

Werk

Titel: A program for linguistics

Autor: Dixon , R. M. W.

Ort: Wiesbaden

Jahr: 2003

PURL: https://resolver.sub.uni-goettingen.de/purl?666048797_0007 | LOG_0021

Kontakt/Contact

[Digizeitschriften e.V.](#)
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

✉ info@digizeitschriften.de

A program for linguistics

R. M. W. Dixon

Dixon, R. M. W. 2003. A program for linguistics. *Turkic Languages* 7, 157-180.

An agenda is presented for the discipline of linguistics. For every question that asks “how”, there should be a related enquiry into “why”. The “why” questions are generally shunned by linguists; only some of them can presently be provided with informed answers. Building on these matters, linguistics must attempt evaluations of the relative worth of different languages (in terms of specified criteria). And, like other sciences, linguistics should essay predictions concerning languages, both about what structural possibilities may be found in a newly studied language, and about how a given language is likely to change over time. Sadly, few of the people who currently call themselves “linguists” attempt to deal with more than a small fraction of what is outlined here. In essence, the discipline of linguistics needs to reinvent itself in order to properly explain the nature of human language.

R. M. W. Dixon, Research Centre for Linguistic Typology, Institute for Advanced Study, La Trobe University, Bundoora, VIC, Australia 3086.

The editor of this journal asked me to write a paper about general linguistics which might be of interest to scholars of Turkic languages, among others. I enquired whether he wanted me to describe what people who call themselves linguists presently do, or rather what I think linguists should do. He opted for the latter, and suggested that I could, at the end, relate this to the current situation. So, here is the idealised program that I describe.

1. Languages in context

Just as one would not consider studying an animal without reference to its habitat and habits, so a language can only properly be studied in terms of the situation in which it is used. We need to know about the speech community—its social organisation, political system, economic basis, religious predisposition, laws and beliefs. We need to know what kind of relations (political, trade, social) the speech community has with its neighbours; and we need to know what kind of terrain they inhabit.

Some of these questions may seem inappropriate for a language such as English or Spanish, spoken in many nations by peoples with a variety of cultures and religions, and living in a variety of ecological zones. English is spoken, worldwide, by around four hundred million people, and Spanish by more than three hundred million. However, of the four thousand or so languages currently spoken, only about a hundred and twenty have more than one million speakers. The great majority of lan-

guages have (and always have had) less than ten thousand speakers. That is, a typical language is spoken by a fairly homogeneous social group, in a particular type of terrain. The fact that some languages have a very large number of speakers is a recent development (within the past couple of thousand years). In fact, languages with many speakers probably do show recurrent differences from languages with relatively few speakers; it may be that languages with many speakers tend to have a relatively simple word structure, and place more reliance for the communication of meaning on the ways in which words are combined (what is called syntax). This is an important topic for study by linguists.

Just as each language must be considered in terms of its community of users, so each utterance in a language must be considered in terms of its context of use. That is, for any utterance we must take account of who said it, to whom, where, when, and in what circumstances (for example, there is a difference between something said in a job interview, and the same thing occurring within a tête-à-tête conversation). We need to know not just what was said but also whether there was a between-the-lines message; and what the pragmatic effect might be.

To consider language out of its social context (as a kind of algebra, or as a truth-value logic) will lead to no real insight into the intrinsic nature of language—as a social habit and as a social skill. Similarly, quoting and attempting to analyse a sentence without reference to its context of use is a lazy and misguided course of action. Speakers' opinions about what they can say do not fully accord with what they actually do say. And what may be said varies with the context of use. Unfortunately, much of the work currently called “linguistics” does consider languages without reference to their social context, and does examine sentences outside of an actual context of use. I will return to this in §7.

2. The questions to be asked

Linguistics consists in the systematic study of languages as social phenomena. This involves posing—and attempting to answer—a series of interconnected questions concerning the nature of language and of language use.

As in any discipline, the first step is to pose the appropriate questions. These fall into two broad types. The “how” questions are the most basic, investigating the nature of language, its interaction with its social context, and its historical development. The “why” questions build on this basic knowledge, enquiring into the reason for a particular language being organised in the way it is. At an advanced stage of investigation, linguists should also attempt evaluation and prediction. That is, it is relevant to enquire whether one technique of linguistic organisation is more effective than another (for example, it may make a language easier to learn). Or, indeed, whether one language is better than another, for some specific purpose. And a linguist should be able to predict that if a language has a certain feature then this is likely to implicate some further feature. Along the temporal axis lie predictions concerning how a given language is likely to change in the future, assuming certain social conditions.

3. The “how” questions

HOW-1. Investigate how each language is organised—both within itself, and in relation to the situations in which it is used

The primary function of language is to pass a potage of meaning from speaker to hearer. The meaning has first to be organised; this is done by the grammar of the language. Different aspects of meaning are coded by members of large open word classes (noun, verb, adjective). The words are then joined together in various kinds of sentence pattern. Each language requires a number of specifications from grammatical systems such as gender and tense. The systems involved vary from language to language; one language may specify tense but not gender, another gender but not tense, another both, another neither. And the nature of the system also varies. In one language the tense contrast is past versus non-past, in another future versus non-future, and in another past versus present versus future, while a further language distinguishes three futures, one present, and five pasts.

The two components of a linguistic description directly relating to meaning are the lexicon, a listing of the word bases in open classes, and the grammar, dealing with the analysis of word forms (morphology) and the ways of combining words (syntax). These two components are directly linked—the meaning of a word will determine which grammatical system may apply to it, and which sentence types it can occur in. For example, *hear* and *see* have similar meanings and can both be followed by either a *that* clause or by an *-ing* clause (*I heard / saw that she was laughing*, and *I heard / saw her laughing*). *Conclude* and *infer* have similar meanings and can be followed by a *that* clause but not by an *-ing* clause (we can say *I concluded / inferred that she is clever* but not **I concluded / inferred her being clever*). This can be explained in terms of the meanings of *that* and *-ing* clauses, and of the meanings of types of verbs (full details are in Dixon 1991).

It is an unfortunate feature of the present-day scene that, for the major languages, different groups of practitioners—with different methodologies and priorities—construct dictionaries (or lexicons) and write grammars. Ideally, these activities should be integrated.

To convey meanings (coded in terms of grammar and lexicon) to a hearer, they must be accorded physical expression. This can involve sounds, or signs. Like meanings, sounds and signs are, in their raw state, fluid. Each spoken language has a system of phonological representation, which relates to lexicon and grammar, on the one hand, and to speech sounds, on the other. A sign language has a similar system, which could be called signology. That is, a sentence (whose slots are filled by words) will have a phonological or signological representation, and this will be realised by a stream of sounds (or of signs) that a hearer can understand and interpret.

Returning to the HOW-1 question. The basic business of linguistics is to investigate the organisation of each language and produce a lexicon, grammar and phonology (or signology) for it. Some parts of the lexicon and grammar have a meaning in terms of the world at large (for example, the ideas of past and present time, and of

male and female sex). But other parts may have a meaning in terms of the particular social organisation of beliefs and habitat of the community of speakers—for example, terms referring to colours of cattle and to bride-price, and categories of possession relating to kinship organisation.

HOW-2. Investigate and explain the ways in which (how) languages differ

This can be briefly exemplified for negation. There are various ways in which negation can apply. In English these include: (a) negation of a noun or noun phrase, as in *No honest person would tell a lie*; (b) negation of a complete sentence, as in *He can't write the review* (he may be illiterate, or too busy); (c) negation of a verb, as in *He could always not write the review* (he could decline the invitation, if he were afraid that his review might offend the author). Note that these can be combined, as in *I didn't (b) dare not (c) buy it* (I didn't want to buy it but my wife had told me to and I was scared of how angry she might be if I didn't comply), and *No (a) honest person would not (b) tell the truth*.

There are various ways of marking these—and other—types of negation. As can be seen, English uses *not* for (b) and (c) and *no* for (a). In Punjabi, there is just one negator, *naii*; this is either placed after the element which is being negated, or *naii* can be in its normal position just before the verb, with the element that is to be negated being stressed (Bhatia 1993: 120).

Languages vary in how many strategies of negation they employ. The only one found in all languages is sentence negation, like (b) in English. And there are different ways of marking negation. Whereas English has a particle *not*, which—for sentence negation—follows the first word of the verbal component (as in *He might not have been enjoying it for all I know*), in Jarawara (see Dixon 2004)—spoken in the southern Amazonian rainforest—negation is shown by a suffix to the verb. Compare *o-tafa* 'I'm eating' (where the subject 'I' is shown by prefix *o-* to the verb *-tafa* 'eat') with *o-tafa-ra* 'I'm not eating' (where the negative suffix *-ra* is added to the verb). In Fijian, negation can only be shown through the verb *sega*, which means 'it is not the case'. The sentence 'I'm not eating' has to be translated as, literally, 'That I am eating is not the case'.

Most languages have an interjection 'no' which can be used to provide a negative answer to a polar question (such as 'Are you eating?'). Note that in English this has the same form as the noun negator (*no*), different from the clause and verb negator (*not*). In others it has the same form as the sentence negator. And in some languages there is no simple word 'no'. In Jarawara, for instance, where negation is marked by an affix to the verb, a negative answer to a question (such as 'Are you eating?') must be a full sentence with the verb bearing the negative suffix (*o-tafa-ra* 'I'm not eating').

The paragraphs above provide a sample of how one studies the ways in which languages differ. For the typology of negation, one must first establish the different mechanisms of negation that human languages show, and the scope and marking of each. Besides noun, sentence and verb negation, there may also be negation in a sub-

ordinate clause (*I said that he didn't lie* is quite different from *I didn't say that he lied*), a negative tag (as in *He lied, didn't he?*) and a negative marker of disjunction (*I like neither apples nor pears*). We also find that some verbs (e.g. *forbid, fail, forget*) are inherently negative; that is, they behave like a positive verb plus a negative marker.¹

We need also to examine how the system of {positive, negative} interacts with other systems within the grammar of a language. In some languages the same distinctions are made for person, number and tense in positive and in negative clauses. In others there is variation, and it is a universal fact that—where variation occurs—there will be less choices available in a negative than in a positive sentence (never the other way round). For example, in Estonian, in a positive clause a verb has separate forms depending on whether the subject is first, second or third person and, if it is first or second person, whether it is singular or plural. In the negative, each verb has a single form, used for all persons and numbers. In Amharic, the verb of a positive sentence has separate forms for past ('did') and for preterite ('has done'); in a negative sentence, one form covers both past and preterite meanings.

I have tried here to illustrate a microcosm of the broad field of linguistic typology, the study of how languages differ.

HOW-3. Study how each language is used by its speakers

Two communities may utilise the same language but deploy it in quite different ways, according to their interpersonal perceptions and social niceties. Scientists in western countries follow the convention of couching any statement in the passive (*An experiment was devised to investigate ...*, rather than *I devised an experiment to investigate ...* and *It was observed that ...* rather than *We observed that ...*). Taking the agency out of an action is supposed to make the science more objective. Other varieties of English freely use active sentences, with the participants taking personal responsibility for what they do.

All languages have ways of asking questions, but the actual use of questions varies a lot. In English, a question can be used as a mild form of command, e.g. *Would you mind opening the window?* In Russian and in Danish, a negative question is more polite than a positive one, and is generally preferred (one would ask 'Isn't there any bread?' rather than 'Is there any bread?'). In Aboriginal communities of Australia, people tend to avoid direct questioning of a visitor ('Where are you from? Are you married? How many children do you have?'). Instead, an interlocutor might volunteer this kind of information about themselves, and the visitor would then be expected to reciprocate.

¹ It is an interesting feature of English that *some* in a positive sentence corresponds to *any* in a negative one. Compare *I ordered him to eat something* with *I ordered him not to eat anything*. The inherently negative verb *forbid* behaves like *order-plus-not* in taking *any*, as in *I forbade him to eat anything*.

Plainly, the way in which a language is used interrelates with how the language is structured. One can only avoid attributing agency by using a passive if the language includes a passive in its inventory of construction types (not every language does). And so on.

HOW-4. Study how each language is learnt—most centrally, by children, but also by adults

There are some recurrent features that apply in all instances of language acquisition. For instance, if a language has a trilled rhotic (like the *r* sound in Scots English), this is one of the last sounds to be mastered by a child, especially when it occurs in word-initial position. Other observations are language-specific. For examples the fricatives *ð* (as in *the*) and *θ* (as in *thin*) pose special difficulties for an adult learner of English (if these sounds do not occur in their native language).

A good deal of work has been done on topics such as when and how relative clause constructions are acquired. Unfortunately, the great bulk of this study relates to well-known languages, with a large number of speakers. There is a great need for detailed studies of child language acquisition in small tribal societies, in the Americas and in Africa. Little information is available on how children master some of the more unusual grammatical categories, which are predominantly found outside the mainstream languages of Europe and Asia.

For example, some languages require that each statement include a marker of the type of evidence on which it is based—whether the speaker saw it happen, or heard it, or inferred it from indirect evidence, or had it reported to them by someone else. Little work has yet been done on how children acquire such an evidentiality system. Would they start with a binary contrast (say, visual versus non-visual, or perhaps reported versus non-reported) and gradually build up to the four-term system? (Languages with grammatical marking of evidentiality are predominantly found in North and South America, in languages of the Tibeto-Burman family from East Asia, and in some of the languages of Eastern Europe and Central Asia—including the Turkic family—and the Caucasus.)

HOW-5. Study the ways in which languages change over time

A language is always in a state of change, as each generation uses it in a slightly different way from the previous one. Change is, as a rule, slow enough that the oldest and youngest speakers in a community can understand each other.

There are two conflicting principles at work in language change. One is the inherent laziness of humankind which tends towards least effort in pronunciation. The prefix *in-* (as in *inapplicable*, *indubitable*) becomes *im-* before *p-* (as in *implausible*) simply because *-mp-* (both sounds being made with the lips) is easier to say than *-np-* (a tongue sound followed by a lip sound). As another example, *want to* is reduced to *wanna* in certain environments.

In Old English, each noun had endings showing case and number. A final nasal ceased to be pronounced, and then the preceding unstressed vowel (which had now

come into word-final position). Originally we had forms like nominative singular *nama* and accusative singular *naman*; both of these reduced simply to *name* (pronounced [neim]) in modern English. The failure to pronounce final segments of words led to the loss of case and number marking on nouns.

The other principle relates to the need for efficient communication; to achieve this, certain fundamental distinctions must be maintained. In south-western France, sound changes occurred such that some words fell together; both 'cock, rooster', from Latin *gallus*, and 'cat', from Latin *gattus*, became *gat*. This homonymy might lead to confusion. To avoid it, *gat* was kept for 'cat' but replaced in its meaning of 'cock, rooster' by *pul*; this is a development from Latin *pullus* 'chick', and is here shifted in meaning to fill a lexical gap (Bloomfield 1933: 396-398).

When the erosion of word endings in English led to the loss of cases, new ways had to be found for indicating the function of a noun phrase in a sentence. Object function, previously shown by accusative case ending, was now indicated by placing the noun phrase immediately after the verb; benefactive function, previously shown by dative case, came to be marked with the preposition *to* (as in *John gave a book to Mary*).

The evolution of a language is a constant balance between laziness, which tends to lose contrasts, and the need for clear communication, which requires distinctions to be made between words, between grammatical specifications, and between construction types.

Language change happens over time. To properly study this, we need to observe a language over a fair period—at least several centuries, ideally several millennia. One person's lifetime is too short to achieve more than a glimpse of change in progress. What are required are sets of language data gathered over a long period. These are, sadly, lacking for the great majority of speech communities. There are records of Ancient Greek and Sanskrit going back about three thousand years, but for most languages the time depth is a fraction of this. For Slavic languages, for instance, the oldest records are from the tenth century CE. In North and South America and in Australia, the white invaders came and conquered and—in the great majority of cases—autochthonous languages had died out within about three generations of intensive contact. In Australia, there are a handful of reasonably good grammars from the nineteenth century, but these languages have long passed into extinction. Of the languages still actively spoken, only one (Arremte, or Aranda) was accorded a reasonably full documentation more than fifty years in the past.

There is, however, a branch of linguistics which can recover past stages of languages. This leads into the next question.

HOW-6. Investigate how languages are genetically related

Prototypically, each political group has its own language. If the group expands in size, it may split into two (as with Spain and Portugal, for example). Each of the new political groups will develop its own identity, customs and laws. At first they will speak mutually intelligible dialects of one language, but these will change in-

dependently of each other, getting more dissimilar until they are no longer mutually intelligible—that is, they have become distinct languages.

Now when a number of languages have evolved from a common ancestor (called the ‘proto-language’ of the group), and provided they have not diverged too much, it is generally possible to perceive this genetic relationship and to reconstruct (i) a good deal of the lexicon, grammar and phonology of the proto-language, and (ii) the systematic changes by which each modern language developed from the proto-language. To achieve this, one must have a fair-sized collection of word bases and of grammatical elements which have similar form and similar meaning between the modern languages. For example, Kulina and Jarawara (small tribal languages spoken in Brazil and Peru) can be shown to be genetically related through cognates such as: Kulina *phaha*, Jarawara *faha* ‘water’; Kulina *wapha*, Jarawara *wafa* ‘woolly monkey’; and Kulina *ephe*, Jarawara *efe* ‘leaf’. These examples (and other similar ones) demonstrate a recurrent association between *ph* (an aspirated labial stop) in Kulina and *f* (a labial fricative) in Jarawara. This provides a tiny sample of the extensive correspondence sets between languages (which involve every type of lexical and grammatical element) needed to provide a genetic connection.

It has been possible to prove, beyond doubt, that all the Indo-European languages make up one genetic family (there are over a hundred of them, ranging from Welsh to Albanian, from Italian to Persian, and from Latvian to Maldivian). Similarly, it has been shown that Turkish, Uzbek and Uyghur all belong to the Turkic family; that Hebrew, Arabic and Amharic all belong to the Semitic family; that Crow, Dakota and Winnebago belong to the Siouan family; and so on.

There can be a number of reasons for similarities between languages. Two of the minor ones are chance, and some universal tendency (for example, the word for ‘blow’ typically begins with a labial sound like *b* or *p* or *ph*, followed by a back vowel such as *u* or *o*). The major ones are shared genetic inheritance—as just outlined—and borrowing from neighbours. This leads into the next “how” question.

HOW-7. Study how languages in contiguity influence each other

There is always some degree of contact between individual speech communities living in a given geographical area. These links are likely to involve trade and intermarriage, and may also extend to music and dance, sport, and technical collaboration. As a consequence, there will be a degree of bilingualism or multilingualism between members of the speech communities. This can lead to the languages becoming more like each other. A limited number of words are likely to be borrowed; habits of pronunciation may become more similar; and—most pervasive of all—grammatical patterns are likely to be adopted or modified. The languages will gradually become more similar in their structural organisation, effectively converging towards a common grammatical pattern.

Thus a “linguistic area” is established. A well-known instance of this is South Asia, where languages of the Indo-Aryan subgroup of the Indo-European family, of the Munda subgroup of the Austroasiatic family, of the Dravidian family, and of the

Tibeto-Burman family, have converged in terms of certain critical features. These include retroflex consonants, two kinds of causative, and the trait of subjects in some types of sentences being marked with dative case (rather than the normal nominative case). In fact, South Asia is more than just a linguistic area; it can be described as a “translation area”. That is, a sentence in, say, Bengali, can be changed into a sentence in, say, Tamil by merely substituting Tamil lexemes and grammatical elements for those in Bengali; the grammatical organisation of the sentence stays unchanged. We do, of course, still have quite separate languages since the actual forms used (the words and affixes) are different.

Some work has been done on language contact—on what kinds of grammatical features are most likely (and which are least likely) to be borrowed between languages, and under what social conditions; but much more is needed. Work on genetic connections between languages has a longer history and there has been a tendency to misconstrue similarities due to contact as being proof for genetic relationships. The fact that two languages share some typological feature—say, each has a system of four genders—cannot in itself constitute genetic evidence. For this we would need close correspondences of form and of meaning between the markers of gender in the two languages (plus many more instances of form / meaning correspondence).

As noted above, the establishment of genetic relationships is the one area of linguistics where we can offer definitive proof, which is accepted by all practitioners in the field. This applies to perhaps two hundred language families across the world. Unfortunately, a small number of people have muddied the waters by suggesting unsupportable links between some of these established language families. They may illicitly take typological similarities (without any extensive correspondences of form and meaning) as genetic indicators. For instance, it has been suggested that the Turkic, Tungusic and Mongolic families make up a macro-family, which has been called Altaic. In fact the three families have considerable typological similarities, but lack the systematic correspondence of forms necessary for justifying a genetic link.

Some investigators are overly generous in recognising similarities of meaning, such that—in fact—anything could be related to anything (for example, the identification of a word meaning ‘egg’ in one language with that meaning ‘an owl species’ in another, supposedly linked through an unattested intermediate meaning ‘eye’, see O’Grady (1990: 457)).

There is, of course, no general approval by linguists of such work. But, sadly, the media is enamoured of illusions such as “proto-world”, with the consequence that great publicity is accorded to the idiot fringe of linguistics, while the solid but unsensational work being done by the core of workers is passed by.

4. The “why” questions

There is, naturally, a “why” question associated with each “how” question. Why is a given language organised in the way that it is (and not in some other way)? Why do languages differ in the ways that they do (and not in other ways)? Why do languages

change in the ways that they do (and not in other ways)? Why do languages in contact influence each other in the ways that they do (and not in other ways)? And so on.

Let me present a sample of six specific “why” questions.

(a) Why does Lahu (a language of the Tibeto-Burman family whose territory spans the borders of China, Burma and Thailand) have five spatial demonstratives, translated as ‘here’, ‘there’, ‘way over there’, ‘up there’ and ‘down there’, whereas Chinese and English have just two, ‘here’ and ‘there’? As a corollary, we can ask whether Chinese or English might develop a system similar to that in Lahu.

(b) Why does Turkish have no genders, while French has two, German has three, and Swahili has eight?

(c) Why does Japanese have a developed system of honorific speech registers, while French has two forms (*tu* and *vous*) of the singular pronoun ‘you’, whose use depends on social distance, and English has nothing of this sort?

(d) Why does Jarawara have three past tenses (one extending from a few seconds to a few weeks ago, another from a few weeks to a couple of years in the past, and the third from a couple of years back to the beginning of time), while Russian just has one past tense?

(e) Why does the “Pennsylvania German” dialect spoken by Old Order Mennonites in Canada differ from standard German in that the verb *wotte* ‘want’ has changed its grammatical profile so that it can no longer be used in constructions like ‘I want to come’?

(f) Why does Igbo (spoken in Nigeria) have a small class of just eight adjectives (which are similar to nouns in their grammatical properties) whereas Korean has a large open class of adjectives (which are similar to verbs in their grammatical properties)?

In our present state of knowledge, useful answers can be provided for only some of these questions. Taking them one at a time:

(a) The Lahu speech community lives in scattered villages in mountainous country (Matisoff 1973). Whether something is ‘up’ or ‘down’ with respect to the speaker is as relevant as whether it is ‘near’ or ‘far’ (shown in ‘here’, ‘there’ and ‘way over there’). The grammatical coding of the height of something referred to thus assists efficient communication. (A similar feature is found in many—but not in all—smallish speech communities inhabiting hilly country.)

It is unlikely that Chinese or English would develop a similar system, simply because these languages have hundreds or millions of speakers living in every kind of terrain. That is, a grammatical feature relating to a particular aspect of the environment could not be adopted into a non-local language.

(b) I have no explanation for the varying number of genders (or the lack of genders) in these languages.

(c) The Japanese people have a highly stratified society, with strict conventions of “politeness”; the relative social status of speaker and hearer determines the speech register that is employed (and this, in turn, serves to mark their relative status). In

the history of Japanese, the use of honorifics has waxed and waned according as social ranks were accorded more or less importance (Shibatani 1990: 123-124).

The French people recognise levels of formality: *tu* is used when addressing a close relative or an intimate friend or a colleague of similar age (and also for talking to children, servants and animals), with *vous* being employed in other circumstances. The ripening of a relationship towards intimacy is marked by shift from *vous* to *tu*. Social relationships in England lack any such clear demarcation, and there is no corresponding linguistic device. (Originally, English had *thou* for singular and *ye* or *you* for plural address. Then *thou* became restricted to addressing friends, children and people of inferior rank, with *you* employed as a mark of respect or when addressing a superior; this is rather like the current situation in French. Finally, *thou* dropped out of use, with *you* becoming the only second person pronoun, in both singular and plural number.)

(d) I have no explanation for why Jarawara should have three past tenses while Russian only has one. Note that there are other languages used by a small speech community (with just a few hundred members) which have a single past tense. And there are languages with millions of speakers that have several past tenses (e.g. Swahili).

(e) The Old Order Mennonite community in Canada which uses Pennsylvania German has strong religious beliefs, which subordinates self to the will of God. The idea of an individual wanting something for themselves is in direct conflict with their faith. This would seem to have motivated the change whereby *wotte* 'want' can no longer be used with an infinitival complement, as in 'want to come' (Burridge 2002).

(f) I have no explanation for why Igbo has just eight items in its adjective class (they are 'large' and 'small', 'new' and 'old', 'good' and 'bad', 'black, dark' and 'white, light') while Korean has a large open class of adjectives. Nor for why in Igbo adjectives are similar in their grammatical properties to nouns, while in Korean they resemble verbs (for example, they take tense suffixes like verbs).

In fact, linguists only occasionally ask "why" questions. I have heard the opinion that each language is ideally suited to its speech community, and vice versa—"that's the way it is". In this view it is politically incorrect to ask why Turkish should have no genders while Swahili has eight (or even to enquire "how" a particular language is suited to its speech community).

If linguistics is to fulfil its function—of fully explaining the nature of human language—then the "why" questions must at some stage be faced. Plainly, the "how" questions come first, and a great deal of work is needed before these can be considered to have been dealt with in a satisfactory way. But, while dealing with the "how" questions, a linguist ought to at all times bear in mind that these are a way-stage that should lead up to the "why" questions. The value of information gained through asking "how" lies ultimately in the light it can shed on "why".

Questions (a-f) were selected to illustrate the magnitude of the task. For (a), (c) and (e), I was able to point towards an answer (which should be pursued in much

more detail). For (b), (d) and (f), I was at a loss to provide any sort of answer. And most of the “why” questions that could be asked belong to this category.

Linguists simply do not pose “why” questions like those given above. In any scholarly endeavour, the first task is to recognise the questions that should be considered, and articulate them as clearly as possible. A great deal of time and thought and research will then have to go into providing answers. Linguists have yet to pose the appropriate questions.

Here is a further small sample of the kinds of “why” questions that need to be asked (for most of them I have, at the present time, no answer):

(g) Languages include a number of words with shifting reference, depending on who the speaker and hearer are, what the place is, and what the time is, for an utterance. There are first and second person pronouns, ‘I’ and ‘you’. John says to Mary ‘I’m hungry’ and Mary replies ‘You’re hungry’; John is referred to by ‘I’ when he is speaking and by ‘you’ when Mary is speaking. There are spatial demonstratives, ‘here’ and ‘there’. John is standing by the door and Mary by the table; he says ‘Put the parcel on the table there!’ and she replies ‘I’ve put it on the table here’. What is ‘there’ for John (distant from him) is ‘here’ for Mary (near to her). And there are time shifters, like ‘yesterday’, ‘today’ and ‘tomorrow’. What is ‘today’ today becomes ‘yesterday’ tomorrow.

From examination of a wide range of languages, it appears that there are always pronouns ‘I’ and ‘you’. Why is this? I can imagine a language without ‘I’ and ‘you’ (indeed, young children’s language sometimes passes through a stage like this)—why does every human language include these pronominal shifters? There are some languages which have a single nominal demonstrative ‘this / that’ (one word corresponding to both *this* and *that* in English. But it appears that every language has at least two spatial demonstratives, ‘here’ and ‘there’. Why is this?

There are languages which have time words ‘yesterday’ and ‘tomorrow’, but lack anything corresponding to English *today*. And there are some languages which even lack ‘yesterday’ and ‘tomorrow’. Jarawara is like this. One can of course refer to ‘yesterday’ with a circumlocution such as ‘before the sun went down’. There is a verbal suffix *-mina* ‘morning’; when used with past tense it refers to ‘this morning’ and with future tense to ‘tomorrow morning’ or just ‘tomorrow’; but there is no lexeme ‘tomorrow’. Now why do some languages, like Jarawara, have words for ‘I’ and ‘you’ and ‘here’ and ‘there’ but lack words for ‘yesterday’ and ‘tomorrow’? It is not that they lack any perception of time—recall that Jarawara distinguishes three past tenses.

(h) The answer I provided for question (a)—concerning why Lahu has demonstratives ‘up there’ and ‘down there’—was that this speech community lives in mountainous country. Some other groups living in similar terrain include specification of ‘up’ and ‘down’ in their grammar. But not all. Why not? Why should some groups of mountain-dwellers code information about their geographical habitat in their grammar, while other groups omit to do this?

(i) It has been the accepted doctrine of linguistics, for a couple of hundred years, that each language has only one parent. Suppose that one political group, speaking language X, merges with another political group, speaking language Y, so that they become a single community. This community will either speak language X (with quite a few loans from Y) or else they will speak language Y (with quite a few loans from X). It will not speak a language which is a blend, equally related to X and to Y.² Why is this so?

There are many examples of this. For instance, when the Normans (speaking a dialect of French) invaded England in 1066, the new Norman-plus-Saxon community soon came to speak just English (with some admixture of French words, but virtually no French grammar). When the English invaded New Zealand, in the nineteenth century, the common language of the new English-plus-Maori community came to be English, with just some words (no grammar) taken over from the Polynesian language Maori.

(j) Why is it that contiguous languages tend to become more like each other in certain ways but not in others? For instance, if a language without tones comes into contact with one or more languages that have tones, it is likely to develop its own system of tones. This applied to Vietnamese, under contact with Chinese; and it has applied across a large area of Africa. If a language with no gender system comes into contact with one or more languages that have gender, then it is rather likely to develop its own system of genders (by change within itself). The attributes of 'having tones' and 'having a gender system' tend to diffuse between languages within a geographical area. But it appears (on the basis of preliminary study) that, for instance, the attribute of 'having a complex tense system' typically does not diffuse in the same way. Why not?

5. Evaluation

It is normal in any enquiry to assess and evaluate. Scholars enquire which is the preferred alternative—of economic systems, political systems, taxonomic models, ideas concerning the origin of the universe. Similar questions should surely be asked concerning languages. Given two languages, X and Y, which is the most complex? Which is easiest to learn? Which language is—overall—better?

Such questions are never asked by linguists. Indeed it is considered subversive to even think of them. There is a reason.

² This scenario applies to spontaneous language development. There are examples of "mixed languages", but in each case this is the result of deliberate and conscious construction of a language, generally by an ethnic group which wishes to have a language of its own, or to have a special language style for certain in-group purposes. See Dixon (1997: 11-13). Creoles constitute a special case, which it is not pertinent to discuss here.

The self-aggrandisement of the white race has had a number of consequences. Beginning in the sixteenth century, the white race (previously confined to Europe and adjacent parts of Africa and Asia) colonised almost the whole world. By 1910 the only countries that were not governed by white people were Liberia, Ethiopia, Thailand, China, Tibet, Japan and Korea. The situation began to reverse after the second world war, when the indigenous peoples of Africa, Asia and the Pacific Islands re-established their own control over their territories. (This reversal happened too late for the Americas, Australia and New Zealand, where the white invader had established a superiority of numbers, with the indigenous peoples being left as a marginalised minority.)

Hand-in-hand with this political imposition came a total lack of respect for the religions, laws and customs of the indigenous population. And, most especially, for their languages. It was believed (as an article of faith) that the languages of non-white people—whether in Africa, in Brazil, in New Guinea, or in Australia—were simple affairs, with just a few hundred words and a paucity of grammar. These were looked upon as primitive peoples, and it was natural that they should speak primitive languages.

Attitudes of this sort persist today. In Australia, for example, people talk of Aboriginal “dialects” but of European “languages”. They evince surprise and disbelief when told that there were originally about 250 distinct languages spoken in Australia, as different from each other as are Turkish and Hungarian. When I began field work in Australia, in 1963, a white farmer asked what I was doing. “Writing a grammar of the local Aboriginal language” brought the response “That should be pretty easy.” When I enquired why, he replied: “They haven’t got any grammar”. (My published grammars of two of the languages of that region run to 420 and 563 pages respectively—Dixon 1972, 1977.)

The first task of the writer of a textbook or the teacher of an introductory course in linguistics is to combat this mistaken belief. We stress that there is no such thing as a primitive language in the world today. All languages are roughly equal in complexity. The *Oxford English Dictionary* may comprise half-a-million words but many of them lie in specialised fields (such as law and medicine) or are archaic. It is likely that the average speaker of any language (whether English or Turkish or Chinese or Lahu or Jarawara) has about the same size working vocabulary—about ten thousand words. And the grammar of every language has about the same level of complexity; it takes about three to five hundred pages to explain the grammatical organisation of any language, in moderate detail. Some languages do, of course, have complex word structures while others have rather simple word structures; but those with simple word structure tend to have complex syntax, and vice versa. The areas of complexity differ from language to language but—according to the received doctrine among linguists—the overall level of complexity is about the same. There are no primitive languages.

This is true. No language is a hundred times or even ten times as complex as any other (which is what the European colonisers believed). All languages spoken in the

world today are *roughly* similar in their lexical and grammatical wealth. But surely they are not all *precisely* equal? Some may be a little better than others—perhaps ten or twenty percent better. This is a question that I have posed in academic debate and been told—by some—that it is quite out of order.

There are three historical stages. Firstly, the belief that non-white-skinned people speak primitive languages. Secondly—to rebut this—the statement that no present-day language is in any way primitive; they are all roughly equal in complexity. I suggest that we are now ready to proceed to the next step. Agreeing that all languages are *roughly* equal, we can suggest that they are not all *exactly* equal, and then enquire whether some might be better than others (with specification of what is meant by “better” in this context).

Two reasons have been given to me for not proceeding to the third stage. One is that it might be confused with returning to the first stage. This is a matter of genuine concern, since there are many people who do still (as a matter of ignorance, or prejudice) subscribe to this belief. The other reason—given me by a linguist of Polish origin—is that one shouldn’t, as a matter of political correctness, attempt to evaluate languages. She was scandalised when I gave a talk entitled “Are some languages better than others?” and exclaimed “But I could never evaluate or criticise my native language, it would be like criticising my mother, which I would never do.”

Nevertheless, it clearly behoves linguists to naturally move beyond the “how” and “why” questions, to the matter of evaluation. This must relate to some specified parameter(s). I will briefly mention some of these.

(1) How easy is it to learn a given language, (a) for a child, and (b) for an adult?

It is undoubtedly the case that a language with simple word structure is easier to acquire a basic competence in, than a language with complex word structure. This applies for both child and adult learners.

One thing that militates against easy acquisition of a language—especially by an adult learner—is a plethora of irregularities. Before embarking on linguistic fieldwork in Amazonia, I endeavoured to learn Portuguese. A high proportion of lesson time involved memorising irregular verb after irregular verb. Irregularities are a needless complication, which contribute nothing to the communicative power of a language. A student learning a language like Turkish, with rather little irregularity, can devote all their efforts to achieving an understanding of word structures and of the basic syntactic patterns, which serve to code and transmit meanings. In similar vein, a language like Turkish with an agglutinative structure (where each meaning element has a segmentable form) is easier to learn than a language like Latin with fused morphemes (for example, the *-t* of *amat* ‘he/she loves’ is a portmanteau coding all of: third person, singular number, present tense, active voice, and indicative mood).

How easy it is for an adult to learn language X depends on what their first language is, and how similar it is to X. For example, someone who already speaks a tone language is likely to find it easier to learn X, if X has tones, than will a speaker of a non-tone language. As another example: if, in a multilingual region, all the languages have a simple consonant-vowel-consonant-vowel structure (as in *bali*

'head'), except for one language, Y, which has much more complex structures (as in *smbarg* 'head'), then people will find it hard to learn Y, and Y will be less and less acquired as a second language and less used as a language of inter-group communication than the other languages of the area. Similar remarks apply to a language with complex word structure, in an area where other languages have a simple word structure.

(2) How does a language measure up with respect to explicitness?

The ideal language should have the means to refer to any situation explicitly and unambiguously, with the amount of detail that is required. To this end, words and grammatical elements should each have a specific meaning. In English the word *hot*, when applied to food, can mean either 'of high temperature' or 'very spicy'. This can and does cause confusion. German is a better language in this respect, having distinct adjectives *heiß* 'hot in temperature' and *scharf* 'spicy hot'. When comparing two languages, one has to evaluate such instances over the whole grammar and lexicon.

(3) How does a language perform in terms of succinctness and cohesiveness?

If there are two ways of saying something, which differ only in length, then the shorter way is better. A two-syllable word is superior to one of six syllables, all else being equal. It is good to communicate a message as quickly and efficiently as possible. As a corollary to this, a simple word should be preferred to describe a common property, rather than some compound or circumlocution. It is easier to say *blue* rather than something like *sky-coloured*, and to say *cruel* rather than something like *hard-hearted*.

In similar fashion, communication is more effective if two clauses which are linked semantically are joined together in grammar. For example, *John jumped off the wall and twisted his ankle* is more cohesive than *John jumped off the wall and John twisted his ankle*. And *I don't like that man who tormented the cat* is superior to *I don't like that man; that man tormented the cat*.

(4) How does a language perform in terms of richness and subtlety of expression?

If a language has varying ways for referring to a given thing or action or property, this will provide wide possibilities of description. For example, corresponding to English verbs *cover* and *cook*, the Australian Aboriginal language Dyirbal has, in each instance, both a verb and an adjective. There is a verb *guban* 'cover' and its participle *gubangu* 'covered' can refer to something covered in any way; for example, a blanket over just half a bed, or crumpled up on it. There is also a non-cognate adjective *ngulguñ* and this means 'covered properly, covered all over'. The verb *ñadjun* 'cook' has a participle *ñadjungu* which can describe something being cooked a bit or a lot, not enough or too much. In contrast, the non-cognate adjective *ñamu* means 'cooked to perfection'. (This is just a sample, Dyirbal has several dozen more verb-and-adjective pairs, where English has just a verb.)

(5) How does a language measure up in terms of richness of grammatical systems?

Some languages have three nominal demonstratives. Their meaning can be 'this, near speaker', 'that, near hearer' and 'this, distant from both speaker and hearer'; or

else ‘this, near speaker’, ‘that, mid-distant from speaker’ and ‘that, far from speaker’. These provide more information than the two nominal demonstratives in English, *this* and *that*, and will facilitate the transfer of more exact information. Similar remarks apply to a tense system which distinguishes several degrees of past and / or of future. A pronoun system which recognises three numbers (for example, ‘you one’, ‘you two’ and ‘you more than two’) is more informative than one with just two numbers (‘you one’ and ‘you more than one’) and this is in turn superior to the second person pronoun system in formal English which just has one term, *you*, referring to any number of people. (An invitation such as *Would you like to come to dinner?* leaves the hearer not knowing whether or not their spouse is also invited, sometimes with embarrassing consequences.)

The evaluation of a language (and comparison of the relative merits of two languages) must relate to one or more of the functions for which the language is intended. For example (this list is far from exhaustive):

- Just describing events, such as a battle or a sporting contest or a tea party.
- Talking about things that are specific to a given community. For instance, if some group has a classificatory kinship system (where everyone in the community is regarded as related to everyone else, by applying a set of algorithms) then an appropriate set of nouns and verbs will be needed to talk about this.
- For technical description; for example, of how to build a machine and how to operate it.
- For philosophical (or religious) discourse, ruminating on the nature of matter or the purpose of life.
- For song and poetry and various kinds of literature.

The value of a language can have consequences for the people who use it. Suppose that there are two small communities, A and B, at the opposite ends of a large island with other communities between them. In each of A and B, a charismatic leader emerges who is determined to conquer all other communities and rule the entire island. But whereas A has a language with complex word structure, considered difficult to learn by outsiders, B has a language with simple word structure, similar to that of most other languages in the island and considered easy to learn. Which leader is more likely to succeed in controlling the island? Plainly, many factors will be involved—weapons available, military organisation and strategy, how conquered groups are treated and whether they will ally themselves with the conqueror and join his army. But language will be one factor. The greater accessibility of the B language, in this particular situation, will provide the leader of the B people with a definite advantage.

My comments here have been tentative and exploratory. There is no tradition of evaluating languages (and of asking questions such as “Are some languages better than others?”) on which I can draw. But if linguistics is to tackle the analysis and assessment of its subject matter in the way that other disciplines do, these are questions which must be systematically investigated by linguists.

6. Prediction

The basic profile of science is to describe, to explain, and then to predict. For example, in 1915 it was suggested by Percival Lowell that perturbations in the orbit of the planet Uranus could be explained in terms of the gravitational attraction of an outer planet in a certain position. The prediction was confirmed when Pluto was actually observed, in 1930.

There are a number of ways in which linguists could make predictions about languages, which might be confirmed or falsified. I will briefly mention three. Firstly, it is possible to predict what the semantic content of a given grammatical category will be. Secondly, it is possible to say that if a grammar includes a certain feature, then this is likely to entail the presence or absence of some further feature. And, thirdly, it should be possible to predict how a given language is likely to change over time.

One caveat is in order. In any discipline (like linguistics) dealing with some aspect of the behaviour of humans, the results will be in terms of probabilities rather than of certainties. Linguistics does not feature “laws” like those of physics; it deals instead in strong tendencies, in likely correlations, and in predictions that have a high chance of proving to be correct.

(i) Predicting the content of a small adjective class.

Under (f) in §4, I admitted that I did not know why Igbo (and other languages) have a small class of adjectives, while Korean (and other languages, including English) have a large adjective class. What *can* be done is to *predict*, for a language with a small adjective class, what its semantic content is likely to be. The prototypical members of an adjective class are words referring to dimension, age, value, and colour. Recall, from §4, that Igbo has two of each: ‘large’ and ‘small’ (dimension), ‘new’ and ‘old’ (age), ‘good’ and ‘bad’ (value), and ‘black, dark’ and ‘white, light’ (colour). Slightly bigger classes may include more words from these semantic types (for example, ‘long’, ‘short’, ‘red’) and also some physical property items (for example, ‘raw, green, unripe’, ‘heavy’, ‘light’, ‘sharp’, ‘hot’). Only when an adjective class is much bigger (with at least a few score members) is it likely to include terms referring to human propensities (for example, ‘happy’, ‘jealous’, ‘kind’, ‘clever’).

If you tell me that, working on a previously undescribed language, you have recognised a small class of, say, fifteen adjectives, I will be able to predict what their meanings are likely to be. I wouldn’t expect my predictions to tally exactly, but I would expect to have a high measure of success. (If I had a bet on the outcome, I’d expect to make money.)

Similar predictions can be made in many other areas of grammar. Suppose you discover that your newly described language has two varieties of complement clause (which can function as object of a main verb), one like English ‘to’ as in *I decided to go*, and one like English ‘-ing’ as in *I like eating*. Now I would predict that verbs like ‘want’ and ‘hope’ and ‘try’ and ‘tell’ would take a ‘to-type’ clause, while verbs

like 'see' and 'hear' and 'imagine' would take an '-ing-type' clause. These predictions are likely to be predominantly—but not absolutely—correct.

(ii) Predicting associations between the components of a grammar.

In a sentence like *Woman sees man*, there is one verb and two nouns. The grammar must be able to specify whether it is the woman who does the seeing and the man who is seen, or vice versa.

There are basically three ways of showing this. Some languages have what is called "dependent marking": for example, what is subject is shown by nominative case ending and what is object by accusative ending. Thus, Latin has *mulier virum videt* for 'woman sees man'; since the nominative ending on *mulier* 'woman' shows that it is subject, and the accusative ending on *virum* 'man' shows that it is object, the three words in this sentence can occur in any order—for example *virum videt mulier* or *virum mulier videt*—without any difference in meaning.

Other languages (including Swahili and Navaho) have what is called "head marking"; this requires obligatory subject and object pronouns as part of the verb. A sentence such as 'woman sees man' would be rendered as, literally, 'woman man she-him-sees'. Again, the words 'woman' and 'man' can occur in any order, since what is subject and what is object are shown by the pronoun indicators 'she' and 'him' on the verb.

And there are languages which simply use the order of words in a sentence to indicate their function, as in English where the subject precedes the verb and the object follows it.

Now if a language has head marking we can predict that it is unlikely to have dependent marking, and vice versa. If it has neither dependent marking nor head marking, then it is likely to use the order of words to show syntactic function. As with most linguistic generalisations, this is not a hard-and-fast prediction. There are some languages which show both head marking and dependent marking (for example, Classical Arabic). And there are languages which lack head marking and dependent marking but also permit words to occur in any order in a sentence (for example, Lao); here, what is subject and what is object has to be inferred from the surrounding discourse and from the situational context. But the number of exceptions to our statement that a language is unlikely to have both dependent marking and fully-developed head marking are relatively few, and do not affect the value of this generalisation concerning a strong and pervasive tendency across human languages.

Other types of prediction are possible about the interaction between components of a grammar. As mentioned under "HOW-2" in §2, if there are different numbers of tense choices available in positive and in negative sentences, then the greater set of choices will be in positive sentences. Just a little work has been done on dependencies between grammatical systems. As a further example, the number of gender specifications available may depend on the tense that is chosen, but we never find the number of possible tense specifications depending on the gender that is chosen (see

Aikhenvald & Dixon 1998). Generalisations like this can be the basis for predictions concerning how some newly described language is likely to be structured.

(iii) Predicting how a language will change.

The ways in which a given language is likely to change will depend on (among other factors) the social circumstances in which it is used, the languages it is in contact with, and the speakers' attitudes towards their own and towards other languages.

The ways in which a given language is likely to change in the future would surely be a fruitful field of enquiry. However, linguists have scarcely ventured into it. I will here just put forward some elementary speculations concerning likely future directions for English.

Old English had a rich array of irregular verbs, and also some irregular nouns; these are gradually being replaced by regular forms. The plural of *cow* used to be *kine* and has been replaced by *cows*. The plural of *brother* was *brethren* but has now been replaced by *brothers* (the form *brethren* is retained just for members of a religious order). In due course, *oxen* will certainly be fully replaced by *oxes* and, I predict, *children* will—not soon, but in the fullness of time—be replaced by *childs*.

Irregular verbs are also gradually being regularised; for example, *dreamed* is now often used in place of *dreamt*, *kneeled* in place of *knelt*, *lighted* in place of *lit*. Occasionally a new irregular verb is introduced (for instance, *dove* can be used in American English in place of *dived*, on the pattern of *drive* and *drove*, *ride* and *rode*) but the overall tendency is towards reducing the number of irregularities. It takes little imagination to suggest that this trend will continue, and only a little more to suggest which irregular forms may fall out of use over the next centuries (I'd opt for the elimination of, among others, *crept*, *fled*, *bled*, *flung*, *strung*, *ground*, *swore* and *blew*.)

Let me venture to predict another kind of change which has not yet (to my knowledge) commenced. English is replete with what are called "phrasal verbs". These each consist of a verb plus a preposition, but with the combination having a meaning that cannot be inferred from the meaning of its parts. For example, *take after*, as in *She takes after her mother*; and *pick on*, as in *The teacher picked on John*; and *sum up* as in *He sums up the situation in one sentence*. At present each of these is one lexeme which consists of two words; the tense is shown on the first word, the verb. I predict that, in the not too distance future, these may be reanalysed as each being a single complex word, with tense added at the end. That is, instead of *He sums up the situation* we will hear *He sum-up-s the situation*; and also *She take-after-s her mother*, and then—eventually—things like *The teacher pick-on-ed John*.

A serious examination of how languages are known to have changed in the past—and why, and under what conditions—should be allied with a study of present-day dialect variation. From this, it should be possible to predict how a major language like English is likely to change over future time. And also how a small tribal language (most of whose speakers also have some proficiency in the lingua franca of the nation in which they live) is likely to change. (If, indeed, it does not move into ex-

tion, as the vast majority of small languages are likely to, in the fairly near future.)

7. The present situation in linguistics

I will now, as promised, look briefly at the concerns of people who are on the payroll of departments of linguistics. By and large, these are quite different from those discussed above.

Chomsky is by far the best known linguist, both within and without the field. He has spawned a succession³ of formal theories that are held to be “simple and elegant ... with fundamental principles that have an intuitive character and broad generality”. Chomsky does state that “language is a biological system, and biological systems typically are ‘messy’”. His *modus operandi* is to investigate the “human language faculty”, focussing on what is called “logical form”. This appears to be a kind of computational mechanism which relates to just a part of the grammar of a language (leaving aside the messy bits), considered quite apart from any cultural context. There is no attempt to provide a framework in terms of which a comprehensive grammar of a natural language could be formulated; rather, bits of individual grammars may be taken account of (without any reference to how they relate to the rest of the grammar), as they may be perceived to help explicate some aspect of the “human language faculty”.

Chomsky states that the “task at hand” for his Minimalist Program “is to show that the apparent richness and diversity of linguistic phenomena is illusory and epiphenomenal, the result of interaction of fixed principles under slightly different conditions” (quotations from Chomsky 1995: 29, 8). His program operates in terms of constraints, rule-systems and principles that are inviolable. This is in stark contrast to the substantive generalisations in linguistic typology (as briefly mentioned above), which are—due to the nature of language—almost always probabilistic in nature, describing strong tendencies but very seldom absolute correlations.

A plethora of other formal theories have emerged, briefly flourished, and then faded (these have mostly been initiated by Chomsky’s students). One current preoccupation is Optimality Theory, which was first introduced in phonology. There are a number of possible “optimality constraints” and each language utilises those most appropriate for its structure. For example, if in language X each word ends in a vowel, then “having each word end in a vowel” is an optimality constraint for the language. (There may be odd violations, such as a few unassimilated loans which end in a consonant.) This is rather like saying, at the end of a football match, that the team which won was destined to win. Optimality Theory has recently been extended to

³ Since about 1970 the names used by Chomsky for his sequence of formal theories include: Standard Theory, Extended Standard Theory, Revised Extended Standard Theory, Government and Binding Theory, Principles and Parameters Theory, and the Minimalist Program (I may have blinked and missed a couple).

syntax. Instead of constraints, rule schema and principles which are inviolable, we now have constraints, rule schema and principles which are violable, with a form that shows least violations being preferred (Barbosa et al. 1998: 1-2). Unsurprisingly, the problems which could not be solved within the earlier Minimality Program, have since reappeared in different guise within Optimality Theory.

Approaches of this type are espoused by more than half of the people employed, across the world, in departments of linguistics.⁴ There is a particular concentration of formalists in the eastern part of the USA, with considerable pockets in the centre and west, and also in Europe. And there are camp-followers in South America. In Australia, formalists are fairly scarce.

The only way to become a competent practitioner in any craft or discipline is to actively engage in it. A watchmaker must actually make and mend watches, a surgeon must perform operations, a geologist must examine rocks (both in the field and in the laboratory), an organic chemist must conduct experiments in the laboratory. In the same way, a linguist must describe languages. The ideal apprenticeship is to undertake fieldwork on some previously undescribed (or scarcely described) language—recording, transcribing and analysing texts; observing how people use the language in the daily round; writing a grammar and phonology; compiling a dictionary; and publishing a volume of annotated texts.

Once a linguist has mastered the fundamentals of the discipline in this way, they are equipped to go on to the next stage, working on linguistic typology. This involves studying the grammars of a selection of languages, inductively inferring some putative generalisation (for example, that mentioned under (ii) in §6, that gender specifications available may depend on the choice from a tense system, never the other way round). The next step is to thoroughly check a generalisation, examining as many languages as possible (not, at this stage, just a sample of languages).

Only some of the people who call themselves typologists follow this path. Besides the great mass of formalists (who are explicit about what they are doing) there is also a group who can be called “armchair typologists”. They have not undertaken the apprenticeship which is necessary for mastering the fundamentals of linguistics—writing a grammar of a language (ideally, based on fieldwork in a community where the language is actively spoken). The armchair typologists simply try to do typology “in the raw”; this is rather like someone attempting to make generalisations in geology without ever having examined in detail a single rock. Their work is, by and large, slight and unenlightening. Indeed, they share with the formalists a predilection for pretty little, nice-and-neat statements, which mock the untidy and messy nature

⁴ A few formalists have done a little work on languages—for example, Baker (1988, 1996) includes data on Mohawk from his own field work. However, this is based on elicitation from English, rather than on analysis of running texts in Mohawk, examined within the cultural context in which they occur. And Baker’s transcription is inadequate in that he does not distinguish vowel length. (This is rather like failing to distinguish between *sleep* and *slip*, *deep* and *dip*, etc. in work on English.)

of the totality of a human language (a messiness which they have never exposed themselves to).

Moving beyond the formalists and the armchair typologists, there are some hundreds of people describing languages, learning the discipline, and then going on to judicious typological work on an inductive basis. These people—a minority of those employed in university departments of linguistics—are involved in the true business of linguistics, adding to the objective stock of knowledge concerning spoken languages, and working together to refine our understanding of the basic nature of human language.

They ask some of the “how” questions but—as yet—scarcely any of the “why” questions. Nor do they pay attention to evaluation of whether some languages are—in certain defined respects—better than others. And they are almost never concerned with prediction. Linguistic scholars of today are, effectively, nibbling at the surface of an apple of knowledge concerning the nature of language; we need to take a measured bite, towards the core.

In essence, the discipline of linguistics needs to reinvent itself, in order to tackle the questions I have outlined. This can only be done by a new generation dissociating itself from most of its elders, and setting out to really investigate the nature of language, as it is used.

References

- Aikhenvald, Alexandra Y. & Dixon, R. M. W. 1998. Dependencies between grammatical systems. *Language* 74, 56-80.
- Baker, Mark C. 1988. *Incorporation: a theory of grammatical function changing*. Chicago: University of Chicago Press.
- Baker, Mark C. 1996. *The polysynthesis parameter*. New York: Oxford University Press.
- Barbosa, Pilar & Fox, Danny & Hagstrom, Paul & McGinnis, Martha & Pesetsky, David 1998. *Is the best good enough? Optimality and competition in syntax*. Cambridge, Mass.: MIT Press.
- Bhatia, Tej K. 1993. *Punjabi: a cognitive-descriptive grammar*. London: Routledge.
- Bloomfield, Leonard 1933. *Language*. New York: Holt, Rinehart and Winston.
- Burridge, Kate 2002. Changes within Pennsylvania German grammar as enactments of Anabaptist world view. In: Enfield, N. J. (ed.) *Ethnosyntax. Explorations in grammar and culture*. Oxford: Oxford University Press. 207-230.
- Chomsky, Noam 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Dixon, R. M. W. 1972. *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. 1977. *A grammar of Yidiñ*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. 1991. *A new approach to English grammar, on semantic principles*. Oxford: Clarendon Press.
- Dixon, R. M. W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. 2004. *The Jarawara language of southern Amazonia*. Oxford: Oxford University Press.

- Matisoff, James A. 1973. *The grammar of Lahu*. Berkeley & Los Angeles: University of California Press.
- O'Grady, Geoff N. 1990. Prenasalization in Pama-Nyungan. In: Baldi, Philip (ed.) *Linguistic change and reconstruction methodology*. Berlin: Mouton de Gruyter. 451-476.
- Shibatani, Masayoshi 1990. *The languages of Japan*. Cambridge: Cambridge University Press.