

---

## **A TYPOLOGY OF LINGUISTIC FIELD RESEARCH**

---

**ASMAH HAJI OMAR**  
Faculty of Languages and Linguistics  
University of Malaya

### **INTRODUCTION**

The term "field research" (i.e. in linguistics) has come to mean research that is not conducted in the library. A trip to a remote settlement to record the language spoken by the people there is called field research. So is research done in a car assembly plant to see how the people in the various levels of hierarchy communicate with each other in a horizontal or vertical relationship, or research done in the classroom to see how students perform when given a certain set of language usages.

On close examination, instances of research conducted outside the library differ from one another according to the following parameters:

- (i) focus;
- (ii) purpose;
- (iii) objective;
- (iv) method of collecting data;
- (v) location,
- (vi) uses.

On the basis of these differences, field research in linguistics can be divided into two types: *upstream field research* and *downstream field research* (Asmah Haji Omar, 1993).

### *Upstream field research*

Upstream field research is older than its downstream counterpart. It was this type of research that opened the way for modern linguistics in the tradition of Bloomfield in the USA and J.R. Firth in the U.K. to name a few

### *Focus*

Upstream field research can be equated with Samarin's field linguistics (Samarin, 1967). The focus of this research is on hitherto unknown languages or dialects and their communities. That is to say, these languages or dialects have never had any description in terms of their systems and structures. Nothing has been written about them. If at all there is any written document on them it may be in the form of wordlists, usually produced by interested travellers or visitors to the place. These people recorded the words usually out of casual interest. If they were researchers who were non-linguists, the list took form because of a need to know the native nomenclature for artefacts, plants, animals, etc. All the lists are usually bilingual, i.e. with the native words which are given equivalents in the languages of the compiler.

Some of what I call "hitherto unknown languages" may have phrase books for foreigners visiting the community. The phrases given are simple everyday expressions that people may use when they encounter one another in a social or business situation.

As the compilers of the wordlists and the phrase books are not linguists, the compilations are usually not done according to any system. Nevertheless, such lists have their usefulness, especially to a first-time visitor or researcher to the community

### *Purpose of Research*

The purpose of research is to give a written form of the language/dialect to its own speakers and to have documentation on the language/dialect as a step to placing it in the inventory of the world's languages.

When a community has its own language in written form, it is able to go a step further in teaching the language formally, i.e. in the school, to its own children. With this comes manuals for teaching the language, dictionaries and literature books and books of other genres.

The availability of documentation on the language/dialect will be useful to linguists in general, or those working on the typology of languages in the area or on a comparison of languages. Besides the linguists, people from other disciplines may find such documentation useful for research, for example the anthropologists and the ethnobotanists working in the area.

### *Objectives of Research*

Arising from the purposes mentioned above, the linguist in upstream research has to set his/her objectives. These can be summarised into descriptions of the systems and structures of the phonology, morphology, syntax and semantics of the language/dialect concerned.

Phonological description is necessary in terms of giving the languages/dialects concerned a written form with a spelling system that can be used by others. These "hitherto unknown languages/dialects" usually do not have a writing system or an orthography. They usually adopt a writing system from another language, usually a major language of their country. The aboriginal languages in Peninsular Malaysia and the indigenous languages of Sabah and Sarawak fall into the group of languages which call for upstream research. So do most of the Malay dialects.

To put all these languages and dialects into writing, the easiest way is to adopt a writing system which is already used by a major language. There are many examples of such languages which we can cite. Among these are Iban and Kadazan Dusun, which have adopted the Rumi (Roman) writing system through Malay. There have been cases where languages tried to create their own writing systems from symbols which seem to occur in their culture. An example is Bidayuh. However, this has not proven to be a success. It goes to show that creating a writing system is not as easy as "discovering" a phonological system. With the writing, phonological and spelling systems available, the language can already be written, i.e. the spoken word can be transferred into a written form.

A more systematic picture of the written form can be produced if the language has a general morphological and syntactic description. With this resource, words in all their simplicity and complexity can be

written in a systematic and consistent way. This is so because units of the words can be inventorised and the method of writing them can be codified. A syntactic description is needed for the writing of phrases (as opposed to words) and sentences. Phonology, morphology and syntax are the basic aspects of a language which make it possible for putting the language into writing not just in terms of words but also in terms of sentences and discourse.

At the earlier means knowing the meanings of words that a researcher collects from his/her informants. A full semantic study of the language is better carried out when the researcher becomes more familiar with the language and has acquired a certain level of proficiency in communicating via the language concerned.

#### *Method of Collecting Data*

The upstream field research makes use of informants. This is why it is also called the *informant method* (Samarin, 1967: 1). The informant is a native speaker of the language/dialect under study. As he/she lives in the community of the language/dialect concerned, the researcher has to go to the field to collect data from him/her as well as from other native speakers.

Upstream field research also allows the researcher to use an informant available at the researcher's work place, without the researcher going to the community. That is to say, a researcher in Kuala Lumpur is able to collect data on Iban if he gets an Iban native speaker in his place of work. In such a situation he takes the informant, i.e. the Iban speaker concerned, to his place and does a recording of the language, followed by a description or an analysis of the language. Research which involves bringing the informant to the work place is also known as *field-type research*. (Samarin, 1967:1). This means that the collection of data is still the same as that involved when the researcher goes to the field. The only difference is that the researcher does not have to go to the field.

Collection of data through informants has a set of techniques, each with its purpose and objective. However, these techniques will not be discussed here.

*Location*

The languages/dialects which are the targets of the upstream field research are found in the upstream regions of the country, hence the nomenclature given to this type of research. Examples of research done in such places are those of Asmah Haji Omar (1969, 1981, 1983) and Eric Wong Tat Koon (1994).

*Uses of Upstream Field Research*

Upstream field research has practical as well as academic uses, as discussed below

*Giving the language its infrastructure*

This is the most immediate use of the research, i.e. in giving writing and spelling systems to the language, and compiling a lexicon for it. These two efforts can lead to the writing of grammar. All these are what I call the "infrastructure" of a language.

*Identification of language subfamilies and their membership*

Language families sufficient data on the "members" which can be grouped together. Such identification needs descriptions of the phonology, morphology and syntax of these members.

In the absence of such descriptions people are prone to making conjectures. For example, until lately, Bidayuh (of Sarawak) was assumed to consist of four different dialects, i.e. Biatah, Bukar Sadong, Bau and Jagoi. When more and more research had been done, these so called dialects proved to be heterogeneous languages. Arising from this, Bidayuh cannot be considered a single language (with dialects) but rather a subfamily of heterogeneous though closely related languages. Apart from the four mentioned above, there are other "speech systems" within the Bidayuh area which can be considered languages in their own right but within the Bidayuh subfamily.

The same can be said of Melanau. It is not a single language but rather a subfamily of languages, among which are Mukah, Oya-Dalat, Matu Daro, and Rejang.

### *Realignment of subfamilies*

People are in the habit of aligning ethnic group names to names of languages. In most cases they are right. However, this method of alignment may not always give a true picture of the linguistic versus group situation.

An example is classifying the language spoken by the Melanau people of Bintulu as a language of the Melanau subfamily. Hence, Bintulu Melanau was grouped together with the other members of the Melanau subfamily. With field research into the Bintulu speech system, it was found out that this system had features that necessitated placing it in a different group. This speech system is now known as the Bintulu language, although its speakers are still referred to as Melanau. Pending further field research, Bintulu now stands on its own as a subfamily.

### *Determining the degree of genetic relationship between languages*

Field research has been able to show the proximity or distance in genetic relationship between members of the same subfamily. This is done by comparing wordlists and linguistic Dusun (also known as Kadazan) subfamily where Lotud is considered to have features which make it distant to Penampang (or Tanggara), Ranau, Tambunan, etc. The latter group of languages are considered to have a closer relationship with each other than any to Lotud. In turn, all these are closer to each other than they are to Rungus.

### *Determining language typology*

When languages belong to a single family or subfamily they are said to share certain features in their systems and structures. However, such languages may differ from one another due to the presence of features in some but not in others. This gives rise to a typology of the languages concerned. For example, using numeral classifiers as a variable, the languages of Sarawak may be classified into two types: those with a system of numeral classifiers and those without (Asmah Haji Omar, 1985: Chapters 4 dan 5).

Linguistic typology can also be done across subfamilies and families of languages. Looking at the indigenous languages spoken in Sabah and Sarawak, specifically at their phonological and morphological features, one can posit two main types: the Sabah-type and the Sarawak-

type of languages. This does not mean that the former are all found in Sabah, and the latter in Sarawak. A Sabah-type language is found in Sarawak, and that is Bisayah which is spoken in Limbang. Conversely, a Sarawak-type language is found in Sabah, and that is the Lun Dayeh language of Nabawan. Bisayah has features which make it more akin to Paitan, Murut and other Sabah-type languages than it is to those of the Sarawak-type. In the same way, Lun Dayeh is very much closer to Kelabit and Lun Bawang of Sarawak than it is to Murut of Sabah. In earlier writings on Sabah, the Lun Dayeh were called Murut, and their language, Murut. This is because these people's closest neighbours are the Murut of Sabah.

#### *Discovering "new" languages and their names*

Upstream field research usually opens the door to a "new" speech system or language. The language has always been there, only that it had not been "discovered" before, due to factors which have already been discussed above, e.g. no fieldwork had been done on it and as such there had not been any documentation on it, or it had always been subsumed under the name of another language or subfamily of languages.

Below is my personal experience with regard to such a discovery. It was the "discovery" of the Narom language. In 1978 I had stopped over for 2 days in Marudi on my way from Miri to the Tutoh area. I had gone to the District Office to find out about the demography of the area. I was told that there was a *Kampung Melayu* nearby I took the opportunity to go to the *kampung* with my tape recorder and managed to collect language data from some informants. What came out was not any dialect of Malay but a totally different

The people at first said that they were Malays. (Their houses were of the Malay-type and they were Muslims). But at my probing, they informed me that they were *Narom*. I have given a chapter to the Narom language (not the Malay dialect of Marudi) in my book *The Malay Peoples of Malaysia and Their Languages* (1983).

#### *Rectification of folk classification*

People tend to classify an ethnic group or its language(s) according to their general perception of those people as well as their general impression of the language. Iban (once known as Sea Dayak) was considered to be a dialect of Malay (Cense and Uhlenbeck, 1958). A well-

known Malay figure once said in his speech that "with a little bit of imagination", a Malay is able to speak Iban, because it was only a dialect of Malay just like the Kedah, Johor, etc. dialect. His impression arose from lists of words which showed words in Iban and Malay which are similar or almost similar to one another, e.g.

Iban	Malay	English
rumah	rumah	house
makai	makan	eat
panjai	panjang	long
aku	aku	I

Apart from words such as the above, the numerals in Iban are carbon copies of the Malay ones. However, there are more aspects to these two languages than those lists of words, and these necessitate the learning of Iban by Malays as they would any other language. Field research involves more than the collection of such wordlists, and the results confirm the fact that Iban and Malay are two heterogeneous languages.

The Naron case can be cited as another example of the rectification of folk classification. Naron is not a Malay dialect but a language in its own right which shows a close affinity with Kenyah.

Another example which I wish to cite here pertains to the Selakau people and their language. *Selakau* (pronounced with the diphthong *au*) is the name used by the people concerned (found in Lundu, Sarawak) to refer to themselves as well as to their language. Along the way they have come to be known as *Selako*, a pronunciation probably originated by the Malays, and this name has stuck on with the non-Selakau. There is no *o* in the language concerned, and with this evidence alone one can conclude that the actual name is *Selakau*, and this name should be chosen above *Selako*.

#### *Formulation of linguistic theories*

Data is needed in order to formulate theories especially in general linguistics. This can only be obtained through field research. As an example, descriptions of the languages of Sabah have shown that allomorphic variations of morphemes may not be caused exclusively by features of the immediate environment, but also by the non-immediate ones, for example, two syllables apart (Asmah Haji Omar, 1978). Various languages in the Kadazan/Dusun subfamily show that the vowel of the prefix or the suffix simulates its quality from the vowel of the root



form. This results in prefixes having variations according to vowel quality. However, the vowels concerned (the simulator and the simulated) may not occur in close proximity with one another.

#### *Downstream field research*

This type of research concerns itself with aspects of language which are already "known" and of their functions in society.

#### *Focus*

The focus of downstream research is on languages/dialects which have already had descriptions on their systems and structures as well as reference and prescriptive grammars and dictionaries. This type of research also looks at varieties of such languages as well as "derivatives" of the mixing of different languages which usually take place in cosmopolitan areas, i.e. the pidgins and the creoles. The focus also centres on the choice of speech systems in society and in various situations (such as the work place, the market area, the academic setting, places of worship, the law court etc.), the status given to languages in multilingual societies and the changes and developments that affect these languages.

In short, one can say that downstream research encompasses such research that may be categorised as sociolinguistics (macro- as well as micro-sociolinguistics) and psycholinguistics. With these come language attitude, language maintenance and shift, domain study, code-mixing and code-switching, language standardisation, pidginisation and creolisation, conversation analysis, etc.

#### *Purpose*

Each of the fields of research cited above has its own specific purpose. However, the overall purpose of downstream research is to give a picture of the actual situation pertaining to language use in society from various aspects, at the macro-as well as the micro-level.

#### *Objective*

The micro-level downstream research looks at the changes and development in the systems and structures of the languages and varieties

under consideration. It also looks at the use of linguistic items for effectiveness of communication, and in dealing with this topic, downstream research places a great deal of importance on paralinguistic features and their role in communication.

In language planning, which can be considered as part of macro-sociolinguistics, the research objective may consist of an examination of factors which may or may not sustain the status allocated to languages in the community, e.g. that of national and official language and the language used in education.

### *Method of Collecting Data*

In downstream field research, there are various methods of collecting data, and these are:-

- (i) the downstream informant method,
- (ii) the language-in-action method,
- (iii) the survey method,
- (iv) the test method.

### *Downstream informant method*

This method is reminiscent of the original or upstream informant method already discussed, in the sense that language users are made use of for the purpose of getting linguistic information by the researcher. However, there is a slight difference between the *upstream informant method* and the *downstream informant method*. The former provides a part to the discovery of language systems and structures, and the informant is always a native speaker. On the other hand, the latter provides data which is not concerned so much with the "discovery" of systems and structures but rather with choice and change which is linked with situations and attitudes. In the downstream informant method, the informant may consist of a native speaker or otherwise depending on the purpose and objective of research that is being done.

An example of the use of the downstream informant method may be seen in the research by Halimah Haji Ahmad (1994). She used informants (native and non-native Malay speakers) in order to gather information on their attitude towards loanwords used in books on management written in Malay

### *Language-in-action method*

This method gathers information/data right at the time and place of the language in action, i.e. when the speakers are actively interacting with each other. The researcher may choose to be part of the action or stay in the background. Modern technology, in the form of audio and video facilities, has helped a great deal in bringing this type of research to the fore. Examples of this type of research are the works of Nor Hashimah Jalaluddin (1994), Elaine Morais (1994), and Jamaliah Mohd. Ali (1995).

With the advancement of computer technology and the widespread use of the internet, one cannot dismiss from this type of study the type of language-in-action that goes in the e-mail. As shown by Mani Le Vasan (1996), discourse via the e-mail is closer to spoken discourse than to the written. If telephone conversations can be considered data for conversation analysis, then the e-mail discourse can be placed under this category.

### *Survey method*

The survey method in field research is applied in the collection of sociolinguistic data relating to the population size of speakers, speaker attitude, language choice, maintenance and shift, etc. Although all methods may have a quantitative and qualitative approach, the survey method usually places a major emphasis on the former.

In this method, the researcher not only chooses the sources of information but prepares a questionnaire to facilitate data gathering. The people who form the source of information are known as *respondents*. Examples of this type of research are found in the work of Maya David (1996) and Nor Hisham Osman (1991).

### *Test method*

This method is used in applied linguistics in the gathering of information on language usage as well as on the ability of language learners to use linguistic forms. In this method, the language users, who are known as *subjects*, are required to answer tests in various forms, e.g. filling in the blanks, substituting one item for another, constructing sentences, writing compositions, etc. Examples are found in the works of Ng Keat Siew (1992) and Kalpana Ponniah (1993).

### *Location*

Downstream field research can be done anywhere: in the work place, in an office situation, in the classroom, in places of worship, etc.

### *Uses of downstream field research*

The uses are variegated based on the purpose and objective of the research itself, but most of them are directed towards the relationship between the human linguistic and extralinguistic behaviour. It should also be noted here that downstream field research is not concerned with establishing the infrastructure of the language under study. As the focus is language and its speakers (and society), the results of the research can be more or less utilised by professionals, e.g. in management, teaching, etc.

The following are some of the examples of uses which can be cited. Elaine Morais' research, for example, is useful in getting an understanding of the interactional behaviour between people who work in a car assembly plant, on both the horizontal and the vertical axes. Such data is useful for management training.

Nor Hashimah's research provides useful information on human behaviour when it comes to bargaining. The data and the analysis as found in this work is also useful in training people for interviews as well as for constructing conversational exercises for language teaching.

Jamaliah Mohd. Ali's work gives a picture of strategies people use in discussions and debates. Her data and analysis are also useful for the understanding of human behaviour in a conflict situation.

Norah Mohamad's work (1991) on communication among the Malays and the Iban in Betong, Sarawak provides information on language preferences as well as on accommodation or otherwise that feature in the relationship between members of the in-group and the outgroup.

### *Conclusion*

What I have attempted to do in the previous sections is to typify the various types of research which require the researcher to get first-hand information from the source(s). In this way, I hope I have been able to give an idea of what field work in linguistics is all about. This type of research is different from the library type where the source is written materials found in books, journals, etc.

## BIBLIOGRAPHY

- Asmah Haji Omar 1969. *The Iban Language of Sarawak: A Grammatial Description*, Ph.D. thesis submitted to the University of London.
- Asmah Haji Omar 1973. *Linguistik Deskriptif dan Tradisi Lisan*. Dewan Bahasa, 17(18), 1973, pp. 357-366.
- Asmah Haji Omar 1978. Verbal Affixes in Standard Kadazan, *Sarawak Museum Journal*, Vol. XXIV, No. 47 (New Series), July-December 1978, pp. 161-194.
- Asmah Haji Omar 1981. *The Iban Language of Sarawak: A Grammatical Description*, Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Asmah Haji Omar 1983. *The Malay Peoples of Malaysia and Their Languages*, Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Asmah Haji Omar. 1985. *Susur Galur Bahasa Melayu*, Edisi Kedua, 1993. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Asmah Haji Omar. 1990. *Aspek Bahasa dan Kajiannya*, Kuala Lumpur: Dewan Bahasa dan Pustaka; Edisi Kedua 1992.
- Asmah Haji Omar. 1992. An Overview of Linguistic Research on Sarawak, in Peter W Martin (ed.), *Shifting Patterns of Language Use in Borneo: Papers from The Second Biennial International Conference*, Kota Kinabalu, Sabah, Malaysia, July 1992, Borneo Research Council Proceedings Series, Volume Three, pp. 69-80.
- Asmah Haji Omar. 1993. Linguistic Field Research in Malaysia: The State of the Art, in *Proceedings of the XIVth Scandinavian and VIIIth Conference of Nordic and General Linguistics*, August 16-21, 1993, Volume 1, pp. 1-16.
- Cense, A.A. and E.M. Uhlenbeck. 1958. *Critical Survey of Studies on the Languages of Borneo*, s-Gravenhage - Martinus Nijhoff.
- David, Maya. 1996. Language Shift Among the Sindhis of Malaysia, Ph.D. thesis submitted to the Faculty of Languages and Linguistics, University of Malaya, Kuala Lumpur
- Halimah Haji Ahmad. 1994. Penggunaan Istilah Pengurusan Organisasi: Satu Kajian Sikap, Master of Modern Language Studies dissertation, Language Centre, University of Malaya, Kuala Lumpur
- Jamaliah Mohd. Ali. 1995. Malaysian Student Seminar: A Study of Pragmatic Features in Verbal Interaction, Ph.D. thesis, Faculty of Languages and Linguistics, University of Malaya, Kuala Lumpur
- Kalpna Sinha Ponniah. 1993. Improving Academic Reading Ability of EFL Students at Tertiary Level, Ph.D thesis, Language Centre, University of Malaya, Kuala Lumpur
- Le Vasan, Mani. 1996. System and Process in Computer Mediated Discourse in a Malaysian Corporation, Ph.D thesis submitted to the Faculty of Languages and Linguistics, University of Malaya, Kuala Lumpur
- Morais, Elaine. 1994. Malaysian Business Talk: A Study of Patterns of Conflict and Non-conflict in Verbal Interactions, Ph.D. thesis, Language Centre, University of Malaya, Kuala Lumpur
- Ng Keat Siew 1992. Learning Strategies: A New Dimension in the Preparation of Instructional Materials for ESL Learning, Ph.D. thesis, Language Centre, University of Malaya, Kuala Lumpur

- Noor Hisham bin Osman. 1991. Kajian Pemilihan Bahasa: Kajian Kes di Kawasan Perumahan Ulu Dedap, Seberang Perak, Master of Arts Dissertation, Malay Studies Department, University of Malaya, Kuala Lumpur
- Nor Hashimah Hj. Jalaluddin. 1994. Bahasa Jual Beli Dalam Perniagaan Runcit: Satu Analisis Semantik dan Pragmatik, Ph.D. thesis, Faculty of Languages and Linguistics, University of Malaya, Kuala Lumpur.
- Noriah Mohamed. 1991. Penggunaan Bahasa Dalam Komuniti Melayu-Iban di Betong, Sarawak: Satu Kajian Etnososiologi, Ph.D. thesis, Language Centre, University of Malaya, Kuala Lumpur.
- Samarin, William J. 1967. *Field Linguistics: A Guide to Linguistic Field Work*, New York: Holt, Rinehart and Winston.
- Wong Tat Koon, Eric. 1994. A Systemic Phonological Study of Penampang Kadazan, Ph.D. thesis, Faculty of Languages and Linguistics, University of Malaya, Kuala Lumpur.