow.

The Second Law applies to closed systems. Speakers and speech communities are open systems, however, and receive informational (and energetic) input from a rich, living environment. The inputs feed the constant processes of structuration that offset the tendency toward a spread of the system into more and more equally probable, independent microstates. So certain events (and certain correspondences between events) become more predictable and law-like, to the effect of maintaining systemic identity. This structuration of probabilities is the 'life work' of living systems.

'We can isolate a crystal, but cities and cells die when cut off from their environment. They form an integral part of the world from which they draw sustenance, and they cannot be seperated from the fluxes that they incessantly transform... It is difficult to see how Boltzmann's order principle can be applied to such situations... It is the opposite of disorder, a priveleged state to which only a comparatively small number of complexions may correspond. In Boltzmann's terms, it is an 'improbable state.' (Prigogine and Stengers 1984:127)
As speakers interact, make selections, accommodate, and transmit information by strategic use of significant social variables they are organizing their community away from equilibrium. In such systems variety is high, variety reduction is high, information flow is high, signals are not equiprobable, and they are highly interdependent. We find redundancy chasing variety in a kind of inflational spiral of ever-increasing informational capacity.

Individuals

This improbable, high-flow, non-equilibrium sort of order is the result of the aggregation and amplification of purposeful individual behavior. In speech communities it is constructed from the situated interaction of speakers, negotiating their identities and relative advantages.

'A system far from equilibrium may be described as organized not because it realizes a plan alien to elementary activities, or transcending them, but, on the contrary, because the amplification of a microscopic fluctuation occurring at the 'right moment' resulted in favoring one reaction path over a number of other equally possible paths. Under certain circumstances, therefore, the role played by individual behavior can be decisive.' (Prigogine and Stengers 1984:176)

Above, we referred to the bootstrapping by which style is created by manipulating variables, which in turn are constrained by style. This constructive connection between levels of functioning according to Prigogine, is characteristic of all living systems.

'One of the most important problems in evolutionary theory is the eventual feedback between macroscopic structures and microscopic event: macroscopic structures emerging from microscopic events would in turn lead to a modification of the microscopic mechanisms. Curiously, at present, the better understood cases concern social situations.' (Prigogine and Stengers 1984:191)

The integration of the macro and micro levels of analysis is at center stage in sociolinguistics and discourse analysis (Knorr-Cetina and Cicourel 1981). In this connection, we see the need for an understanding of information theory, and of the variable rule, that emphasizes decision and control. The traditional structural interpretation won't do.

Actuation

Labov talked about how our active manipulation of the frequency of the redundant sociolinguistic variables can define styles. Minor variations don't count, but patterns do.
'If contrast exists between casual and careful styles, and the variables which we are using play a significant role in that contrast, they do not seem to operate as all-or-none signals. The use of a single variant ... does not usually produce a strong social reaction; it may only set up an expectation that such forms might recur, so that the listener does begin to perceive a socially significant pattern. Every speaker occasionally begins a (dh) word with a sharp onset, which can be interpreted as an affricate, [dθ]. However, in the prestige form of speech, these forms recur so seldom that they are negligible. Any pattern of expectation set up by them dies out before the next is heard.' (Labov 1972:108)

David Lewis long ago described how small fluctuations can grow into stable structures, noting that the source of the fluctuation is unimportant.

'A convention is produced when a big enough fluctuation meets strong enough amplifying forces.' (Lewis 1969:86)

What sorts of social interaction can amplify variation? How does the structure of the community affect the interactional dynamics of groups of speakers in it? How big of a fluctuation is big enough? These questions have long intrigued Labov and he has written of them as the 'actuation problem' (Weinreich, Labov and Herzog 1968, Labov 1980).

In discussing the social origins of sound change Labov relies on four explanations for why change does or does not result from fluctuations: population influx, imitation, standardization, and interaction networks (Labov 1980:260-64).

The first consideration is the influx of new speaker patterns. Usually such an influx serves as a source of fluctuation and the efforts of the incoming groups to retain their identity may amplify these distinct patterns. In this case, the openness of the system allows for greater variety and differentiation.

'...it is the entrance of new ethnic and racial groups into the community that provides the motivating forces behind this renewed diversification [in Philadelphia].' (Labov 1980:263)

Speakers who share identity and solidarity will tend to imitate the core members of their group. The 'emblematic function' of a pattern of variation (Labov 1980:262) encourages speakers to value certain peculiarities. The identity of a group is expressed in these patterns, as is the membership of individuals in that group. Such autocatalytic (self-perpetuating) accommodation will amplify fluctuations to the point where even the wider community grammar can be restructured.

'Once we are willing to refine our notion of prestige to give full weight to the local prestige associated with the Philadelphia dialect, Tarde's laws of imitation gain in respectability. But we must be ready to recognize that such a local prestige ... is powerful enough to reverse the normal flow of influence, and allow the local patterns to move upward to the upper middle class and even to the upper class.'
Variations, however, also tend to be dampened by the system. This is what Labov referred to in the above passage as the 'normal flow of influence'. The power of macro-level prestige standards to inhibit participation in divergent patterns is well known. Labov provides an example.

'Young black speakers do not participate at all in the evolution of the vowel system that is described here; instead, they clearly show their allegiance to a nationally based black English vernacular that is extraordinarily uniform in all the cities of the North.' (Labov 1980:263)

Fourth is the reference to the nature of the communication networks of innovating speakers.

'We have approached the problem of why sound changes take place at a particular time by searching for the social location of the innovators: asking which speakers are in fact responsible for the continued innovation of sound changes, and how their influence spreads to affect the entire speech community.' (Labov 1980:252,261)

He goes on to confuse innovation with hypercorrection, but the valid point is that the channels between people and groups in the speech community will certainly have effect on whether changes and variables are going to be amplified or dampened. The general principle is this: to the extent that boundaries are effective, internal fluctuations will be protected. On the other hand, where boundaries are ineffective, the effects of the larger system, or the environment, will be overpowering.

Thus in Philadelphia, as in Labov's earlier study of Martha's Vineyard, there were clearly bounded and rather autonomous subsystems that tended to pursue changes and amplify the variants that marked their identities.

But in situations where subgroups are highly integrated into the larger system, fluctuations will be dampened. This has been the case with immigrant speech communities that disappeared into the American 'melting pot' (Ferguson and Heath 1981:Part III). The parallel in chemical systems is described by Prigogine.

'... theoretical studies and numerical simulations show that the critical nucleus size increases with the efficacy of the diffusion mechanisms that link all the regions of systems. In other words, the faster communication takes place within a system, the greater the percentage of unsuccessful fluctuations and thus the more stable the system. This aspect of the critical-size problem means that in such situations the 'outside world', the environment of the fluctuating region, always tends to damp fluctuations. These will be destroyed or amplified according to the effectiveness of the communication between the fluctuating region and the outside world.' (Prigogine and Stengers 1984:187-88)
There are various forces at work in the speech community, dampening and amplifying the fluctuations. The balance between them determines the state of the system. Influx from outside can be a source of new fluctuations. The imitation that arises out of in-group prestige will amplify them. Standardization imposed by the larger system will dampen them, by means of rich channels of communication with the subgroups.

All these mechanisms depend on information flow, generated by selecting signals out of a range of variation. The attitudes and interactions that prevail between groups are cumulative results of these same informational processes. The degree of variation at different points in the actuation process can be measured using the concepts of information theory.

Summary

We have looked at the history of information theory in linguistics and tried to draw a picture of a new view of language and communication. The frame for this picture has been borrowed from the 'hard' sciences, but it has been painted using the palate of linguists themselves. Our approach has been an exercise in the creative mixing of metaphors, revealing the common concepts between how Prigogine describes self-organizing chemical systems and how Labov has described the dynamics of the speech community.

The informational interpretation of the variable rule has been used to unify discussions of style shifting, accommodation, variation, language change and the relation between individual and community grammar. This approach enriches the issues and provides sociolinguistics with tools, both conceptual and instrumental, for productive research.

To provide examples of the application of these ideas to actual cases of variation would demand another paper. But first the ideas need to be laid down. If the concepts are sound, the numbers solve themselves. The constructs of information theory are not just terminology, but are empirically testable in the field.

Sociolinguistics amounts to more than just using a tape recorder instead of intuition. By looking at how people use language, we arrive at a different idea of its essential nature. As in all sciences, the future lies in an appreciation of how order arises out of the interaction of indeterminate elements. Compare these two statements about science; one a prediction, the other a partial fulfillment.
'Some things never happen in the physical world because they are impossible, others because they are too improbable. The laws which forbid the first are the primary laws: the laws which forbid the second are the secondary laws. It has been the conviction of nearly all physicists that at the root of everything there is a complete scheme of primary law governing the career of every particle or constituent of the world with an iron determinism ... One would not be surprised if in the reconstruction of the scheme of physics, which the quantum theory is now pressing on us, secondary law becomes the basis and primary law is discarded. In the reconstructed world, nothing is impossible though many things are improbable.' (Sir Arthur Eddington 1958:75,98)

'... our vision of nature is undergoing a radical change toward the multiple, the temporal, and the complex. We were seeking general, all-embracing schemes that could be expressed in terms of eternal laws, but we have found time, events, evolving particles ... A new unity is emerging: irreversibility is a source of order at all levels. Irreversibility is the mechanism that brings order out of chaos.' (Ilya Prigogine 1984:292)

I would like to offer this stochastic, indeterminate, and informational view of language as an alternative to the static view of the formalists.
References.


Miller, George A. 1956. The magical number seven plus or minus two: Some limits on our capacity for processing information, Psychological Review, 1956, 63, 81-97.


Within the last fifteen years, much interest and attention have centered around the concept that children develop an adult lexicon by acquiring in their definitions of words semantic features or components progressively more similar to the adult lexicon. E. Clark (1973) synthesized various aspects of lexical development into her Semantic Feature Hypothesis, tracing the development of lexical meaning through stages of overgeneralization, starting when the child's perception of objects may be limited to one or only a few of their semantic features, through a gradual adding of features more and more specific to the meaning of a word. This finally results in the adult lexicon.

As part of her hypothesis, E. Clark predicted the confusion of antonyms and words that overlap in meaning within a particular semantic field. Based on her own research and others' (Donaldson and Balfour, 1968; Donaldson and Wales, 1970; H. Clark, 1970; E. Clark, 1971, 1972, 1973), she postulated that relational concepts which are expressed in binary pairs, such as more and less and dimensional adjective pairs, are overgeneralized by the young child to the effective early exclusion of appropriate meaning for the negative or marked member of the pair. (For discussion of positive/unmarked and negative/marked pairs, see H. Clark, 1970; E. Clark, 1973; Klatzky, et al., 1973.) For example, children between the ages of 3:0 and 5:0 are reported to fail to differentiate between the word less and the word more. In the studies cited, less was treated as if it were synonymous with more (Donaldson and Balfour, 1968; E. Clark, 1973). It was concluded that the feature of 'amount' was the salient feature, adopted first by children. In other words, some comparative adjectives can be characterized as consisting of a dimensional feature and a polarity feature, and the dimensional features are acquired before the polarity features. When the dimensional feature has been acquired but not the polarity feature, children interpret both members of a pair to refer to the positive pole. (See Clark, 1973; Townsend, 1976.)

Clark, based on the evidence for more/less and her own studies of before/after (1971), postulates that children can be expected to have similar difficulty distinguishing between the polar adjectives in such pairs as big-little, high-low, long-short, tall-short, fat-thin, thick-thin and wide-narrow. Children are expected to interpret big-little as 'size,' high-low as 'height,' long-short as 'length' and so on, with the marked or negative member of the pair swallowed up in the unmarked member, at least in younger children, before complete proficiency is achieved. Clark reported that in this semantic field, children demonstrated understanding
of the positive-pole dimensional adjectives when they responded incorrectly to stimuli relating to the negative member of the set. (E. Clark, 1973.) Donaldson and Wales (1970) reported a similar ambiguity, citing a 'consistent and obviously statistically significant finding that responses to the positive pole are superior to responses to the negative pole (p. 259).'. They, therefore, concluded that children clearly operate in terms of the polarities. H. Clark (1970) reported similar results in his studies, with subjects comprehending positive adjectives in comparative sentences more quickly than negative adjectives. He further stated that the child first learns the nominal use of polar adjectives, so that both members of a pair mean 'having extent.' He then acquires the subordinate properties of the antonyms, their contrastive use.

The other aspect considered here of the Semantic Feature Hypothesis predicted by E. Clark relates to the order of acquisition of the complete dimensional adjective pairs, from most general to more specific. This consists of an ordering, basically, from big-little through the other pairs. Big-little, she suggests, serves as a cover for the other adjective pairs, particularly in the marked member of a dimensional pair, and is, therefore, acquired early. Her order of acquisition predicts learning of the big-little pair first, concurrently with the acquisition of the pair's polarity feature, followed by tall-short, high-low and long-short, and finally by wide-narrow, thick-thin and deep-shallow, the most difficult to acquire (E. Clark, 1972, 1973). Her predictions were based upon percentages of correct responses to stimuli in each of the adjective pair categories.

Support for E. Clark's hypothesis, or parts of it, and reiteration of her predecessors' results were reported by Klatzky, et al. (1973), Layton and Stick (1979), Richards (1979), Bartlett (1976), Eilers, et al. (1974) and Ehri (1976). However, most of these observers could confirm only part of Clark's hypothesis. Bartlett (1976) and Eilers, et al. (1974), for example, both confirmed Semantic Feature Hypothesis predictions about the order of acquisition for dimensional features: more general features are acquired first. However, neither found evidence to support the notion that marked members of dimensional pairs are interpreted as having the meaning of the unmarked members. Richards (1979) also found no asymmetry in comprehension of the polar adjectives used in his 1979 study. Similarly, Townsend (1976) found no evidence for a marking explanation of children's acquisition of dimensional adjectives. He rejects a 'strong form of the marking theory, in which the polarity feature is acquired only once and then applied to all marked/unmarked pairs.' He also rejects a weak form of the marking theory 'in which the polarity feature is acquired separately for each marked/unmarked adjective pair (p. 392).'

The Current Study

As certain aspects of Clark's Semantic Feature Hypothesis are shown to be less than universally accepted, the purposes of the current study emerge. This study was designed to determine whether children interpret negative or marked dimensional adjectives as their positive counterparts.
Seven polar adjective pairs were used, some defined as positive or negative and others more strictly defined as unmarked or marked. (See Clark, 1972, 1973; Townsend, 1976.) That distinction is not dealt with in the scope of this study, and positive or negative is often used to refer to both positive and unmarked or negative and marked, although the reverse is not appropriate.

Subjects. The children used in the study included 36 children between the ages of 2:6 and 6:9. The 6:9 year old was tested out of curiosity, although it was expected that he would have mastered the pairs. He was included in the results of the study because he did not have a perfect score. The only two children who did have perfect scores were ages 4:8 and 5:10. Most of the children were found and tested at a nursery/day care center or a church nursery, but a few were tested in their homes. Most of the children were anglo-Americans, but a few were Hawaiian- or Mexican-Americans, all native speakers of standard English.

Experimental Design. Over 3000 individual responses were received from the group of 36 children. Each of the subject children was given 12 opportunities to respond to dimensional identification problems on each of the adjective pairs: big-little, high-low, fat-thin, tall-short, long-short, wide-narrow and thick-thin, a total of 84 responses per child. Although original provision was made for a 'don't know' category, all the children responded with either a definite right or wrong answer. A slight amount of vacillation was observed in less proficient children, but even they made a final, definite choice in each case.

The study was designed to measure comprehension only, without regard to production. Since the debate continues over whether performance or comprehension serves as the best indicator of competence (Donaldson and Wales, 1970; Ehri, 1976; Cocking and McHale, 1981; Layton and Stick, 1979), an arbitrary decision was made to test comprehension. Certainly both must be understood to contribute to overall competence. However, a simple comprehension task might be considered uniformly comparable over a broad range of ages, whereas complex production tasks might tend to favor older children. This study was designed so that no verbal responses were required. In fact, most of the children, both older and younger, did not choose to respond verbally at all, but pointed to or touched their choice of answer. This gave no advantage to the older children. Care was also taken in avoiding advantages for the older children in how the responses were elicited. For example, words such as choice, which might not be familiar to the child aged 2:6, were not used. Instead, the child was asked to 'show me,' 'point to' or 'touch' his choice of answer. The task was also approached as though it were a game and most of the children seemed to enjoy it. Many, in fact, were eager to repeat it. Some of the very youngest, however, were taken to the limits of their attention span, and a few had difficulty concentrating through the last set of questions.

Materials. In designing the experiment, questions as to the appropriateness of two- or three-dimensional testing materials were considered.
Both object and picture stimuli have been used in linguistic elicitation activities. Consideration of this subject has suggested that actual objects, or naturalistic stimuli, were more likely to elicit correct responses from children than were pictures of objects (Layton and Stick, 1979). In a recent study, Cocking and McHale (1981) designed an experimental setting in which they measured the performance of children on production tests using objects, object-choice, and picture stimuli. They also evaluated children's responses on comprehension tests using the same three kinds of stimuli. Cocking and McHale reported that with respect to picture-object differences, children showed their best comprehension performance when pictures were used.

The present study was designed to utilize pictures as the only stimuli in the comprehension testing situations developed. The pictures were divided into three major categories: fruit, houses and clowns. Simple line drawings of the objects were manipulated to feature the desired dimensions. When the specified dimension would not normally be used to describe the particular object, additions were made to the picture and explained to the children. For example, since long-short would not normally be associated with individual clowns, the child's attention was directed toward the clown's hair, which was distinctly either long or short.

Methods. Forty-two drawings were used to elicit the 84 responses, 14 in each object category. In the first dimensional adjective category, the children were shown a line drawing of two fruit and asked, 'Which fruit is bigger?' or 'Which one is bigger?' If the child showed any hesitation in responding, he was told he could point at or touch the one he chose. After a response was given, the child was shown a second picture, different from the first, and asked 'Which one is smaller?' The testing continued through higher, lower, fatter, thinner, taller, shorter, longer, shorter, wider, narrower, thicker and thinner, all treated individually and all in the 'fruit' category. The same sequence was then followed in the 'house' category and then the 'clown' category, referring each time to a single dimension, in the comparative. Then the entire procedure was repeated in the superlative mode, starting again with fruit and asking the child, 'Which fruit is biggest?'

The procedure of using the superlative structure while referring to only two items has correlates in the literature (Townsend, 1976). It is also similar to the grammatical license used by Donaldson and Balfour (1968) and Donaldson and Wales (1970) and subsequent others in their more-less testing, using less where the grammatically correct structure would be fewer rather than less. The strict interpretation of grammatical appropriateness was not deemed necessary in their studies because of the linguistic immaturity of the subjects.

The questions were asked in exactly the same order each time, but the location of the correct responses on the right or left side of the paper varied randomly. Also, in each case, the children were given encouraging but non-commital feedback. There was considerable distance between repetitions of questions in each adjective category, each repetition separated by 12 stimuli relating to the other adjectives.
A smaller, eight-item survey of more-less was attached to the beginning of the study for purposes of comparison. A total of 288 responses were elicited regarding the more-less phenomenon. Each child was given four opportunities to respond correctly by pointing to an object with more items and four opportunities to point to an object with less items. Again, the children were shown pictures. Results were expected, according to Clark (1973) to correspond quite closely with the results on the dimensional adjective tasks.

Results

The most obvious and striking statistic to emerge from analysis of the accumulated 3024 responses on dimensional adjectives was the overall average of 78 percent right on all positive pole or unmarked dimensional adjectives, compared to an overall average of 70 percent right on all negative pole or marked adjectives. This finding, in direct contradiction to Clark's predictions and the results of some earlier tests, suggests a more detailed analysis of the data, given so much variability in age and the relatively large sample in terms of children and questions.

An arbitrary division of the 36 children into three groups of twelve was made. Data were arranged in order of proficiency with the twelve most proficient in Group I, the twelve least proficient in Group II and the intermediate children in Group III.

The purpose for the division along proficiency lines rather than according to age was that proficiency obviously was not determined solely by age. (See Townsend, 1976; Ehri, 1976.) For example, in the most proficient group, children ranged in age from 2:11 to 6:9, the child aged 2:11 missing only four of the 84 possible points. She obviously was more proficient than some near three-year olds, and her responses would have biased the results had she been categorized with them.

Other subjects in Group I were 3:8, 4:0, 4:2, 4:8, 4:11, 5:1, 5:9, 5:10, and 5:10, with a mean age of 4:8. Proficiency ranged from 0 to 11 errors out of 84 possible. Children in Group II ranged in age from 3:1 to 5:0, with a mean of 4:0 years. Scores ranged from 12 to 21 wrong. Group III ranged in age from 2:6 to 3:8, with a mean of 3:1 years. Their scores ranged from 23 to 41 wrong. The analysis for average correct positive pole or unmarked and average correct negative pole or marked dimensional adjectives is illustrated in Table 1. It will be noted that nothing approximating Clark's (1973) prediction of comparability to the more-less data (91 percent positive pole correct, 72.7 percent negative pole incorrect) was observed.

To further refine the analysis, however, it is appropriate to consider the data on each individual adjective pair, across groups and within groups. Within each group, the responses are analyzed in terms of correct responses to elicitations using both the comparative and the superlative form. The
Table 1. Average correct positive/unmarked and negative/marked responses across 14 dimensional adjectives, according to proficiency groups.

<table>
<thead>
<tr>
<th></th>
<th>Group I</th>
<th></th>
<th>Group II</th>
<th></th>
<th>Group III</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Pole/</td>
<td>% right</td>
<td>S.D.</td>
<td>% right</td>
<td>S.D.</td>
<td>% right</td>
<td>S.D.</td>
</tr>
<tr>
<td>Unmarked</td>
<td>95.2</td>
<td>7.9</td>
<td>84.5</td>
<td>18.5</td>
<td>55.5</td>
<td>20.8</td>
</tr>
<tr>
<td>Negative Pole/</td>
<td>% right</td>
<td>S.D.</td>
<td>% right</td>
<td>S.D.</td>
<td>% right</td>
<td>S.D.</td>
</tr>
<tr>
<td>Marked</td>
<td>91.5</td>
<td>13.7</td>
<td>75.5</td>
<td>23.1</td>
<td>43.0</td>
<td>14.6</td>
</tr>
</tbody>
</table>

responses in Group II approximated those of Group I for five out of the seven dimensional adjective pairs (big-little, high-low, fat-thin, tall-short, long-short), so they will be grouped together for analysis and illustration with Group I. They will be grouped for analysis with Group III for the final two pairs (wide-narrow, thick-thin). Table 2 contains the illustration of this comparison.

It will be noted from Table 2 that the cluster group 1&2 (Groups I and II) differed less than 10 percentage points between the correct responses for the positive/unmarked and negative/marked polar adjectives in the first ten adjectives (five pairs), in the comparative, although the fat-thin pair came within 0.3 percentage points of a 10 percent difference. Similarly, the only pair separated by more than 10 percentage points in the superlative was fat-thin, with a difference of 19.4 percent. That is still considerably less than the difference of 63.7 percent predicted by Clark. Interestingly, the only other notable discrepancy between positive and negative adjectives in the most proficient group was for thick-thin, in which the negative adjective is the same as in the fat-thin pair. With 83.3 percent for positive compared to 59.7 for negative, however, the difference of 23.6 percent is still significantly narrower than predicted.

Ehri (1976) observed that where the negative adjective of a pair has a dual role, like the use of thin in association with both thick and fat, its attachment to one of the positive adjectives might be delayed. She observed that with dark-light and heavy-light, light is delayed in its attachment to heavy, presumably because of its dual role. Such an explanation might account for the delayed attachment of thin to fat and more particularly to thick, but then a similar phenomenon might be expected with regard to long-short and tall-short, where no delay is noted. The obvious difference between the thin pairs and the short pairs is their placement in Clark’s order of acquisition, with thick-thin being acquired near the end, while long-short and tall-short are acquired at about the same time.
<table>
<thead>
<tr>
<th>Adjective</th>
<th>Group</th>
<th>% Right Comparative</th>
<th>S.D.</th>
<th>% Right Superlative</th>
<th>S.D.</th>
<th>% Right Total</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big</td>
<td>142</td>
<td>100.0</td>
<td>0</td>
<td>100.0</td>
<td>0</td>
<td>100.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>58.0</td>
<td>8.3</td>
<td>97.2</td>
<td>4.8</td>
<td>76.4</td>
<td>21.5</td>
</tr>
<tr>
<td>Little</td>
<td>142</td>
<td>95.8</td>
<td>4.6</td>
<td>98.6</td>
<td>3.4</td>
<td>97.2</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>41.7</td>
<td>8.3</td>
<td>66.7</td>
<td>14.4</td>
<td>54.2</td>
<td>17.3</td>
</tr>
<tr>
<td>High</td>
<td>142</td>
<td>95.1</td>
<td>9.7</td>
<td>94.4</td>
<td>6.8</td>
<td>95.8</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>47.3</td>
<td>17.3</td>
<td>80.6</td>
<td>4.8</td>
<td>63.8</td>
<td>21.5</td>
</tr>
<tr>
<td>Low</td>
<td>142</td>
<td>94.4</td>
<td>8.6</td>
<td>94.4</td>
<td>6.8</td>
<td>94.4</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>41.7</td>
<td>14.4</td>
<td>63.9</td>
<td>33.7</td>
<td>52.8</td>
<td>26.2</td>
</tr>
<tr>
<td>Fat</td>
<td>142</td>
<td>97.2</td>
<td>6.8</td>
<td>98.6</td>
<td>3.4</td>
<td>97.9</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>30.6</td>
<td>4.8</td>
<td>66.7</td>
<td>14.4</td>
<td>48.6</td>
<td>21.9</td>
</tr>
<tr>
<td>Thin</td>
<td>142</td>
<td>87.5</td>
<td>11.5</td>
<td>79.2</td>
<td>22.8</td>
<td>83.3</td>
<td>17.8</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>36.1</td>
<td>26.8</td>
<td>44.4</td>
<td>25.5</td>
<td>40.3</td>
<td>23.8</td>
</tr>
<tr>
<td>Tall</td>
<td>142</td>
<td>95.8</td>
<td>4.6</td>
<td>95.8</td>
<td>4.6</td>
<td>95.8</td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>38.9</td>
<td>9.6</td>
<td>50.0</td>
<td>7.5</td>
<td>46.3</td>
<td>9.4</td>
</tr>
<tr>
<td>Short</td>
<td>142</td>
<td>95.0</td>
<td>7.4</td>
<td>95.8</td>
<td>7.0</td>
<td>95.8</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>36.1</td>
<td>4.8</td>
<td>55.6</td>
<td>4.8</td>
<td>45.8</td>
<td>11.5</td>
</tr>
<tr>
<td>Long</td>
<td>142</td>
<td>98.6</td>
<td>3.4</td>
<td>100.0</td>
<td>0</td>
<td>99.3</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>41.7</td>
<td>8.3</td>
<td>75.0</td>
<td>8.3</td>
<td>58.3</td>
<td>19.7</td>
</tr>
<tr>
<td>Short</td>
<td>142</td>
<td>97.2</td>
<td>6.8</td>
<td>94.4</td>
<td>6.8</td>
<td>95.5</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>25.0</td>
<td>8.3</td>
<td>58.3</td>
<td>22.1</td>
<td>41.7</td>
<td>23.6</td>
</tr>
<tr>
<td>Wide</td>
<td>1</td>
<td>79.2</td>
<td>5.9</td>
<td>86.0</td>
<td>9.6</td>
<td>85.3</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>263</td>
<td>40.3</td>
<td>15.3</td>
<td>57.0</td>
<td>17.0</td>
<td>48.6</td>
<td>17.7</td>
</tr>
<tr>
<td>Narrow</td>
<td>1</td>
<td>86.1</td>
<td>4.8</td>
<td>86.1</td>
<td>4.8</td>
<td>86.1</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>263</td>
<td>43.1</td>
<td>23.8</td>
<td>50.0</td>
<td>21.1</td>
<td>46.5</td>
<td>21.8</td>
</tr>
<tr>
<td>Thick</td>
<td>1</td>
<td>83.3</td>
<td>8.3</td>
<td>83.3</td>
<td>14.4</td>
<td>83.3</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td>263</td>
<td>47.2</td>
<td>17.6</td>
<td>70.8</td>
<td>14.7</td>
<td>59.0</td>
<td>19.0</td>
</tr>
<tr>
<td>Thin</td>
<td>1</td>
<td>58.3</td>
<td>22.1</td>
<td>61.1</td>
<td>26.8</td>
<td>59.7</td>
<td>22.0</td>
</tr>
<tr>
<td></td>
<td>263</td>
<td>29.2</td>
<td>18.8</td>
<td>30.5</td>
<td>8.6</td>
<td>29.7</td>
<td>14.0</td>
</tr>
</tbody>
</table>

Table 2. Comparison of percentages of correct responses—comparative, superlative, and total—in clusters of children Groups I, II and III, for fourteen individual dimensional adjectives, positive and negative poles.
For the least proficient group, analysis is not so simple. Differences between positive and negative pole adjectives, in the comparative, exceeded 10 percent in only two cases: long short at 41.7 percent and 25.0 percent, and thick thin at 47.2 percent and 56.1 percent. The negative pole occasionally exceeded the positive in correct responses: fat thin at 50.6 percent and 56.1 percent, and wide narrow at 40.3 percent and 15.1 percent, in the comparative. In the superlative, tall short negative also exceeded the positive at 50.0 percent and 53.6 percent. However, in the superlative, the positive polar adjectives exceeded the negative adjectives by more than 15 percent in five of the seven adjective pairs. In every case, the superlative adjective elicited more accurate responses in the least proficient group than did the comparative adjectives. Clark (1973) and Townsend (1976) mildly anticipated this trend, while Layton and Stick (1979) predicted the opposite. Correctness was improved from 50.0 percent to 97.2 percent in the most extreme case (47.2 percentage points). The average improvement across the 14 adjectives was 24.3 points. Naturally, this significantly affects the total percentage correct, both in the positive and negative adjectives.

The comparative and superlative percentages correct for this group, along with the total percentages, are illustrated in Table 3.

<table>
<thead>
<tr>
<th>big-litt.</th>
<th>high-low</th>
<th>fat-thin</th>
<th>tall-short</th>
<th>long-short</th>
<th>wide-narr.</th>
<th>thick-thin</th>
</tr>
</thead>
<tbody>
<tr>
<td>C*</td>
<td>50.0/41.7</td>
<td>47.2/41.7</td>
<td>30.6/36.1</td>
<td>38.9/36.1</td>
<td>41.7/25.0</td>
<td>40.3/43.1</td>
</tr>
<tr>
<td>S*</td>
<td>97.2/66.7</td>
<td>80.6/63.9</td>
<td>66.7/44.4</td>
<td>50.0/55.6</td>
<td>75.0/58.3</td>
<td>47.0/50.0</td>
</tr>
<tr>
<td>T*</td>
<td>73.0/54.2</td>
<td>63.8/52.8</td>
<td>48.6/40.3</td>
<td>46.3/45.8</td>
<td>58.3/41.7</td>
<td>48.6/46.5</td>
</tr>
</tbody>
</table>

* C=comparative, S=superlative, T=total

Table 3. Comparison of percentage correct for each of fourteen dimensional adjective responses elicited in the comparative and superlative forms, and total percent right for each adjective, for children in the least proficient group.

Tables 1, 2, and 3 also illustrate, based on percentage right of each adjective pair, the naturally occurring acquisition order of the dimensional adjectives. As predicted by Clark (1972, 1973), big little has the largest percentage of correct responses, both for the more and less proficient groups. Her prediction that the next sequence includes tall short, high low, and long short also seems to be confirmed by the data, at least in the case of the more proficient group. Her prediction that wide narrow and thick thin will be mastered last has similar verification. (Also see Layton and Stick, 1979.) In other words, it appears that terms which describe overall size are acquired before terms which describe height and length, and these are learned before those which describe width. (See Bartlett, 1976.)

Differences between positive pole correct and negative pole correct were in each case considerably less than predicted. In thick thin responses,
both for the more and less proficient groups, the greatest variation was noted. However, even in the most extreme case, 59.0 percent compared to 29.7 percent, the difference was less than half that predicted by Clark. A look at each individual child's responses also failed to produce even one who substantiated Clark's predictions. Although one child's responses favored the positive pole by around 40 percent, another favored the negative pole by over 35 percent. Most children were within 10 or 15 percentage points difference on most adjectives, even in the least proficient group.

A thorough analysis of the limited more-less data was not attempted. However, an impressionistic analysis suggests little correlation between performance on the two tasks. Approximately 33 percent of the children performed close to predicted on the more-less task, and some confusion was noted in others. However, often those who missed 100 percent of the less questions did well on the adjective questions. Only two children seemed to suggest a definite correlation between performance on the two tasks.

Conclusion

The data accumulated in this study do not confirm Clark's Semantic Feature Hypothesis with respect to the polarity concept of feature acquisition. The accuracy scores for positive and negative polar adjectives are so similar that this data would tend to disconfirm Clark's notion of the importance of polarity acquisition. Where confusion was observed, it seemed to be more general than polarity related. However, it must be conceded that a slight preference for the positive pole was observed. The data do, however, tend to support Clark's prediction that more general features are acquired first. The nature of the study, however, was not such that it can be interpreted as a strong confirmation of that aspect of the Hypothesis, either; especially considering the scores in Group III.

It was assumed when the study was being designed that no substantial difference between comparative and superlative forms would be noticed in the performance on the dimensional adjective task. The strong statistical finding in this regard was unexpected. Whether the children's responses were better because the superlative was presented last, or whether the responses are attributable to the form in which the adjective was presented is not known (Layton and Stick, 1979). It seems obvious that the form, not the base lexicon, affected the difference. Future studies seem indicated in order to test this phenomenon more specifically.

Finally, it would seem from the results of this study that the polarity prediction of the Semantic Feature Hypothesis does not take into account the individual learning differences noted in children, nor does it allow the child sufficient flexibility in sorting and narrowing his lexical meanings. While a strategy employing definition in terms of polarities may be employed by some children, in others it seems to be missing. As suggested by Eilers (1976), while confusion based on polarity does apparently occur, it cannot be considered a universal phenomenon. In any case, adoption of the polarity aspects of the Semantic Feature Hypothesis is not consistent with the results of this study.
References


