Lenguas Modernas 12 (1985), 84-104
Universidad de Chile

ON SPEECH FORMULAS AND LINGUISTIC COMPETENCE

ANDREW PAWLEY
University of Auckland

> These talks pointed me in the right direction but they did not prove as successful as I had hoped... My project demanded a richer and more subtle mastery of the language.
> (Marjorie Shostak, *Nisa: The Life and Works of a !Kung Woman*)

> There are many matters of interest to our society and to our species in which language is of central importance ... I do not believe that the current paradigm of linguistics provides a good basis for dealing with some, or indeed with most, of these matters. I particularly want to call into question the view of the nature of language which underlies that paradigm.
> (George Grace, *The Linguistic Construction of Reality*).

"Linguistics! How interesting! And how many languages do you know?"[1] A conversational problem for which every linguist craves a simple solution. The questioner expects, in a single word, an answer that is at once friendly, honest and informative. I must admit that I have never managed to work out a reply combining all these Gricean virtues. One reason is that I am not sure when it is true to say that one 'knows a language'. We speak of various *kinds* of linguistic knowledge or competence: "Reading knowledge required", "conversational fluency essential", "X understands Spanish but can't speak it", "Y speaks a dialect, not the standard", and so on. We recognize *degrees* of competence: "P speaks pretty good Spanish", "Q understands a little German", "R speaks with a foreign accent". And in linguistics special technical notions of linguistic and bilingual competence have developed.

This, then, is the first question that I will be concerned with here: What does it mean to know a language? In particular, what sorts of knowledge or skills make up the command of a language shown by mature native speakers? I refer here to spoken not written language.

The answer is, of course, going to depend on what is meant by 'know' and 'language'. Vernacular understandings of these terms differ from the definitions given or assumed in many studies in theoretical and descriptive linguistics. I will be primarily concerned here with senses of 'know' and 'language' which I believe to be widely

[1]This paper is based on lectures given at the 1985 Institute of the Linguistic Society of America, Georgetown University, June 24 - August 2. A version was read at the 6th Congress of the Linguistic Society of New Zealand, Wellington, August 1985. Many students and colleagues have contributed ideas and criticisms; in particular I am indebted to Dwight Bolinger, Colin Bowley, Ralph Bulmer, Wallace Chafe, Chris Corne, Peter Crisp, Charles Fillmore, George Grace, John Herbert, Robin Hooper, Paul Kay, Chris Lane, Frank Lichtenberk, Alzira Macedo, Ann Peters, Wolfgang Sperlich and Frances Syder.

accepted by ordinary users of English, and which encompass certain linguistic capacities that are excluded from the grammar-lexicon treatment characteristic of structural linguistics.

My argument will be that mastery of a language in this broad, vernacular sense rests largely on command of a lexicon of conventional ways of talking about particular subject matters. The phrase 'ways of talking' is borrowed from George Grace, for whom it is a technical term in his account of how we construct conceptual worlds by means of language. In order to speak or understand a language properly, Grace says, it is necessary to know "how to talk about the subject matter of which our effective worlds are constituted" (1984:131); we must know how things are said[2]. The 'how' falls under the heading of idiomatology, as well as grammar, of forms of expression characteristic of native speech. The 'things' we say may belong to the realm of idiomatology, too, but on the side of ideas, meanings and subject matters that are characteristic of a particular cultural tradition.

I suspect that this argument will come as no surprise — may even sound old hat —to people whose job is teaching foreign languages or doing translation. I would like to think that the general idea is close to being an obvious truth. Certainly it is an idea or assumption which is implicit in the practice of lexicography, as exemplified in the great dictionaries of European languages —a point which I will return to later.

The problem is to understand the details of this 'truth' and to see if it can be married to the 'truths' of structural linguistics. Whether these things can be done, and how they might be done are the second and third questions which will be addressed.

The grammar-lexicon model of language

I believe the problem is partly one of understanding different views of what a language is.

The terms 'language , 'competence' and 'lexicon' each has a technical meaning in structural linguistics, standing for a more or less well-defined theoretical construct. The reader may feel that I have begun to play fast and loose with these terms, by using them in non-technical senses. However, my intention is simply to propose that certain metalinguistic concepts of the vernacular culture be developed into technical constructs. Obviously, it is desirable to avoid using the same term for a pair of different though related concepts. In some cases, e.g. that of 'dictionary' versus 'lexicon', a distinction between vernacular and structuralist labels can easily be made. But not always. It is hard to find a more convenient name for what laymen mean by 'language'

[2]Grace (1981, 1984) discussed at some length what is involved in saying something in a human language. It is an achievement of quite a different order from the hinting or indicating which is the nearest other mammals come, in nature, to saying things.

Central to an act of saying something is analytically specifying a conceptual situation, e.g. by a clause which says who does what where and to whom. But it is not the case that the situations we describe are given by nature, and that our utterances are simply word-maps of actual scenes and events. Even when speaking of the most concrete situations, e.g. *A man was bitten by a dog*, our utterances follow processes of selective perception, interpretation and structuring of information. Most of the situations which we talk about, however, are better described as products of human imagination and language.

itself. It is, however, easier to find a suitable substitute for the sense of 'language' which has been constructed by theoretical linguists, namely, 'grammar'.

Leaving aside phonology, it is fair to say that linguistics has throughout its history been largely concerned with the study of grammar— either morphology or syntax. Around this focus on grammar has developed a certain idea of what a language consists of. This idea — call it the grammar view of language — has its roots in Greek and Roman scholarship and has persisted, with formal refinements, in 20th century work. In this view a language is essentially a system of conventions for pairing the form of sentences with their intrinsic meanings — those which belong to the forms without reference to wider discourse context or to extralinguistic factors.

In recent decades, Chomsky's definition of 'linguistic competence' as the native speaker's knowledge-in-principle of a generative grammar for a particular language has been widely accepted. A generative grammar specifies all the grammatical sentences of a language, while excluding ungrammatical strings, and it assigns to each of an infinite range of sentences a structural description, indicating how competent speakers understand the sentence. A major component of the grammar is a lexicon, which contains all and only the unpredictable pairings of form and meaning, i.e., morphemes and various sorts of idioms. Lexical items (or lexemes) each have certain grammatical features or categorizations which determine whether or not they can occur in certain syntactic environments. Call this the grammar-lexicon model, the dominant exponent of the grammar view.

While other schools within structural linguistics use different terms the essentials of the grammar-lexicon model are characteristic of the structuralist tradition as a whole. Until recently, a descriptive linguist who had compiled both a grammar and a lexicon of language X could sit back with a satisfied sigh and say "My job is done; I have described the language".

The grammar-lexicon view of language has been an extremely fruitful one, providing a framework for making sense of a vast amount of data. However, it is important to remember that it and its various ingredients are no more than constructs. 'Language', like 'grammar', 'lexicon'. 'syntactic rule', etc. as defined by any school of linguistics, or anyone else, is simply a hypothetical entity, a theoretical notion. The construct should not be reified and consecrated. It is true that there is some agreement between diverse observers as to what bits of actual behaviour count as linguistic or language-related. But the domain of language may be defined narrowly or broadly, and each definition may yield a universe that can be made sense of in many ways.

Vernacular understandings of knowing a language

In *Nisa, The Life and Works of a !Kung Woman*, the anthropologist Marjorie Shostak writes as follows of her first months adjusting to Kalahari life and trying to learn the !Kung language.

> I learned two important phrases: "What is the name for that?" and "when I do this, what is it called?" Using these two questions, combined with some obvious gestures, I was able to elicit a large vocabulary in a few weeks. ... I struggled to write down the answers, which bore no relation to any sounds I had ever heard before.

> By the time three months had passed, the !Kung sounds had become a little less strange, the clicks, glottal stops, and fricatives a little more manageable, and I could finally hear the tones (although it took close to twenty months before I started using them correctly). Polite formalities could now be exchanged: "How are you this morning, my niece?" "Why I'm fine, my uncle."
>
> But more was required than finding the !Kung equivalents for English words; it was, of course, necessary to say them correctly, at the appropriate times, and in the culturally accepted manner. ... By the time I had lived there six months ... I was more at ease with the language and occasionally succeeded in asking questions that elicited personal answers. At that point I started to "interview" anyone who was willing and friendly, trying various ways to ask them about the things they perceived as important in their lives... These talks pointed me in the right direction, but they did not prove as successful as I had hoped... My project demanded a richer and more subtle mastery of the language. (Shostak 1983: 17-18).

Shostak goes on to describe a progression in her quest of a "richer and more subtle mastery of the language", of becoming more fluent, of learning to frame questions appropriately, of when to be direct and indirect, of how to interpret the meanings and motives of utterances framed in a certain way or spoken in a certain context. After some two years she was able to elicit and translate the narratives in which the woman Nisa vividly recounted her life and philosophy.

From this and other linguistic anecdotes, we can gain impressions of ordinary language users' understandings of what it takes to know a language. I believe that, in English-speaking communities and elsewhere, the following elements are widely held to be part of knowing a language properly. A proper speaker should show nativelike standards in respect of:

1. a) grammaticality,
   b) pronunciation of consonants and vowels,
   c) musical conventions: intonation, stress and rhythmic patterns, voice quality, modulations of volume, etc.
   d) productive fluency: conforming to norms of tempo, structure and quantity for chunking utterance elements into fluent units,
   e) hearing fluency: being able to decode fluent speech,
   f) idiomaticity: the selection of familiar, nativelike ways of saying things, as opposed to ways that are merely grammatical,
   g) lexical knowledge: including the ability to distinguish between those expressions that are lexicalised (standard designations) and those that are ad hoc descriptions,
   h) contextual appropriateness: saying the right thing at the right time,
   i) coherence: saying things that make sense in terms of normal understandings of the world shared by a particular speech group, and in terms of standard procedures of inference,
   j) inference: being able to make sense of ordinary discourse, to work out conversational implicatures, to understand the communicative functions or intentions of particular utterances,
   k) creativity, of various kinds, including:
      i) phonological - making up new word forms,
      ii) syntactic, (iii) semantic, (iv) contextual — apt matching of expression with

situational context in a non-routine way. A distinction (not sharp) may perhaps be drawn between rule-governed creativity and special kinds of creative use of language in which conventions are broken or manipulated to achieve special effects; as in Pig-Latin, puns, metaphors, etc.

I do not wish to imply that these elements are all sharply distinct and independent, merely that each is an aspect of linguistic competence that is nameworthy — distinctive enough to be categorized separately for purposes of discussion.

Speech formulas

The suggestion was made earlier that mastery of a language, as this is understood by ordinary language users, rests to a considerable extent on knowing conventional ways of talking about particular subject matters. Can we make this suggestion into a more precise claim? While we have defined some of the characteristics of native command of a language, we have so far said little about what is meant by 'conventional ways of talking'. How do these differ from, say, words, idioms or grammatical constructions?

I will be concerned here with a particular class of ways of talking which I will call *speech formulas* or, simply, *formulas*[3]. A speech formula is a conventional pairing of a particular *formal construction* with a particular *conventional idea* or *idea class*. An idea is conventional in the sense of being part of a cultural tradition — belonging to the body of conceptual elements shared by members of a speech community. I use 'idea' rather than 'meaning' or 'function' in order to emphasize that the conceptual element is not simply a byproduct or function of the occurrence of a particular linguistic form. Not only is the idea familiar, but it has a certain independence from form, perhaps being expressible by a number of different forms without losing its status as 'the same idea'. For example, *go in* and *enter* may express the same idea. So, too, may *He pleaded guilty* and *He entered a plea of guilty*, or *John asked Mary to marry him, John asked Mary to become his wife*, and *John proposed to Mary*. In these cases the idea itself is a social institution. In a speech formula a conventional idea is expressed in a certain standard way, i.e., by a construction form which is recognized as being the usual way of designating the idea, in contrast to other less usual or ad hoc ways of referring to it.

A (formal) construction consists of two or more syntactic units standing in construction, e.g. as a complex word, phrase, clause, etc. In a formulaic construction, the syntactic units may be (i) lexically specified in all cases, i.e. the construction is made up of a fixed sequence of morphemes, such as *A stitch in time saves nine*, or *Every man has price*, (ii) lexically specified in some cases, as in the formulas *Long live NP*! and *The* ADJ-*er* (X), *The* ADJ-*er* (Y), discussed below, or (iii) lexically unspecified, i.e. made up of syntactic categories such as V, N, NP, etc. In the last case, however, one or more of the

[3] 'Formula' or 'speech formula' has been used in roughly the present sense by many scholars including Jespersen (1924), without ever becoming a major analytic construct in linguistics. Recently the term 'construction' has been used in a similar sense. I prefer 'formula' to 'construction' —chiefly because, whereas 'construction' carries strong associations with a syntactic type, a string of abstract syntactic categories, 'formula' suggests a recipe for putting things together, where the 'things' may be ideas, syntactic categories or particular words.

syntactic categories is lexically very restricted; it represents a small class of lexemes, usually one that forms a semantic set, as in the time-telling formula M PREP H, discussed below.

A particular realization of a formula may be termed a *formulaic expression*, or, loosely, a *formula*.

As an example of a formula, Jespersen (1924: 30) gives *Long live the King*! and its congeners *Long live the Queen/the President/Mr Johnson*, etc., which are used to publicly wish someone well. The formula *Long live NP*! (cf. *Long may she reign*!) has the unusual structure ADVERB SUBJUNCTIVE-VERB SUBJECT; this construction type is no longer productive in English, except in having a subject constituent which can be rather freely varied.

Easy does it! is another formulaic expression, belonging to a small class in which only the first element is variable (*easy, gently, slowly* and perhaps a few others), the whole being used as a directive to someone at the critical phase of a delicate operation. It is not only the placement of the adverb first that is unusual here; it is also the use of the third person singular form of the verb as part of a command. We do not say **Quickly runs it*! or **Carefully cuts this page*!

A formula whose syntactic constituents are wholly or largely filled in by particular morphemes or words is a *lexicalised* or *partly lexicalised formula*. *Long live NP*! and [*Easy/gently/slowly*] *does it*! fall into this class. A formula whose syntactic elements wholly or largely consist of syntactic categories may be called a *schematic* or *general formula*.

An example of a schematic formula in English is the simile formula NP_i V *like* NP_j, as in *She danced like an angel*. Another is the comparative formula *The* ADJ-*er* (X), *the* ADJ-*er* (Y) as in *The bigger the better, The bigger they are the harder they fall, The sooner she arrives the sooner we can go*. Ways of telling the time and reckoning weight provide other examples. In my dialect of English, for instance, one says that the time is M *to* H, or M *past* H, where M is the number of minutes and H is the hour. M may be reckoned in terms of minutes or in units of 'a quarter (*to* or *past*)' or 'half (*past*)', but not in, say, 'sixths' or 'thirds' of an hour. The class of prepositions linking M and H is highly restricted in English; in some places people say *before* and *after* instead of *to* and *past*.

Formulas have something in common with the familiar analytic constructs 'idiom', 'lexical item' and 'grammatical construction' but are not adequately covered under any of these three headings[4]. All idioms are formulas but not all formulas are idioms (in the strict sense of a construction with an unpredictable meaning or irregular form); most are not idioms. *Long live NP*! and *easy does it*! happen to be syntactic idioms (with odd phrase structure), of which there are relatively few. *Drop a brick*, in the sense of 'commit a social gaffe, behave tactlessly', happens to be a semantic idiom, of which there are many thousands.

The bulk of formulaic expresions, however, are syntactically and semantically well-formed. For example, *tell the truth* is a well-formed expression which names a

[4]On 'construction' see fn. 3. For a discussion of various senses and kinds of 'idiom' see Weinreich 1969, Makkai 1972. Weinreich points out that the idioms of a language are not a well-defined class; rather, 'idiom' is a theory-dependent notion. It can be argued that speech formulas are all, at least, 'encoding idioms' (Makkai 1972), if not 'decoding idioms'. That is, while the meaning of a given formulaic expression may be transparent to the hearer who knows the grammar, the grammar provides no clue that that particular expression is *the* conventional one for that meaning.

conventional idea. It happens that in English there is no single word antonym of *lie*, and while it is possible to say *speak truthfully, give a true account, say what is true, state the facts*, etc., the phrases *tell the truth* and *speak the truth* have greater status as official antonyms of *lie* — indicated not only by their frequency of use in everyday speech but by their required use in courtroom and other rituals. That the meaning of *tell the truth* is part of our cultural tradition cannot be doubted when a man can be sent to jail for lying to the court.

Many formulas have a grammar that is partly unique. There are special restrictions as to how constituents may be moved, inserted, expanded, etc. without destroying the formula. For example, in the formula *Who* (EXPLETIVE)

Who do-TENSE$_t$ NP$_i$ *think*-TENSE$_t$ PRO$_i$ *be*-TENSE$_t$!

the tense cannot be varied from SIMPLE PRESENT or PAST without changing the force of the expression. We can say, as an indignant reaction to someone's presumptuous conduct, *Who do you think you are*!, *Who the devil do they think they are*!, but not **Who do you think you will be*!, or **Who did they think they used to be*! The last two expressions are grammatical as ordinary questions, craving an answer, but they do not carry the sense of a strong objection to someone's behaviour, which is characteristic of the formulaic usages. Again, one cannot insert modifiers freely (as in *Who, in this particular case, did they apparently think they were*?) without losing the formulaic sense.

Writers on formulas, and related ideas such as 'collocation' and 'compound lexemes' sometimes define these units solely in terms of their syntactic unity, their fixedness, in contrast to 'free constructions'. But in most cases such a formal definition is meaningless, in fact invalid, unless the formal restrictions in question are associated with a particular meaning or class of meanings or speech functions. *Good morning* is a formulaic expression only when it serves as a greeting rather than as a description of the weather.

The role of formulas in linguistic competence

We turn now to the relationship between speech formulas and the various aspects of linguistic competence distinguished in (1 a-k) above. To deal in depth with any one of these aspects would take a lengthy paper, and here I must confine myself to summary remarks on a few of them.

An important function of speech formulas is captured in *Webster's* definition for one of the senses of 'formula': "any conventional rule or method for doing something, especially when used, applied or repeated *without thought*" (my italics: AP).

Nativelike fluency, as has been argued elsewhere (Bolinger 1975, 1976, Pawley and Syder 1976, 1983, Wong-Fillmore 1976, Kuiper and Haggo 1983, Peters 1983) is probably unattainable until the language learner has memorized a repertoire of formulas—I would argue, clause and multi-clause length formulas—including many in which most of the constituents are lexically specified. There is a good deal of evidence that the human capacity for encoding novel linguistic sequences is limited to a few words in a single focus of consciousness. A focus of consciousness such as takes place in planning a novel linguistic sequence is usually associated with a break or slowdown in the flow of speech (Goldman-Eisler 1968, Pawley and Syder 1976, Chafe 1979, 1980).

In English and probably many other languages, the connected discourse of native speakers typically comes in fluent bursts of about three to ten words. A fluent chunk is usually a single clause, occasionally a multi-clause construction. However, it is rarely, if ever, the case that fluent chunks of this order are mainly composed of novel material. The introduction of more than one or two novel words into a phrase is usually marked by a break in the flow[5]. When we examine fluent connected speech closely it turns out to be composed mainly of formulaic expressions, each of which realizes a schematic or a (partly) lexicalised formula.

The following extract from a narrative illustrates some of these points. The speaker is a New Zealander, about 70 years old, recalling scenes from his childhood. Pauses (in seconds) and shifts of tempo are indicated below each line of text. / indicates stress on the following syllable; ⌞ marks the main or tonic stress in a tone group.

2. /we /had a /fan |tastic /time - - -
[slows] (1.1)
/there /were /all |kinds of re/lations /there/
[accel] [slows]
/I dun/no |where /they /all |come /from/
[accel], [slows]
I didn't |know /'alf o' them - -
[accel] (0.9)
and' ah - the |kids /sat on the |floor - - -
(0.2) (1.5)
and ol' /Uncle |Bert /he /ah/
o' /course /he was the |life and soul of the |party
[acel] [slows]
/Uncle /Bert 'ad a /black |bottle - - -
[accel] [slows] (1.5)
an ah - 'e'd t /tell a /few |stories
(0.2) [accel] [slows]
an 'e'd /take a /sip out of the /black |bottle/
[accel] [slows]
n' the /more /sips he /took /outa /that |bottle - - -
[accel] (1.6)
the |worse the /stories |got - - -
(1.6)

In (2) the narrator uses the following sequence of clause-length constructions, most of them partly lexicalised formulas. Formulas within formulas are marked off by square brackets and labels.

[5]Goldman Eisler (1968, esp. chs. 3-4) describes experiments suggesting that hesitations in speech "are the delays due to processes taking place in the brain whenever speech ceases to be the automatic vocalization of learned sequences, whether ocassioned by choice of an individual word, by construction of syntax, or by conception of content" (Goldman-Eisler 1968:58). In a recent unpublished paper Chafe finds evidence that in narrative discourse speakers usually introduce only one new concept per burst of fluent (pause-free) speech, other concepts being previously mentioned or part of a familiar schema.

3. NP *have* - TENSE *a* [ADJ *time*]$_{NP}$ (ADJ = *wonderful, fantastic,* etc.)
   there be - TENSE [*all kinds of N*]$_{NP}$ *there*
   NP *do* - TENSE *not* Vt [*where* NP *come*-TENSE *from*]$_{NP}$
   NP *do* - TENSE *not* Vt [*half of them*]$_{NP}$
   NP *sit* - TENSE *on* NP
   NP *be* - TENSE *the life and soul of the party*
   NP *have* - TENSE *a* NP
   NP *would* [*tell a few stories*]$_{VP}$
   NP *would* [*take a sip out of* NP]$_{VP}$
   the more NP *the worse* NP *get*-TENSE

Extensive use of formulas, then, has the advantage of freeing speakers from concentration on the mechanics of speech production — the tasks of finding and articulating words for their thoughts and ensuring that utterances are grammatical and idiomatic. Using ready-made phrases and clauses as building blocks, speakers can turn their attention to other and higher-order tasks — framing these blocks into larger structures, fine-tuning the music of each utterance, taking extra pains with a particular word choice, and so on.

Let us turn now to idiomaticity. Compare the fragment of actual narrative in (4) with the paraphrase given in (5):

4. I had four uncles.
   They all volunteered to go away.
   And ah that was one Christmas - that I'll always remember.
   Because ah-my four uncles came round, they were all in uniform, an' ah they are goin' to have Christmas dinner with us.
   And what was more important, they're goin' to provide it.
   And that was really something.

5. The brothers of my parents were four.
   Their offering to soldier in lands elsewhere in the army of our country had occurred.
   There is not a time when my remembering that Christmas will not take place.
   Because of the coming of the brothers of my parents
   to our house,
   having put on their bodies the clothes of the army.
   The eating of Christmas dinner by them in our
   company was to happen,
   and above that thing in importance,
   the buying of the food by them was to occur.
   A thing that was indeed unusual and indeed good.

What we have called elsewhere (Pawley and Syder 1983) the puzzle of nativelike selection is how the language learner knows which of the well-formed sentences of a language are nativelike. How are normal usages distinguished from those grammatically correct sequences which are unnatural or highly marked? Grammaticality does not ensure idiomaticity. Many sentences are called by the grammar; few are chosen as fit for normal use.

The narrative 'translation' in (5) is grammatical but strikes native speakers of English as bizarre, something like a literal translation from another language. To sound like a native we have to know 'how things are said'. This seems to be partly a matter of knowing the kinds of form-meaning pairings treated here under the heading of speech formulas. Indeed, formulas are by definition idioms of a language-culture system, that is, standard ways of talking about familiar ideas.

Here the question of creativity comes up. Do we have to talk exclusively in clichés in order to speak idiomatically? Can we say nothing new? Clearly such is not the case. But much as people like novelty, we like it in limited doses, and we like it to come wrapped in familiar packaging. Effective creativity is highly constrained, comprising variations on familiar themes. The rules of syntax provide certain constraints but evidently they allow a freedom that is often too great for the purposes of ordinary speech exchange, which demands that ideas be packaged for a fast ride. Formulas seem to provide the right combination of familiarity and flexibility. English has, for example, tens of thousands of schematic and partly lexicalised formulas whose highly restricted grammars allow partly new ideas to be presented in a familiar form. Some idea of the size of the formula lexicon may be gauged by counting the number of formulas which individual words participate in. The most common verbs, such as *have, get, take, go, come* and *do* each figure in several hundred formulas. Verbs of the second and third rank in frequency, such as *know, see* and *think*, each figure in several score. The following, for example, are among the formulaic expressions which *think* occurs in:

6. Come to think of it, ...
   What do you think?
   I thought better of it.
   Think nothing *of* it!
   Think it over.
   I hardly dare think about it.
   It doesn't bear thinking about.
   (Just) think about it for a [moment, second, minute, while, etc.]
   I'd think none the worse of you ...
   I think I'd have done the same thing in his [shoes, situation, etc.].
   Do you think I came down in the last shower?
   Do you think I was born yesterday?
   Who do you think you are?
   I don't think much of that [suggestion, idea, etc.].
   Can you think of a better one?
   I (just) can't think straight.
   I'll need a few days to think it over.
   I haven't stopped to think about it.
   Think twice before you PREDICATE PHRASE.
   Who (ever) would have thought it!
   I don't think NP will like that.
   I think a *lot* of
   NP_i thinks the *world* of NP_j
   NP_i thinks the sun shines out of NP_j's [bottom, arse, etc.].

NP_i thinks PRO_i is really somebody.
NP_i thinks PRO_i's shit doesn't stink.
NP_i thinks he's the cat's pyjamas.
NP_i thinks *noth*ing of NP_j.
NP thinks *noth*ing of V-ing NP (e.g. walking 50 miles).
Think what that could mean to NP.
I though you'd never ask!
[S, NP, ADJ], I don't think. (ironic tone).
I was (just, only) thinking aloud.
Think before you [open your mouth, speak].
I couldn't think of [a single thing, anything] to say.
I don't know *what* to think!
I thought you knew better [than that, than to S].
I thought you knew!
I think so.
I thought I told you not to do that!
What I think is, ...
Do you really think so?
He only thinks of himself.
He thinks highly of you.
Think again!

The question of what is involved in saying things that make sense in a particular language community, is discussed by Keesing (1979) and especially by Grace (1984)[6]. It appears that coherence of discourse, as this is perceived by members of a speech community, is very closely tied to conventional ways of talking about particular subject matters. That is why translation is difficult — whether across languages or from specialist discourse to discourse that will be understood by a layman. If the target speech tradition lacks conventional ways of talking about the subject matter in question coherent translation is extremely difficult, if not impossible. Even if the tradition has such conventional usages translation may still be difficult; it is necessary that the ways of talking match in conceptual particulars. This is so because to a large extent subject matters do not exist 'out there' but are created by language (see fn. 2). Our discourse about 'things' is in fact discourse about a conceptual world that is linguistically constructed. And such worlds are to some extent specific to particular groups of people (within or across language communities) who share a way of talking.

For examples of linguistically-constructed conceptual worlds that are not readily intelligible to outsiders we do not need to look beyond our own language. The ways of talking characteristic of any specialist group, such as the medical profession, or linguists, or linguists of a particular school, illustrate the point. Thus, the Chomskyan

[6]This question has also been a major concern of certain schools of sociology, especially ethnomethodology, and of ordinary language philosophers, and has begun to occupy linguists investigating conversation and other forms of discourse. However, with some exceptions (e.g. Grice 1975), the focus in these investigations has been on decoding rather than encoding, on how people make sense of what is said rather than how speakers decide, from among the many ways they might say things, what to say.

picture of language which became widely adopted during the 1960s was constructed with metaphors, flow chart diagrams and a whole new way of talking derived partly from mathematics. 'A language is an infinite set of sentences', 'surface structures are labelled bracketings of formatives', 'structure A is derived from structure B', 'transformations convert deep structures to surface forms', 'rule A is ordered before rule B', and so on. A good deal of the subject matter of Chomskyan linguistics simply did not exist, for all practical purposes, before it became part of a particular theory of language. This situation may be generally as true of scientific theories as it is of religious cosmologies.

I do not wish to say that all conventional ways of talking are speech formulas as defined here. However, I am saying that speech formulas make up a large part of what Grace means by conventional ways of talking.

A final remark about the distinction between designations and ad hoc descriptions. It is part of both speakers' and hearers' competence to be able to distinguish standard, i.e. lexicalised names for things from ad hoc descriptions. Knowledge of formulaic expressions is constitutive in this distinction. However, it should be noted that lexicalisation is a matter of degree — expressions which describe or characterize a thing may be more or less lexicalised; in English there are more than 20 criteria which indicate whether or not an expression is a standard form-meaning pairing (Pawley in press).

Can the grammar-lexicon model accommodate speech formulas?

This final section touches on the question of whether the grammar-lexicon model of language can accommodate speech formulas, or the sorts of facts which the speech formula construct attempts to make sense of.

As it stands now, the discipline of linguistics does not have much to say about most of the kinds of competence listed in (1). To be precise, structural linguistics has a great deal to say about (a), (b) and certain ingredients of (c), and about ingredient (k-ii), and little or nothing to say about the others.

These observations should not be regarded as censorious of modern linguistics. They refer to omissions, things excluded from analyses. But you should regard these omissions as shortcomings, as *regrettable*, only if you think descriptive or theoretical linguists presume to be centrally concerned with understanding linguistic competence in its broadest, vernacular sense. Or if you think linguistics *ought* to be centrally occupied with this task. On this matter opinion within the discipline seems to be divided. Quite rightly so — there is no reason why all linguists should study the same things.

When I entered linguistics, more than 20 years ago, grammar was the big deal — either compiling grammars of particular languages or working in grammatical theory. Along with most linguists, I continue to believe that these last two enterprises should remain among the central concerns of our discipline. However, certain experiences have disturbed my initial naive belief that a comprehensive account of a language can

be achieved by compiling a grammar and a lexicon according to grammarians' principles[7].

The single most powerful experience was attempting to describe Kalam, a language spoken by about 15,000 people living in the Bismark and Schrader Ranges of Papua New Guinea. In the mid-1960s I spent about a year among the Kalam-speaking community of Kaironk, and wrote a grammar of the language for my doctoral thesis. Later, an anthropologist, another linguist and I put together a dictionary.

The grammar, in the tradition of the time, provided a fairly detailed account of the formal structure of Kalam sentences, but left me feeling dissatisfied. Kalam is an unusual language — verbs, for example, are a closed set, with fewer than 100 members. Some of the most characteristic and striking features of the language were not captured in the formal description. Compared to other languages I was familiar with —Indo-European and Malayo-Polynesian— Kalam has distinctive ways of talking about everyday experiences such as seeing, hearing, tasting, feeling cold or sick, gathering firewood and going hunting.

The following comparisons of translation equivalents illustrate some of the differences between English and Kalam ways of talking:

7. *English*	*Kalam*
see	*wdn* *nŋ-* eye perceive
hear	*tmwd* *nŋ-* ear perceive
dream	*wsn* *nŋ-* sleep perceive
think	*gos* *nŋ-* mind perceive
taste	*ñb* *nŋ-* consume perceive
feel (deliberately)	*d* *nŋ-* hold perceive

[7]Perhaps it would be more accurate here to speak of 'assumption' rather than 'belief'. I might argue, with many fellow linguists, that I have never realy *believed* that grammar-lexicon descriptions provide a comprehensive account of a language. But actions speak louder than beliefs. What really counts in the formation of a scholarly tradition (as Grace 1984: 19-22 points out) is not the personal beliefs of individual scholars; what matters are the assumptions underlying the prevailing practices which make up the tradition— what subject matters its adherents attend to or ignore, their ways of talking about the subject matters, and so on.

like, approve	*gos tep | nŋ-* mind good perceive
dislike	*gos tmey | nŋ-* mind bad perceive
8. I heard the plane	*balws agek (twmd) nŋbyn* plane it-sounded (ear) I-perceived
I smelt the pig	*kaj kwy apek | nŋbyn* pig odour it-came I-perceived
9. bring	*d ap-* get come
take	*d am-* get go
fetch	*am d ap (ay-, ñ-)* go get come put, give
10. fetch firewood	*am mon pk d ap ay-* go wood break get come put
hunt game-mammals, catch game-mammals	*am kmn pk d ap ad ñb-* go game kill get come cook consume
gather pandanus nuts	*am alŋaw kab tk d ap* go pandanus nut cut get come *ad ñb-* cook consume

(7) and (8) give an idea of how Kalam speak about events of perception and cognition. Each such event is analysed into components, such as the part of the body which does the perceiving, or the sequence of acts which are necessary for perception to occur. Thus, tasting involves, first, consuming something by mouth, then perceiving; feeling involves touching something, then perceiving. In Kalam one cannot hear or smell an *object*, such as a plane or a pig. The plane must make a *sound*, which one then hears; the pig must have or give off an *odour* which *comes* to one.

(9) and (10) refer to action events. Here again Kalam shows a highly analytic pattern. In (9) we see an approximate equivalence in the overall information given in the translation pairs, the difference being that Kalam explicitly breaks each event down into a series of acts while English packs the same information into one verb. In (10) we begin to see another kind of difference between the two languages, having to do with what kinds of information must be mentioned when giving an account of an event.

In English it is normal to refer to a complex event, made up of several component events occurring in sequence, by mentioning only one or two components, which stand for the whole. For example, in answer to the question *What did you do today?* people might answer *We went to the market/theatre/swimming pool/park* (and did the normal things one does there), or *We caught three fish* or *We went and gathered a lot of firewood.* When talking about any action in Kalam, however, it is normal to indicate at least the following:

11.

1	2	3	4
MOVEMENT OF ACTOR FROM POINT OF REFERENCE TO SCENE OF ACTION	ACTION AT SCENE	MOVEMENT OF ACTOR AND AFFECTED OBJECT TO POINT OF REFERENCE	FATE OF ACTOR/ OBJECT AT POINT OF REFERENCE

The first component tells whether the actor was already at the scene of the action (indicated, e.g. by the verb *md-* 'be, stay') or whether he had to move there (e.g. *am-* 'go'). Component 2 says what happened there. In (10) the actions each involve something being done to an object: killing game-mammals, cutting down nuts, etc. Component 3 tells whether or not the actor stayed at the scene or moved, and whether or not he took the affected object with him. The sequence *d ap* 'get come' which recurs in (10) indicates that the actor returned to base (his home or camp or the narrator's previous point of reference) and that he carried the affected object with him. Component 4 rounds off the description of the event by telling what the actor did with the transported object, or with himself. In the case of firewood one stacks it or burns it, according to need; in the case of pandanus nuts and game mammals one usually bakes and eats them.

Why are Kalam ways of talking about events difficult to describe in a grammar? Haven't I just described some of the conventions above?

First, remember that structural linguistics is primarily a formal discipline, by which I mean that its strengths have been in dealing with linguistic form rather than with meaning or function. In writing grammars the tradition is to consider formal syntactic categories and relations as basic. Structural grammarians write mainly about such things as transitive constructions, relative clauses, nominalizations, heads and attributes, etc. It is not customary to begin with function and say "Language X has conceptual categories 'Action Event' and 'Cognitive Event', which involve the following elements". "If a speaker wishes to talk about hunting he must phrase things as follows; if he is talking about harvesting root crops he must say such and such". It may well be possible to write grammars in this way, but I have never seen such a grammar[8].

[8]Of course, many insightful papers have been written about the semantic styles of particular languages. For European languages, especially, there is a large literature on idioms, proverbs, etc., and linguists have recently turned to the study of metaphor (e.g. Lakoff and Johnson 1980). In the collected essays of Whorf (Carrol 1956) and Dixon (1984) —to name just two anthropological linguists— we find pioneering investigations of some aspects of the semantics of lesser-known languages. Talmy (1985) develops a typology of patterns by which languages map semantic patterns onto surface forms. Wierzbicka has broken new ground in semantic analysis (e.g. Wierzbicka 1985). However, comprehensive grammars, taking complex

Although there must be many languages in the world which, like Kalam, have a semantic style that differs radically from Indo-European norms, it seems that other grammarians have also found it hard to make the semantic style of a language shine through a form-based description.

In any case, a grammar consists of generalizations. Particular expressions, instantiations of the rules of grammar — such as expressions meaning 'perceive with the eyes', 'go get come', 'go get wood, bring it and stack it' — are not singled out for mention. All potential expressions are equal in the sight of Grammar.

How then do we single out those well-formed sequences which are institutionalised from those that are not? Or should we make no attempt to single them out? I chose to list many well-formed clause-length expressions (or more exactly, to list the lexicalised and partly lexicalised formulas which they instantiate) in the Kalam dictionary, whenever I felt that they were the standard way of expressing a familiar idea.

I felt somewhat guilty about doing so. It seemed that this procedure was breaking three descriptive principles cherished by structural linguists. One is the principle of precision. There were no clear grounds for a cutoff point between expressions to be put in the dictionary and expressions to be left out. We included the clause-length formulas by which Kalam refer to 'enemy' and 'market'. 'Your enemies' are 'people you fight with':

12.

byn-b	penpen	ñapm
man-woman (people)	reciprocally	you-fight pl.

A 'market' is a 'place where they exchange (= buy-and-sell) things

13.

tp	tap	tawpay
place	thing	they-exchange

But what about other well-formed expressions for less familiar ideas, expressions which are something less than common usages and something more than ad hoc descriptions? Our rule of thumb was that common usages were in, possible but unusual ways of saying things were out, and everything in between remained problematic in status.

A second principle is that of economy. Linguists are a cost-conscious lot, ever-ready to apply Occam's Razor to their products. Lees (1966: xxxi) states a view which has been prevalent in descriptive linguistics for much of this century:

> A constraint to maximum generality (measured hopefully by some notational property called "simplicity") is not special to linguistics; it characterizes any scientific enterprise.... For us it would mean, roughly, that given two descriptions... there is no point in preferring that one which contains superfluous assumptions or apparatus...

One kind of superfluity disdained by grammarians is the duplication of information in two different parts of a description, as when the same form-meaning

---

conceptual constructions (e.g., types of events and situations) as the point of departure, in the way that syntacticians take S, NP, etc. as theirs, have not yet (as far as I know) been compiled.

pairing is specified once by the rules of syntax and semantics and once again in a lexical entry. One specification is enough — the question of whether a given pairing is to be handled by one component or another is deemed to be an empirical issue. Twenty years ago, Chomsky (1965:184) commented on the handling of words formed by productive derivational processes:

> Clearly, the words *destruction, refusal*, etc. will not be entered in the lexicon as such. Rather, *destroy* and *refuse* will be entered.... A nominalization transformation will apply at the appropriate stage of derivation.

A few years later he revised his views, finding grounds for treating derived nominals as lexical items and for dropping the transformational derivation in this case. The details of this example do not matter. What is significant is that the grammarian is expected to make a choice between one analysis or another.

The third principle, already alluded to above, is that the object of a linguistic description should be grammatical competence, and that this competence is independent of knowledge about whether a particular well-formed sequence is a common usage, occurs rarely or has not yet been observed to occur. It is also independent of knowledge about the rules of pragmatic use and interpretation of products of the grammar. Call this the autonomy principle: a language, in the grammar-lexicon sense of grammarians, is a system of knowledge separate from knowledge of the world, cultural and pragmatic. The operation of the autonomy and economy principles can be seen in Lees' influential work on complex nominal expressions in English, where he proposed that many nominal compounds be treated not as lexical items but as products of regular syntactic processes nominalizing verbal structures, and in Lees' critical remarks on the attempts of Jespersen and others to distinguish 'compounds' from free expressions:

> [T]here simply *is* no neat physical or semantic criterion for compounds, ... there need not be any such, and ... the point of linguistic research is to find grammatical descriptions, not to classify physical or semantic "objects" (Lees 1966: xxiv).

In this view of linguistic description there is no place for a concept such as 'name of a thing', 'nameworthy thing' or 'familiar idea'. There are only form-meaning connections which are either predictable or unpredictable (lexical) according to the grammar.

I do not believe that speech formulas can be accomodated within a descriptive framework which adheres to these three principles and to the conception of language which they impose. If one were to write a description of Kalam accordingly, the list of lexical items referring to events would contain only the 90 or so verb stems plus a handful of idioms. If your goal is to describe what is grammatical in Kalam that may be reasonable. But if you want to describe what it takes to speak Kalam like a native, it is plainly absurd.

However, the guilt which I felt in this matter was surely unnecessary. Because the descriptive principles outlined above do not reflect the way language users structure their knowledge. Instead, they represent a set of aesthetic principles and assumptions which have been adopted by a particular group of scholars, and which, in association

with particular analytic procedures, allow sense to be made of a certain range of linguistic data.

There is, within the field of language description, a separate tradition of research with a long and honorable history, which tackles a somewhat different (but overlapping) range of data in a different way. What we were struggling to do in the dictionary of Kalam was in essence the same as conventional dictionary-makers have been doing for generations.

The compilers of the *Oxford English Dictionary*, say, or of *Webster's New Twentieth Century Dictionary of English*, deal with a different body of data from that treated by the grammar-lexicon model. The differences stem not only from their distinctive purposes but from having different notions of what a language is. A dictionary such as Webster's treats a language as a system of conventions by which particular groups of people communicate about particular worlds. The dictionary is a kind of encyclopaedia of a culture (or collection of sub-cultures)[9]. Let us call any expression included in the dictionary a *lexeme*. What counts as a lexeme for the compilers of *Webster's* is any expression which has a high degree of institutionalisation in any English-speaking community as either the name of a thing or a way of saying something. Semantic and grammatical regularity are no bar to inclusion[10]. Thus, *Webster's* lists such well-formed common usages as *bloodstain, blood-stained, blood-colored, bottle-holder, forgettable, forgettableness, forgetful, forgetfulness, forgetfully* and *forgetter*. However, it does not list other possible forms such as *dyestain, paint-stain, sword-holder* or *unicorn-holder*. In respect of complex expressions which are in common use practice is somewhat inconsistent. Proverbs and idioms are likely to be recorded, but generally, the longer an expression is, the less likely it is to be found in *Webster's*.

In this view of language, there is no cutoff point between lexemes and non-lexemic expressions, between common usages and nonceforms, between institutionalised and ad hoc ways of talking about things. There is a continuum. And this continuum exists at, say, the clause level no less than at the level of derived words or nominal compounds. In these matters Kalam is not different from English, except in its greater reliance on complex lexemes. The cutoff point that we actually make in dictionaries is determined by practical limitations or purpose rather than principle.

While the most highly regarded dictionaries include many formulaic expressions, they hardly provide systematic descriptions of these, still less of the formulas which underlie them. Formulas present numerous problems of description which we cannot expect the traditional tools of lexicography or grammatical analysis to cope with. Each formula is a sort of mini grammar, a body of conventions connecting form and meaning, some of which are particular to that formula. We have already noted that the rules and usages which make up the 'grammar' of formulas are diverse in nature,

[9]That dictionaries are not in principle different from encylclopaedias is argued by Haiman (1980). I read Keesing (1979) as saying that a description of what it takes to speak a language should not be sharply distinguished from an ethnography of the language community.

[10]Matthews (1974: 193) distinguishes 'lexicalised' from 'institutionalised' expression without claiming that there is a sharp boundary. Here, however, he adopts a grammarian's view of what counts as a lexeme (indivisible compounds, like *dishwasher*) versus non-lexical institutionalised collocations (such as *sick joke*, where the modifier *very* may apply just to *sick*). Laymen and lexicographers do not make such a distinction; the grammatical criteria for the proposed distinction are in any case problematic, as Matthews acknowledges.

representing not only grammatical constraints in the traditional sense but most or all of the elements of linguistic competence listed in (1 a-k): idiomaticity, musical conventions, contextual appropriateness, etc.

We may *extend* structural notions of grammar and lexicon to accomodate formulas, but in doing so we necessarily change the character of the description in rather fundamental ways[11].

Conclusion

The most fundamental change is in the view of what language is and how it is organised. In order to make sense of speech formulas it is necessary to regard a language as a collection of ways of talking about things (as Grace puts it), expressing ideas (old or new) in a manner that is conventional (grammatical, idiomatic, etc.).

It is the 'things' that are hardest to marry with the structuralist tradition. 'Ways of talking', 'idiomatic forms of expression' —these notions, while distinct from 'ways of talking grammatically', are at least partly amenable to description using the tools for analysing forms which are the chief glories of structural linguistics. But linguistics has no technical construct corresponding to the vernacular 'idea', 'subject matter', 'thing talked about', as an independent entity rather than as what a particular utterance or expression means. For example, while 'sentence', 'clause', 'verb phrase' and the like are posited as types of formal constructions which can be analysed independently of meaning, we are not used to talking about 'events', 'situations' and 'acts' as kinds of conceptual constructions, analysable into conceptual elements independently of form. Only at the level of lexemes is there a place for talking about particular concepts as part of a language-culture system.

Yet there are various linguistic abilities which hinge on knowledge of conceptual contructions and elements as entities separable from particular forms. When fluent bilinguals translate, for example, it is not (usually) equivalence of form that they seek but equivalence of meaning; they try to match ideas, to capture the thing that was said. Grace (1981, 1984) discusses this point at some length. The ability to paraphrase, to say the same thing in different words and different styles, depends on being able to isolate the conceptual elements. Speaking idiomatically, or recognising that someone's utterance is intended to be understood as a request formula and not as a statement, call for the language user to know that particular subject matters are talked about in particular ways.

In dealing with speech formulas, then, we are led to much the same view that Marjorie Shostak took of her task of mastering the !Kung language: it is necessary to know not only how things are said, but the kinds of things that are said and done with words. This kind of linguistic competence is inseparable from a broader cultural and social competence.

[11]Over the last 10 or 15 years theoretical linguists have begun to develop a richer conception of lexicon, including the awarding of lexical status to expressions formerly regarded as transformationally derived. However, the changes envisaged as needed for describing speech formulas are of a more fundamental nature than a trade-off of material between lexicon and other components of a grammar.

I do not suggest that every linguist should try to mine this broader linguistic competence. Ways of talking about things constitute a more complex ore than we are used to extracting or analysing. Our machinery is mainly designed for obtaining pure syntax, of which the supply is not yet exhausted. But there is a much larger supply of the complex aggregate that is language, in the vernacular sense, awaiting excavation. If linguists are able to mine this rich material they may be doing something that is of quite general value, as well as intellectually challenging.

REFERENCES

BOLINGER, DWIGHT. 1975. *Aspects of Language* (2nd ed.) New York. Harcourt Brace Jovanovich.

BOLINGER, DWIGHT. 1976. "Meaning and Memory". *Forum Linguisticum* 1(1): 1-14.

CARROL, JOHN B. (ed.) 1956. *Language, thought and reality: selected writings of Benjamin Lee Whorf*. New York, MIT Technology Press and John Wiley.

CHAFE, WALLACE. 1979. "The flow of thought and the flow of language." In: *Syntax and Semantics*. Vol. 12: *Discourse and syntax*. New York. Academic Press pp. 159-181.

CHAFE, WALLACE. 1980a. "The development of consciousness in the production of a narrative." In Chafe, ed., pp. 9-50.

CHAFE, WALLACE. (ed.) 1980b. *The pear stories: cognitive, cultural and linguistic aspects of narrative production*, Norwood, NJ, Ablex.

CHOMSKY, NOAM. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts, MIT Press.

DIXON, R.M.W. 1984. *Where have all the adjectives gone? and other essays in semantics and syntax*. The Hague, Mouton.

GOLDMAN-EISLER, FRIEDA. 1968. *Psycholinguistics: experiments in spontaneous speech*. New York, Academic Press.

GRACE, GEORGE W. 1981. *An essay on language*. Columbia, South Carolina, Hornbeam Press.

GRACE, GEORGE W. 1984. The linguistic construction of reality. Printout pp. 235. Honolulu, University of Hawaii Linguistics Department.

GRICE, H.P. 1975. "Logic and conversation". In *Syntax and Semantics*. Vol. 3 *Speech acts* (ed. by P. Cole and J.L. Morgan), pp. 44-58.

HAGGO, D. and K. KUIPER. 1983. Review of F. Coulmas (ed.) *Conversational Routine*. *Linguistics* 21: 531-551.

HAIMAN, JOHN. 1980. "Dictionaries and encyclopaedias." *Lingua* 50: 329-357.

JESPERSEN, OTTO. 1924. *The philosophy of language*. London, Allen and Unwin.

KEESING, ROGER. 1979. "Linguistic knowledge and cultural knowledge: some doubts and speculations." *American Anthropologists* 81: 14-36.

KUIPER, KOENRAAD and DOUGLAS HAGGO. 1983. "Livestock auctions, oral poetry and ordinary language." *Language in Society* 13: 205-234.

LAKOFF, GEORGE and MARK JOHNSON. *Metaphors we live by*. Chicago, University of Chicago Press.

LEES, ROBERT B. 1966. *The grammar of English nominalizations*. The Hague, Mouton. (First published by Indiana University 1963).

MAKKAI, ADAM. 1972. *Idiom structure in English*. The Hague, Mouton.

MATTHEWS, P.H. 1974. *Morphology: an introduction to the theory of word-structure*. Cambridge University Press.

PAWLEY, ANDREW. 1980. On meeting a language that defies description by ordinary means. Paper read to 13th Congress of Linguistic Society of Papua New Guinea, Lae. (To appear in a volume published by Pacific Linguistics).

PAWLEY, ANDREW. In press. Lexicalization. In: *Georgetown University round table on language and linguistics* 1985, ed. by D. Tannen. Washington, D.C., Georgetown University Press.

PAWLEY, ANDREW and FRANCES SYDER. 1976. The one clause at a time hypothesis. Paper read at lst Congress of the Linguistic Society of New Zealand, Auckland.

PAWLEY, ANDREW. 1983. Two puzzles for linguistic nativelike fluency. In: *Language and communication* (ed. by J. Richards and R. Schmidt). London, Longman. pp. 191-225.

PETERS, ANN. 1983. *The units of language*. Cambridge University Press.

SHOSTAK, MARJORIE. 1983. *Nisa: the life and words of a !Kung woman*. New York, Vintage Books. (First published 1981 by Harvard University Press).

SYDER, FRANCES HODGETTS. 1983. *The fourth R: spoken language, English teaching and social competence*. Auckland, Talanoa Press.

TALMY, LEONARD. 1985. "Lexicalization patterns: semantic structure in lexical forms." In: *Language typology and syntactic description*. Vol. 3. *Grammatical categories and the lexicon*. Cambridge University Press.

WEINREICH, URIEL. 1969. "Problems in the analysis of idioms." In: *Substance and structure of language* (ed. by J. Puhvel). University of California Press, pp. 23-81.

WIERBICKA, ANNA. 1985. Lexicography and conceptual analysis. Ann Arbor, Karoma.

WONG-FILLMORE, LILY. 1976. The second time around: cognitive and social strategies in second language acquisition. Stanford University doctoral dissertation.