ISSN 0970-0277

OSMANIA PAPERS IN LINGUISTICS

Volume 24 1998

Editor
K. RAMESH KUMAR

DEPARTMENT OF LINGUISTICS
OSMANIA UNIVERSITY
HYDERABAD 500 007
INDIA

OSMANIA PAPERS IN LINGUISTICS

Editor
K. Ramesh Kumar

Osmania Papers in Linguistics (*OPiL*) is an annual refereed research journal devoted to the publication of articles in the field of general and applied linguistics. It also includes book reviews and review articles.

The journal publishes work primarily by the staff, students, and visiting faculty of the Department of Linguistics, Osmania University. However, articles may occassionally be invited from scholars outside Osmania on special topics. All Contributions must be in English. Views expressed in *OPiL* are only those of the authors.

This publication is supplied to institutions, societies, and Departments of Linguistics in exchange for similar publications. Others may send their subscription orders and matters relating to payments, exchange, and change of address to the Head, Department of Linguistics, Osmania University, Hyderabad, 500 007 India. The annual subscription rate for *OPiL* (one volume) is Rs. 25 or US $ 6 (excl. postage).

Articles for publication, review copies, and communications relating to editorial matters should be sent to the Editor, *OPiL*, Department of Linguistics, Osmania University, Hyderabad, 500 007, India.

ISSN 0970-0277

ISSN 0970-0277

OSMANIA PAPERS IN LINGUISTICS

Volume 24 1998

Editor
K. RAMESH KUMAR

DEPARTMENT OF LINGUISTICS
OSMANIA UNIVERSITY
HYDERABAD 500 007
INDIA

OSMANIA PAPERS IN LINGUISTICS

Volume 24, 1998

CONTENTS

Printed at: Apollo Graphics, Tarnaka,Secunderabad.
Ph : 7015195.

OPiL 24: 1-11(1998)

DIRECTIONS FOR RESEARCH IN LINGUISTICS*

H.S. Ananthanarayana
Osmania University, Hyderabad

ABSTRACT

Tracing the background for the establishment of Departments of Linguistics in Indian Universities and noting the areas of interest in teaching and research, in the beginning, the author feels concerned as to how the educational planners and administrators be made to realise the usefulness of linguists and their discipline. Towards achieving this goal, a few suggestions for research are given, viz. 1. to take up exhaustive descriptive studies of various Indian languages to bring relevant information for a better linguistic theory ; 2. to explore the areas of semantics and pragmatics of modern Indian languages by undertaking linguistic analysis of literary texts; and 3. to bring in their linguistic expertise for improving literacy programme to achieve functional democracy.

Linguistics had enjoyed the attention of Indians from the earliest period and its methodology had reached the finest point in the hands of Paaṇini and Patañjali. The ancient Indians were mainly concerned with the analysis and description of their own language; Comparative and Applied linguistics were probably unknown to them. There came upon a long period of slumber after this glorious achievement and only in this century, the Indian scholars became again active in linguistic studies.

The advent of Modern linguistics in India may be said to have been made in 1954 with the organization of an Autumn School of Linguistics on the campus of Deccan College. Grants

* An earlier version of this paper was presented at a National Symposium on Linguistics and Social Relevance, organised by Centre of Advanced Study in Linguistics, O.U., Hyderabad, on March 5, 1992.

from the Rockefeller Foundation in America and later from the University Grants Commission enabled the holding of several Summer Schools and Autumn Seminars on different university campuses which attracted language students from all parts of this country. A few scholars received intensive training at various American and European Universities.

Following the recommendation made at a Conference of Vice-Chancellors and Linguists held on the 7th and 8th of January 1958, a few universities took steps to start full-fledged departments of linguistics of which the Departments at Annamalai University, Osmania University, and Kerala University are the oldest. As recommended in the same Conference, these departments at various centres provided for instruction and training in all aspects of linguistics, viz. General, Historical-Comparative, Descriptive and Applied in particular reference to the two major language families of India, Dravidian and Indo-Aryan. Teaching and research went hand in hand and one supported the other.

Research in the early years was mainly in Descriptive linguistics. Recording hitherto unwritten languages and the analysis of the grammatical nuances occupied the interest of the newly initiated Dialect surveys, Etymological and Comparative dictionaries of individual languages were undertaken at a few centres. With the establishment of the Central Institutes of English, Hindi, and the other Indian languages, Applied linguistics, especially, methods of teaching of these languages were developed. The format was mainly structural in the beginning and later, Transformational-Generative. As inter-disciplinary fields like Sociolinguistics, Psycholinguistics came on the scene, they too found many practitioners and research started flowing in these areas, involving Indian languages.

It is increasingly felt that our discipline is not enjoying the popularity it deserves among the students and that its importance is not fully realised by the administrators. The popularity of the subject in the early days was due partly to its newness and partly to the opportunity it provided for going abroad for higher studies. I do not think that the newness has in any way diminished. It is true, however, that opportunities for going abroad have become less since there is no need for it as there are well-equipped departments now in many parts of the country where higher studies may be pursued. But, how to make its usefulness felt by educational planners and administrators ? It is only by properly educating them and convincing them that we can achieve this in due course. The situation does not seem to have changed much since 1958 since a resolution drawn at that time had given vent to the same feeling : 'It is the considered opinion of this conference that linguistic studies, which at present do not occupy a place which their importance and usefulness require, commensurate with that in other countries, be given greater attention and a more central position in university education in India'. We alone are to be blamed for the continuance of this state of affair and every effort should be made to bring about the desired change.

Now, what should be the direction for research in Linguistics in the present Indian context ? I consider myself too small to suggest any particular direction ; and linguists are too individualistic to take kindly any suggestion from others. They continue to do research in the areas of their special interest. However, I may put forward a few ideas for consideration by fellow linguists. It is upto them to accept these ideas or reject them.

1. Some of the notable achievements in modern linguistics is attributed to the work of the transformational grammarians under the brilliant direction of Noam Chomsky.

Such achievement is often accomplished in a discipline by narrowing of aims. The narrowing of aims by Transformational Grammar was clearly stated by Chomsky in his famous book, 'Aspects of the theory of Syntax'. Chomsky stated that 'Linguistic theory is concerned primarily with an ideal speaker listener, in a comparatively homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors in applying his knowledge of the language in actual performance'. Yet every linguist knows that such a situation is a fiction. There are no completely homogeneous speech-communities. Every language and every community has a history. This aspect needs hardly to be emphasised in the Indian context. Linguistic theory must accordingly be concerned not with speech communities, but with speech communities using language which include characteristics often at variance with each other because of their history.

Prof.Lehmann of the University of Texas at Austin has argued in one of his papers that Chomsky's above position must therefore be abandoned. He has clearly demonstrated with illustrations from German that it leads to erroneous views and to wasteful activities. He concluded that language cannot be understood or adequately observed, described or explained unless it is regarded as a human phenomenon subject to the effects of time, that is, as an activity with a history. Moreover, linguistic theory cannot limit its concerns to a single language, no more than to a single speaker-listener. It may be true that all humans share a set of abstract rules which forms the basis of their language and of every other language as well. And linguistics may ascertain these rules and the elements which are directed by them, such as grammatical and lexical features. It must be pointed out, however, that the determination of such phenomena is warped when linguists limit their attention to a single language (e.g. English in Transformational

grammar). Moreover, an accurate understanding of language is thwarted if investigators limit their attention to a restricted sample of languages (e.g. Indo-European family by Neo-grammarians). Further, it is shown that linguistic theory has also been hampered by primary attention to one type of language (e.g. Indo-European).

While discussing the structure of German and the category referred to as 'middle', Prof. Lehmann has demonstrated again how the expression of underlying patterns in the several language types must be known to achieve an adequate linguistic theory and understanding of language. However, he concedes that we know too little about too few languages to make definitive statements about linguistic theory. As an instance, he cites Indo-European syntax which, according to him, is still poorly understood, because the historical syntax of the dialects are not well investigated.

It is here that Indian scholars can contribute to a better understanding of linguistic theory. The investigation of the historical syntax of Indo-Aryan languages has scarcely begun; historical syntax of the Iranian languages is virtually unknown. I would like to suggest therefore to some of our younger scholars to direct their attention towards filling this gap by their research in the near future. I would like to stress again that our work must be based on the study of language rather than on psychological constructs that supposedly correspond to the representation of language in the mind. Because, it is well known that science operates by dealing with the observable, by proceeding from the known to the unknown rather than vice versa. Further, it is not enough on our part to base generalizations on knowledge of one language ; it is dangerous. For better perspective, linguists master the structure of different types of language. They have a wonderful opportunity in India as there are many languages, some of

them with long history, and belonging to four different linguistic families. There are languages with much sociological variation as well as geographical differences. Our linguists should bring into the building of theory information from as many of these languages as possible. They should invariably equip themselves with the knowledge of the history of the languages they treat since a language in use is constantly modified. For proper perspective, they need to know the background of their discipline and of its theories. I am afraid that some of our younger scholars appear to lack the knowledge of the history of our science since such courses are no more given in the curriculum.

2. Another field in which Indians are better equipped to contribute, because of their heritage, is semiotics or the study of communication. For the sake of our understanding, it may be briefly described as comprising three elements, viz. (i) the user, (ii) the universe outside language, (iii) the means for communication. For understanding any system of signs, three sets of relationships must be known: A) relationships between (i) and (iii); B) relationships between (ii) and (iii); (c) relationships between entities of (iii). Study of these is carried out under three sub-divisions of semiotics: (a) pragmatics; (b) semantics; (c) syntactics, roughly, what is generally called grammar.

Linguists in the modern tradition have concentrated their attention on (c) grammar, that is, phonology, morphology and syntax, leaving outside their sphere concern with meaning. This concentration has been found to provide inadequate understanding of language. Attention to semantics and pragmatics leads naturally to dealing with texts rather than sentences as units.

When we examine the linguistic tradition in India, we note that both grammarians and philosophers were here concerned with language. While the former evinced interest in the analysis and description of the formal aspect of language, the latter dealt with the semantic aspect of it. The earliest extant work which was specifically devoted to a discussion of meaning is the *Vaakyapadiiya* of Bhartṛhari (450-510 A.D.) who combines in himself both the grammarian and the philosopher. Even a cursory examination of the contents of this work reveals that Bhartṛhari was a great exponent of semantics. It should be a matter of pride for us to realise that our ancestors had more than 1500 years ago reached such greater heights in the areas in which the West has only begun to explore in recent decades. Unfortunately, most of our modern linguists do not have access to this tradition since it is available in Sanskrit. But, as inheritors of the great tradition, I would like to recommend to our modern linguists to undertake a close study of the early works and equip themselves with this knowledge so that they can explore in the area of semantics and pragmatics of modern Indian languages and give a lead to others in this field.

Text linguistics has brought together the study of language and literature. The division between linguistics and literary study, formally ignored in academic arrangements such as language departments, is recent. The founders of modern linguistics (e.g. the Grimm brothers, the Schlegel brothers, Franz Bopp) dealt with both the study of literature and the study of language. Bopp's work on the 'system of conjugation of Sanskrit' devoted nearly as much space to concern with early Sanskrit literature as to the Sanskrit language. Yet, this was the work that sent linguistics on its course throughout the 19th and much of this century. Linguists concentrated increasingly on the forms of language. Recently the

concentration has been modified. The impulse towards modification did not come from text linguistics but its effects are clearly evident in the approach to the study of language.

Among structural linguists, Prof. Archibald Hill was one of the pioneers who provided models for literary investigations. His essays include attention to 'meaning'. Generative Transformational grammar, rooted in cognitive psychology, seemed to promise closer cooperation with literary scholarship. Language was viewed here from a mentalist rather than mechanist standpoint. Generative grammar allowed poetic language to be distinct in usage. Maintaining concentration on the sentence as the 'primary unit of understanding', they proposed further transformations in literary language. Text linguistics embraces all these approaches. It has contributed to the study of language by extending its scope to cover all elements of language in their relations to their references and their users. It contrasts in this way with other approaches to language which seek to illuminate language by concentrating on selected aspects.

In Indian tradition, the study of language was always combined with the study of literature. There arose in India the subtle art of literary criticism which was carefully nurtured into a respectable science by scholars like Aanandavardhana and Abhinavagupta. We may note that linguists of the first generation were all students of literature and some of them made equal contribution to linguistics as well as to literature. It is the present day students of modern linguistics who have restricted themselves to the knowledge of language without realising that they will be depriving themselves the use of language in society. Language can not be treated in isolation. All linguistic data must be treated with knowledge of environment that produced it. Since language in use is

constantly modified, proper understanding of linguistic problems requires knowledge of the history of the language which is revealed only in the texts. Indian linguists, I think, should be interested in seeking some knowledge of literature and attempt literary analysis with principles of Linguistics. By doing so, they can better justify their Indian scholarship.

3. Linguists should realise that they have an obligation to society. They cannot continue as ivory tower academicians involving themselves in describing and theorising about the languages of the world. They should look for areas in which their knowledge could be put to practical use. They have already played an important role in the area of language teaching. The results of their analysis have been usefully utilised by the language teachers in drawing up better textbooks and for grading the material. There are other areas too where a linguist can richly contribute and thereby impress on administrators and planners the importance of their subject to the society. I may mention one such area of social relevance and suggest that Indian research may be directed towards that.

Though India is a democratic country, its entire population is not fully participating in the decision making since a large portion still remains illiterate who lack the ability to read and write. Thereby, they will be lacking in social awareness. The Government is spending a huge amount to spread literacy and it is the responsibility of every educated person and more so of every linguist to lend a hand in making this important venture a success. There are several things in the programme which may be decided only by a linguist. There are, for instance, language specific conventions of writing and reading. These are better analysed and interpreted by a linguist in the literacy curriculum. It is often said that literacy is most effectively achieved in one's own mother tongue. It may

therefore be appropriate to employ the dialect of the prospective literate in the primers being prepared. But, if his literacy should prove fully functional, he should know the equivalents in the standard language also. In continuing literacy programme then there should be at some point a switchover from the dialect to the standard variety. These distinctions are understood by the linguist and he should therefore get himself included in the programme. The dialect may be lacking in a script which then needs to be devised. This becomes possible only when a complete analysis of the phonemic system of the language is available.

The literacy programme should not end in teaching the person to write only his name, as it is made to sound sometimes. It must be a continuing programme and the learner's interest must be sustained for long. He must become competent to understand fairly complex structures of the language so that he can easily read newspapers and such other material in his language. The linguistic structures need to be presented in graded manner which could be done with the assistance of a linguist.

Further, it may be noted that in a large country like ours it would be expensive to have separate projects for each language area. It may prove useful to employ transfer strategies to get primers prepared in one Indian language and get them translated into other Indian languages. I believe that our linguists have an important role to play in this socially relevant field and they should involve themselves in eradicating illiteracy. Accordingly, their research should confront the above questions.

I believe that I have suggested for our younger scholars a few directions for their research in our discipline. I hope that they will rise up to our expectations and will strive hard to get for their subject the importance it rightly deserves.

REFERENCES

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax. Cambridge. Mass*: MIT Press.

Hill, Archibald A. 1976. *Constituent and Pattern in Poetry.* Austin: University of Texas Press.

Lehmann, Winfred P. 1980. Language as a human phenomenon. The importance of history for the understanding of languages. *Folia Linguistica* I/1, 5-18.

OPiL 24: 13-47(1998)

AREAL LINGUISTICS AND THE LANGUAGES OF SOUTH ASIA*

Nikita V. Gurov
Leningrad State University, Russia

ABSTRACT

This paper deals with the concept *Areal Linguistics* and shows how it differs from others such as *Areology* and *Linguistic Geography*. The paper takes into account the structural features of South Asian Area to make it a *Model Object* for the *Areal studies*. It describes the explored and unexplored fields of the Areal Studies on the languages of Indian subcontinent and set some methods and approaches for further investigation. The author proposes the two/three layers structure of nominal paradigm and the categorial unification of nominal paradigm in South Asian Languages in general. The problem of inner structure of South Asian Area; its cultural and historical implications are discussed.

1. INTRODUCTION

In the present paper we have chosen to use the term *areal* as it was primarily proposed in 1956 by M.B. Emeneau in his pioneering article *India as a linguistic area*[1] and later adopted by the most specialists in Indian linguistics. M.B. Emeneau defines linguistic area as a region 'in which there are languages belonging to more than one family but showing traits in common which are found not to belong to the other members of (at least) one of the families'.[2]

The definition requires certain clarifications. In the first place, the term *family* is not to be treated in a strictly terminological way: it can equally denote – from the point of view of taxonomic hierarchy – a smaller unit of related languages, i.e. *branch*, *group* etc. (In the South Asia we find languages which belong to three groups of one and the same Indo-European family: Indo-Aryan, Iranian and Dardic. At the same time heterogeneity of the languages in contact- at

least, recognised as such from the point of view of the ethnical identity of their speakers– should be considered as *condition sine qua non*). Second, as it is evident from the studies by Emeneau and other linguists who dealt with this problem, common features in the neighbouring languages can (and must) reveal themselves at different levels: phonemic, morphological and syntactic. Linguistic contacts are, therefore, to be historically long-lasting, while the speakers are to be connected not only geographically but also by some clone cultural relations.

2. DEFINITION : LINGUISTIC AREA

One can easily see that the definition of a *linguistic area* adopted here is principally different from the one usually found in the works on linguistic geography and *areaology* (the latter being founded on the basis of *linguo-geographical* studies). where *area* can be defined by a certain single feature (or a number of features) of a certain single linguistic level. A study of the convergence of genetically unrelated and typologically remote languages requires a totally different methodological approach. Nevertheless, methods common to linguistic geography could (and perhaps should) be applied to the areal studies, particularly for delineation of certain *microareas* or for mapping of linguistic phenomena.

From the definition of *area* quoted above it is evident that *areal linguistics* should be treated, along with comparative linguistics and general typology, as one of the three main trends in modern linguistics focused on the studies of linguistic continuum. The main difference between areal linguistics and the latter two lies, in the first place, in the general choice of the objects for study along the lines of genetic origin and contact. While comparative linguistics studies related languages irrespective of their geographical distribution and contacts, areal linguistics, on the contrary, treats neighbouring

languages regardless of their genetic affinity. For linguistic typology which deals with the problem of structural isomorphism, neither is essential. The three trends differ also in the prospective approach to the processes under study. If comparative linguistics concentrates on the divergencies of languages emerged from the common source, areal linguistics draws its conclusions from the convergence of languages derived from different sources. For the typological studies the problem of historical comparison of data does not seem to be relevant at all, though for historical typology, chronological arrangement of linguistic facts cannot possibly be excluded.

Areal linguistics is often identified with *Sprachbundtheorie* (theory of linguistic unions). Such an identification, can be accepted only with certain restrictions. First, the term *Sprachbund* is obviously polysemantic. It refers both to the so-called *one-level* unions like the *Eurasian Union* proposed in 1931 by R. Jakobson[3] (such unions are in fact nothing but *typological* or *structural* zones and consequently, are to be treated within the frame work of general , not areal typology) and to the genuinely areal entities e.g. Balkan, Caucasian, Central-Asian and other *unions*. Besides, even the general concept of *Sprachbund* as it was stated in 1928 by N.S. Trubetskoy,[4] ontologically seems to be somewhat elusive. While admitting that even the languages in contact do not always form a *language union* universally, the advocates of Trubetskoy's theory fail to answer the most important question: what degree of structural proximity enables us to ascertain that the union really exists.[5] Recording a certain state of linguistic contact and yet not strictly defining the main characteristics inherent in it the *Sprachbund* theory turns out to be, therefore, static in character, devoid of chronological perspective and fails to reveal the dynamics of linguistic interrelations.

The inconsistencies of this theory might be, to a certain extent, eliminated, if we admit that an *area* (a linguistic area in particular) is not a purely linguistic, but essentially an ethno-cultural concept as well, characterised by a definite location in time and space. Thus, the aim of areal linguistics is a typological study of any forms and stages of contacts between the languages encountered in the area since the time of its formation.

So we can see that areal linguistics appears to be directly related to the theory of linguistic contacts which was treated by U. Weinreich, E.Haugen, V.Ya. Rosenzweig and other scholars.[6] At first sight both theories appear to deal with the same set of problems: contacts between two or more languages, leading to the interference of linguistic systems at different levels– phonological, morphological, syntactical and lexical. For the areal studies it is essential that, according to the principles of the *theory of contacts*, we should take into account not only the *direct* interference but the *indirect* one as well, the latter meaning that the influence of one language stimulates development of certain processes and mechanisms potentially existing in the other.[7]

At the same time, the Areal Linguistics (AL) differs from the Theory of Contacts (TC) in some general characteristics relating both to the ontological definition of object of study and methods of its description: 1) For AL genetic heterogeneity (or remoteness) of contacting languages is substantial; for TC it is irrelevant; 2) TC predominantly deals with the contacts leading to the state of bilingualism or diglossia; AL is mainly concerned not with the actually existing forms of bilingualism but rather with the consequences of a bilingualism in the past, resulting in some kind of direct or indirect interference; It means that; 3) TC studies languages in a purely synchronic aspect while for AL the data of both synchronic and diachronic analysis are equally important; 4) Unlike TC where as a rule

one language is always regarded as the donor and the other as a recipient, in the framework of AL we most frequently deal with the mutual multilateral interrelations of languages (including the relations of *substratum*, *superstratum* and *adstratum* types); 5) While analysing various types of linguistic contacts unlike TC, AL pays more attention to the cultural background of a given language than to its modern social status.

All these differences by no means deny close connection between AL and TC, Thus, many analytical procedures and principles worked out by TC (e.g. U. Weinreich's representation of interference patterns of phonological systems with the rules of *super* and *sub-differentiation* of phonemes)[8] must certainly be used in the areal studies, - depending on their ability to elucidate mechanisms by which an area originates and functions.

In areal studies the specific conditions of the formation and development of a linguistic area may require the use of different methods adopted in different branches of linguistics including comparative linguistics, general typology along with a broader approach suggested by the TC, sociolinguistics and linguistic geography. Diachronically the history of a given linguistic area can be represented by a gradual change of *areal situations*: AS_1-AS_2-AS_3- ...–AS_n. The change of the situation means some alteration in the constituents of the area and/or in the areal structure. One can observe, for instance, disappearance of certain constituents (languages, dialects, groups of dialects or languages) which remain preserved only in the form of a substratum; splitting of a constituent to form separate linguistic entities (language, dialects); appearance of new constituents; change in the direction of the *interference vector* reflecting the new functional status of some language or languages etc. The pattern of the *areal structure* can be changed by a shift of its geographical *centre* (i.e. zone of

particularly intensive multilateral contacts), by weakening or, on the contrary, strengthening of contacts which may lead sometimes to a split of the area into a number of sub-areal zones or even separate areas etc.

It should be noted that areal situations, when they change, do not disappear completely; on the contrary, earlier situations are preserved in the later periods as groups of iso-glosses distributed along the area or connecting separate languages.

3. SOUTH ASIA AS A LINGUISTIC AREA

The South Asian (Indian) subcontinent providing invaluable data for comparative linguistics is now being intensively studied on typological grounds.[9] For areal linguistics this broad region is no less valuable than the classical *linguistic areas* like Balkan, Caucasian etc. Moreover, there are sufficient grounds to believe that, given a proper stimulus, the areal approach to South Asia can in fact serve as a reliable model for similar studies in other regions. This is not impossible considering the fact that literary tradition which has existed on the Indian soil since early times and which is evidenced by a vast body of texts in many languages enables us to follow the history of Indo-Aryan languages beginning from 13-12 cent. B.C. (the date of earliest Vedic texts), Dravidian languages - from 2-1 cent. B.C. (earliest Old Tamil texts and inscriptions) and Dardic (Kashmiri) - from c. 14-15 Cent. A.D.

We can rely here on considerable number (exceeding 200) of languages belonging to four different families – Indo-European (Indo-Aryan, Dardic, Iranian), Dravidian, Austro-Asiatic (Munda, Khasi) and Sino-Tibetan, to say nothing about several *isolated* languages (Burushaski, Nahali and Kusunda spoken in Central Nepal). In this respect the situation in South Asia differs favourably from that in the

Balkan region, where at present one can find mostly Indo-European languages. Equally important is the fact that certain languages within our area are genetically related to languages beyond their boundaries. Any attempt to establish possible extra-areal connections for the languages of uncertain genetic affinity would be rather fruitful for the study of linguistic history of the area. Special attention should be paid to the genetical identification of *isolated* languages which evidently are the relics of the earliest 'language strata' here.[10]

Up to now the attention of the specialists dealing with the areal problems was drawn mainly to the languages covering the Indian mainland, – Indo-Aryan, Dravidian and Munda. The Sino-Tibetan languages of the region remain practically a *virgin soil* in this sense. The lack of information on Sino-Tibetan group (especially on the so-called *Pronominalized Himalayan* languages) might of course be viewed as a serious drawback, though convergent influence of Sino-Tibetan on the other languages of the area is apparent only in its periphery. Practically no notice was paid to the convergent tendencies in Iranian languages spoken in South Asia (Balochi, Pashto), yet areal approach here seems to be rather promising.

4. SOUTH ASIAN AREAL FEATURES

In the initial stages of areal pursuit the aim of the scholars was to prove the very existence of South Asian area and to delineate its boundaries. The task, thus formulated, (which seems to have been generally accomplished at present[11]), was in many ways decisive in the choice of linguistic material and in the mode of its presentation. So, in spite of the fact that in the previous studies the linguists in their efforts to find cross-points of areal convergences at different levels of languages under study, turned to their phonological, morphological, syntactical and lexical data, they have failed to perform a thorough description at any one of the levels and therefore, could not draw a convincing picture of the nature of linguistic interrelations.

In phonology attention has been directed rather to the comparison of phonemic inventories (laying stress first of all upon the existence of retroflex consonants) and only then turning to the general principles of paradigmatic and syntagmatic structures of the phonological systems.[12] Although a great deal of information has been provided on the lexical borrowings in the languages of the area, no attempt has been made to establish a system of phonological correlations between *donor* and *recipient* languages (e.g. between Indo-Aryan languages at various stages of their history and Dravidian). Such a system could be constructed on basis of E.Haugen's *diaphony* or on the rules of phonetic interference suggested by U. Weinreich.[13]

In the field of morphology the linguists have been principally interested either in the material correspondences between the formants in heterogenous languages – correspondences, as a rule, casual, fragmentary and, therefore, devoid of any diagnostic significance (like the coincidence of Dravidian format of the Dative *-ku/-ki* and the Indo-Aryan postpositions e.g. H. *ko,* Or. *ku*), or in the separate structural isoglosses taken out of the grammatical context, or, finally, in the most general features, significant at the initial stage of the areal investigation but deserving to be rearranged systematically in further study. (absence of prefixation; the same set of case-markers for different numbers; morphological causative, etc.).

The same can be said about syntactic features taken by some scholars as areal characteristics.

Having accepted the definition of a *linguistic area*, one should admit that the general surveys of areal events do not usually go beyond a brief consideration of the collected data. In their attempts to establish areal specificity of Indian languages with the utmost degree of precision, C. Masica, H. Vermeer and other scholars take into account only those

phenomena which are represented virtually in every language on the Indian subcontinent and have no counterparts in other languages of the world. Consequently, from an earlier list which comprised 27 *areal features*, C.P. Masica has chosen an unquestionable only four isoglosses (retroflex consonants echo-words, emphatic and other enclitical particles, Dative subject construction), which he thought reasonable to acknowledge as characterising the given area as a whole.[14] Naturally, this *strict* selection cannot be viewed as methodologically indisputable. Even if one admits that certain linguistic phenomena can be considered unique at all, which seems to be doubtful the very fact that a separate isogloss may reach beyond the areal boundaries gives no reason to exclude it from the list of areal features. The approach, outlined in the previous studies, according to which only continuous isoglosses were taken into consideration, tends to be too simplistic.

Our working definition of an *area* as basically an ethno-cultural concept requires the complex and systematic approach to any form of linguistic convergence within the boundaries of a given geographical territory. Most important, perhaps, at the present stage of investigation, is to proceed from empirical observations to the problems of interference of systems as a whole, from the registration of correspondences between separate elements and structures to the linguistic (typological) interpretation of these phenomena.

Thus, in phonetics and phonology particularly at issue are analysis of systemic relations of various classes of phonemes, the combinatory factors, types of suprasegmental phonemes (stress, tones, intonation contours) and syllabic structure. (For such studies vast experimental and field material is required, which is, unfortunately, very scarce.)

In morphology, attention should be focused not on the direct material correspondences, but on the analogies of the

more general type – both structural and categorial. As one of the primary aims of such an analysis we should indicate the basic principles of the construction of grammatical forms and their organisation into the system of the grammatically relevant oppositions. Thus, in establishing areal characteristics of the nominal declension the most important issue is not occurrence of postpositions (in the Indo-Aryan, Munda, Dravidian and Burushaski) as such, but hierarchical arrangement of postpositional formations (together with primary case-affixes) according to three structural layers, leading to the construction of complex, analytical, or secondary (tertiary) synthetic forms.

5. LAYERED STRUCTURE APPROACH OF NOMINAL PARADIGM

Two (or three) layers' structure of nominal paradigm in the New Indo-Aryan languages was discussed at length by Zograph.[15] Analogous structure is exhibited by the other South Asian languages (Dardic, Dravidian, certain Sino-Tibetan languages, e.g. Newari; Burushaski, Nahali). In certain languages, groups or families derived nominal forms may differ from New Indo-Aryan by the degree of grammaticalisation of the formants (affixes : syntactic words), by the pattern according to which affixes are linked (agglutination : fusion) or, finally, by the character of distribution of grammatical information among the formants (flexion : agglutination): still the general principle – distribution of nominal formatives between several structural layers – remains universal along the whole territory of the region (or greater part of it).

The structural pattern for the nominal forms suggested for the New Indo-Aryan languages, in its generalised variant, which takes into account data provided the languages of the area, looks as follows:

$B + (n) + Obl\ (=Gen)_{(1)} + (P'/F_{(2)} + P''_{(3)}$, where B is base of stem; n- number marker; Obl - oblique case ending ;

G - 'genitive' ending; P'/F - primary postposition or inflectional case ending; P"- secondary postposition.

Cf. following examples: H. *saheeeliy* - (B) + *õõ* (Obl Pl)$_1$ + *kee* (P' = Gen)$_2$ + *saath* (P")$_3$ - 'with girl-friends'; S. *(ghaNane) uuiha* - (B) + -ne (Obl p l)$_1$ + *khaa*(P' = Abl)$_2$ + *pooe* (P")$_3$ - 'after (many) days'; M. *bhint* - (B) + *-ii* (Obl)$_1$ + *-laa* (Adit)$_2$ + *laagun* (P")$_3$ -'near the wall'; B. *śikha* -(B) + *-r* (Gen)$_2$ + *uparee* (P")$_3$ - 'on the (mountain) top' ; K. *cur'* - (B) + *ɑvˑ* (Obl II pl = Instr- Abl)$_1$ + *səət-* (P')$_2$ + - n (P" = Obl I)$_3$ - 'by means of theft'; *Ta. aaɽ*-(B) + *-ɽu* - (Obl = Gen)$_1$ + *-(k)k(u)* (Dat)$_2$ + *appaal* (P")$_3$ - 'across (on the other bank of) the river'; Te. *nuu-* (B) + *-ti-* (Ob l=Gen)$_1$ + *-ki* (Dat)$_2$ + *mundu* (P")$_3$ - 'in front of the well'; Pj. *mer* - (B) *-t-* (Ob l)$_1$ *-o* (Gen)$_2$ *kan* (P")$_3$ - 'on the tree'; Brah. *ɖaa* - (B) + *-ɽ-* (Obl)$_1$ + *-aan* (Abl)$_2$ *-baar* (P")$_3$ - 'like him'; New. *che* - (B) + *-yaa* (Gen)$_{1-2}$ + *pakhe* (P")$_3$ -'towards the house'; Bur. *gus-* (B) + *-mu* (Obl = Gen)$_{1-2}$ + *ka* (P")$_3$ -'with the woman'.

In the Austro-Asia languages of the area another type of nominal inflexion, is dominant, since an explicit form of the Oblique case is missing as a rule, while the Subject – and Object – markers are either included into the verb forms or joined to the noun forms as enclitics. Yet, in the Munda group in a number of languages (Kharia, Korku, Sora, certain dialects of Santali) there is, again, a tendency to produce certain 'case' forms by combination of respective postpositive elements which, contrary to existing rule, are not added to the nominal stem, but to the base with the adjectivising (Genitive) formant, e.g. Kharia: *lebu*-(B) + *-aa*-(Gen)$_1$ + *-te-* (Loc) -i (Abl)$_{2-3}$ - 'from a man'; Korku: *koon* - (B) + -a- (Gen)$_1$ + *-ten* / -tan (Abl)$_2$ -'from the son'; Sora: *dumbaa*- (B)+ *-naa* (Gen)$_1$ *maŋ* (P")$_{2-3}$ - 'before the Dom'; Santali (Kar-mali): *IN* -(B) + *-akˈ* - (Gen anim)$_1$ + *-maatraan*- (P')$_2$ + *-re* (Gen non-anim=Loc)$_3$ - 'before me'; Birhor: *(tiimiin) ɖiiɳ*- (B) + *-rii-(n)* (Loc)$_2$ *iicˈ* (Pos)$_3$ - 'of (how many) days (old)'.

Thus, we have every reason to include the hierarchical structure of the nominal formants into the list of distinctive features, marking languages of the South Asian area. Further studies are needed to clarify whether this phenomenon is due to the one-sided influence of substratum or adstratum type or whether it should be treated as an instance of reciprocal influence of contacting languages.

Formation of the layer-hierarchy in the structure of nominal paradigm is directly associated with the tendency for the unification of the number of its grammemes as well as of their significance which is manifested in various languages of the area. The most evident demonstration of such process gives the form of the Oblique in the New Indo-Aryan languages. The older generation of scholars (G.A. Grierson, J.Bloch, S.K.Chatterji) associated this phenomenon, which had been traced directly in Apabhraṁśa (late 1st - early 2nd mill. A.D.), to the influence of the non-Aryan (Dravidian) substratum.[16] An illustration of the opposite influence, i.e. that of the Indo-Aryan languages on the other languages of the subcontinent, is usually seen in the merger of grammemes of the Dative and Accusative cases in some Dravidian languages (Naiki of Chanda, Gondi of Adilabad, Konda, Pengo, Manda, Brahui and partly, Kurukh and Malto), the contamination of the Instrumental and Ablative in Pengo, Konda and Kui, elucidation of the grammeme of the Aditive case in the paradigm of certain Munda languages (Korku, Mundari)etc.

We are probably the last people to deny the influence of Dravidian languages on Indo-Aryan (from the Vedic times onwards) and vice versa. But the linguistic data collected in last three decades show that the process of interference between the languages of South Asia is not (and probably never was) so simple and straight – forward as one might have imagined at the beginning of the century. It was, on the contrary, long and complicated, involving many kinds of

linguistic interrelations, such as direct influence, convergence and contamination. In this particular case one must take into account the fact that functional interference of the Dative and Accusative can be sporadically observed not only in the aforementioned modern Dravidian languages, but also, in the Old Tamil - in the texts of the Sangam period (*'kuṉuntokai'*, *'puṟanāṉūṟu'*- 1st -2nd cent. A.D.) i.e. much before the proposed time of the development of this phenomenon in Indo-Aryan.[17]

The merger of the grammemes of the Instrumental, Ablative and Sociative, likewise, appears to be immanently inherent in the Dravidian nominal declension. This can be observed, apart from the languages of the Gondwana group mentioned above, in the South Dravidian languages like Kannada and Badaga, in the former from the earliest texts (5th -8th cent. A.D.).

So, the question is still open: when and in which languages did these phenomena originate? Can we attribute the unification of the nominal paradigm to the process of the accommodation of the Indo-Aryan languages on the Dravidian territories or should it be ascribed to the influence of some unknown 'third power'?

The answer to this question (as well as to many others) can be found through the complex study of the system of nominal inflexion in South Asian languages, in particular in one of its most characteristic traits - the system of postpositions. The study in question should comprise a thorough analysis of this group of formants (syntactic words) in different languages, including a classification of postpositions according to (1) their functions, (2) genetic links with definite morphological classes of words (nouns, adjectives, adverbs, pronouns, non-finite verb forms etc.) and (3) their correlations with certain *case* forms. There are sufficient grounds to believe that the isoglosses obtained by means of this approach may

help not only to clarify the general pattern of the areal processes, but may also give some criteria for structural division of the area into subareal and microareal zones. It is only at this stage of investigation that one will be able to estimate, for instance, the areal *validity* of the wellknown and fully evident fact – i.e. the semantic and functional proximity of a group of postpositive elements which were derived from the non-finite verb forms and which exist in the unrelated languages, never before (at least in historical times) linked by direct contacts, cf.: H. (se) *lee-kar*, B. *laiya*, Ta.,Ma.-*koɳɲu*, Kod.-*koɳḍi*, Te.-*tiis* (u)*kuni*-'with', 'together with', 'by' (the conjunctive participle of the verb 'to take'); B. *haitee*, Ta. -*niɲɲu*, -*iruntu*, Ma. *niɲɲu*, *iɲɲuntu*, Kod. - *iñji*, Te. -*uɳḍi* -'from' (the conjunctive participle of the verb 'to stay somewhere', 'to be'; most interesting are the combinations of postpositions: B. *nikaɭ* 'near' + *haitee* Te. -*loo-n* [<**u ɭ* -inner part'] *uɳḍi* 'out of'; cf. also H. *hoo- kar* 'through'); M.karuun, Old Te. *ceesi* 'by means of' (the conjunctive participle of the verb 'to do').

6. CATEGORIAL FEATURES OF NOUNS: AREA-WIDE ANALYSIS

As for the categorial features of the class of nouns, the area-wide analysis should be concentrated on the following phenomena: (1) reorganisation of the sub-classes of nouns according to the categories of gender, animateness and personality; (2) the opposition of definite vs. indefinite forms and its manifestation; (3) some characteristics of pronominal system, such as presence or absence of inclusive pronouns or sub-grouping of demonstrative pronouns on the line of proximity vs. remoteness; (4) the category of *subordination* expressing various degrees of politeness through the forms of the noun (pronoun) and the verb ; (5) the place of finite and non-finite verb forms in the conjugation system, etc.

It should be noted, however, that many such features have no area-wide distribution. Of particular interest to areal linguistics, though, are the features expressed in the non-related languages representing each of the main genetic groups, which is indicative of convergence tendencies. Certain features of categorial distinctions of various classes and subclasses of words are expressed not only morphologically, but syntactically as well. Thus, the opposition of personal and non-personal names could be expressed not only through the declensional paradigms, (in some Dravidian languages), but also in the form of a noun in the position of direct object (Indo-Aryan languages).[18]

On both grammatical levels are manifested, in particular, the various forms of possessive relations. This phenomenon serving as one of the main characteristics of a linguistic type in general is presented in the Indian languages by a variety of forms and constructions. So we think it worth to be discussed in details.

As it has been already noted, many times by various scholars, verbs with the common meaning *have*, *possess* are either absent in the South Asian languages or occur extremely rarely. Pronominal possessive affixes are also quite uncommon. Their occurrence along with other ways of expression of the same relations is limited to certain languages and language groups - as enclitics (Sindhi, Siraiki, Assamese, Pashto, Baluchi, Brahui, Dardic and, Munda languages) or - as proclitics (Burushaski). The commonest way of expressing possession in the South Asian area is, evidently, a coordinative structure: *subject of possession* (possessor - Ps) and *object of possession* (relator - R) used within the limits of two types of syntactic constructions, nominal (NPC) and verbal (VPC) respectively. Areal uniformity is expressed both structurally and through the inventory of syntactical and morphological features characteristic to relations (r) between the two main constituents.

As is typical of many languages in the world, the nominal possessive construction in the South Asian languages is formed according to the general, characteristic for the region, pattern of attributive syntagm: *modifier + modified word*. The position of modifier is usually filled with a noun or a pronoun in the Oblique case (most often in combination with a *Genitive* postposition of the 2nd layer as cited above), or by a numeral (cardinal or ordinal), a non-finite (nominal or adjectival) verb form and also, in the languages where this morphological class is strictly opposed to other nominal classes, - by an adjective. Noun and verb formations can be included into the NPC either as separate words or, together with dependent words, as nominal or verbal phrases. In terms of deep structure we can distinguish between two subtypes of NPC: endocentric (NPC-I, Ps r R) and exocentric (NPC-II, R r Ps) where the word naming possessor is used in the position of the attribute and modified word respectively, e.g. H. *baap kaa ghar* 'father's house', Te. *naa kuTumbamu* 'my family': but H. *lakṛii kee makaan* 'wooden houses', Te.*iddaru pillalu* 'two children', Sant. *ooṛ-ak'-rɛ n kisā r'* the owner of the house', etc. The two types of NPC differ from one another by the sets of *fillers* of 1st position: in the NPC-I the attribute cannot be expressed by a numeral, while in the NPC-II - by a personal pronoun.

Among the areal characteristics of South Asian NPC (the exocentric type) one should certainly mention a special class of *attributive qualifiers* (affixes, auxiliary words), which indicate that the preceeding word (or combination of words) is related to the following one as R to Ps. Certain derivational suffixes of nominal adjectives are generally used as 'qualifiers' e.g. H. *-ooṛaa, -auṛaa, -iilaa, -eel*; As. *-ulaa*, etc. Yet this type of possessive formants, even in those languages, is restricted to a definite group of stems and cannot be considered as productive. A much more numerous group, having a wider distribution in the languages of the area, is represented by

analytical constructions with another type of 'qualifiers' - syntactic 'words which either belong to the class of postpositions' (as in the Munda languages), or, being close to postpositions in their characteristics, are yet distinct from them in certain distributional features.

In the Dravidian languages of South India (Tamil, Malayalam, Kannada, Kodagu, Telugu, Tulu) the attributive position in the NPC-II-type structures can be taken by analytical forms with qualifier markers, meaning 'possessing of a certain feature/quality' or 'devoid of certain feature'. These qualifiers are genetically related to the non-finite (participial) forms of a link-verb are – verbs with the meaning 'to be' 'to become' 'to exist' etc.

Ta. *-aaṉa*, *-aakiyaa aṟak-aaṉa (piraaṟi)* 'handsome (creature), cf. *aṟaku* 'beauty'; Kā. *-aadaa agala-v-aada* (nadi) 'wide (river)', cf. *agala-vu* 'width' ; Te. *-ainaa nijam-aina* (*maaṭa*) 'true (word)', cf. *nijamu* 'truth', 'true' <**aa*(kV) 'be', 'become';

Tu. *-iṭṭi – kaṭṭale-iṭṭi* (*koṇe*) ; 'dark (room)', cf. *kaṭṭale* 'dark-ness' <**irV* 'be';

Ta. *-uḷḷa – putti-y-uḷḷa* (paiyan 'clever (boy)' , cf. *putti* 'mind', 'wit' ; Tu. *-uppuu kaṭṭale-uppu* (koṇe) 'dark (room)' <**uḷ-* 'be', 'exist';

Te. *-gala – telivi-gala* (*baaludu*) 'clever (boy)' <**kala-(kV)* 'exist', 'appear';

Ta. *-aṟṟa – irakkam-aṟṟa* (manitaṉ) 'merciless (person)' <*al- 'not to be' ;

Te. *-leeni – buddhi-leeni* (pilla) 'silly (child)' <**il(l)ay* 'not to be', etc.

There are other instances indicating that the same qualifiers in the same languages may adjectivize not only single

words, but also nominal phrases with the noun in an indirect case as the central element. cf. e.g. Ta. *iṉṟiyaa-v-il* (Loc) – *uḷḷa carekkaar* 'Indian Government', *meeyvataṟ kk-* (Dat) – *aaṉa nilam* 'pasture' ('land for pasturing'), etc.

Thus, the ability to adjectivize – either separate words or verbal and nominal phrases – is to be regarded as the main distributional feature of these qualifiers, which enables us to distinguish them from other types of syntactic words and affixes.

Qualifiers represented in the South Indian languages have no direct analogies either in the Dravidian languages located northward, or in the Indo-Aryan languages contacting Southern Dravidian. Still the most interesting, both historically and typologically , is the fact that there are analogous constructions in Burushaski, remote as it is, from any contacts with Dravidian. Here, according (to C.A. Klimov and D.I. Edelman, attributes can be expressed by participial constructions built according to the *bahivriihi* pattern[19]: *iiɲi* (1) *burum* (2) *manum* (3) *hir* (4) 'a man whose beard has turned white' (lit. 'beard (1) white (2) turned (3) man (4)').

The formant-*vaalaa* in Hindi (cf. its variants in the western group of Hindi dialects –*haaraa*,-*haarau* along with its correlates in other Indo-Aryan languages - P. -*vaalaa*, S. -*vaaroo* cf. also M., G. -*aaraa*) is usually described in normative grammars either as a derivational suffix of relative adjectives (H. *śahr-vaalaa* 'urban', *matvaalaa* 'drunk') or as a formant of substantive derivation, which serves to form words, denoting professions (H. *rooṭiivaalaa* 'baker', *khōceevaalaa* 'hawker', S. *bungaaleevaaroo* 'door-keeper') or, finally, as a suffix of 'habitual tense participles' (H. *brij koo aaneevaalii saṛak* 'a road leading to the bridge', P. *kaam karnevaalee mazduur* 'workers doing their work', cf. M. *lihi-ṇ-aaraa* 'writer'). Meanwhile analysis of semantic structure of formations with -

vaalaa (and its correlates) enables us to sort out, in every instance, an invariant meaning, i.e. to define *-vaalaa* as a determinant of possession of some object (or quality) by a certain subject or, on the contrary, of the involvement of the subject in a certain situation. This, along with the possibility of adjectivization of nominal and verbal phrases (e.g.H. *niilii aakhõõ-vaalii laṛkii* 'a girl with blue eyes', P. *piilii ʧoopii-vaalaa munnaa* 'a boy in a yellow cap') allows us to define the formants of the *-vaalaa/-haaraa* type as functional analogoues of Dravidian *attributive qualifiers*.

In modern spoken Hindi one can encounter pleonastic combinations of *-vaalaa* with qualitative adjectives identical in their structure to the above mentioned Dravidian phrases: *meeraa garm suuʧ nikaaloo . niilaavaalaa* 'Take out my suit for cold weather. The blue one'.

Atributive qualifiers and other formatives of possession in the New Indo-Aryan languages – e.g. 'adjectivizing postpositions' H.*kaa*, P. *daa*, S. *joo*, M. *ca*; formatives of the H. *saa* ('-like:, 'resembling') type, etc. -follow an adjectival paradigm and agree with the dependent member in gender, number and case.

In both types of NPC represented in Dardic languages the position of the attribute can be filled by a noun in the Genitive case (formant of : – **(V)s(V)* type - Dameli, Shumashti, Pashai, Kalasha), in the Direct case (Which is less common -Tirahi, Shina) or Oblique case (Khowar). At the same time, in a number of languages at the southern border of Dardic geographical zone, in the endocentric variant of NPC possessor may be expressed by a noun with a special possessive marker, which does not coincide with the Genitive affix and which is added to the oblique form (*kati -stə* , Wa. *-bä*, *-bə*, Ash. *-oa*, *-wa*, Pr. *-ur*, *-uuri*, Gaw. *-ana*, WoT. *-(a)n*, Bshk. *-an*, *-aa*, Mai. *-aa*, Ky. *-aa*, ʃir. *-ma*; e.g. *kati kat-ó-stə* gul 'the

land of Kati', Ash. *nuristan-wa do* 'Nuristan mountains', etc.). In Kashmiri, alongside with the formations of this type (*-un* for the proper names,*-uk* for the common inanimate nouns), we can find (usually with common animate nouns) postpositional formations *-hund*, *-sund* 'of ', 'belonging to' (*naraan-un* 'belonging to Narayan', 'Narayan's'; *gar-uk* 'belonging to the house' ; *mɔɔl'* -sund 'of the father'<*mɔɔl'* -s (Gen) + *hund*).

In Gawar, Kanyawali, Tirahi, Kashmiri, Dardic possessive formants, like attributive qualifiers in the New Indo-Aryan languages agree with the determined member in gender, number and case: Ky. *maaʈ-aa mahaala* 'the boy's father', *maaʈ-ei mahaal ĩ* the boy's mother', *dhiiy-oo mahaali* 'the mother of the daughters'.

By analogy with Indo-Aryan (and Dravidian) attributive qualificators Dardic possessive formants can adjectivize not only separate words, but also phrases: K.*məəl' -is ti- maaj -e -hund kaar* 'father's and mother's business'. And yet, unlike Indo-Aryan, Dardic qualifiers are added to nominal and not to verbal stems.

Specific ways of expressing possession in Munda are determined by the general typological characteristics of these languages. Possessive meanings in Munda are expressed, along with the system of pronominal enclitics (Sant. *hɔpɔ n-iñ* 'my son',*ə pu-m* 'your father'; Koda *beʈaa-t* 'his son'), by postpositional formations which, at the same time, may, express common attributive relations too. In some languages of the North-Eastern group (Santali, Mundari) attributive possessive formants are classified according to their compatibility with names denoting animate (Sant. *-r ɛ -n*, Mu. *-rɛɛn*) and inanimate (Sant., Mu. *-(re)-ak* '/*-an* objects (cf. Sant. *am-r ɛn giḍə* 'your child', *am-rɛ - akʼ katha* 'your story'). But this principle does not seem to be applied consistently even in these languages (cf. also Mu. *ja hoṭo-akʼ bar-hoṛhon-*

kiŋg tai-ken-a 'there was a man who had two sons', lit. 'some man-of two-men sons were"). The most common type of attributive possessive formations in Munda are, evidently, formations akin to Sant. *-ak'* (Mu. *-ak'*, Kharia. *-(y)ag*. Korku. *-a*, Sora. *-ə* , etc.).

A majority of Munda possessive formants can be used both in the endocentric and in the exocentric variants of NPC (e.g. Sant. *ooɽ-ak' -rɛn kisaɽ* 'the master of the house', *paɳḍu-rɛn hɔpɔ n-kin* 'both sons of Pandu'). With the exception of *-ic'* usually marking the possessor (Mu. *aapɔɔ-t-in-iic' khurji* 'the property of (his) father', Sant. *uunii-iic' hɔ pɔ n* 'his son').

Lack of formal boundaries between grammatical classes of words and also between autosemantic and syntactic words, characteristic of the Munda structure, in a considerable degree determines the ability of possessive affixes to join freely with root morphemes of different semantics - substantival, 'adjectival 'and verbal' (e.g. Sant. *hu ḍiñ-iic' hɔ pɔ n* 'the younger son', *ra-ran-iic'* 'sorcerer', *bek'nao-iic'* 'door'). Combinations of these affixes with 'noun' and 'verb' phrases are formed with equal freedom (Birhor. *tiimiin-diin-iic'* 'of how many days?', 'how old?').

We have mentioned earlier that in the New Indo-Aryan languages attributive qualifiers of the *-vaalaa* type and other possessive formations are inflected according to the adjectival paradigm. At first glance this phenomenon has no analogy in Dravidian. The qualifiers derived from the participial forms of auxiliary verbs, like all Dravidian participles, are morphologically invariable. Yet, as has been repeatedly stressed in many words in Dravidian linguistics,[20] participial endings – *-a, -i, -ni*, marking these qualities, turn out to be identical – both materially and structurally - to the derivational suffixes of the so-called 'primary adjectives' (cf. Ta. *-aaṉ-a*, Ka. *aada* Te. *ain-a* lit. 'having become' Ta. *ariy-a* 'rare', Ka. do*ḍḍ*-a

'big', Te. *nall-a* 'black', Tu. *itt-i* lit. 'having been in', Kui. *payii-i* 'beating' Kui. *neg-i* 'good', Te. *lee-ni* lit. 'non-existing' Te. *nall-an-i* 'black', *pacca-ni* 'green'). In other words, in spite of all paradigmatic differences, in both Dravidian and Indo-Aryan languages there is a universal tendency to mark the attributive element of NPC (irrespective of its contents, i.e. including noun or verb phrases) according to the pattern of adjective and forming a kind of 'adjectival phrase'.

Data from Dardic (Kafir) languages support the view that this tendency to be area wide; e.g. the formant *-stə* denotes possession in Kati, but act as a derivational adjective formant in Ashkun and Waigali (Ash. *Astriimalii kaaṇ-stə seei* 'the woman is blind'. Wg. *saarastaa-oi* 'he is well'). In Munda the formant *-(a)n*, which normally marks 'adjectives' (Sant. *ḍaṛe-an* 'strong', Mu. *bugi-n* 'good'), can at the same time serve to form possessive attributes (Sant. *h ɛ r ɛ l-an* 'having a husband') or to be used in complex formants denoting possessivity (Sant. *paɳḍu-r ɛ -n hɔ pɔ n* 'son of Pandu', Birhor *In-rI-ɳ-Ic*' 'my').

Thus, the NPC structure in the languages of South Asia can be presented as following:

N(Nph)-//V(Vph)- (Obl =Gen)-(Qual)-(M_{adj}), N

Ps // R + r + R // Ps

Where NPh is noun phrase, VPh - verb phrase, M_{adj} - adjective marker Or, in a simplified way:

AdjPh + N

Ps // R (r) R // Ps

None the less important categorial feature of possessivity in the South Asian languages is the substantive transformation, i.e. transformation of a possessive phrase (NPC -I/II) into a single lexico-morphological complex ('possessive noun' meaning 'possessing something', 'belonging to somebody or something'). Substantive transformation is most clearly represented in the Dravidian languages where possessive nouns are formed by the gender and number affixes derived from the pronouns of the 3rd person (e.g. Ta. *-(a) vaṉ -ooṉ*, Te. *-vāāḍu* -masc. sing.; Ta. *-(a) vaḷ, -ooḷ* -fem. sing. ; Ta. *-atu* - neut. sing.; Te. *-(a)di* - fem. - neut. sing.; Ta. *-(a) var, -oor* - masc. -fem. pl.; Te.-*vaaru*, *-vaaḷḷu* - masc. -fem. pl.; Ta. *-(a)vai*, Tel. *-avi* - neut. pl.). These affixes may be added to any type of nouns and verbs as well as to phrases which can be used attributively within the NPC:(1) nouns in the Oblique or Genitive case ('genitive nouns': Ta. */eṉ/ makaḷ-iṉ* 'of /my/ daughter' -- */eṉ/ makaḷ-iṉ-atu* ' that which belongs to /my/ daughter'; Malto. *aḍḍa* 'house' -- *aḍḍa-taa-we* 'those who own the house'), (2) adjectives, including compound adjectives formed by attributive qualifiers ('adjectival nouns': Kan. dial. *doḍḍa* 'big' – *doḍḍ-avā* ' a great man'; Te. *nijam-aina* 'true' -- *nijam-aina-di* 'that which is true') and (e) participles ('participial nouns' : Te. */ikkaḍa/ nivasiñc-ee* 'living here' – */ikkaḍa/ nivasiñcee-vaaru* 'those who live /here/') etc. Possessive nouns, thus formed, are provided with all syntactic functions of nouns (e.g. are used as subject, direct object and nominal predicative), retaining in their paradigm peculiarities of pronominal declension.

Specific formation of possessive names in the Dravidian languages is undoubtedly conditioned by the rules of sentence structure (obligatory agreement in the predicative syntagm and its absence in the attributive syntagm) and the restrictions inherent in the 'morphological potential' of adjectives. In the New Indo-Aryan languages, where an attribute almost

universally agrees with the determined word (the languages of Eastern group being an exception) and where the adjective paradigm is not essentially different from the nominal one, substantive transformation is generally realized by means of conversion, i.e. without special formants. Such is the case with the substantivization of possessive attributes marked by qualifiers of *-vaalaa/-haara*a type, as a result of which *-vaalaa* and its correlates are considered either as derivative nominal suffixes (see above) or as the formants of the agent noun (H. *deekhnee-vaalaa* 'looking', 'spectator', M. *kar--ɳ-aar* 'doer', S. *paḍhi na-vaaroo* 'reader', etc.). The situation in Dardic languages is obviously similar. It seems rather interesting that in Kati, Ashkun, Waigali the adjectives with the possessive formant *-stə* (cf. above) are normally used, (as it was stated by D.I. Edelman,) in an independent (usually predicative) position.[21] In Munda languages, where the formal boundaries between classes of words are vague, the transformation is accomplished mainly syntactically – by the ellipsis of the second component of NPC (Sant. *paɳḍu-ak'* 'that /which/ belongs to Pandu', Mu. *huuṇiŋ-n-iic'* 'the younger /son/', etc.). At the same time there are reasons to believe that omission of the second component of NPC in Munda is not a mere ellipsis; here we can observe something similar to the substantive transformation. This is indicated by the constant use of certain possessive attributes without a dependent word (Mu. *uraaōō-tan-iic'* '/a man/ from the Uraon tribe', Sant. *bek'nao--iic'* 'doer'), by a combination, in predicative position, of 'genitive formant' and 'abstract nominalizer' *-t'a* (Parengi. *ar'i non? ming-ə n-t'* a 'this object is mine') and most convincingly, by partial lexicalization of the NPC components (Sant. *paɳḍu-rɛ-n-iic'* 'Pandu's wife' < 'she, who /belongs/ to Pandu').

Locally mixed patterns of substantive transformation are formed in agreement with the underlying mechanism of areal interrelations between various groups of language. In Brahui,

which has been developing in the course of millenia in the heterogeneous milieu, adjective, genitive and other types of possessive nouns characteristic of Dravidian languages are formed without pronominal formatives (*naaiilum-naa* 'of your brother '=' that which belongs to your brother'), i.e. nearly in the same way as they are formed in Indo-Aryan, Dardic and Munda. Yet the indirect cases of these names are formed according to the pronominal declension, i.e. in full agreement with the Dravidian pattern (*naaiilum-naa-ṭ-aṭ* 'that which belongs to your brother'- cf. *daa-ṭ-aṭ* 'to them', 'to these').

The Verbal Possessive Construction (VPC) also follows an universal pattern along the whole South Asian area with the exception of the region where some Sino-Tibetan languages and Burushaski are spoken. General structural features of this model are as following: (1) the link verb meaning 'be', 'exist', 'take place' is used in VPC as the predicate; (2) the subject of possession (PS) is marked by a formant of an indirect case; (3) the object of possession or relator (R) is used in the Direct (Nominative) case and the predicate agrees with it in the languages for which agreement is common; (4) from the point of view of standard South Asian order of words (SOV) the possessor stands invariably in the position of a grammatical subject. While the relator occupies the position of the object.

Preliminary analysis shows that possessor is marked differently depending on the semantics of the word used in the function of a relator – whether it denotes parts of one's body, close relatives, certain spiritual qualities etc. (i.e. objects which cannot be alienated) or, on the contrary, objects that are freely alienated. In other words, those are reasons to believe that in a number of languages of the area two types of possessive relations are opposed in one or another way, viz. alienable and non-alienable possession. Occurrence of two variants of VPC is reported for the whole bulk of the Indo-Aryan languages, for at least some Dardic and, also, for the

overwhelming majority of Dravidian languages.

In the construction of non-alienable possession in the Indo-Aryan languages the possessor is most often expressed by the Genitive case form)or by the Oblique case with an adjectivizing post -position): H. *meeree doo bhaaii hãi* ' I have (lit. ;'my are') two brothers', P. *admiide: do kann han* 'a man has two ears'. Sometimes the subject of possession in this type of construction may also be marked with the formant of the Dative (Dative-Accusative) case (M. *mahaaraajaa dhrtaraaṣṭraa-s śambhar muulee hoot̃i:* 'king Dhritarashtra had one hundred sons'). Yet a majority of Indo-Aryan languages use the Dative case markers only to denote possession of some immaterial object, some quality etc. (P. *ḍaakṭar nuu aapṇe aap vic visuvaas hai* 'the doctor has confidence in himself' - 'has self-confidence').

On the contrary, in the Dravidian languages the Dative case is most widely used to express possessor in the given variant of VPC (Ta. *aṭiiye:ṉ avarukku-neer-t-tampi alla* 'I, humble, am not his brother', Prj. *oonug saat jan cindul aaṭ jan papkul* 'he /had/ seven sons, eight children', Konda. *aya guruyeɳ budi siled* 'this teacher has no sense'). But certain examples can be found where the possessor is expressed by the oblique case as well (Kol. *okkon anden ba'la paṭlaknet* 'there was /only/ one son of the headman').

In the construction denoting alienable possession in the Indo-Aryan and Dravidian languages the possessor is marked with a postposition of locative semantics ('at', 'near', 'beside'), which is as a rule genetically related to the pronominal adverb 'here' or to a name with the meaning 'place', 'side', 'proximity' (H. *kee paas, kee hã:*, P. *dee kool*, S. *vaṭa*, B. *-kaachee*, Ta. *-iṭam*, To. *-kiḍs*. Ka. *-hattira*, Tu. *kaitaḷu*, Te. *-daggara*, Prj. *-kan*, Go. *-agga*, etc.), e.g. S. *hunane vaṭa vaḍoo gharu ahee* 'they have a big house', B. *eemaar kaachee churi aachee* ' I have

a knife', Ka. *ninna hattira eṣṭu duḍḍ' ide* ' how much money have you got?', Prj. *an kan gurrol cilə* 'I have no horse'.

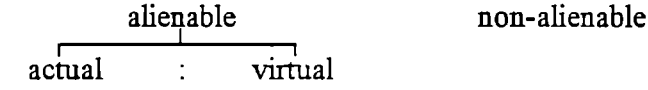
In the Dardic languages possessor in VPC is usually marked by the Oblique case form sometimes combined with the possessive formant (Tir. *myaa-na malə brok mazduraano waana* 'my father had many servants'). There are though some examples indicating that in certain Dardic languages affixes with Locative (or locative-aditive) semantics are also used as formants of possessor, e.g. Wg. *kiti rupai tuukaaooṭaa* 'how much money have you got?' (cf. *braawââ-kä* ' to the brother's', *amââ-kää* 'in the house').

The opposition of alienable vs. non-alienable possession seems to be lacking in the Munda languages where in a majority of cases the subject of possession is marked by a group of 'genitive' post positions (*-ak'*, *-rɛn*, *-ic'*, etc.) or by functionally identical pronominal enclitics. Nevertheless, a tendency to express, in one way or another, this opposition appears in these languages too. It can be traced, for instance, in Santali where possessor in alienable VPC is sometimes marked by the locative postposition *-rɛ(iñ-rɛpaṛaok' dhan-daulat bahka* 'my (lit. 'in me') part of the property'). Besides in Mundari and Korku the subject of non-alienable VPC is marked by the Dative case formant (Kor. *miia kooroo-ken baaria kon-king daan* 'one man had two sons').

Detailed study of possessive relations in South Asian languages (including the alienable vs. non-alienable opposition within the frame of VPC) is the subject which undoubtedly deserves special attention. Of particular interest is interference of the two variants of VPC in certain Indo-Aryan languages (e.g. in Spoken Hindi: *aarmii-aafisar see preem karnee keeliyee faulad kaa dil caahiyee, aur vah kitnii laṛ kiyookee paas hai* 'To love an army officer one needs a heart of steel. And how many girls have it with them?'). In great number of Dravidian languages the use of postpositions of locative semantics to mark a possessor seems to indicate the actual

possession, which means that the subject possesses the object at the moment (Ta. *eṉ-ṉ-iṭam paṇam uṇṭu* 'I have got money with me')virtual possession of such an object is expressed according to the pattern of non-alienable possession (Ta. *eṉakku meecai illai* ' I have no table').[22]

The cited examples give no reason to doubt the existence of the alienable vs. non-alienable opposition, but merely serve to indicate its gradual character:



The reality of the opposition is proved not by the mere presence of the two variants of VPC, but by the possibility (or impossibility) of their transformational rearrangement; cf. e.g. Te. *naaku ḍabbundi* 'I have some money'- *naadaggara ḍabb' undi* 'I have some money (with me)', while analogous transformation *naaku bhaarya undi* ' I have a wife' → **naa daggara bhaarya undi* is impossible.

Our survey of paradigmatic and categorial features of the South Asian languages serves to illustrate the general principles of areal studies, formulated above. Though incomplete and fragmentary, it seems to be productive in two ways: in the first place, it sheds some new light on what has been known of the areal convergence in the Indian subcontinent and, second, it develops the approach outlined by the founders of South Asian areal studies M.B. Emeneau and F.B.J. Kuiper in directing attention at the linguistic entities which, being less than rewarding for scientific pursuit, yet have the greatest relevance for the long-term results.

We refer to a set of formal elements which perform the key role in the nominal and verbal inflexion in almost all the languages of the area. Into this set are included postpositions and other formants of analytical derivation discussed in this paper, auxiliary verbs, particles etc., which functionally belong

to morphology (*analytical morphology*), while, by the degree of their semantic value and distributional freedom, to syntax ('minor syntax'). It is within this group of linguistic units which hold intermediary position between the 'morpheme' and the 'word' that semantic homogeneity of genetically heterogeneous and materially alien formants is actually manifested. These data seem to have most obvious implications for our understanding of the universal mechanisms underlying various forms of expression. Apart from the phenomena described above, compound verbs (including combinations with analytical operators of aspect, voice, mode of action etc.) appear most promising to us in the areal study of minor syntax.

All previous attempts at classifying South Asian languages in terms of contensive typology reveal the fact which, at first sight, seems paradoxical : while the areal languages demonstrate obvious structural proximity they do not fit in with the patterns of a single type, but, conversely, according to conclusions by B.A. Zakharyin, 'demonstrate considerable contensive - typological diversity'.[23] We suppose that the reason for this can be found not only in the complex character of historical development of the area; it seems to be rather a case of an approach which tends to do less than justice to the linguistic data other than subject-object relations. Study of categorial features based on the analysis of minor syntax in its elements may reveal new perspectives in the typological studies of South Asian languages.

7. CONCLUSION

The data discussed above along with the material in this subject provided by our predecessors enable us to make some conclusions concerning the structure of the area. It seems to be polydimensional and more complicated than it was rendered in the first half of the century, when the lines of division were generally drawn in the latitudinal direction according to the boundaries of the main genetic groups. Recent studies make

it possible to discern new isoglosses including those which run in the longitudinal direction F.C. Southworth was first to direct attention to isoglosses of this type.[24] The maps of the subcontinent, compiled by him indicate that it is possible to distinguish at least three subareal zones: Western, Middle and Eastern. The boundaries of the zones, mapped according to the separate isoglosses, may not coincide, while the general zonal structure remains intact. The Eastern and Western zones invariably form two poles, characterized either by the presence or by the absence of a feature, while in the middle zone this feature may retain, although manifested in a limited number of cases or irregularly, (sometimes it exists only as a relic).

That the zonal division presented by Southworth is not accidental, is supported by a certain number of isoglosses of analogous configuration. Thus, the isogloss, which illustrates categorial features of gender, indicates that two fold systems (masculine : feminine and /or animate : inanimate) are mainly typical of the Middle zone, which embraces a majority of Indo-Aryan languages, Dravidian languages of Central India (minus Pengo) and Telugu, while a three fold system (masculine :femininee neuter) is represented by Western and South-Western languages (Marathi, Gujarati, Kannada, Tulu, Tamil); the Eastern group of languages lacks the category of gender. The same tendency can be observed in the hierarchial structure of the nominal declension and in the main characteristics of possessivity: being most clearly manifested in the west and north-west, the features, advancing further towards the east and north-east, gradually loose clarity of formal expression. Even the field of diachronic phonology is subjected to zonal division. Thus, in both Indo-Aryan and Dravidian the historical initial **v* - is retained mostly in the Western languages (Sindhi, Siraiki, Gujarati, Marathi, Malayalam and Tamil), while in the Eastern and Middle zones it merges (fully or partially) into *b*-.

In a number of cases a zonal structure of a different type (centre : periphery) is superimposed on the above mentioned bipolar structure. Thus the set of features characteristic of the Eastern zone is sometimes represented by enclaves in the remote West of the area (e.g. absence of the category of gender in Brahui, Toda, Malayalam, etc).

None of the subareal zones, either in present time or in the historical retrospective, corresponds with the territory occupied by a linguistic family or a group. Rather, we deal here with the most archaic structure, which may be traced back to the earliest period, long before the ancestors of Dravidians and Mundas migrated into the Indian subcontinent, let alone Indo-Aryans.

All the more interesting are the analogies of the threefold areal structure in the 'historical' (i.e. quasi-historical) Ancient Indian tradition, represented the Mahabharata, Ramayana and Puranas. The opposition of the Eastern kingdoms (*Anga, Vanga, Pundra* etc.) to the states of Midland (*Madhyadeeśa*) and the kingdoms of the North-West (*Sindhu, Sauvira, Madra* etc.) has been central for the ancient Indian epic . (e.g. hostile attitude to the Pāñjab kingdoms as 'barbarous' (Mbh. VII. 93. VII. 40.44; XII. 307); distribution of the tribes in the battle at *Kurukshetra*; specific contacts of the *Ayodhya* kingdom with the remote Eastern, North-Western and Southern states, yet lack of contacts with the neighbouring parts of *Madhyadeeśa* (Ram. I. 13, 21-29), etc.)

Of particular interest for our study are variants of myths connected with the starting point of traditional history – the origin of royal dynasties which is traditionally interpreted as going back to the three sons of *Manu – Iḷă:, Ikshvaku* and *Sudyumna.* [25] A certain geographic area is assigned to each of them and their progeny. Thus *Iḷa's* descendants (*Aila* -'the Lunar race') are stably asociated with the West and North-West, while *Iḷă:* is traditionally named king of *Balhi* (*Balhika*)

in Panjab (Ram. VII. 83. 3, 7) and the lands, belonging to his descendants (*Yaadava, AAṇava*) occupy the same region. *Ikshvaku's* ('the Solar race') are in possession of the lands in the Ganges valley (*Ayoodhyaa, Videeha, Vaiśaalii*) and in Deccan. *Sudyumna's* domains occupy Bihar, Orissa and Bengal.

Thus, the cultural, geopolitical and, perhaps, even linguistic self-existence of the three zones dividing the territory of the subcontinent along longitudinal lines does not seem to be alien to the ethnic consciousness of ancient Indians. Moreover it appears in the myths as something primordial, something that existed long before the historical process originated. Admitting that, according to the Puranas, *Sudyumna* – the king of the East – is nothing but a travesty form (*kimpurusha*) of the Western king *Iḷā*:, we have to acknowledge that the ancient Indian myth provides us with both variants of the structural pattern inherent in the areal configuration of isoglosses: 'bipolar' and 'centric'.

This final example can be viewed as another evidence to indicate genetic and conceptual links between the notion of a linguistic area and the general cultural process. It appears that only on the basis of coordination of linguistic data with those provided by historical, ethnographic and cultural studies can areal linguistics properly define linguistic groups constituting the area and answer the questions a scholar is traditionally concerned with : what, when, where and where from?

NOTES

* The abbreviations of languages names are those used in CDIAL.

1. Emeneau M.B. 1956. India as a linguistic area. *Language*, 32 p.16; cf. also: 1974. The Indian linguistic area revisited. *International Journal of Dravidian Linguistics,*

1974, 3: 1, 92-134.

2. On the so-called 'areology' see e.g. Borodina M.A.1978. Areologie de Jean Séguy. Hommage à J.Séguy. Toulouse. T. I, 165-171.
3. Jacobson R. 1931.Uber diephonologischeSprachbünde; 1938. Surla theorie des affinites phonologiques entre los langues. 1962. *Selected Writings, Vol. I.* The Hague. 137-143, 234-246.
4. Actes du 1-er congrèss internationale de linguistes à la Haye 1928. Leiden, 1930, p. 18; see also: Becker, M.1958. *Der Sprachbund.* Berlin-Leipzig.
5. Birnbaum M.1965. Balkanslavisch and Sü:dslavisch - Zeitschrift fu:r Balkanologie, III, S. 12-63 ; Georgiev V. 1968. Le problème de l'unité linguistique balkanique. - Actes du 1-er congrès des études balkenique et sud-est européennes, VI. Sofia, p. 9-18 ; Schaller H.W., 1975. Die Balkansprechen. Heidelberg.
6. See, e.g.: Weinreich U. 1953. *Languages in Contact.* New York, ; Haugen E. Language contacts. *Reports for the 8th International Congress of Linguistics* (Osilo, 5-9 August 1957), II, p. 263-267 ; Rosenzweig V.Ya.1972. *Yazykovye kontakty* (The Linguistic Contacts). Leningrad.
7. Rosenzweig V.Ya. Op. cit., 25.
8. Weinreich U. 1972. Unilingualism and monolingualism. *Novoye v linguistike, VI.* Moscow. 32-33 (in Russian).
9. See especially: Emeneau M.B. 1980. *Language and Linguistic Areas.* Stanford; Kachru Y., Pandharipande R. 1979. On ergativity in selected South Asian languages. *South Asian Languages Analysis, No 1,* University of Illinois; Andronov M.S. 1965. O tipologicheskom skhodstve novoindiyskikh i dravidiyskikh yazykov (On the typological similarity of the New Indo-Aryan and

Dravidian languages). *Linguisticheskaya tipologiya i vostochnyye yazyki*. Moscow; Elizarenkova T.Ya. 1967. Ergativnaya konstruktsiya predlozheniya v novoindiyskikh yazykakh (Ergative construction in the New Indo-Aryan languages). *Ergativnaya konstruktsiya predlozheniya v yazykakh razlichnykh tipov*. Leningrad.; Zakharyin B.A. 1987. *Tipologiya yazykov Yuzhnoi Azii* (The Typology of the Languages of South Asia). Moscow; Zograph G.A. *Yazyki Yuzhnoi Azii. Opyt arealno-tipologicheskoi kharakteristiki* (Languages of South Asia. An attempt of areal-typological characteristic) - forthcoming.

10. Gurov N.V. 1989. *Kusunda - sinokavkazskiye leksicheskiye paralleli k kharakteristike nachalnogo etapa formirovaniya Yuzhnoasiatskogo areala* (Kusunda -Sino-Caucasian lexical correspondences. Towards the characteristic of the initial stage of the formation of South Asian area). *Lingvisticheskaya rekonstruktsiya i arkheologiya*. Pt 3. Moscow.
11. Besides the works referred to in the notes 1 and 9 see also Kuiper F.B.J. 1974. The genesis of the linguistic area. *IJDL*. 3: 1, 135-153; Vermeer H. 1969. *Untersuchungen zum Bau Zentral-su:d-asiatischer Sprachen*. Heidelberg.; Masica C.P. 1976. *Defining a Linguistic Area: South Asia*. Chicago -L.; Shapiro M., Schiffman H.F. 1981. *Language and Society in South Asia*. Delhi.; Krishnamurti Bh. (ed.).1986. *South Asian Languages. Structure, Convergence and Diglossia*. Delhi.
12. Ramanujan A.K., Masica C.P.1969. Toward a phonological typology of the Indian linguistic area. *Current Trends in Linguistics*, 5 543-577.
13. Haugen E. Op. cit., 246-247; cf. note 8.
14. Masica C.P.Op.cit., 187-190.

15. Zograph G.A. 1976. *Morfologicheskiy stroy novykh indoariyskikh yazykow* (The Morphological Structure of the New Indo-Aryan Languages). Moscow, 79-81.

16. Bloch J. 1965. *Indo-Aryan from the Vedas to Modern Times*. Paris, p. 322-327; Chatterji S.K. 1960. *Indo-Aryan and Hindi*. Calcutta, 31-65.

17. Shanmugam S.V. 1972. *Dravidian Nouns (A Comparative Study)*. Annamalainagar, 259, 265.

18. Gurov N.V. 1972. *Imennoye skloneniye v dravidiyskikh yazykakh i mikroparadigma protoindiyskikh tekstov* (Dravidian nominal declention and micro-paradigm of the proto-Indian texts). Proto-Indica: pt I. Moscow. p. 106-115 ; Zograph G.A. *Morfologiches-kiy stroy ...*, 55-63.

19. Klimov G.A., Edelman D.I. 1979. *Yazyk burushaski* (*The Burushaski Language*). Moscow. 93-94.

20. Andronov, M.S. 1978. *Sravnitelnaya grammatika dravidiyskikh yazykov (The Comparative Grammar of the Dravidian Languages)*. Moscow, pp. 286, 380-381; cf. Shanmugam S.V. Op. cit., 239-240. 384-385.

21. Edelman D.I. 1983. *The Dardic and Nuristani Languages*. Moscow, 48, 98-99.

22. Cf. in Ao Naga: ni ki 'thy house', but ni-Lh ('Genitive' suffix) ki 'a house (which) temporary belongs to thee' (Zakhariyn B.A.Op.cit., 74).

23. Zakharyin B.A. Op. cit., 173.

24. Southworth F.C. 1974. Linguistic stratigraphy of North India. *International Journal of Dravidian Linguistics*, 3: 2, 210-215.

25. Pargitter F.E. 1922. Ancient Indian Historical Tradition. p. 253-255, 287-288 (with references to Vaayu-PuraaNa, 85, Brah-maaNDa-Pura:Na, III.60, Harivamśa, 10, Matsya-PuraaNa, 12, etc.).

OPiL 24: 49-69(1998)

DISTINCTIVE VOWEL QUALITY, QUANTITY AND NASALIZATION IN TELUGU AND HINDI*

K.Nagamma Reddy
Osmania University, Hyderabad

ABSTRACT

This paper deals with phonological analysis and description of vowel quality, quantity and nasalization. Some pairs of vowels are phonetically characterised by both relative duration and of phonetic quality. In such instances, the phonological contrast between the pairs is not confined to one single distinguishing feature (i.e. quality or quantity) but to more than one. This paper examines in detail whether quantity difference is distinctive when it is accompanying a noticeable difference in quality (or vice versa) focussing on Hindi and Telugu Vowel systems. The quality and quantity relationships and the accompanying nasalization are investigated by providing instrumental phonetic evidence in support of the analysis put forward.

1. INTRODUCTION

There are many salient common phonetic and phonological features among the languages of India, despite their different genesis (Nagamma Reddy, 1992). The linguistic data are rich and varied as expected in a multilingual society like India. Any attempt to give an overall account of the phonetic/phonological description, especially of vowels, in Indian languages raises several problems as cited by Masica (1991). The inadequacy of the available data, different analyses of the same data, the inclusion of borrowed elements and their status, the nature of writing systems and consideration of higher level (i.e. grammatical) units are some of the basic issues.

* An earlier version of this paper was presented at the 18th All India Conference of Linguists held during October 1995, Central Institute of Indian Languages, Mysore. Organised by the Linguistic Society of India, Pune.

According to Masica there are languages which have no length or quantity (i.e. tenseness) distinction in vowels. These include, for example, Kharia, Korku, Mundari, Marathi, Oriya, Ollari etc. There are languages which have quality distinction only in mid vowels, for example, Assamese, Bengali, Burmese, Santali and Sora. The quality (i.e. tense/ lax) distinction in all heights is noticed in Hindi, Awadhi, Punjabi, Sindhi etc.; quantity distinction in all three heights includes Sinhalese and South Dravidian languages, Khasi and Lushai. Length distinction is found in Brahui also but only in the high and low vowels and quality distinction in Pashto.

Some Indian languages have identical vowel clusters interrupted by a glottal stop or fricative, for instance, Mundari, Kurukh and Kolami. In some other languages, vowels are modified for quantity as in Tamil, for quality as in Hindi and neither for quantity nor for quality as in Oriya. Masica(1991) observes that Dravidian languages with the feature of length seem not to exchange it for tenseness while a Munda language (Juang) with no length has acquired it, and an Indo-Aryan language (Sinhalese) has lost the tense/lax distinction altogether but developed long/short opposition in every vowel. Tense /lax distinction is found at all the three heights in standard Hindi, Lahnda, Punjabi and Sindhi, but Gujarati has it only in mid vowels and the low vowels. Combined length contrast with tenseness at every height is found in Lhasa. Nicobarese combines length contrasts with tense/lax only in the mid vowels.

Vowels in some Indian languages are characterised further (phonemically) by both length and nasalization. However, there are differences in the phonemicization of the vowels which differ in length or nasalization or both. For example, long vowels are taken either as separate phonemes or as corresponding short vowels plus the suprasegmental phoneme /:/ 'length', alternatively long vowels as basic plus

the 'phoneme of shortness'. The number of phonemes, thus, may be the same but basic inventory would be different. Further in major Dravidian languages like Telugu, Kannada, Tamil and Malayalam, there may be only six vowel phonemes instead of ten when length is treated as a separate phoneme.

Nasalized vowels, as unit phonemes or as corresponding oral vowels plus the supra-segmental phoneme /~/ 'nasalization' also lead to significant differences in the inventory of phonemes similar to the long vowels. For example, all the ten vowels are nasalized in Konda and Kuvi, and hence when the nasalization of vowels is conceived of as a separate phoneme, the number of nasalized vowel phonemes is reduced into eleven from twenty. Nasalization plays a significant role phonologically in some languages. Accordingly, in certain languages, the length and nasalization are treated as separate phonemes when all the vowels function as long and/or nasalized. In such cases, all the Indian languages may have to be treated the same way for the sake of uniformity and or for the neatness of pattern in such studies as typological inventory of vowels.

There are further problems in the phonological analysis and the classificaion of vowel phonemes and their inventory arising from the way length and quality are handled independently or in combination. Some pairs of vowel phonemes are phonetically characterized by differences of both relative duration and of phonetic quality. By making the length the distinctive factor, we obtain vowels *i/i:* or *i/ii, a/a:* or *a/aa* etc., and ignore the qualitative differences. But by making quality the distinguishing factor, we obtain vowels i/I or a/A etc. and ignore the length difference which hides certain phonetic facts of a given language and poses a greater problem. This is due to the fact that in certain languages the phonological contrast in the pairs (of vowels) is not in terms of one single distinguishing feature, i.e. quantity or quality, but both. When both duration and quality are involved, either one or both of

the phonetic differences are utilized as the basis for choosing an appropriate phoneme and its representation in transcription. Some scholars have preferred to give priority to phonetic quality and some others to quantity, and still others to both quality and quantity, for example [*bi:t*] 'beat' vs [*bit*] 'bit' or [*kʰa:t*] 'cart' vs [*kʰat*] 'cat', etc. in English, [*pi:ʈa*] 'wooden seat' vs [*piʈʈa*] 'bird' in Telugu, or [*di:n*] 'poor' vs [*din*] 'day' in Hindi. In English, Jones (1918) gives priority to quantity over quality in the matter of vowel symbolization, whereas Gimson (1967) gives preference to quality over quantity. Phonological symbols themselves are a source of confusion in that different descriptions have used different analysis and symbols for the same sound, or the same symbols for different sounds without implying any analytical-theoretical disagreement. Laver (1994) points out that 'In many languages, including English, some pairs of vowel phonemes are phonetically characterized by differences both in relative duration and of phonetic quality', for which the examples are as provided above. Combining both quantity and quality in the phonological representation, transcribing the pairs accordingly may have a better solution in the designing of pronouncing dictionaries for foreign learners and may be helpful for particular pedagogical applications in teaching a foreign language and to improve the intelligibility and naturalness of artificial speech produced by computer from key-board-typed sentences.

It has been observed that a number of languages show phonological opposition between vowel units which is rarely dependent on duration alone, but is accompanied by certain differences in quality as well (Lehiste, 1970). For instance, Gimson (1970) states that of the two factors, quality and quantity, 'it is likely that quality carries the greater contrastive weight', though it should be noted in an earlier analysis Jones (1967) concludes that in the 'borderline cases', the length difference may be considered. In the case of Telugu the most

widely accepted view appears to be that far from being a 'borderline case', the only difference between a corresponding long and short vowel is one of duration. So far no one has said with certainty that there is also a quality difference accompanying the quantity difference between all the short and long vowel pairs in Telugu.

Nasalization of vowels plays a significant role phonologically in certain Indo-Aryan languages such as Hindi, Punjabi, Gujarati, Marwari, Sindhi, Konkani etc. But unfortunately, there is a great deal of discrepancy in the description of nasalized vowels even of the same language and Hindi can be taken as one of the examples. In some Indian languages, nasalization plays a significant lexical as well as morphological role particularly in some Indo-Aryan languages such as Hindi, Punjabi, Gujarati, Rajasthani, Sindhi, Konkani etc. According to some scholars, for instance Dixit (1963), even for the same language (i.e. Hindi) there are no nasalized mid vowels, but for Misra (1967) and Kelkar (1968) there is phonemic nasalization of all vowels, yet still for others, Narang and Becker (1971) for instance, all vowel nasalization is predictable and for Ohala (1975) predictability is different for short and long vowels either phonetically or morphologically, For details of the phonological systems of Indian languages, see Ramanujan and Masica (1969), Masica (1991) and the phonetic readers published by the Central Institute of Indian languages, Mysore.

The main aim of this paper is to examine whether quantity difference is distinctive when it is accompanied by a noticeable quality difference focusing particularly on Hindi and Telugu vowel systems with respect to their quality and quantity relationship as well as accompanying nasalization by providing acoustic (i.e. instrumental) phonetic evidence in support of the conclusions drawn.

2. QUALITY-QUANTITY DISTINCTION BETWEEN VOWEL PAIRS IN TELUGU

There are ten basic vowels in Telugu which have been traditionally divided into two classes: one consisting of five long and the other five correspondingly short vowels (Nagamma Reddy, 1986). The anonymous author (1918) states that there is no difference other than quantitative one between low vowels and front mid vowels. According to his analysis only the close vowels have qualitative difference which is not the case in the descriptions of, for instance, Kostic et al (1977). In the typology and universals of vowel systems, Crothers (1978) uses same symbols for Telugu for short and long mid vowels, but different symbols (differing in quality) for only close and open vowel pairs.

Kelley opines (1959) rightly that the vowel system has caused the greatest degree of divergence in the description of Telugu phonology and the number of vowel phonemes in Telugu has been a very controversial issue. Studies by Sjoberg (1957), Kelley, (1959) and Krishnamurti (1961) have added an additional vowel phoneme */æ:/* in the vowel system. In a later account of the inventory of Telugu vowel phonemes, Krishnamurti (1962) introduced another short vowel */æ/* as a counterpart of */æ:/* thereby proposing 12 vowel phonemes for modern Telugu. These two vowels have a restricted distribution and its contrast is also a matter of dialectal and morphological or grammatical one (see for details Nagamma Reddy, 1981, 1986).

All vowels in Telugu, except */o/*, can occur and contrast in all three positions of a word as exemplified in Tables III for initial, IV for medial and V for final in Nagamma Reddy (1986;1988). However, a simple statement of contrast between long and short vowels in word-final position, reveals certain facts about the nature of the distribution of long vowels in

Telugu since they occur relatively infrequently in that position of polysyllabic words as compared to other initial, medial positions. (For details, see Nagamma Reddy, 1986 : 127-128). Even phonetically, only two degrees of vowel length, short and long, are distinguished by several scholars with the exception of Kelley (1959) who considered more than two (in that he observes four degrees of length in vowels: short, lengthened short, long, and lengthened long in final position, but only two degrees, short and long, in non-final position).

In view of the above discrepancy in the statements of various scholars in the field and with regard to quality and quantity differences between the vowel pairs in Telugu, a preliminary investigation of vowel duration (quantity opposition) and its phonetic quality was carried out using several instrumental acoustic and articulatory techniques such as spectrography, kymography, labiography, palatography, intensity metre and x-ray photography. The following are the findings (Nagamma Reddy, 1986).

X-ray tracings of vowels in Telugu (see Fig. 1) reveal that the tongue hump position in the front vowel is somewhat further forward for long vowels than for short ones. Similarly, in the back it is more retracted for long vowels than for short ones. The results of the palatographic investigation suggest similar conclusions and spectrographic study also reveals that F3 is considerably higher in long vowel series than in short vowel series. F1 is slightly higher for short vowels (except for /a/) than for long vowels. The F2 is also higher for front long vowels than for corresponding short vowels. The details regarding the minimal pairs of words used for the measurement of F1, F2 and F3 in initial, medial and final positions along with their durational values can be found in Nagamma Reddy (1986). That all short vowels move towards the centre when compared to the corresponding long ones (see Fig. 2). Labiographic investigation shows that the opening between

the lips (i.e. lip-height) is slightly larger for long vowels than for corresponding short vowels (see Nagamma Reddy, 1981 : 238).

The measured duration of vowels of both kymograms and spectrograms revealed that word-initial and word-medial vowels have a short to long ratio of more than 1:2, which supports the claim that the relative vowel length is distinctive in the pairs under discussion. There is also a great deal of variation in the ratio of short and long vowels depending on different structures of the same disyllabic words used. The long vowels, thus have one and a half to three times the duration of the corresponding short ones in comparable contexts. The relationships are very similar to those described by Abramson (1962) for Thai and Velayudhan and Howie (1974) for Malayalam, but different from those revealed in the American studies of English where the difference is less pronounced, the ratio of short durations to long in related pairs being 1 to 1 1/2 at its maximum value instead of its minimum.

Instrumental recordings of all kinds used in the study reveal a considerable difference in all vowel pairs irrespective of their height differences, contrary to the statements made earlier by some scholars. Format frequency charts and x-ray tracings of long and short vowel pairs in Telugu show higher or lower tongue position. The long vowels have a relatively higher intensity than the short ones.

All short vowels fall more in the centre of the vowel area and the corresponding long vowels more on the periphery (see Nagamma Reddy, 1981; 1986). The difference in vowels both duration and quality are maximized in non-final position and minimized in final position of words. There is a correlation of vowel quantity and quality in Telugu.

3. QUALITY-QUANTITY DISTINCTION BETWEEN VOWEL PAIRS IN HINDI

The number of vowel phonemes in Hindi also appears to be controversial. For example Kostic et al (1976) list only 8 phonemes including the (nasalised vowels) whereas many others list at least 10 vowel phonemes (Ohala, 1983). Moreover, some scholars have preferred to represent the vowels only in terms of quality differences (e.g. Ohala) and others in terms of quantity differences (e.g. Kostic et al). The selection of the feature that distinguishes a phoneme can be of quality or quantity or even both.

Accordingly depending upon the choice of the feature, the phoneme inventory varies. Rudin (see Masica, 1991) considers the so-called 'short' and 'long' vowels in Hindi to be different phonemes, some other scholars are of the opinion that only some pairs have quality differences but not all and in some others opinion there is only a quantity difference. Yet there is a difference in the representation of vowel quality and quantity relationships between the pairs. Kostic et al (1976) use the same symbols for both short and long vowel pairs (differing only in length) except for the low vowel pair, whereas Dhamija's (personal communication, a phonetician and a native speaker of Hindi) opinion is that the high and low vowel pairs differ mainly in quantity rather than in quality, for example in /*din*/ 'day' vs /*di:n*/ 'poor', /*u ʈʰ*/ 'get up' vs /*ū: ʈ*/ 'camel' or /*guɽ*/ 'jaggery' vs /*gu:ɽʰ*/ 'deep' and /*jal*/ 'water' vs /*ja:l*/ 'net'. The mid vowels differ in quality and quantity in such words as /*be:l*/ 'creeper' vs /*bɛ:l*/ 'bullock' and /*o:r*/ 'direction' vs /*ɔr*/ 'and'. But Kostic et al state when 'contrasted with Hindi short vowel i, Hindi long vowel i: is phonemic, not only due to its quantity but its quality as well'. The acoustic studies of Hindi vowels, for example by Ganesan et al (1985), consistently show a difference in quality as well as in quantity

of all vowel pairs. The vowel format frequency and its durational measurements represented by Ganesan et al (1985) reveals a significant difference in quality and quantity of each vowel pair irrespective of their tongue height difference. Ganesan et al (1985) represent the vowels only in terms of quality difference (*i/I, u/ʊ, e/ɛ, a/ə; o/ɔ)* and ignore the quantity difference.

The acoustical measurements show that voiced consonants are inherently and significantly shorter than voiceless ones whether they are aspirated or unaspirated in both the languages mentioned above, which fact does not, however produce a compensatory difference in the duration of the preceding vowel. The aspirated consonants are also distinguished by a considerable difference in their duration from the corresponding unaspirated ones. The difference in vowel length before aspirated (voiced or voiceless) and unaspirated is much larger than when the same vowel is followed by a voiced or voiceless plosive. In Telugu, contrary to Hindi as claimed by Maddieson and Gandour (1977), there is no lengthening of the preceding vowel before the aspired plosive (see Nagamma Reddy, 1992).

4. VOWEL NASALIZATION IN HINDI AND TELUGU

Theoretically it is possible to produce all segments, other than the stop consonants, with velic opening (i.e. with a lowered velum) such that they can all be nasalized. However, the use of nasalization as a phonetic feature may or may not reflect phonological contrast. For example, in English or in Telugu, the nasalization of vowels is a phonetic phenomenon in that it occurs only when they are adjacent to a nasal consonant. On the other hand, in French (Delattre, 1962) or in Hindi (Ohala, 1975 reproduced in 1983) the vowel nasalization is contrastive and manifests thus as a phonoligical feature. Furthermore, various studies on phonetic/phonological

nasalization, particularly with regard to vowels (Delattre, 1962) reveal that languages in general have a tendency to nasalize certain segments only. Languages also vary in the degree and direction (i.e. the spread) of nasalization of vowels (progressively or regressively). That is, the nasalization of vowels adjacent to a nasal consonant can occur either left-to-right direction from nasal-to-vowel or from right-to-left, or both depending on the nature of language.

Indian languages have both phonetic and phonological nasalization. There are two kinds of nasalization in Indian languages: (1) the predictable phonetic nasalization (in the context of nasal consonant) and (2) the contrastive phonological nasalization. It is common for one to talk about vowel nasalization rather than the consonant nasalization. In fact, the consonant itself might be produced in Telugu as nasalized intervocalically or word-finally as in [*ma: w̃a*] 'father-in-law' and [*po:da:w̃*] 'let us go'. Certain other consonants, in fact, have the same tendency as vowels to be nasalized when they occur adjacent to a nasal consonant. For example */ɾ,r,y,h/* can at least be partially nasalized. In fact, /h/ is heavily nasalized even when it is not immediately adjacent to the nasal consonant, as in [*mõ:h̃ãm*] 'love'. Moreover, the use of nasalization as a phonetic feature may or may not reflect a phonological contrast. In most of the Dravidian languages nasalization of the vowel is a phonetic phenomenon, which occurs only when they are adjacent to a nasal consonant. However, in some cases word-final nasal consonant is deleted, for example, the formal */avan/* is represented by colloquial */aṽa/* 'he' in Tamil; or the word final syllable */-m(u)/* in Telugu is pronounced as */po:dã:(w̃)/* 'we will go', and this in turn contrasts with */po:da:/* 'won't it go?'.

In most of the Indo-Aryan languages, unlike Dravidian, the vowel nasalization is contrastive. But in such cases nasalized vowels seem to be historically derived from the nasal

consonants when followed by a homorganic consonant unlike Tamil or Telugu. The distinctiveness of the oral-nasalized vowels in such examples as */sas/* 'mother-in-law' and */sãs/* 'breath' is only a matter of phonetic process. From Ohala's (1983) point of view, the nasalized vowels in Hindi arose due to a phonological change in the development of old Hindi to Middle Indo-Aryan (MIA). MIA word-medial consonant clusters were simplified by compensatory lengthening of the preceding vowel, if the cluster consisted of a nasal followed by a consonant the nasal was deleted, and the preceding vowel was lengthened and nasalized. All vowels may be nasalized phonetically, but phonologically there is a disagreement regarding Hindi (see Masica, 1991). According to Dixit (1963) there is no nasalized /*ẽ*/ or /*õ*/ in Hindi, whereas Misra (1991) and Kelkar (1968) feel that, there is phonemic nasalization of all vowels, and for Ohala, however, the 'long' vowels are nasalized finally and before voiceless stops only, but there is a predictable nasalization of long vowels before voiced stop.

Nasalization of vowels in Telugu is not contrastive. However, there are some vowels which may be nasalized in certain colloquial paralinguistic expressions. They have only lexicogrammatical function. For instance, /*ẽ*/(with raised intonation) 'Why didn't you' or /*ã:*/ 'What is it',etc., are often used by the speaker in day-to-day conversational Telugu. Such vowels with nasalization are used as complete utterances. Each of them could be thought of as contrasting of an oral vowel in isolation as a member in the Telugu alphabet. There is no contrast of length in these expressions. They are always long. As Michael Garman (personal communication) has suggested these vowels in Telugu may be called 'nasal vowels' rather than nasalized vowels. The nasalised vowels that are historically derived due to phonological process in a given language, where the nasal consonant is deleted or lost and the preceding vowel gets nasalized and acquires a phonemic status,

may also be referred to as nasal vowels in order to maintain a distinction between phonetic and phonemic (i.e. phonological) nasalization. There are a number of phonetic, phonological and grammatical factors that may lead to nasality of a vowel. Some forms such as */ẽ:gu/* or */wã:ḍu/* are written as *'e:gu'* or *'wa:ḍu'* in inscriptional Telugu representing historically *'e:ngu'* or *'wa:ŋḍu'*, respectively. Printing system of classical Telugu texts introduced *arasunna* 'half circle' meaning a half nasal different from *niŋḍusunna* 'zero' representing a full nasal.

The vowels occurring after a nasal consonant in Telugu as in */ne:nu/* 'I', have a much stronger tendency to be nasalized than that of vowels occurring before a nasal consonant in such examples as */pa:mu/* 'snake', */pani/* 'work', */ani/* 'having said', and */aŋḍa/* 'protection'. The same would appear to be the case in some other Indian languages, for instance, in Bengali (see, Pathak, 1976; 314). This seems to be contrary to what has sometimes been stated as a language universal tendency, for instance, by Delattre (1962:1142), 'Non-distinctive nasalization of vowels, whenever it occurs, in any language, is often due to the following consonant, never to the preceding one', by Fromkin ad Rodman (1974:80)- 'A vowel is nasalized when it occurs before a nasal consonant..... Very common in languages of the world; it is probably a universal rule', by Ferguson (1975:181)- 'Nasality may spread either regressively or progressively from a nasal consonant to a neighbouring vowel, but regressive spread is more common', and by Hyman universally present in all languages... These universal tendencies are as such not part of the phonologies of individual languages but rather belonging to the realm of universal phonetics'. Thus Ferguson et al, (1975:189) state 'such factors as the difference in speech between raising and lowering the velum result in the universal tendency of regressive nasality'. Another statement by Ferguson et al, on universal tendencies and 'normal' nasality is that: 'There is a universal tendency for

nasality to spread from one segment to another, and this is particularly apparent in vowels becoming nasalized preceding a nasal consonant. Although the tendency is apparently universal. And may be explained by the articulatory phenomenon of timing the velic opening, the details vary greatly from one language to another, and the pattern may thus be phonologically specificative'.

On the contrary, nasalization in Telugu is perservative rather than anticipatory. The first vowel in */ma:ku/* 'to us' is more nasalized that the same vowel in */a:me/* 'she', and no nasalization at all for the same vowel in */ta:mu/* 'they'. The nearest nasalization is when the vowel is both preceded and followed by a nasal consonant as in */mannu/* 'earth', */maɳɭa/* 'fire'. This clearly is different from Hindi (cf. Ohala, 1975:323) where the vowel in */kan/* and */nam/* has about the same amount of nasalization but in */nak/* less than in those two. Kelkar (1968) also states that there is greater and stronger nasalization of vowel before a nasal consonant than after in Hindi. This is, as stated earlier, quite opposite from Telugu. The nasalization of vowel is thus determined differently by its languages specific phonological rules even when the environment (i.e. the surrounding sounds) remains the same, as shown for Hindi and Telugu. In Telugu, we need a rule pf nasalizing vowels generally after a nasal consonant (V→ ṽ/N-) than before the nasal consonant (ie. V→ ṽ/N-), through the latter is the most natural rule (Hyman, 1975) in phonology.

In brief, Telugu exhibits both anticipatory and perservatory nasalization of all vowels. The latter is obligatory and the former (i.e. anticipatory nasalization), is optional. Thus Telugu is different from English and Hindi. In English and Hindi, there is greater assimilatory nasalization if the nasal consonant follows rather than the nasal consonant precedes

the vowel, while in Telugu there is greater assimilatory nasalization if the nasal consonant precedes rather than follows the vowel. Given all these conditions and /or restrictions, vowel nasalization in Telugu may be considered as a sort of intersegmental assimilation, whereby the lowered velum perseverates into the articulation of the following vowel.

5. CONCLUSION

The phonological contrast in both the languages, Hindi and Telugu, in the vowel pairs is not in terms of one single distinguished feature, but more than one, (i.e. quality, quantity and nasalization). Some pairs of vowel phonemes are phonetically characterised by difference of both relative duration and of quality. Since both duration and quality are involved, either or both of the phonetic difference can be used as the basis for describing or choosing an appropriate phonemic symbol. But, given this, there must be a consistency in the representation of facts for the same language and also across the languages. The two languages, Hindi and Telugu, differ further in the degree of difference in quality and quantity depending upon the opposition used in each language. The difference in quality is greater in Hindi than in Telugu and the difference in quantity is greater in Telugu than in Hindi.

Since the length difference in vowel is always accompanied by an equally noticeable quality difference in Telugu and Hindi, we may ask the question as to which of the two – quantity and quality – the listeners might respond to. This needs testing using tape cutting and tape splicing or perceptual (synthesis) techniques to determine the importance (or primacy) of the one over the other. I presume that the native speakers of Telugu would react more to duration than to phonetic quality as in Thai since the duration is considerably large in Telugu. On the other hand, in Hindi or Punjabi and in some dialects of English where the length difference is less

pronounced than the quantity, quality may be one of primary importance. Research into these details across languages and in individual languages is essential for the advancement of speech Technology (Fant, 1993).

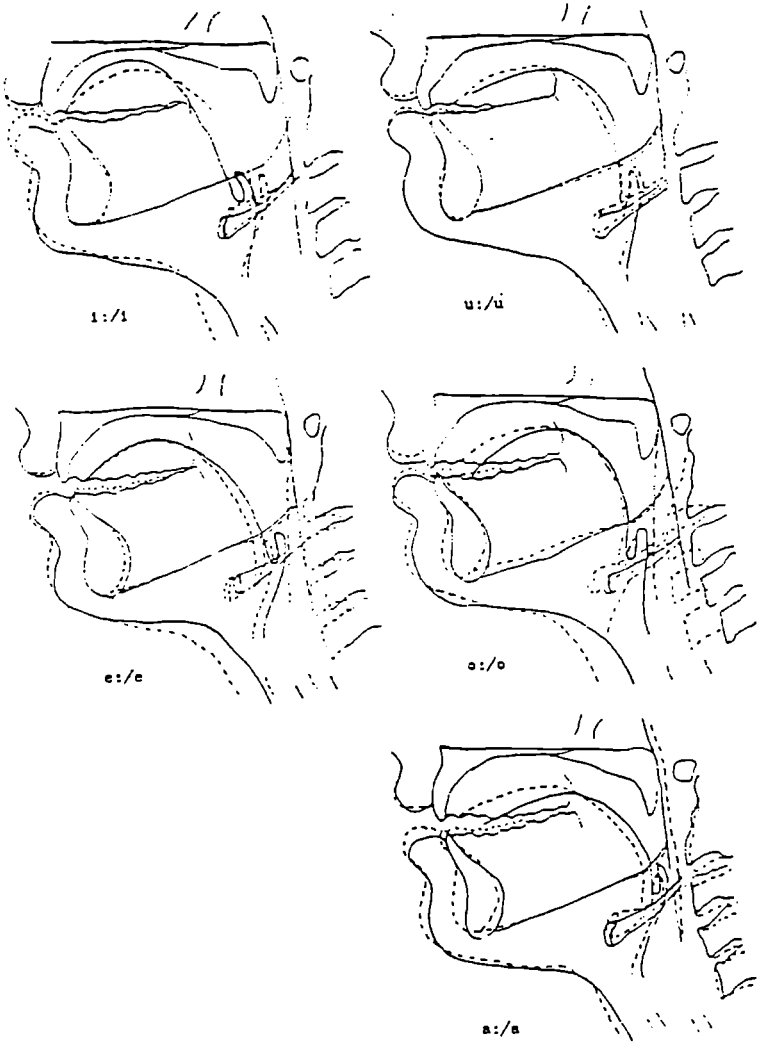


Fig. 1. Superimposed tracings of x-rays of short and long vowels.

– – – – short vowels
_____ long vowels

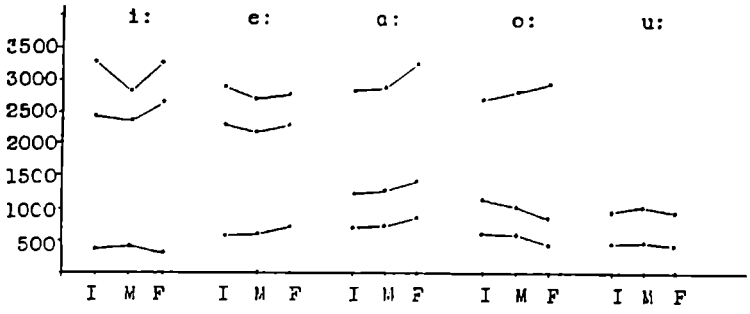


Fig 2 (a). Average formant positions for word-initial, medial and final long vowels

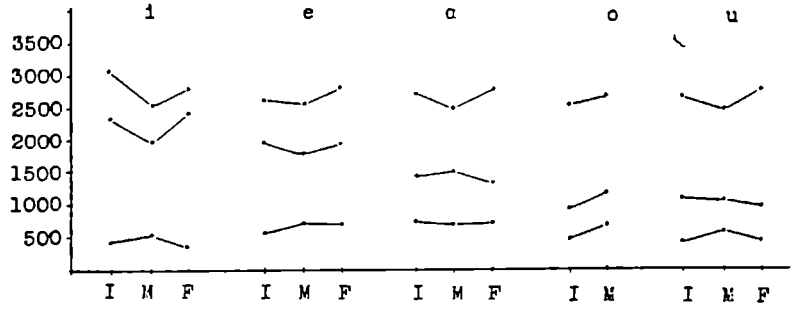


Fig 2 (b). Average formant positions for word-initial, medial and final short vowels

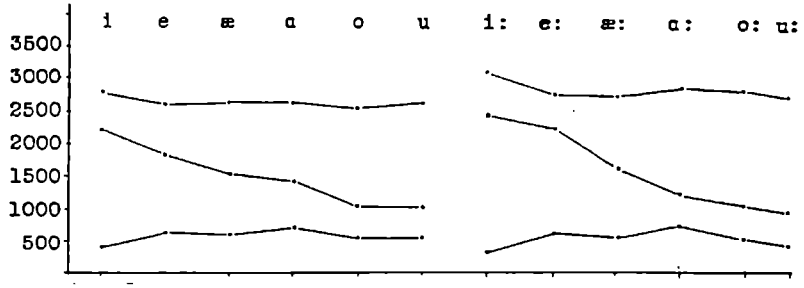


Fig 2 (c). Average formant positions for short and long vowels

REFERENCES

Abramson, A.S. 1962. *The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments*. Bloomington: Indiana University.

Anonymous Author, 1918. *A Manual of Telugu Grammar and Phonetics*. Madras: The Christian Literature Society of India.

Crothers, John. 1978. Typology and Universals of Vowel Systems. In Joseph, H. Greenberg(ed.), *Universals of Human Languages. Vol. 2.: Phonology*. Stanford: Stanford University press.

Delattre, P. 1962. Some Factors of Vowel Duration and their Crosslinguistic Validity. *JASA* 34. 1141-1143.

Dixit, R.P.1963. The Segmental Phonemes of Contemporary Hindi. University of Texas Master's thesis.

Fant, G. 1993 *The role of speech Research in the Advance of Speech Technology. Speech Technology for Man-Machine interaction*. New-Delhi:Tata MaGraw- Hill Publishing Company Limited.13-19.

Ferguson, C.A., Hyman, L.M.and J.J.Ohala (eds.). 1975. *Nasal Fest*. Stanford: Stanford Language Universals Project.

Fromkin, V.A. and R. Rodman. 1974. *An Introduction to Language*. New York: Holt, Rinehart and Winston.

Ganesan, M., Agarwal, S.S., Ansari and Pavate, K.D. 1985. Recognition of Speakers Based on Acoustic Parameters of Vowel Sounds. *Journal of the Acoustical Society of India, Vol. XIII, No. 4*. 202-212.

Gimson, A.C. 1970. *An Introduction to the Pronunciation of English*. Bristol: J.W. Arrowsmith Ltd.

Hyman, L.M. 1975. *Phonology: Theory and Analysis*. New York: Holt, Rinehart and Winston.

Jones, D. 1918. *An Outline of English Phonetics*. Cambridge: Heffer.

Jones,D. 1967. *The Phoneme: Its Nature and Use*. Cambridge: W. Heffer&Sons.

Kelkar, A.R. 1968. *Studies in Hindi-Urdu : Introduction and Word Phonology*. Pune: Deccan College.

Kelley, Gerald. 1959. Telugu Vowel Phonemes. *Indian Linguistics. 19* : 146-158.

.....1969. Telugu. In *Current Trends in Linguistics, Vol. 5* . 382-393. The Hague: Mouton.

Kostic, D.Mitter, A. and Krishnamurti, Bh. 1977. *An Outline of Telugu Phonetics*. Calcutta: Indian Statistical Institute.

Kostic, D.Mitter, A. and Rastogi, K.G. 1976. *A Short Outline of Hindi Phonetics*. Calcutta: Indian Statistical Institute.

Krishnamurti, Bh. 1961. *Telugu Verbal Bases*. Berkeley: University of California Press.

.....1962. A Survey of Telugu Dialect Vocabulary. In *A Telugu Dialect Dictionary of Occupational Vocabularies* : *Agriculture*. Hyderbad: A.P. Sahitya Akademi.1-93.

.....and Gwynn, J.P.L. 1985. *A Grammar of Modern Telugu*. New Delhi: Oxford University Press.

Laver, J. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

Lehiste, Ilse, 1970. *Suprasegmentals*. Cambridge, Mass : MIT Press.

Maddieson, I. and Gandour. 1977. Vowel Length before Aspirated Consonants. *Indian Linguistics. 38.* 6-11.

Masica, Colin P. 1991. *The Indo-Aryan Languages.* Cambridge: Cambridge University Press.

Miltner, V. 1969. Hindi. In *Current Trends in Linguistics, Vol. 5.* 55-84.The Hague: Mouton.

Misra, B.G. 1967. *Historical Phonology of Modern Standard Hindi.* Cornell University, Ph.D. dissertation.

Nagamma Reddy, K. 1981. Telugu Consonants and Vowels : An Instrumental Study. University of Edinburgh, Ph. D. Thesis.

.....1986. An Instrumental Phonetic Study of Long and Short Vowels in Telugu. In the *Proceedings of the XIII All India Conference of Linguists*, A.R. Kelkar and P. Dasgupta(eds). Pune: Linguistic Society of India. 127-138 plus 10 Figures.

.....1988. The Duration of Telugu Speech Sounds: An Acoustic Study. *Journal of Institution of Telecommunication Engineers. Vol. 34*, No. 1: 36-57. New Delhi: Special Issue of JIETE on Speech Signal Processing.

.....1991. Nasalization in Telugu: An Electro-Kymographic Study. In B. Lakshmi Bai and B. Ramakrishna Reddy (eds), *Studies in Dravidian and General Linguistics: In honour of Bh. Krishnamurti.* 141-166 + 166a and b Hyderabad: CASL, Osmania University.

.....1992. Phonetics : Typical Sounds of Indian Languages. *Souvenir.* Hyderabad: CIEFL in collabration with Osmania University and Telugu University.

.....and Kanakavalli, A. 1994. Perspectives on Nasal Consonants in Indian Languages. In S.S. Agrawal and

Subhash Pani(eds.), *Information Technology: Applications in Languages, Script & Speech.* 181-192. New Delhi: BPB Publication.

Narang, G.C. and Becker, D.A. 1971. Aspiration and Nasalization in the Generative phonology of Hindi-Urdu. *Language 47.* 646-667.

Ohala, M. 1975. Phonetic Explanation for Nasal Sound Patterns. In C.A. Ferguson, L.M. Hyman and J.J. Ohala(eds.), *Nasal Fest.* Stanford Language Universals Project. 289-316.

.....1983. *Aspects of Hindi Phonology.* Delhi: Motilal Banarasidass.

Pathak, R.S. 1976. Nasalization in Bengali: An Electrokymographic Study. *Indian Linguistics*, 37. 312-315.

Velayudhan, S. and J.M. Howie. 1974. Acoustical Measurements of Distinctive Vowel Quantity in Malayalam. *Language and Speech* 17-95-101.

OPiL 24: 71-84 (1998)

IDENTITY AND COMMUNICATION IN A PLURILINGUAL MILIEU: ROLE OF THE STATE*

Lachman M. Khubchandani
Centre for Communication Studies, Pune

ABSTRACT

Indian polity is marked by plurality of cultures and languages in one space. There are more than 200 languages which belong to different families showing parallel trends of development and characterizing a single composite region in a plural society. In this paper I argue that if this multiplicity of languages is handled with proper sensitivity; can lead to cohesion and avoid friction during the post-Independence period of language politics.

Indian polity is marked by plurality of cultures and languages in one space. This experience provides a unique mosaic of verbal communications which has withstood the test of time over the centuries. Indian census records 200 and odd languages which belong to different families -- Indo-Aryan, Dravidian, Austro-Asiatic and Tibeto-Burman -- which show parallel trends of development during a long history and characterize a single **composite** region.

1. DEFINING PLURAL SOCIETY

In spite of a wide spectrum of linguistic and cultural variation in everyday life, Indian masses through sustained interaction and common legacies have, by and large, developed a common way of interpreting, of sharing experience, of thinking -- a sort of 'communication ethos'. Many speech groups in the region associate diversity of speech (whether

* A revised version of the paper was presented at the Indian Social Science Institute, New Delhi in February 1997. The author acknowledges the assistance given by the Sindhu Documentation Centre, Pune in preparing the manuscript.

treated as languages, vernaculars, dialects/varieties, *lingua francas*) around it with differential values in social interaction. These considerations provide a strong evidence for treating India as a plural society, where boundaries of neither socio-cultural nor socio-linguistic traits are clearly marked.

Individuals in a plural society belong to different identity groups, clustered around cultural, linguistic and social traits (such as nationality, religion, caste language/ dialect) and share only a *core* of experience, criss-crossing in more than one manner, hardly co-terminating within the same boundary. Each of these differences may be important in that it would operate to distinguish one group from another in all traits. 'Individuals joined by a single trait are generally marked by their *variety*, their *lack of unity*, and their tendency to act as fairly *discrete groups* relative to the pulls and pressures of time and space' (Khubchandani, 1983 : 6).

In the context of culture, the traditional Indian concept of *kshetra* (approximate translation, *'region'*) covers a wide spectrum of linguistic and cultural variation in everyday performance. It helps to foster the feeling of oneness among diverse peoples in the region, creating in them 'a sense of collective reality'. This concept is markedly different from the modern western model of *region* defined as a 'cohesive and homogeneous area', created by *arbitrary* selection of transient features such as religion, language, tradition (Saraswati, 1988; Khubchandani, 1991). In this paradigm, the distinction between the categories such as majority/ minority communities and developed/ developing languages need to be critically evaluated.

2. ORGANIC PLURALISM

Pluralism on the Indian subcontinent is characterized by a stratificational network of primary groups governed by a varying degree of boundedness signifying lineage, language,

occupation, and religion (Khubchandani, 1995a). Different identity groups are involved in a complex web of relationships with one another, presenting a kind of mosaic, and are averse to their being rigidly identified with a particular 'insulated' group. Diverse groups, thus related to as an integral part of the whole under the label 'we', can be characterized as :

(1x1x1 = 1): multiplication (x) signifying an *integral* relation.

This phenomenon is identified as a case of *organic* pluralism, where multiple identities are strengthened by a measure of fluidity in their manifestation. Indian heterogeneity in speech, marked by implicit 'etiquette' and flexibility, can best be viewed within an overall 'organic unity' of communications (Chatterji, 1943). This proposition questions the assumption that many languages in one space (i.e., city, state, nation) pose a communication problem, and leads us to consider that *variation* in speech could be a significant contributing factor to the richness in verbal and non-verbal skills.

In contrast, different identity groups, when combined under the umbrella of a common structure sharing the same space and/or same interests and are proportionately balanced in a structural whole, characterize the label 'we' as :

(1+1+1=3): addition + signifying a *combined* relation.

This phenomenon has been refered as *structural* pluralism. In such a society harmony among diverse primary groups is sought by containing their rival aspirations through safeguards provided within the parameters of equality and social justice (Gordon, 1981). Pluralism in many contemporary societies is generally based on the co-existence of different primary groups structurally separated by ethnic/nationality boundaries insulated through traits such as colour, religion, and language territory (in the case of migrants, their ancestral languages).

The two models of pluralism -- organic and structural-- are sharply distinguished by their relation to the whole :

a) *Integral* relation, where diverse primary groups form an integral part of the 'organic' whole.

b) *combined* relation, where diverse primary groups are proportionately balanced in a 'structural' whole.

One further notices two major cross-currents characterize both models of pluralism, one favouring conditions for *homogenization* and the other promoting the process of *differentiation.* These cross-currents affect, in a significant manner, directions in the maintenance or shift of socio-cultural diversity.

Organic pluralism in South Asia is, by and large, supported by differentiating characteristics of heterogeneity, federality, and so on. This phenomenon is unlike the homogenizing traits as found in the 'melting pot' pluralism (such as in the United States); it accepts variations within universal ideals. Under structural pluralism, the liberal pluralism (as in the Scandinavian countries) favour conditions for *voluntary* homogenization where diversity is subtly tempered with individual preferences and individual rights. Corporate pluralism, on the other hand, contributes to the accentuation of socio-cultural identities through mandatory safeguards for 'group' rights, as typified by the erstwhile Soviet Union, Switzerland, Belgium, and Canada. Recent trends in India, and the United States point to the processes favouring corporate pluralism.

This schema gives us a *Plurality Square*, where the distinction between organic and structural is regarded as of primary order, and the distinction favouring homogenization versus differentiation is treated as of secondary order:

Plurality Square

	Homogenizing	Differentiating
Organic	Melting pot	Stratificational
Structural	Liberal	Corporate

Diverse profiles of speech communication in different countries and at different times make us realize the futility of pursuing illusionary goals of universal order in the name of 'efficient' communication. In this framework, India, as a sociolinguistic area, is not a collection of fragments which the State holds together, but it presents a series of mosaics--religious, linguistic, regional and covering other socio-cultural dimensions -- which fit together in a whole as in a jigsaw puzzle, and no single constituent, however small numerically, is marginalized. Identities of Panjabi, Hindi, Urdu, and Sindhi speakers in both India and Pakistan regions provide a vivid example of the phenomenon. Inter-language boundaries in many of these regions have remained fuzzy and fluid (Khubchandani, 1979, 1991). Plural speech communities tend to organize their repertoires through diglossic patterning, grassroots 'folk' multilingualism (distinct form 'elegant' bilingualism or trilingualism learned through conscious effort), code-switching, pidginization, and other such processes of language contact.

Indian pluralism is being viewed as a 'centrifugal' devise by which different groups attempt to retain and preserve their unique cultural attributes while developing common institutional participation at the national level (Schermerhorn, 1970). This trait can also be regarded as a good example of 'ethnic arithmetic' where particularist loyalties are allowed to be nourished without eliminating the subordinate cultures (Weiner, 1972).

3. CONSTRAINTS

The role of State and of language-elite tends to make language identity, hitherto a cultural trait, more political. This has led to a shift in language identity, an upsurge away from a low-key *instrumental* role in a framework of stratificational pluralism, to a top-gear *defining* characteristic in the new emerging order of pluralism. We see the signs of India turning away from an organically "accommodating' *plurilingual* nation into an institutionally 'assertive' *multilingual* nation.

Lately Indian academia has been groping with the issues of cultural identities and linguistic plurality in communication networks of the federal polity. On the one hand, many social and political thinkers have been paying glorious tributes to India's colourful diversity in search of the post-modern paradigm of socio-cultural development; on the other, a large body of them view such diversity as pre-modern, inefficient, lending to political instability and technological backwardness. Under the influences of purists' in philology and pedagogy, agencies concerned with social planning have largely been treating linguistic heterogeneity as a serious problem of human adjustment. In their dilemma, many linguists seem to have not yet come to terms with fully grasping the *composite* and integral character of plurilingual ethos in many Oriental and African Societies.

In this regard, Indian scholarship seems to suffer from another serious constraint, namely that of the lack of the cognitive level of defining the plurilingual situation relevant to, Indian reality, that is, many of the transformations sought through modernization in these societies (such as targets of literacy, of language development) are virtually drawn from Western experiences. Consequently the changes sought are *externally* induced, rather than *internally* generated; unlike classical European modernization processes which were

stimulated during the Renaissance and the Reformation. Hence most language development programmes in the developing world are characterized by Western norms and values as a further intensification of modernization based upon both methods and substance overtly borrowed from successful models.

4. LANGUAGE POLICY AND LANGUAGE PLANNING

In formal communication, literate societies generally treat language as an autonomous system with its distinct history and tradition. It has either acquired or invented a writing system: it is associated with a distinct personality, a language genius inherited through tradition and literary creations; it is evaluated through the conventions as standardized in its spelling rules and other prescriptive devices of correction in grammatical and lexical usage along with the 'etiquette' rules of choosing styles and idioms from its heritage. In short, language is regarded as an autonomous 'institution'.

This notion is reflected in the unending debates over nation's language policy. These debates assumed that each language is a monolithic 'crystallized whole', and claim equal privileges for different languages in education, administration, etc. Various pressure groups aspire for parallel development of their 'pet' languages -- be it mother tongue, national language, or international language -- as a vehicle of creativity and thinking, or for their use in technology and other needs of the contemporary society. These claims have resulted in dichotomies such as developed 'rich' languages versus undeveloped 'poor' languages, or majority/minority languages.

At the same time, speech activity as an ongoing process, responding to a variety of communication settings and the labels such as Hindi, Urdu, Panjabi of the Fluid Zone, are as much products of *environment* as of *tradition.*

When probing into diverse facets of human communication, we can identify at least three distinct functions of language;

1) as an expression of creativity in literature, an artefact, transcending to aesthetic heights as in visual arts, music and dance.

2) as a vehicle of cultural identity, fulfilling the human urge of gratification, of belongingness to a particular heritage; in the context of insecurity, such affinity to language could find expression in emotive terms, as in the case of one's religious loyalty.

3) as a medium of communication in everyday life, in the realms of education, media, administration, judiciary, and other occupational needs.

It is, therefore, necessary to break the *monolith* of promoting and developing language *X* or *Y* in its entirety when determining its functions in interpersonal, societal, aesthetic and political domains. As an illustration the language needs for literacy, for general education, for occupational skills, and for specialization (such as in scientific enquiry) may be at variance in a plurilingual society, and may require different strategies to cope with the challenge. In plural societies different dimensions interflow simultaneously into one another, as in an osmosis.

Reflecting over the experiences of five decades on the language issue, today we are in a better position to appreciate different strands of language in a plural society -- its cultural, societal and political dimensions. Following Independence, language consciousness has grown in the country and loyalties based on socio-cultural identities have acquired political salience such as:

a) There has been an upsurge of 'language movements' which become a vehicle of different interest groups to assert their cultural heritage through maintaining, upgrading, or even mythologizing language. How can this upsurge be channelled in a positive direction which promotes cohesion in plurilingual societies, instead of generating negative tendencies of alienation, fragmentation, domination, discrimination, exploitation, friction, mistrust, and so on?

b) Given the mosaic of conflicting elites in the country, there is no dearth of language solutions. Instead of discarding or inducting different languages for specific domains (in education, administration, judiciary, etc.), a system of proper checks and balances need to be developed by which *malignant* functions associated with certain languages are curbed (such as the imperial domination of English, trans -- created *here* but ' evaluated' *there!* and the *constructive* role of others get proper incentives (such as, the communicative assets of lingua franca Hindustani, spread in the entire South Asian region and beyond in the Gulf, Afghanistan, Mauritius, and in the countries touching the rim of Indian Ocean).

Language plays a crucial role in the reorganization of institutions and it is inseperable from such activities as planning, propaganda, and evaluation. Every culture/ language, irrespective of being 'big' or 'small' or being treated as 'strong' or 'weak', serves as a *bridge* between others and as an instrument of interaction which is humanly universal. The overall guiding spirit should be to serve language environment which makes sense and meets with the demands of social justice (Khubchandani, 1995a; 1997).

4. ROLE OF THE STATE

In a democratic society primordial identities such as religion, language, tradition need not be allowed to intrude unduely into the domains of public life. Nevertheless, the role of the State in the contemporary set up has a vital *instrumental* role to play in facilitating and coordinating the aspirations of primordial interest groups through diasporic bodies *(bira:dari:)* and professional agencies), such as the academic pursuits of developing language, literary promotion, preserving and vitalizing folk heritage, religious freedom, selection processes and creativity in communication, and so on (Khubchandani, 1991).

The state is further obliged to deliberate on the matters concerning policies and implementation in the spheres of education and communication (e.g. Three-Language Formula), press and mass media, tourism, etc. which touch upon primordial sensitivities in a significant manner. At times, the State is even required to play judicial and policing role to remove imbalances and mitigate hardships, concerning any particular primordial activity and in emergency situations when such conflicts could disturb peace and endanger national security.

Inter-cultural and inter-lingual communications, as a means of harmonizing primordial identities are indeed of great significance in promoting respect for cultural and linguistic diversity and injecting vitality into the great nation. In the contemporary world one generally notices that the relationship between small and big cultural entities are not organized on the basis of *reciprocity*. The 'big' cultures all round the world, supported by enormous political and economic advantages, tend to pursue one-sided policies which radically go against the basic intentions of a fair communication. We have yet to come to grip in formulating a coherent policy for knitting complex pluralities and making available equal opportunities

to different sections of societies which can lead to an integral cultural development. The litmus test of national integration in the Indian context will be creating the climate to the *sublimation* of cultural identities, rather than submitting to the aggressive assertion of narrow identities.

If we want to guard against the sweeping currents of fundamentalist ideologies, the State has to exercise utmost care when it comes to *defining* and *regulating* primordial identities in the name of socio-cultural development. State's attempts to resolving primordial issues in an adhoc and arbitrary manner often lead to aggrevating the sense of deprivation and to the growth of insular tendencies among different cultural groups. Setting up of a permanent autonomous commission on language and communication, on the lines similar to the French and Scandinavian language academies, can provide a useful forum for the intellectual articulation of the issues, and guide the Executive and the Parliament in framing the policies with a long term perspective. This can have a significant bearing on the issues concerning the *quality* of human communication in a changing society.

5. CONCLUSION

It is necessary to bring a pluralist vigour in making policies for fair play in communications though the flowering of cultural diversity (as environmentalists' respect for bio-diversity). It will require a substantive shift in the concerns of social scientists to take seriously the fuzzy reality and transactive domains of language(s) as a 'live force' in the contemporary milieu, recognizing the fact that language remains in perpetual flux along with the usage just as the reality keeps changing. Language boundaries in plurilingual settings merely serve as markers to construct the fuzzy reality, as manifested in a verbal repertoire through a variety of 'blending' processes such as relexification, code-switching, encouraging the use of lingua franca, and so on.

Such a breakthrough is possible by stretching autonomy in defining time-and-space-bound reality as a manner of conviviality so that Oriental societies are not reduced to mere objects to be studied in terms of Western concepts and categories (which are treated not as culture-bound but universals).

Recent debates over human rights in international forums poignantly reflect an awareness of this malady, There is a pervading tendency in the Third World to apply externally induced 'universal' models to explain contemporary Indian phenomena to the West in its own idiom. In this process pertinent enquiries about plurilingual societies (such as the fuzzy demarcation of language boundaries in the case of Marathi and Konkani, of Panjabi and Dogri; issues of identity versus communication in defining languages as in Hindi, Urdu and Panjabi; grassroots and elitist tensions in prescribing a 'representative' standard as in Indian Angrezi) do not merit serious attention.

Our conceptualization of the prevailing diversity in communication in everyday life needs to be treated as a 'humane enterprise' with somewhat lighter overtones. Among the tribals in India, a plurilingual repertoire rests rather lightly on their shoulders (Khubchandani, 1992). They do not get distracted by grave intellectual stands, often rigid, taken by the academia regarding the psychological and sociological theories of language acquisition and language contact.

In the realm of social planning, intellectuals very often lose track by weighing grains (such as speech in everyday life transactions) on the 'precision' scale used for wieghing gold (such as language in a 'sophistic' discourse). The Indian experience tells us that a genuine understanding of plural societies will largely be guided by viewing language as a synergic network inspiring *trust* in cross-cultural settings, along with the complementarity of empowering the 'particular'.

The multiplicity of languages in a plural society, if handled with proper sensitivity, can lead to cohesion, instead of generating friction as has been the bane of language politics in the country during the post-Independence period.

REFERENCES

Chatterji, Suniti Kumar. 1943. *Language and the Linguistic Problem.* London: Oxford University Press.

Gordon, Milton M.1981. 'Models of pluralism'. *Annals of the American Academy of Political and Social Science.* 454 (March), 178-88.

Khubchandani, Lachman M. 1979. Demographic typology for Hindi, Urdu, Panjabi speakers in South Asia. In W.C.McCormack and S.A.Wurm (eds.), *Language and Society: Anthropological Issues.* The Hague: Mouton. 183-94.

.....1993. *Plural Languages, Plural Cultures : Communication, Identity and Sociopolitical Change in Contemporary India.* Honolulu: University of Hawaii Press. East-West centre.

.....1996. Language communication filtered through technology. *Sociolinguistics. Special issue : Orewall and the Language Question. Vol. XVI: 2. Rotterdam.*

.....1991. *Language, Culture and Nation-building : Challenges of Modernisation.* Shimla: Indian Institute of Advanced Study.

.....1992a. *Tribal Identity: A Language and Communication Perspective.* Shimla: Indian Institute of Advanced Study.

.....1992b. Exercising " Right To Communicate": Interpretation of News. In A. Dutta and M.M. Agarwal

(eds.), *The Quality of Life*. Shimla: Indian Insitute of Advanced Study.

.....1995a. 'Minority cultures' and their communication Rights. In T.Skutnabb-Kangas and R. Phillipson (eds.), *Human Rights: Overcoming Linguistic Discrimination.* Berlin: Mouton de Gruyter.

.....1995b. India : A case of organic pluralism. In S.Leger (ed.), *Towards A Language Agenda: Futurist Outlook on the United Nations*. Ottawa: Canadian Centre for Linguistic Rights, University of Ottawa.

.....1997 *Revisualizing Boundaries : A Plurilingual Ethos.* New Delhi: Sage Publications.

.....1999 Philosohical issues of contact language planning: Response to the prevailing tendencies of linguistic hegemonism. In A.F. Christidis (ed.), *'Strong' and 'Weak' Languages in the European Union: Aspects of linguistic Hegemonism. Vol 1 Proceedings of an International Conference (March 1997).* Centre for the Greek Language. Thessaloniki. 43-57.

Saraswati, Baidya Nath1988. 'Thinking about "region". In S.B. Chakrabarti (ed.), *Social Science and Social Concern.* Felicitation Volume in Honour of Prof B.K. Roy Burman. Delhi: Mittal Publications. 410-14

Schermerhorn, R.A. 1970. *Comparative Ethnic Relations.* New York: Random.

Weiner, Myron. 1972. Political integration and political development. In P.F. Techua (ed.), *Developing Nations: What Path to Modernization*? New York : Dedd, Mead & Co.

OPiL 24: 85-101 (1998)

CLEFTS IN MALAYALAM : A FOCUSSED MOVEMENT APPROACH

K. Srikumar
The M.S. University of Baroda, Vadodara.

ABSTRACT

This paper presents an account of Clefts in Malayalam from a Focussed Movement perspective. The proposed analysis for clefts also explains the derivation of a hitherto unnoticed construction in Malayalam. Further the difference in wh-clefting between Malayalam and other languages is shown to follow from a difference in the relative strengths of the verb *be*, in conjuction with the other properties of these languages.

Adopting essentially an idea from Hoh and Chiang (1990), in Srikumar (1992, and 1994), I proposed a Focussed Movement approach for deriving clefts in Malayalam. Besides considering this proposal, in the present paper I shall attempt to recast the same in terms of the Minimalist assumptions (Chomsky, 1992). The Focussed Movement approach we adopt, largely conforms to the movement - based analyses for clefts arrived at in a number of recent works (cf. Horvath, 1986; Rochemont , 1986; Kiss, 1987; Ortiz de Urbina, 1989; and Brody, 1990).

The organization of the paper is as follows. In Section 1 I shall dwell briefly on clefts in Malayalam with a view to highlight the reasons why I adopt the Focussed Movement approach for it. Section 2 details the Focussed Movement approach as proposed in Srikumar (1992). In section 3, I will try to look at the Focussed Movement approach from the point of view of Minimalist Theory (Chomsky, 1992). And in section 4, I shall draw attention to a parametric variation that is probably at work distinguishing clefting in languages resorting to focussed movement from those that do not do so.

1. INTRODUCTION

The sentences in (1) illustrate that the, clefts in Malayalam have two sub-parts: one, the *at*ə-ending clause which is the cleft clause and two, the clefted phrase along with the copula *aaɳə* 'be'.[1]

1a. *raaman vaayicca-tə [takaZiyuʈe pustakam aaɳə]*
read nomn. Thakazhi's book be
'It is Thakazhi's book that Raman read'

b. *kuʈʈiye Sakaaricca-tə [raaman aaɳə]*
child acc. scolded nomn. be
'It is Raman who scolded the child'

On analogy with the apparently similar pseudo-cleft construction in Malayalam (cf. 2), Madhavan (1987) proposed a bipartite structure for the clefts in (1), comprising the sentential subject and the VP constituted by the clefted phrase and the copula *aaNə*.[2]

2. *[[paalattinə tiivecca-var] [tiivravaadika[aaɳə]]*
bridge dat . fire set hum.pl. extremists be
'Those who set fire to the bridge are extremists'

This hypothesis was however shown to present an empirical difficulty in Srikumar (1992), given the possibility that clefts in Malayalam permit certain scrambled variants such as the following.

3 a. [$_{IP}$ [$_{CP}$ *raaman vaayiccatə*] [$_{VP}$ [$_{NP}$ *takaZiyuʈe pustakam*] [*vaaɳə*]]]

b.[$_{IP}$ t_i [$_{VP}$[$_{VP}$[$_{NP}$ takaZiyuʈe pustakam] [$_{V}$ aaɳə]]
[$_{CP}$ *raaman vaayiccatə*]$_i$]]

c.[$_{IP}$ *raaman*$_k$[$_{IP}$ t_i [$_{VP}$[$_{VP}$[$_{NP}$ *takaZiyuʈe pustakam*] [$_V$ *aaɳə*]] [$_{CP}$ t_k *vaayiccatə*]$_i$]]]

To state briefly, the analysis recommended in Madhavan obtained the variant (3 b) by postposing the sentential subject onto the VP, and then , the variant (3 c) by a further extraction from the postposed subject which left adjoins to the IP. But given the assumption that V must raise to INFL [Jayseelan, 1989; and also Madhavan, 1987), the S-structure so formed would have the postposed sentential subject intruding between the clefted phrase and the copula aa*ɳə*, which is however not a legitimate surface output (cf. 4 a and b)[3].

4 a. *[$_{IP}$ t_j [$_{I'}$ [$_{VP}$[$_{VP}$[$_{NP}$*takaZiyuʈe pustakam*] [$_V$ t_i]]

[$_{CP}$ *raaman vaayiccatə*] $_j$] [$_I$ *aaɳə*$_i$]]]

b. *[$_{IP}$ *raaman*$_k$ [$_{IP}$ t_j [$_{I'}$ [$_{VP}$[$_{VP}$[$_{NP}$ *takaZiyuʈe pustakam*]

[$_V$ t_i]] [$_{CP}$ t_k *vaayiccatə*]$_j$] [$_I$ *aaɳə* $_i$]]]]

To do away with the undesirable consequence thus, Srikumar(1992) argues for a different analysis for clefts which is termed there as the Focussed Movement approach. This analysis crucially capitalises on the fact that there exists a construction in Malayalam which has constituting elements identical to those in clefts but no clefting. Moreover akin to clefting, the proposition which precedes the verb *be* element *aaɳə* becomes focussed. Therefore it was given the name Propositional Focus Construction(PFC)[4]

To illustrate the parallel in constituent elements in a PFC and its corresponding variant with a clefted constituent, one may consider the contrast exhibited by the sentences in (5).

5a. *raaman takaZiyuʈe pustakam vaayiccatə aaɳə*

Thakazhi's book read nomn. FOC

'It is the case that Raman read Thakazhi's book'

b. *raaman vaayiccatə takaZiyuɻe pustakam aaɳə*

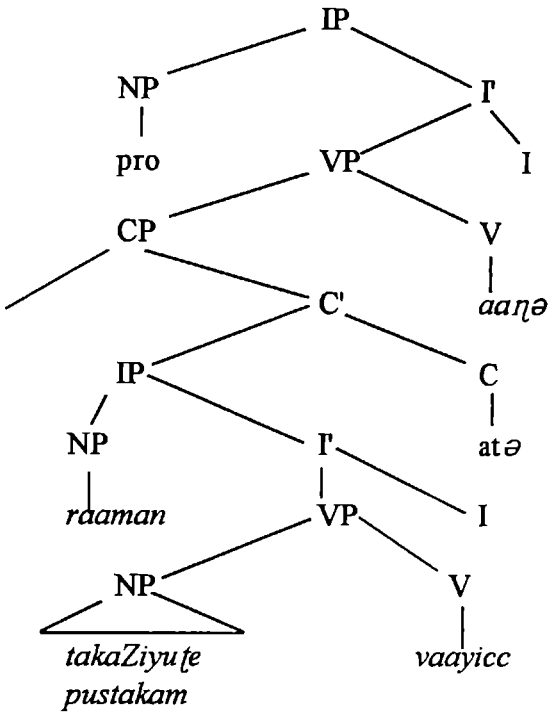
'It is Thakazhi's book that Raman read'

As noticeable from these sentences, it is clear that the PFC differs from the cleft only in that the former involves no instance of clefting. But, nevertheless, the whole proposition preceding the verb *aaɳə* is as though it were focussed in some intuitive sense. Hence the conjecture follows that clefts could be derived out of the underlying structure for PFC by movement of the constituent to be clefted to the pre-*aaɳə* position as indicated below informally:

6. *raaman t vaayiccatə [takaZiyuɻe pustakam] aaɳə.*

To obtain the above result we shall be relying upon the biclausal structure (7) for clefts which is non-distinct from the one motivated as the underlying structure for clefts in English by Rochemont (1986).[5] Here the verb *aaɳə* 'be' is treated as a monadic predicate with a clausal complement. Therefore the extraction of the clefted phrase can proceed without the intervention of any barrier on the way for the complement is L-marked by *aaɳə*.[6]

7.



The subject of the matrix clause in (7) is, we assume, an expletive *pro* which is presumably also the null subject of other constructions with the verb *be* in Malayalam such as those given below in (8) and (9).

8 a. pro *maZa aaɳə*
rain be

'It is raining'

b. *veɭiyil* *maZa* *aaɳə*
outside rain be

'It is raining outside'

9 a. pro *iruʈʈə* *aaɳə*
dark be

'It is dark'

b. *viiʈʈinə* *akattə* *iruʈʈə* *aaɳə*
house inside dark be

'It is dark inside the house'

2. FOCUSSED MOVEMENT PERSPECTIVE

As mentioned earlier, for instantiating clefts, we shall be assuming the Focussed Movement approach. This approach, as proposed by Hoh and Chiang (1990), holds that FOCUS (FOC) as an abstract syntactic feature is base-generated in some constructions in a language. When base-generated, as is the case with any other syntactic feature, FOC must also be realized on some constituent at the relevant level of derivation. This requirement was sought to follow from the FOCUS CRITERION given as below in Srikumar (1992).

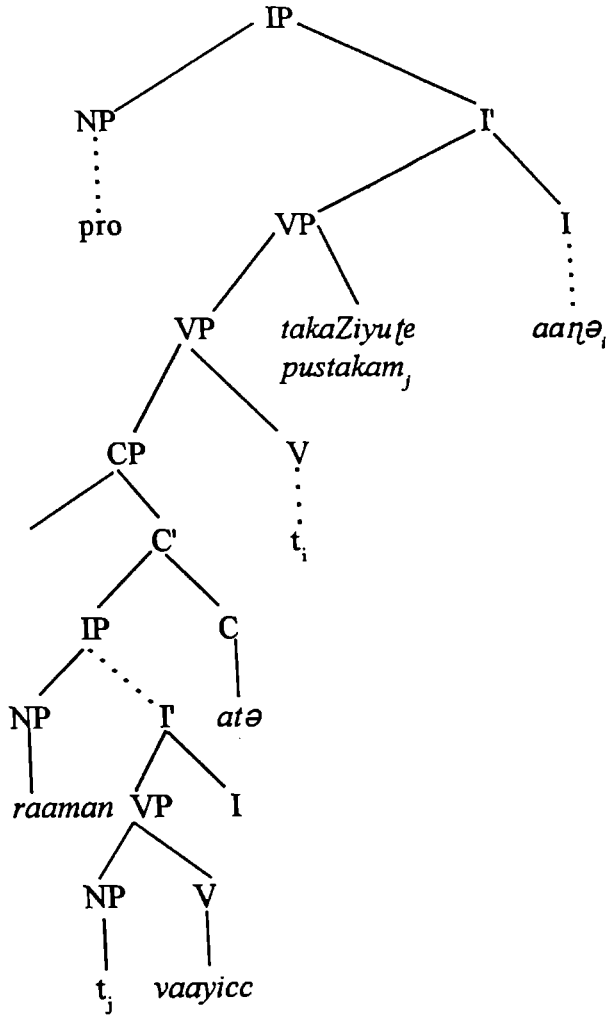
10. Focus Criterion (FC)

If FOC feature is base generated in a construction, it must be discharged to an adjacent constituent at the S-structure.

The verbs *aaɳə* and *shi*, which are the equivalents of the verb *be* in Malayalam and Chinese, are construed as the overt manifestations of the abstract syntactic feature FOC. FOC-assignment by such a verb could then be ensured if the constituent to be clefted moves out to the verb adjacent position by the S-structure so that the canonical configuration necessary for feature-assignment in the language is obtained, i.e., leftward for Malayalam, for the language is head-final, and rightward for Chinese which is head-initial. The movement thus resorted to could be subsumed under the generalized transformation Move-Alpha. Since the verb *aaɳə* in the matrix clause in the structural representation(7) has to undergo raising to INFL to collect tense (cf. Madhavan, 1987; and Jayaseelan, 1989), the pre-*aaɳə* position thereby obtained turns out to be

the right-adjoined position on the matrix VP. Thus the S-structure (suppressing irrelevant details) corresponding to the D-structure (7) of the sentence (1 a) would be as indicated in (11).

11.



Given FC (10) FOC-assignment on the clefted phrase would follow as predicted. However consider the case if the movement resulting in the S-structure (11) from the D-structure (7) were not to apply, FC would still require FOC to be discharged to some constituent at the S-structure. Hence by default *aaɳə* may be assumed to assign FOC to the entire proposition preceding it whereby the structure for the PFC reading is obtained.

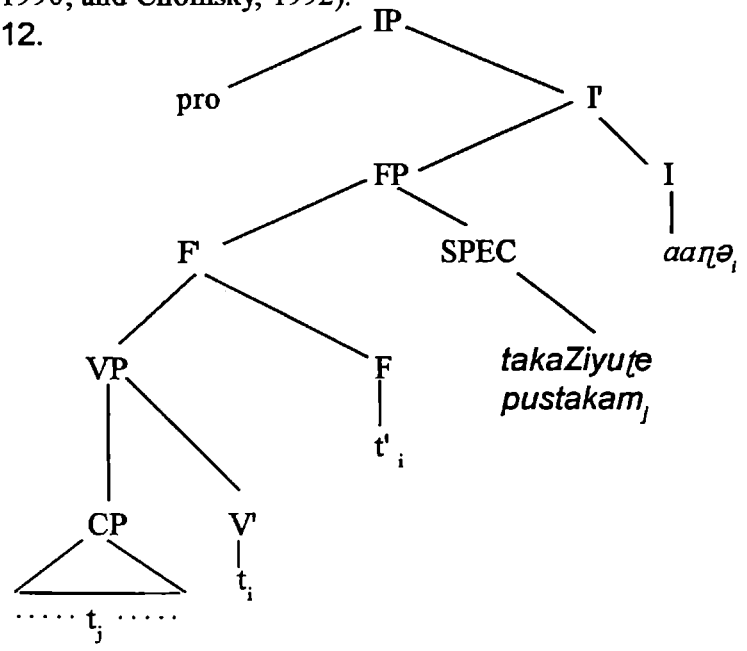
3. A MINIMALIST APPROXIMATION FOR FOCUSSED MOVEMENT.

Making a departure from my earlier account regarding Focussed Movement sketched out in section 2, let us consider how the same may be obtained under the Minimalist assumptions of Chomsky (1992). Under this approach to linguistics theory "each linguistic expression is an optimal realisation of interface conditions expressed in elementary terms ..." (cf. Chomsky, 1992: 27). Keeping within the 'domain of virtual conceptual necessity', according to Chomsky (ibid.), PF and LF 'generated in the most economical manner' are the only interface levels which need to be considered within a simple design for language. Given such an assumption, the clefted word-order arrived at above by deriving the S-structure (11) from the D-structure (7) for the sentence (1 a) must be available at the interface level PF for convergence to obtain there. So the movement necessary to produce it could be construed as a fallout of the 'principle of last resort'.

To execute the above idea, let us recall from the earlier discussion (cf. section 2) that we have assumed the verb *aaɳə* as the overt bearer of the abstract syntactic feature FOC. In most current works (cf. Mahajan, 1990; Chomsky and Lasnik, 1991 ; and Chomsky, 1992), feature-assignments like those of Case and Agreement are realized as a consequence of Spec-head agreement in the maximal phrase hosting them. In terms

of the Minimalist theory (Chomsky, 1992), this amounts to bringing phrases into the 'checking domain' of the head for feature matching , i.e., the specifier or an adjoined position within a maximal projection of the head. To adopt this shift, let us make two assumptions : one, assume the VP-internal subject hypothesis (cf. Kitagawa, 1986; Fakui and Speas, 1986; and Speas, 1990) whereby all the arguments of a predicate are generated within its own projections; and two, following Jayaseelan (1989), consider the possibility of FOCUS hosting a maximal projection FOCUS PHRASE between the IP and VP.[7] The spec of the FP must however be stipulated to occur on the right periphery for the correct word-order to follow.[8] Although this step is out of tune, with the occurrence of the spec to the left in the general X-bar schema adopted for Malayalam, we assume this to be likely for the spec of FP may be construed to belong to the class of broadly L-related positions whose details are still controversial (cf. Mahajan, 1990; and Chomsky, 1992).

12.



Given a structure like (12) then FOC-assignment can be made to follow from Spec-head agreement in FP consequent to the raising of the clefted phrase to [Spec, FP] concomitant with the raising of *aaɳə* to F. Further, the verb *aaɳə* must also raise to the INFL in order to receive tense whereby the correct word order expected in the case of clefts follows without much ado. However, if clefting does not take place, the embedded clause in the structure (7) may itself have to undergo raising to the [Spec, FP] to effect FOC assignment by default yielding the PFC interpretation. The default mechanism may follow naturally under the Minimalist scheme because the FOC feature which is perhaps strong in Malayalam may cause the structure to crash at PF, if it were to survive until then.

The above analysis can also be extended to clefts in Chinese, provided we remove the stipulation that the [Spec, FP] branches to the right. Thus the structure corresponding to the sentence (13 a) will be as shown in (13 b).

13 a. *Shi Zhangsan Wangwu shuo Xiaoming yao le*

FOC said bite ASP

'It is Zhangsan that Wangwu said Xiaoming bit'

b. $[_{IP}$ pro $[_{I'}$ *shi*$_i$ $[_{FP}$ *Zhangsan*$_j$ $[_{F'}$ t'$_i$ $[_{VP}$ t$_i$ $[_{CP}$ *Wangwu shuo Xiaoming yao* le t$_j$]]]]]]

Further Brody (1990: 206) also reaches a similar conclusion for the well-known fact of the pre-verbal focus in Hungarian with the additional assumption though that the inflections are generated together with the verbs in Hungarian which implies that there is no separate IP in the language.[9] The structural analysis for the Hungarian sentence (15 a) would therefore be as represented in (15 b).

15 a. Janossal vittem le a szemetet

John with took-I down the rubbish

'I took down the rubbish WITH JOHN'

b. $[_{FP}\ [_{SPEC}\ Janossal_j]\ [_{F'}\ vittem_i\ [_{VP}$ le t_i a *szemetet* t_j]]]

4. A PARAMETRIC VARIATION

As noted in the foregoing discussion, clefts in Malayalam, Chinese and Hungarian are plausibly derived through Focussed Movement which is an instance of Move-Alpha. Although the derivation of clefts in English suggested in Srikumar (1992) which is an improvement upon Rochemont's (1986) analysis, is similar to my previous account for clefts in Malayalam and Chinese described in section 2, there is however a crucial respect in which the former differs from the latter. That is, unlike English, in Malayalam, Chinese and Hungarian, both questioned and non-questioned constituents when clefted behave the same way, i.e., they are moved to the [Spec, FP], as we assume is this paper. Thus the structural analyses of the clefted question-word questions in Malayalam and Chinese given in (16 a & 17 a) are as represented in (16 b & 17 b) below.

16 a. *kuʧʧi* *vaayiccatə* *aaruʦe* *pustakam* *aaɳə*

child whose book

'Whose book is it that the child read'

b. $[_{IP}$ pro $[_{I'}\ [_{FP}\ [_{F'}\ [_{VP}\ [_{CP}$ *kuʧʧi* t_j *vaayiccatə*] t_i] t'_i] *aaruʦe pustakam*$_j$] *aaɳə*$_i$]]]

17 a. *shi shei Wangwu shuo Xiaoming yao le*

FOC said bite ASP

'Who is it that Wangwu said Xiaoming bit'

b. $[_{IP}$ pro $[_{I'}\ shi_i\ [_{FP}shei_j\ [_{F'}\ t'_i\ [_{VP}\ t_i\ [_{CP}$ *Wangwu shuo Xiaoming yao le* t_j]]]]]]

In English, on the other hand, while non-questioned items cleft by moving to Spec of FP, question-words end up invariably in the [Spec, CP] as is usually the case even otherwise (cf. the contrast in (18)).

18 a. $[_{IP}$ It $[_{I'}$ is$_i$ $[_{FP}$ *John*$_j$ $[_{F'}$ t'$_i$ $[_{VP}$ t$_i$ $[_{CP}$ that you hit t$_j$ on the way]]]]]]

b. $[_{CP}$ *who*$_j$ $[_{C'}$ is$_i$ $[_{IP}$ it $[_{I'}$ t" $_i$ $[_{FP}$ t $_j$ $[_{F'}$ t' $_i$ $[_{VP}$ t$_i$ $[_{CP}$ *that you hit* t$_j$ *on the way*]]]]]]]

This distinction perhaps is due to a difference in the function of the verb *be* in Malayalam and Chinese as opposed to that in English. That is, it appears to be the case that the verb *be* functions as the overt bearer of FOC in the former, whereas in the latter it does not do so. In terms of the Minimalist assumptions, this idea may be given content if we assume, much like what Chomsky (1992), following Pollock (1989), suggests for the difference in the verb movement in English and French, namely, the feature FOC is strong in languages resorting to focussed movement, and weak in languages without it. If the feature FOC is strong, like contrastively focussed constituents, question-words also being inherently focussed need to be raised to the [Spec, FP] in order to get its FOC feature matched with the base generated FOC. For otherwise, the feature will survive to the PF where it is not a legitimate object and hence block the convergence of the resulting structure at PF. But in languages where the feature FOC is weak, arguably the feature is not visible at the PF, and therefore, does not block the convergence of the interface. However the raising of the Verb *be* to INFL via F is forced in English because the verb *be* is a semantically vacuous verb and therefore its raising cannot be 'procastinated', i.e., delayed until LF (cf, Chomsky, 1992). Hence plausibly the contrastively focussed constituents also need to be raised overtly to get its feature checking done by the overtly raised verb *be*, whereas the question-words being inherently focussed are not dependent on the weak FOC and are thus free to be governed

by the independently motivated WH-criterion (cf. May, 1985) requiring overt raising to the [Spec, CP] which the language has adopted elsewhere.

5. CONCLUSION

To conclude, in this paper, we have examined the Focussed Movement perspective for deriving clefts in Malayalam from the point of view of Minimalist program (Chomsky, 1992). For this purpose, we have assumed that in constructions with a base-generated FOC, FOC may host a maximal projection FP. Given then, the configuration necessary for feature-checking is obtainable through the raising of the constitutent to be clefted to the [Spec, FP] together with the raising of the verb *aaɳə*, the overt manifestation of FOC, to the head of FP. And finally the presence vs. absence of Focussed Movement in a language was explained to be reducible to the strength of the base-generated feature FOC.

NOTES

I am grateful to B. Ramakrishna Reddy, Probal Dasgupta, K.A. Jayaseelan, P. Madhavan, and M.Hariprasad for helpful comments and discussion on the earlier versions of this paper.

The abbreviations used in this paper are as follows:

acc.= accusative, ASP= aspect, dat.= dative,

feat.= feature, FOC= focus, hum.= human,

mas.= masculine, mat.= matrix, neut.= neuter,

nomn.= nominalizer, pl.=plural, and sg.= singular.

1. Although the *atə* particle suffixed to a verb in clefts is a nominalizer (hence, the gloss nomn. for it), the resulting clause retains its categorial status as CP. In fact, *atə* is composed of the complementizer *a* plus a feature matrix { } which may be filled in with features corresponding to the nominal the complementizer *a* a demonstrative in origin of (cf. Rajaraja Varma, 1895), modifies (whenever the latter is overtly absent.

Thus we obtain) the following pairs.

i)a. *aa* [ ] *payyan* 'that boy'

that feat. mat. boy

b. *a*[*van*] 'he'

3. mas. sg.

ii)a. *aa* [ ] *pustakam* 'that book'

book

b. *a*[*tə*] 'it'

neut. sg.

Moreover, the complementizer is bound by a constraint that its empty feature matrix cannot be left empty when it does not have an overt nominal to modify (cf. Anandan, 1985; and also Madhavan, 1987 for details).

2. Regarding the particle *avar* on the verb in the free relative in the pseudo-cleft in (2) cf. n. 1.

3. As shown in Srikumar (1992), the postposing operation and the subsequent extraction to yield the scrambled variants can also be obtained under the Focussed Movement approach.

4. As Huang (1982) notes, the placement of the copula *shi 'be'* immediately before the subject in the sentence (i) in Chinese could be interpreted as a cleft on the subject alone or as emphasis on the whole proposition following it.

i. *shi baozi shuo Meiguo zongtong bei bagpiao ie*

FOC newspaper say America President PASS kidnap ASP

a. ' It was the news paper which reported that the President of the United States has been kidnapped'.

b. 'It was the case that the news papers reported that the President of the United States has been kidnapped'.

It is the emphatic sentential reading b) above which we are referring to as PFC.

5. As the verb *be* is generally considered as a semantically vacuous element, according to Rochemont (1986, the base generation of the clefted phrase as complement of the verb *be* results in violation of the THETA - criterion (Chomsky, 1981). Thus he argues in favour of deriving clefts in English by the movement of the clefted phrase from the cleft clause to a position next to the matrix verb *be*. But given the theory of Barriers (Chomsky, 1986), if the verb *be* cannot assign a theta-role, the extraction of the clefted phrase involves movement across a barrier, i.e., the embedded CP. Hence Srikumar (1992) considers the verb *be* in clefts as a monadic predicate with a clausal complement.
6. L-marking (Chomsky, 1986 :14)

   α L-marks β iff α is a lexical category that θ -marks β.
7. In the earlier version I chose the option of locating FP between CP and IP. The superficial word-order would then be obtainable only in the manner suggested in section 2. And the FP would remain a mere artifact to ensure spec -head agreement for semantic interpretation to follow at LF.
8. In Jayaseelan (1989), the [Spec, FP] is on the left, but then to obtain the pre - *aaŋə* position, the clefted phrase was hypothesised to move subsequently to the head of the FP. This however leads to the obliteration of the trace of the raised verb in F. Hence, following Lasnik and Saito (1984), Jayaseelan assumes that the original trace of the raised verb can be gamma -marked for antecedent government by its trace in F, before getting deleted paving way for the clefted phrase to move in from the [Spec, FP].
9. Brody is considering the possibility that the absence of the VP-fronting and deletion operation may suggest the lack of the node IP rather than that of the VP as generally assumed.

REFERENCES

Anandan, K.N. 1985. Predicate nominals in English and Malayalam. Hyderabad: CIEFL. M.Litt. dissertation.

Brody, M. 1990. Some remarks on the focus field in Hungarian. *UCL working papers in Linguistics.* Vol. 2: 201-225.

Chomsky,N. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

.....1986. *Barriers.* Cambridge, MA: MIT Press.

.....1982. *A Minimalist program for Linguistic Theory. MIT Occasional Papers in Linguistics.* No.1.

.....and H. Lasnik 1991. Principles and parameters theory. In J. Jacobs, A. Von Stechow, W. Sternefeld and T. Venneman (eds.), *Syntax : an International Handbook of Contemporary Research.* Berlin: Walter de Gruyter.

Fakui, N. and P.Speas. 1986. Specifiers and projections. *MIT Working Papers in Linguistics.* Vol. 8: 128-172.

Horvath, J. 1986. *Focus in the Theory of Grammar and the Syntax of Hungarian.* Dordrecht : Foris.

Hoh, P.-S. and Wen-Yu Chiang. 1990. A FOCUS account of moved Wh-phrases in S-structure in Chinese. *Lingua 81:* 47-73.

Huang, C. -T.J. 1982. Logical Relations in Chinese and the Theory of Grammar. Cambridge, MA: MIT dissertation.

Jayaseelan, K.A. 1989. *Question-word Movement in Malayalam.* Hyderabad: CIEFL ms.

Kiss, K. 1987. *Configurationality in Hungarian.* Dordrecht: Reidel.

Kitagawa, Y. 1986. Subjects in Japanese and English. MA, Amherst: University of Amherst dissertation.

Lasnik, H, and M. Saito 1984. On the Nature of Proper Government. *Linguistic Inquiry 15*: 235-289.

Madhavan, P.1987. Clefts and Pseudo-clefts in English and Malayalam. Hyderabad: CIEFL dissertation.

Mahajan, A. K. 1990. A/A-bar distinction and movement theory. Cambridge, MA: MIT dissertation.

May, R. 1985. *Logical Form*. Cambridge: MA: MIT Press.

Ortiz de Urbina, J. 1989. *Parameters in the Grammar of Basque*. Dordrecht : Foris.

Pollock, J, -Y. 1989. Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry. 20.3*: 365-425.

Rajaraja Varma, A.R. 1895. *Kerala Paniniiyam*. Kottayam : NBS.

Rochemont, M. 1986. *FOCUS in Generative Grammar*. Amsterdam : John Benjamins.

Srikumar, K. 1992. Question-word Movement in Malayalam and GB theory. Hyderabad: Osmania University dissertation.

.....1994. Clefts in Malayalam : A Focussed Movement perspective. In Davison, A. & F.A. Smith (eds.), *Papers from the fifteenth South Asian Language Analysis Round Table Conference 1993*. South Asian Studies Program, University of Iowa, Iowa City.

OPiL 24: 103-120 (1998)

TOWARDS BUILDING A COMPUTATIONAL TELUGU LEXICON*

S. Bhoopal Reddy
Osmania University, Hyderabad

ABSTRACT

It is evident that the lexicons or electronic dictionaries form an integral component of almost every activity in computational Linguistics and Natural Language Processing (NLP). Word Processing, Automatic Abstracting and Indexing, Concordance, Lexicographers Workbench, Morphological Analysis, Text Tagging, Parsing and Generation, Machine Translation, Natural Language Interface to Data Bases, Information Retrieval, Speech Synthesis and Recognition, Psycholinguistic Studies, etc. to name a few. This paper is concerned with the computational building of Telugu Lexicon and discusses extensively the strategies and techniques in designing the computational Lexicon. This study also tries to bridge the gap between formal theoretical & practical disciplines of NLP in the field of Computational Lexicography & Lexicology with various needs in designing the lexicon.

1. INTRODUCTION

Theoretical developments within Linguistics are increasingly laid emphasis on lexical component. It is central source of semantic as well as syntactic information (Walker et al, 1995). As a result the applications of Natural Language Processing (NLP) make demands for large scale lexicons/ electronic dictionaries. To be practical they must deal with hundreds of thousands of lexical items in various detail. This paper is an attempt in building a computational Telugu lexicon. It discusses the strategies and techniques in designing it. It also endeavours to reveal the various needs in designing the

---

* An earlier version of this paper was presented at the 17th All India Seminar of DLA, held during June 17-19, 1998 at Hyderabad.

lexicon which has not been explored so far in Telugu language. The study adopts the Lexical Functional Grammar Framework (LFG) (Bresnan et al., 1982) in developing a lexicon computationally.

The lexicon which is considered as a repository of words has often proved to be a focal point in the designing of large scale natural language systems, with the constant updating by new words and changes in the meanings of existing words. This is the reason why, there has been a growing interest recently in building large-scale Lexical Knowledge Bases (LKB) automatically, utilizing various on-line resources such as Machine Readable Dictionaries (MRD) and text corpora as a starting point.

2. INFORMATION AVAILABLE IN THE DICTIONARIES

Generally dictionaries are designed for human users by humans. The human users know implicitly how the lexicon of their language is structured. Obviously, the variety of information in any dictionary depends on different factors such as coverage, user variations, specific purposes like monolingual or bilingual dictionaries, specialized areas like law, science and technology etc. But a monolingual dictionary of any language typically contains information as follows (Meigis et al, 1992.):

1. Form of head word (with spelling variants)
2. Parts of speech : NOUN, VERB, ADJ. etc.
3. Grammatical sub category lables
4. More grammatical information like countability, transitivity, reflexive and impersonal etc.
5. Inflectional variants: Number, Gender, Tense, Participles and irregular form etc.
6. Pronunciation and stress (if any)

7. Sense indications (Meaning definitions etc.)
8. Usage of the entry (example usages)
9. Derivatives and compounds
10. Relationship with other words particularly derived words, synonyms, antonyms and compounds
11. Idiomatic and other contexts in both syntactic and semantic levels

Electronic dictionaries unlike those meant for human users posses the information about lexical category.inflectional paradigm, grammatical gender etc. It requires more information in the form of semantic markers or data. It must provide words with proper equivalence in formal linguistic context particularly for the purpose of Machine Translation (MT). For example the verb entry of METAL system Lexicon (Bennette et al., 1985) contains information with regard to Lexical category (Verb Stem), Allomorph (inflected stem, etc), Inflectional class (Paradigm), Onset (Consonantal / Vocalic), Argument Structure, Valency Codes etc, Auxiliary, Voice marker (Passive/ Active) and Transitivity type etc.

The Nouns contain determiner requirements(whether determiner may / may (not) be used, Grammatical gender, sex etc.

The lexical entries for computer use are formally stated, compact and in coded form. (The codes of all grammatical features are mentioned in the respective sections). The computer based dictionaries include on-line dictionaries, LKBs, Electronic Lexicon / Electronic dictionaries, MRDs, Machine Tractable Dictionaries (MTD), etc. all of them contain more or less the same information with regard to their lexical entries, but differ in some aspects in which purpose they have been constructed.

3. STRUCTURE OF TELUGU LEXICON

As it is known, the lexicon is a list of words containing idiosyncratic information associated with individual words (Aronoff, 1982). The structure of the lexicon is represented in two heads.

- Internal Structure
- External Structure

The internal structure of lexicon contains all grammatical features of the language. It deals with two types of information both linguistic and language.

The Linguistic information:

- Lexical information
- Morphological information
- Syntactic information and
- Semantic Information

The syntactic and semantic information represent

- Word class information and
- Sub class information

The word class information includes syntactic classification of major word classes. These classes are given with a lable like

N for Noun
V - Verb
Prn - Pronoun
ADJ - Adjective
ADV - Adverb
Part - Particles etc. are lables of grammatical categories of each lexical entry.

for instance:

neenu	Prn,	PRED[1]	'I'
illu	N,	PRED	'House
waccu	V,	PRED,	'Come'

The Sub -Class information in the lexicon reflects the contextual features of particular lexical item. For example:

maa Prn, CLASSPOSS, PRED 'Ours'

The feature class **POSS** indicates the further sub - classification of pronouns in Telugu like POSSESSIVE pronouns. (The sub-Classification of all grammatical categories have been discussed in the following sections).

The lexical entry particularly verb form consists of predicate argument structure (in both semantic and syntactic information) with grammatical function.

The verb entry appears as follows:

tinu V, Group[2] 5, PRED 'eat'

(SUB, OBJ)

The verb 'V' accepts the nouns of group-5 and it indicates that the verb is transitive and taking two arguments the **SUB**ject and an **OBJ**ect (SUB, OBJ).

The Argumental features of each lexical entry are listed in the lexicon in the form of annotations (LFG -Bresnan, 1982).

neenu Prn, PRED 'I' PERS I NUM SG.

---

1. It is predicate -argument structure of 'Lexical Functional Grammar (LFG) (Bresman 1982), which will be discussed in the next section.

2. Discussed in the following sections. The nouns are grouped around 10 specifications.

nadi N, PRED ' River' GEN FEM,

In the above example - PERS I, NUM SG denotes that the pronoun is in Ist person singular and in *nadi* gender is feminine.

The said annotations may also be represented with the help of feature bundles (collection of feature specifications of grammatical categories by giving certain lables) (Umarani, 1998) to provide economicity in the lexical entry.

SG 1 = (PERS I NUM SG, -GEN)

These feature bundles are included into each of the lexical entry.

neenu Prn, PRED 'I', SG 1

(@/ Where SG1 = PERS I, NUM SG, - GEN)

Thus the general format of the lexical entry is as follows.

Word, Category, Sub Category, Annotations (as listed above)

The information such as semantic contextual conditioning factors, and selectional restrictions are also incorporated into the lexicon.

tinu V, GROUP 5, PRED 'eat'

(SUBJ, OBJ), (OBJ) = EDIBLE +

The verb *tinu* takes an edible object.

The morphological information in the lexicon contains the information of complex words, its structure and morphological rules if any (requires in the case of complex lexical entries).

Externally the lexicon consist of computer tractable tools which require to design the lexicon. For instance KWIC (Key Word In Context) concordance derived from the

computer corpus, and machine readable/online dictionaries from which information required to the lexical entries may be incorporated. The external appearence of the lexicon is also a matter of concern in this context.

4. TYPES OF INFORMATION STORED IN THE LEXICON

The types of information to be stored in the computational Lexicon is as follows:

1. Lexical information
2. Morphological Information
3. Syntactic information and
4. Semantic Information.

The lexical information contains the form of lexical entries as was represented in the lexicon. It is a word of a language as main or sub entry to which all other grammatical and other informations will be attached in a particular format, depending on the situation and context. The lexical information here includes morphological, syntactic and semantic information of particular lexical entry.

The morphological information is that which relates one lexical item to other in the lexicon. Morphological information includes inflectional and derivational morphology.

The syntactic information as mentioned in the previous section will contain.

- Word class information and
- Sub-class information

The sub-Class information includes the sub-classification of Nouns, Verbs, Pronouns, Adjectives, Adverbs, and particles.

The Nouns are divided in to two main categories. Human and Non-human. Further, human nouns are divided into proper and Common Nouns. Whereas Non-human are sub-

categorized into: Non-human Proper Nouns. Non-human Common Nouns and Other Nouns. Here Proper Nouns include all Place names. Non-human Common Nouns further sectioned into Count and Mass Nouns. Count Nouns are - animate and inanimate Nouns. For examples the entry like *kukka* will be represented in the lexicon as *kukka* : + COUNT + ANIMATE apart from all other annotations. Mass Nouns would be of Quality and Quantity. They represent as in, *andam* + MASS + QUALITY , PRED.....

The Other Nouns in the non -human Nouns are Quantity, Numeric and Action.

The numeric Nouns are subdivided into Cardinal and Ordinal Nouns.

eg. *okaṭi, reṇḍu* -Cardinal 'one', ' two', *okaṭawa, reṇḍawa* - ordinal 'first', 'second'.

Overall, the nouns could be divided into 10 groups. They are :

Proper Nouns (Human)	1
Common Nouns(Human)	2
Proper Nouns (Non-Human)	3
Animate (Non-Human-Count)	4
Inanimate (Non-Human-Count)	5
Quality (Non-Human-Mass)	6
Non Quality (Non-Human Other Nouns)	7
Quantity (Non-Human Other Nouns)	8
Numeric (Non-Human Other Nouns)	9
Action (Non-Human Other Nouns)	10

These Numbers are incorporated into each lexical entry of the Nouns depending on the Verb which take this as a Subject or Object.

The verbs are generally inflected for

Person	-	I, II, III
Number	-	Sing, Plural
Tense	-	Past, Non Past/ Furure habitual
Gender	-	Masc, Fem / Non masc.
Aspect	-	Progressive Perspective Definitive Trial Demonstrative Reflexive
Voice	-	Passive, Active

A verb may be attached with any of these suffixes to it.

Verb Stem + Aspect + Voice + Tense + Gender

Morphologically, verbs are finite and non finite. Finite verbs which accept tense, PNG are further subdivided into :

Imperative verbs and Indicative verbs.

Whereas non-finite verbs are sub divided into : infinitive, adverbial participles, conditional and adjectival participles.

Syntactically verbs can be divided into

Intransitive (only with one argument SUB/OBJ) and Transitive requires two arguments SUB + OBJ. But some times with three arguments (Recipient transitive SUB + OBJ + DAT.OBJ).

Semantic Classification of Verbs

Verbs can be classified semantically, depending on the need of the study (cf. Beth Levin,1993, Dixon 1991, George Miller et al., 1993). This study which deals the construction

of Telugu lexicon, classifies verbs into 18 different categories/ classes as follows :

1.	Verbs of Communication	Vcom
2.	Verbs of Animal Communication	Vacom
3.	Verbs of gesture	Vges
4.	Verbs of Dressing	Vdres
5.	Verbs of Emotion	Vemo
6.	Verbs of Motion	Vmot
7.	Verbs of Construction and destruction	Vcond
8.	Verbs of change	Vcha
9.	Verbs of Social interaction	Vsocin
10.	Verbs of exchange	Vexc
11.	Verbs of creation	Vcre
12.	Verbs of Psychological states	Vpsy
13.	Intake /Consumption Verbs	Vint
14.	Hold and Keep Verbs	Vhk
15.	Killing Verbs	Vkil
16.	Existence Verbs	Vexis
17.	Perception Verbs	Vper
18.	Intellectual Verbs	Vint

The Adjectives are divided into the following types:

1. Simple Adjectives
2. Demonstrative Adjectives and
3. Derived Adjectives.

Simple Adjectives are Qualifiers and Quantifiers.

The Adverbs are of two types

verb qualifiers, sentence modifiers. These are further sub-divided into: Time adverbs (TADV), Manner Adverbs (MADV), Place Adverbs (PADV).

Pronouns are categorised as: Personal Pronouns $\mathbf{Prn}_p$ Interrogative Pronouns $\mathbf{Prn}_I$. Personal Pronouns are differentiated according to Person, Number and Gender. Whereas Interrogative Pronouns are distinguished into: 1. Specified and 2. Non-Specified Personal Pronouns. The other Pronouns are Reflexive Pronouns eg. *tanu / taanu* 'self'.

Particles

The particles are considered to be of prepositional forms, and are added to the nouns.

The semantic information which is to be stored in the lexicon with each lexical Item, include various senses. On each lexical entry the meaning in English is given as a **PRED**icate, in a single sense. If an entry has more than one meaning (Sense), all are incorporated into a particular entry. The sense determination methods such as KWIC concordance has been used in this study. Here the concordance is derived from the Language Corpus, which was collected from various sources and fed into the computer. For any given word the usage will be taken with 8-10 words on right and or left of that particular word. The context of the key word is expected to show different senses, in case there is more than one sense.

The Selectional Restrictions are also given on each possible lexical entry. According to Kartz (1964) the restriction is formally expressed as a condition on that lexical Item, which demands for instance **SUB**ject or **OBJ**ect or **DIRECT OBJ**ect and so on. It is to be considered as a semantic conditioning of that word.

tinu : V(erb) demands some **EDIBLE OBJ**ect 'eat',
taagu : V requires an NP of **ANI**mate as its **SUB**ject.

5. STORAGE AND RETRIEVAL

The indexing techniques of data retrieval system helps to locate a specific information (Murthy, 1997). It also helps us to store explicit information of all lexical items and can be stored in a single text file. It is flexible and user friendly.

6. CONCLUSION

Designing the Telugu Computational Lexicon is an attempt to combine various computational techniques into a practical frame work for linguistic description. It is a Preliminary effort in designing the Telugu Lexicon Computationally using the Lexical Functional Grammar Formalisms.

The lexical Component of Telugu grammar consist of simple words, stems and affixes. The words are labled their syntactic category. The stems are 1. inflectional and 2. Derivational. The affixes are inflectional namely, **CASE, NUM, TENSE, PERS, GEND,** and are given on the particular lexical Item. The word classes (Parts of Speech) are identified, as **N**ouns, **V**erbs, **ADJ**ectives, **ADV**erbs, Pronouns, and **P**articles. Based on these classes the words, stems and affixes are structured accordingly in the Telugu lexicon. A lexical form in this study consist of a predicate argument structure with specified grammatical functions. The types of information stored in the lexicon are syntactic, samantic and morphological informations. The word class information and sub class information comes under syntactic and semantic information. Many techniques are adopted in designing the lexicon. For instance: 1. Reduced the repetition of functional forms in the lexicon. e.g. PERS I , NUM SG, CASE ACC etc. 2. Listed a multiple argument structure on each lexical form.

e.g. *iccu* V, PRED 'give'

(SUB, OBJ2, OBJ.),
(SUB, DAT OBJ, OBJ).

Feature bundles, i.e., a collection of feature specification of grammatical categories: e.g. SG 1 = (PERS I , NUM SG - GEN). Its structure on lexical entry would be *neenu* Prn , PRED 'T , SG1.

A sample of overall structure and format of the Telugu Lexicon has been given as an appendix. In the foregoing description of the Bloomfieldian approach to lexicon where "the lexicon is a mere list of irregularities of a Language" (Bloomfield 1933 : 274) has been salvaged by formalizing the (Ir)regularities by designing computational lexicon for Telugu Language.

Sample lexical entries:

1. VERBS

velḷu V, SUB mot, PRED `go (SUBJ)'. CLASS 6

teliyu V, SUB psy, PRED `know (SUBJ, OBJ)'. CLASS 17

ivvu V, SUB exc, PRED `give (SUBJ, OBJ)'. CLASS 10

tinu V, SUB int, PRED ` eat (SUBJ, OBJ)'. CLASS 13

pagulu V, SUB cha, PRED ` break (SUBJ, OBJ)'. CLASS 8

aṅkitamawu V, PRED 'devote(SUBJ, OBJ)', CLASS 5.

aṅkurinṭcu V, PRED 'sprout (SUBJ, OBJ)', CLASS 11.

aṅgalaarcu V, PRED 'grieve-cry (SUBJ, OBJ)', CLASS 5.

2. PRONOUNS

Personal Pronouns

neenu Pm, PRED T PERS I, NUM SG, GEND MASC-FEM.

meemu pm, PRED `we', Excl +, Pers I, Num PL, GEND MASC-FEM.

niivu-nuvvu Pm, PRED `you', PERS II, NUM SG, GEND MASC-FEM.

miiru Pm, PRED `you', PERS II , NUM HON-SG, GEND MASC-FEM.

itamu Pm, PRED`he', PERS III , NUM SG, GEND MASC, PROX +.

Reflexive Pronouns

tanu-taanu Pm, PRED 'he-himself, she - herself', PERS III, NUM SG, GEND MASC-FEM

tamaru Pm, PRED 'you', PERS III, NUM PL, GEND MASC-FEM, HON +.

tamu-taamu Pm, PRED 'they-themselves', PERS III, NUM PL, GEND MASC-FEM.

Interrogetive Pronouns

evaru Pm, PRED 'who', PERS III, NUM SG, GEND MASC-FEM, PROX -, HON +.

eedi Pm, PRED 'which', PERS III, NUM SG, GEND NUT, PROX -, HON -.

eevi Pm, PRED 'which', PERS III, NUM PL, GEND NUT, PROX -, HON -.

3. NOUNS

talli-amma N, PRED 'mother'. GROUP 2

taṇḍri-naana-naayana N, PRED 'father'. GROUP 2

koḍuku-kumaaruḍu N, PRED 'sun'. GROUP 2

kuuturu-kumaarte N, PRED 'daughter'. GROUP 2

anna N, PRED 'elder-brother'. GROUP 2

tammuḍu N, PRED 'yonger-brother'. GROUP 2

4. ADVERBS

Adverbs of Time

appuḍu ADV, SUB TIME, PRED 'then'.

ippuḍu ADV, SUB TIME, PRED 'now'.

Derived Adverbs

aatarvaata ADV, SUB DER, PRED 'thereafter'.

elluɳɖi ADV, SUB DER, PRED 'day-after-tomorrow'.

idivaraku ADV, SUB DER, PRED 'earlier'.

Adverbs of Place

akkaɖa ADV, SUB PLA, PRED 'there'.

ekkaɖa ADV, SUB PLA, PRED 'where'.

bayaʈa ADV, SUB PLA, PRED 'out_Side'.

mundu ADV, SUB PLA, PRED 'front'.

venuka ADV, SUB PLA, PRED 'behind'.

Manner Adverbs

aʈlaa-alaa ADV, SUB MAN, PRED 'that type'.

iʈlaa-ilaa ADV, SUB MAN, PRED ' this type'.

eʈlaa-elaa ADV, SUB MAN, PRED ' what type'.

aapakuɳɖaa-nilapakuɳɖaa ADV, SUB MAN, PRED ' with out stopping'.

tellagaa ADV, SUB MAN, PRED ' whitish'.

Onomotopoeic Adverbs

kilakila ADVA SUB ONO, PRED 'sound of laughter'.

gaɳagaɳa ADVA SUB ONO, PRED 'sound of bell'.

dabadaba ADVA SUB ONO, PRED 'sound of falling'.

Redupulicatd Adverbs

ippuɖippuɖu ADV, SUB RED, PRED 'every now'.

appuɖappuɖu ADV, SUB RED, PRED 'now and then'.

5. ADJECTIVES

Demonstrative adjectives

aa ADJ, SUB DEM, PRED ' that'.

ii ADJ, SUB DEM, PRED ' this'.

ee ADJ, SUB DEM, PRED ' which'.

Qualitative adjectives

kotta ADJ, SUB QUA, PRED 'new'.

gaṭṭi ADJ, SUB QUA, PRED 'hard'.

tella ADJ, SUB QUA, PRED 'white'.

Quantitative Adjectives

cinna ADJ, SUB QUAN, PRED 'white'.

ekkuwa ADJ, SUB QUAN, PRED 'more'.

mottam ADJ, SUB QUAN, PRED 'total'.

Numeral Adjectives

oka ADJ, SUB NUMA, PRED 'one'.

reṇḍu ADJ, SUB NUMA, PRED 'two'.

Post Positional Adjectives

wenuka ADJ, SUB POP, PRED 'behind'.

pakka ADJ, SUB POP, PRED 'beside'.

kinda ADJ, SUB POP, PRED 'under'.

Distributive Adjectives

prati ADJ, SUB DIS, PRED 'every'.

talaa ADJ, SUB DIS, PRED 'per head/each'.

ceri ADJ, SUB DIS, PRED 'each'.

Derived Adjectives

iṇṭikappu ADJ, SUB DER, PRED 'roof'.

nuutinii||u ADJ, SUB DER, PRED 'well water'.

ceppuladukaaṇam ADJ, SUB DER, PRED 'shoe shop'.

Positional Adjectives

seeru ADJ, SUB POS, PRED 'a measure'.

iddaru ADJ, SUB POS, PRED 'two'.

tappu ADJ, SUB POS, PRED 'wrong'.

Aderbial Adjectives

allaaṇṭi ADJ, SUB ADAJ, PRED 'like that'.

iṭuwaṇṭi ADJ, SUB ADAJ, PRED 'like this'.

ninnaṭi ADJ, SUB ADAJ, PRED 'yesterday's'.

Acknowledgements : I would like to thank Prof. B.Ramakrishna Reddy who initiated this research work, when I was in CIIL - Mysore. Thanks also to Prof. K. Nagamma Reddy, Dr. G.Umamaheshwar Rao, Mr. K. Ramesh Kumar and Dr. P.Uma Rani for their kind comments and cooperation over the manuscript.

REFERENCES

Aronoff, M. 1976. *Word formation in Generative Grammar.* Cambridge: MIT Press.

Atkins, B.T.S and A. Zompolli(ed.). 1994. *Computational Approaches to the lexicon.* Oxford: Claridon Press.

Bloomfield, L. 1933. *Langugage.* NewYork.

Beth levin. 1991. Building a Lexicon: The Contribution of Linguistics. *International Journal of Lexicography.* Vol. 4. No. 3. 205-226.

Boguraev, Branimir and Beth Levin. 1993. In J. Putezovsky (ed.), *Semantics and the Lexicon: Models for Lexical Knowledge Bases*,. Netherlands: K.A. Publishers. 325-340.

Bresnen, Joan (ed.). 1982 : *The Mental Representation of Grammatical Relations.* Cambridge: MIT Press.

Brisco, Ted., et al., 1993 (ed.). *Inheritance , Defaults and the Lexicon.* Cambridge: Cambridge University Press.

Kartz, Jerrold J. and Poul, M. Postal. 1964. *An Integrated Theory of Linguistic Descriptions.* Cambridge: MIT Press.

Koul, O.N. et al., 1998. *Computational Analysis of Kashmiri.* Mysore: CIIL, unpublished project report.

Meigis, W. 1992. Computers and Dictionaries. In Christopher S. Butler (ed.), *Computers and Written Texts.* Oxford: Blackwell. 141-161

Miller A. George (ed.). 1993 *Introduction to Word-net: an On-line Lexical Database.* A wordnet Project report as of 1990.

Narayan Murthy, K. 1997. Electronic Dictioneries and Computational Tools. *Linguistics Today.* Vol. 1, No.1. 34-50.

Uma Rani, P. 1996. Computational Analysis of Tamil Sentences. Unpublished Ph.D. Thesis, Mysore: University of Mysore .

Walker E Donald, Antonio Zampolli and Nicoletta Calzolari (ed.). 1995. *Automating the Lexicon, Reasearch and Practice in a Multilingual Environments.* Oxford: Claridon Press.

OPiL 24: 121-139 (1998)

NEWS OF THE DEPARTMENT

1998

FACULTY PUBLICATIONS

Lakshmi Bai, B. Sounds and Words in early Language Acquisition: A Bilingual Account. Monograph submitted to Indian Institute of Advanced Studies, Shimla.

.....Inflection split or not? Evidence from Tamil-Telugu children. In D. Vasanta (ed.) *Osmania Papers in Linguistics (Combined Special Volume on Applied Psycholinguistics) 22-23.*

.....Contributed the following six entries (Adjective, Adverb, Affirmative, Affix, Article, Aspect and Auxiliary for *Encyclopaedia India.* Calcutta: M.P. Birla Foundation, The Research and publication Unit.

Nagamma Reddy, K. Importance of Durational Knowledge for Text-to-Speech System of Telugu. *Indian Linguistics*. 58: 1-4, 119-128.

.....Women in Development (Keynote address) In *Proceedings of the National Seminar on Women in Development.* Tirupati: College of Engineering. S.V. University. 1-9.

.....A report on 'Speech Technology in India'. *DLA News*, 22: 8-10.

Ramesh Kumar, K. (Co-authored with T. Venkataswamy). Agreement in Adilabad Gondi. *Proceedings of the XXV All India Conference of Dravidian Linguists.* Thiruvananthapuram: International School of Dravidian Linguists.

Swarajya Lakshmi, V. Constraints on Feminine Agentive derivation in Telugu in word formation in Indian Languages. In B. Vijayanarayana and C. Ramarao (eds.). *Word Formation in Indian Languages*. Hyderabad: Centre of

Advanced Study in Linguistics, Osmania University and Booklinks Corporation.

.....Convergence in the Semantics of Polite and Non-polite forms in Indian language at Pragmatic level. In the *Proceedings of XVI International Congress of Linguistics,* Paris.

.....Language in Education. In the *Souvenir 'National Symposium on Literacy'*. Hyderabad: Department of Linguistics, Osmania University.

Vasanta, D. Lexicon : Aspects of Development and Disintegration. *Psycho-Lingua* 28:1, 23-28.

.....Reviewed the book *Language Behaviour: Acquisition of Evolutionary History* by R. Narasimham. New Delhi: Sage Publications. In *Osmania Papers in Linguistics (Combined Special Volume on Applied Psycholinguistics) 22-23.*

.....(Co-authored with P. Sailaja) Word Awareness and Word formation: A Study of Compound noun Produced and Segmentation by Telugu Children. In B. Vijayanarayana and C. Ramarao (eds.). *Word Formation in Indian Languages*. Hyderabad : Centre of Advanced Study in Linguistics, Osmania University and Booklinks Corporation.

Vijayanarayana, B. (Co-edited with C. Ramarao) *Word Formation in Indian Languages*. Hyderabad: Centre of Advanced Study in Linguistics, Osmania University and Booklinks Corporation.

FACULTY ACTIVITIES

B. Lakshmi Bai.

- Presented a paper - 'Inflection Split or not? Evidence from Tamil-Telugu Children' in International Seminar on Agreement, Organized by the Department of Linguistics, University of Delhi. Jan 9-11, 1998.
- Served as Dean Faculty of Arts, Osmania University from 1997 August to 1999 August.

- Chaired a session in the Seminar on 'Bhakti Sahitya Mein Narika Swaroop' Organized by devotional literature Research Centre, Tulasi Bhavan, Osmania University, Feb. 8, 1998.

K. Nagamma Reddy

- Delivered the keynote address at the National Lecture-Cum-Seminar on Women in Development, College of Engineering, S.V.University, Tirupati. February 13, 1998.
- Presented a paper - 'Relevance of Duration for Speech System in Telugu'. DLA symposium on Speech Technology, P.S. Telugu University, Hyderabad. June, 1998.
- Presented a paper - 'Quantitative Analysis of Duration Information for Speech Recognition'. Workshop on Methods in Quantitative linguistics. Indian Statistical Institute, Calcutta. March, 1998.
- Presented a paper - (Co-authored with Sakuntala Sharma) 'Statistical Patterns of Phoneme and Consonant Sequences in Some Indian Languages: Their Relevance to Speech Technology'. DLA Symposium on Speech Technology, P.S. Telugu University, Hyderabad. June, 1998.
- Presented - 'Application of Telugu Computer and its need', Seminar on Telugu Computer: The Need, Prospects and Solutions, Press Club, Hyderabad. November 14, 1998.
- Delivered extension lectures - 'Applied Phonetics', UGC Refresher Course in Linguistics, CIEFL, Hyderabad. 1998.
- Delivered extension lecture - 'Production of English Speech Sounds', Department of English, Annamalai University, Annamalai Nagar. March 24, 1998.
- Visiting Professor, Department of Linguistics, Annamalai University, Annamalai Nagar. Delivered a series of six lectures on certain topics of advanced phonetics. March, 1998.

- Chair-person BoS in Linguistics, Department of Linguistics, Osmania University, Hyderabad. 1998.
- Attended meetings as member of BoS in Linguistics and Foreign Languages and Faculty of Arts, Nagpur University, Nagpur. 1998.
- Attended meetings as general council member of Sharada Vidyalaya College for Women, Shamshergunj, Hyderabad. 1998.
- Chief-guest, Silver Jubilee Celebrations at Sharada Vidyalaya College for Women, Shamshergunj, Hyderabad. 1998.
- Chaired Valedictory Session in the Seminar on Telugu Computer: The Need, Prospects and Solutions. Press Club, Hyderabad. November 14, 1998.
- Secretary, Dravidian Linguistics Association, Thiruvanantapuram, Kerala. 1998.
- Honoured as Life Fellow by the Acoustic Society of India, New Delhi. 1998.
- Nominated Advisory Committee Member, UGC Refresher Course in Applied Linguistics, CIEFL, Hyderabad. 1998.
- Organized a National Symposium on Speech Technology in India - XXVI DLA conference, P.S. Telugu University, Hyderabad. June 19-20, 1998.

V. Swarajya Lakshmi

- Presented a paper - 'Literacy Programme in Mahaboobnagar District of Andhra Pradesh - A Preliminary Study', in the National Symposium on Literacy organised by the Centre of Advanced Study in Linguistics, Osmania University, Hyderabad. February 16-17, 1998.

D. Vasanta

- Organizing Secretary, National Symposium on Literacy organized by the Centre of Advanced Study in Linguistics, Osmania University, February 16-17, 1998.

- Presented a paper - (Co-authored with Sonya Gupta & P. Devi). 'Social Movements and Literacy : On the connections between Literacy and the anti-liquor movement in A.P.', in National Symposium on Literacy organized by the Centre of Advanced Study in Linguistics, Osmania University, Hyderabad. February 16-17, 1998.

A. Usha Rani

- Presented a paper - 'Agreement in Broca's aphasics' at the International Seminar on ' Agreement', Department of Linguistics, University of Delhi. January 9-11, 1998.
- Attended the Glow Colloquim workshop CIEFL Hyderabad. January 20-23, 1998.

K. Ramesh Kumar

- Attended Methods in Quantitative Linguistics workshop held at Indian Statistical Institute, Calcutta. March 3-4, 1998.
- Presented a paper - 'A Quantitative Analysis of /ph/ in Telugu, at the XXVI All India Conference of Dravidian Linguists, P.S. Telugu University, Hyderabad. June 18-20, 1998.
- Presented a paper - 'The Pronunciation of /th/ in Telugu: A Quantitative Analysis' in the XIX South Asian Language Analysis Conference 1998, University of York, York. July 18-20, 1998.

NATIONAL SYMPOSIUM ON LITERACY

With a view to develop a debate between theoreticians and practitioners involved in the field of education / literacy in India, the Centre of Advanced Study in Linguistics, Osmania University, Hyderabad organized a two-day National Symposium on Literacy during February 16-17, 1998. Aditi Mukherjee, who served as the Director and D. Vasanta, the Organizing Secretary of the symposium worked for nearly a year to put-together this academic programme. J.V. Sastry, the Head of the Department of Linguistics, O.U. took charge of the financial aspects of the symposium while other members of the teaching staff extended their help in various activities connected to this symposium.

A Souvenir, containing among other things, a detailed profile of the Department of Linguistics, O.U. and some popular articles on literacy as well as `A Select Bibliography on Literacy Studies' were brought out on this occassion. These publications were released by Prof. V. Ramakistayya, the Vice-Chancellor of Osmania University who was the Chief-Guest of the inaugural session of the symposium. The other speakers of this session were; Prof. Bh. Krishnamurti, Founder member of the Department of Linguistics, O.U and former Vice-chancellor of the University of Hyderabad; Prof. Brian Street, Chair of Language in Education, King's College, London and Prof. Rama Kant Agnihotri, Dean, Arts Faculty, Delhi University. In addition to the faculty members and research scholars of linguistics and language-related disciplines from universities and institutes of higher education in the country, members of several Non-Governmental Organizations concerned with the issues of literacy and education participated in this symposium.

Noted literacy theorist, Prof. Brian Street, in his inaugural address spoke of the need to research multiple

literacy practices among the communities around the world in order to rethink received notions about the ideologies surrounding literacy. He argued that the task of researchers involved in literacy studies today was to make visible, the complexity of local everyday literacy practices. He illustrated this point by sharing his own research experiences in countries like Iran, South Africa, Ghana, the U.K. and the U.S.A. Prof. Street argued that it is findings of research based on such 'local literacies' alone that will help us understand better, the reasons for high drop-out rates, in critiquing existing curricula and in evolving teaching materials that have greater relevance to the learners' everyday lives.

In his keynote address, Prof. Rama Kant Agnihotri argued that while there is no doubt that the local (oral and written) literacies and other forms of knowledge have little or no parallels in contemporary models of education and the latter are essentially behaviouristic in their orientation, we must not romanticize local literacies and cultures. Their revival cannot be the main objective of our literacy agendas, because, these knowledge schemes have limited scope as concepts or tools for any meaningful pedagogic enquiry. The economic and political pull associated with the standard language comes in the way of successfully implementing literacy and education in the vernacular languages. Aside from the language issue, he identified several other paradoxes which have made the literacy project an extremely frustrating one. He lamented that all discussions centring around literacy are associated with poverty as if to suggest that the poor need literacy whereas, the rich and the middle classes need education. After examining the history of literacy programmes in India, Prof. Agnihotri stated that many of these programmes lack a vision of social change. He argued that we must abandon the distinction between literacy and education and enter into an active dialogue with teachers and learners in different communities. His paper also provided critical perspectives on Praathmik

Shikshaa Kaaryakram, the primary education programme of a voluntary group in M.P. called, Eklavya, the DPEP and the Hoshangabad Science Teaching Programme.

There were three theoretically oriented papers in the technical session that followed the inaugural session. The papers by L.M. Khubchandani (Director, Centre for Communication Studies, Pune), Prathibha Karanth (Consultant Speech Pathologist, Bangalore) and Aditi Mukherjee addressed issues like the question of universal literacy, the relationship between metalinguistic skills and literacy and the question of language choice in literacy respectively. Lakshmidhar Mishra (Secretary, Dept. of labour, G.O.I. who was the former Director General of the National Literacy Mission) sent his paper on the "National Literacy Mission: the past, the present and the future", but could not attend the symposium in view of the Parliamentary Elections.

The three papers in the afternoon session of the first day were based on actual field experience. They were presented by (1) Sadhana Saxena (National Institute of Adult Education, Delhi) (2) K. Krishna Kumar (Secretary, Bharata Gnana Vignana Samiti, also an active member of the Kerala Sastra Sahitya Parishad) and (3) Sonya Gupta (Faculty member, C.I.E.F.L. and President, A.P. Jana Vignana Vedika), P. Devi (activist, A.P. J.V.V.) and D. Vasanta. The first paper provided a critical perspective on the campaign based literacy programmes and is informed by the author's involvement with voluntary organizations dealing with rural education like the Kishore Bharati in the state of M.P. In the second paper jointly authored with Mr. S.R. Shankaran (President, B.G.V.S.), Mr. Krishna Kumar described the role played by B.G.V.S. in the successful literacy campaigns in States like Tamil Nadu and Kerala. In the third paper, Sonya Gupta et. al discussed the connections between the literacy programmes and social movements like the anti-liquor agitation that took place in Nellore district in A.P. during 1991-92 by drawing on the

content of the literacy primers developed by A.P. J.V.V., and the experiences of the women volunteers who took part both in the literacy classes as well as the anti-liquor agitation.

The second day's first session began with the paper on Kerala by P.K. Ravindran (of Kerala Sastra Sahitya Parishad, Kerala). Prof. Ravindran provided a brief survey of the Indian Government's efforts to promote adult literacy during the past two decades and described the role played by peoples science movements like the KSSP in successful literacy campaigns like the one in Ernakulam and Pondicherry. In the next paper, Kamal Mahendroo described the Hoshangabad Science Teaching Programme (HSTP) implemented in the rural and semi-urban areas in the Government schools in the state of M.P. He talked of the expansion programme underway with the help of groups like Eklavya and mentioned the activities of other organizations in Gujarat who are involved in developing science teaching packages for grades 5, 6 and 7. Rajesh Sachdeva, the next speaker, who came all the way from Shillong (a faculty member of the North-Eastern Hill University) shared his experiences of trying to implement the programme of mother-tongue education (using the materials prepared by the Central Institute of Indian Languages, Mysore) with the rural poor and tribal children in the state of North-East. He provided a sensitive perspective on the socio-political obstacles that the children face in the classrooms in this part of the country. In his paper titled "Practice of pedagogy in literacy: the West Bengal scenario", Sandip Bandyopadhyay (Consultant, Adult education programme of the All India Institute of Hygiene and Public Health, Calcutta) critiqued several non-formal educational materials used in the state of West Bengal and exposed the "hidden curriculum" behind some of the texts. He then described a literacy primer that he developed and used with the sex-workers of Calcutta.

The last two papers in the post-lunch session on the second day dealt with the issue of evaluation of literacy programmes in the state of A.P. One was a joint paper by D. Subbarao and B. S. Vasudeva Rao (Department of Adult and Continuing Education, Andhra University) and the second one by V. Swarajya Lakshmi (Department of Linguistics, O.U.). These two papers listed several factors influencing the evaluation of Total Literacy Programmes launched in the state of A.P. (particularly the campaigns in Nellore district and Mahaboobnagar districts) including those related to the language used in the text books. The idea of having one discussant raise main points of each paper for discussion did not work out because the authors could not submit their papers ahead of time, but there was a great deal of discussion after each paper on both the days.

In the concluding session, Prof. Rama Kant Agnihotri provided a summary of the arguments put-forward by the various speakers and raised several points for discussion. All the participants expressed their views and unanimously proposed that a forum on literacy studies should be started to consolidate different views among groups all over the country and to continue the dialogue that got started during the symposium on various aspects of literacy and education in the Indian context.

Reported by

D. Vasanta

VISITING FELLOWS

Dr. Apoorva Pauranik, a neurophysician from MGM Medical College, Indore, visited the Centre of Advanced study in Linguistics, Osmania University, during september 21-26, 1998. In collaboration with the Southern Regional Centre of the A.Y.J.National Institute for the Hearing Handicapped, Secunderabad, the Centre of Advanced Study in Linguistics organized a series of six lectures in Neurolinguistics. As part

of this series, Dr. Pauranik gave lectures on topics such as Neuro-anatomy of speech, language and hearing mechanisms; language disorders following brain injuries; Techniques of examination of Brain functions; hemispheric dominance, disorders affecting cognition and personality; Acquired language disorders in bilingual adults etc. Dr. Pauranik illustrated the various concepts with the help of a variety of audiovisual equipment such as overhead and slide projectors, and video films based on aphasic patients. The documentary film covered information regarding the nature of breakdown in aphasia, examination of language breakdown at different levels with respect to Hindi speaking patients, rehabilitation strategies in helping the patients recover the language functions.

The lectures held in the University College of Arts & Social Sciences, Osmania University, were attended by over seventy people belonging to the disciplines of Psychology, Linguistics, Special Education and Speech & Hearing Sciences. A majority of the participants were undergraduate and graduate students and faculty from Osmania University, the National Institute of Hearing Handicapped, Secunderabad, Thakur Hariprasad Institute for the Mentally Handicapped, Hyderabad, Central Institute of English and Foreign Languages, University of Hyderabad, Central Hindi Institute and Government Homeopathic Medical College. In addition to delivering these lectures, Dr. Pauraik interacted with the teaching faculty of the Centre of Advanced Study in Linguistics. These discussions resulted in identification of a number of possible research problems in Neuro-linguistics that need to be taken up by teams of researchers belonging to the disciplines of Linguistics, Psychology and Neurology. The Valedictory function on the concluding day was presided over by the Vice-principal, University College of Arts & Social Sciences, Osmania University.

UGC SPONSORED MAJOR RESEARCH PROJECT ON 'SUPRASEGMENTALS IN TELUGU: ACOUSTIC PHONETIC STRUCTURE AND LINGUISTIC FUNCTION' – 1998-2001.

Principal Investigator: Prof.K. Nagamma Reddy
Co-Investigator : Prof.B. Ramakrishna Reddy

The proposed project concentrates on the multifunctional role of suprasegmentals, namely, acoustic phonetic characteristics as well as their linguistic functions. The interaction of suprasegmentals among themselves and with other linguistic levels like phonology, morphology, syntax, semantics and pragmatics will be explored, thereby highlighting the contribution and relevance of suprasegmentals to language as a system of communication. It is well known for instance, that stress and prominence are signalled by many phonetic parameters, the relative importance of which varies with a particular language and context. Special attention will be paid to the interaction phenomena, i.e. the relation between F_0 variation and perceived prominence or intensity and duration.

An attempt will be made first to identify the phonetic conditioning factors that may affect the manifestation of a suprasegmental feature, then the linguistic functions of each feature at various levels will be considered.

The Project is intended to make a contribution to the study of acoustic phonetic theory in general, and to the phonetics of Telugu in particular by keeping the following objectives in mind : (i) Collection of significant data for developing database of Telugu and general acoustic phonetics, (ii) Recording of data and its systematic phonetic transcription to investigate the variation of suprasegmentals used in different contexts, (iii) To observe, analyse and describe the physiological (i.e. articulatory) and physical (acoustic) properties, processes and patterns of production and

perception of Telugu speech sounds,(iv)Systematic analysis of each suprasegmental unit, i.e. duration, intensity, juncture and intonation; and the interrelationship among these features as well as between these features themselves and the various levels of linguistic analysis, (v) Development of methodology and theoretical models, incorporating the relevant factors that determine variation in the use of each suprasegmental feature, (vi) To provide information on suprasegmental patterning and behaviour for the development of speech systems or speech technology (i.e. speech synthesis, speech recognition and speech understanding), and (vii) To investigate the variation in the use of suprasegmentals across Telugu dialects, i.e. regional, social and stylistic distinctions.

The data collection will be followed by analysis, preliminary observations and generalizations. The generalizations so arrived at will be tested and verified with further empirical evidence from the facts gathered in the experimental phonetic fieldwork. Again, the initial hypothesis will be postulated before arriving at final conclusions.

The Methodology to be followed is an empirical one in that the investigation will proceed from facts (of data) to generalizations based on observation, testing and verification. Data are to be elicited with the help of well designed lists of words and sentences. The design will give prominence to the significant phonetic features so that relevant facts of the language can be found out.

Recording field observation of the speech will include natural conversations, discourse and narrations, over and above words and sentences. Based on these materials a preliminary articulatory phonetic analysis of suprasegmentals will be done to arrive at the phonemic inventory of spoken Telugu, including its regional, social and stylistic varieties. The peculiar characteristics of dialects will accommodate the geographical as well as social determiners of Telugu pronunciations. An

extensive spectrographic study of variable pronunciation in different contexts/ situations, such as formal-informal, spoken-written, male-female speech, educated-uneducated and old-young speech, will call for extensive recording and description. For this purpose, the interviewing techniques evolved for socio-phonetic research will be employed to arrive at reliable generalizations.

The modern tools to be used would include tape recordings of the speech, tape-loop-repeater, microphones and other accessories. For acoustic phonetic analysis, a computer software suitable for speech analysis will be utilised. Speech analysis-synthesis packages and software packages designed for speech production and perception will be adopted for the investigation.

The present study explores to provide precise correlations of articulatory and acoustic features of surprasegmentals. Studies of acoustic phonetics provide quantitative information and realisation of suprasegmental phonology of a language. Acoustic knowledge is crucial in speech technology including Automatic Speech Recognition and Speech Output by computers. No NLP programme can be successful without integrating the speech processing techniques discovered and analysed in the context of suprasegmental studies. It is hoped that the present research will contribute significantly to our understanding of supresegmentals as well as methods of approach and will indicate new directions for future research into other languages. Further the acoustic analysis of suprasegmentals and findings of the present project will provide knowledge for establishing text-to-speech and speech-to-text systems across Indian languages in collaboration with NLP systems. In a predominantly oral (as apposed to literate) society like India, the spoken communication has much wider relevance than communication through the written medium.

The research on speech recognition by machine is hampered by lack of knowledge on the nature of suprasegmental phenomena across Indian languages, and thus the progress in machine-aided translation is hampered considerably. Attempts like the present study can help to establish common acoustic phonetic features across Indian languages to facilitate its creation of mechanical devices.

Furthermore, suprasegmental knowledge is essential to the learning and teaching of Telugu to non-Telugus in India and abroad. It has relevance in the area of teaching the deaf and the dumb, and the mentally retarded children by providing basic clues for speech correction, speech training and for pathological problems. Neurologists and physicians can also make use of this knowledge in their treatment of patients of speech disorders.

The current thinking and theoretical understanding of suprasegmentals has evolved out of the research carried on mainly, if not exclusively, in the European languages. The present study, while concentrating on a non-Indo-European (i.e. Dravidian language) brings out fresh facts in the field of suprasegmentals, from an Indian Language., i.e. Telugu.

Secondly, it has been said that European languages have stress-timed rhythm whereas Dravidian languages like Telugu have syllable-timed rhythm, thus differing in their use of prosodic features. The present study will put forward evidence and analysis from the perspective of Indian Languages. It will investigate the interrelationship between the several suprasegmental features like duration, intensity and fundamental frequency.

There is no comprehensive published research work on the topic of suprasegmentals in Indian languages, to date. However, there have been certain attempts of research on individual suprasegmental features like length, stress or

intonation in few languages of India. And these results are mainly in the form of isolated paper presentations in Conferences/Seminars.

The Principal Investigator is a specialist in phonetics and phonology and her co-investigator is a specialist in syntax, semantics, pragmatics and sociolinguistics. Since the project proposed has ramifications which embrace all these areas, and the suprasegmental features play an important role in grammar. These two investigators will bring to bear upon the project the kind of expertise which is a rare combination of complimentary analytical techniques. The findings are thus expected to be a valuable contribution to the study of Indian languages, in general and to Telugu in particular.

Reported by

K. Nagamma Reddy

SYMPOSIUM ON SPEECH TECHNOLOGY

A symposium on 'Speech Technology' (ST) was convened by Professor K. Nagamma Reddy on behalf of the Dravidian Linguistics Association as part of the 26th All India Conference of Dravidian Linguists held at the Telugu University, Hyderabad during 18-20 June, 1998. The rich contents of the symposium were represented by specialists from divergent fields like phonetics, phonology, linguistics in general, computer science, Engineering, Physics and Speech and hearing hailing from different Medical, Scientific, Engineering, Technology and other Academic Institutes and Universities of India.

NEW BOOK AT A GLANCE

B. Vijayanarayana and C. Rama Rao, eds. *Word Formation in Indian languages*. Hyderabad: Centre of Advanced Study in Linguistics, Osmania University and Booklinks Corporation, Hyderabad. 1998. xii plus 148 pp. ISBN 81-85194-51-33 Rs. 200/-; US $ 25/-; UK £ 15/-.

This volume contains a keynote paper and twelve other papers selected from among those originally presented at the National Seminar on 'Word Formation in Indian Languages', Organized by the Centre of Advanced Study in Linguistics at Osmania University on February 9-10, 1994. Through a study of different languages viz., Telugu, Tamil, Dakkhini, Panjabi, Bengali, Sanskrit, and Indian variant of English, the contributors provide descriptions of either word formation processes or issues involving word formation.

Contents and Contributors: List of contributors; Preface; Reflections on morphology and the lexicon with apologies to neo-Paninians: Rajendra Singh; Syntax as a source of word formation: A case of Telugu: C. Ramarao; On certain aspects of word formation processes: A case study of Telugu: G. Uma Maheswar Rao; Constraints on feminine genitive derivation in Telugu: V. Swarajya Lakshmi; Word awareness and word formation: A study of compound noun production and segmentation by Telugu children: D. Vasanta and P.Sailaja; Grammatical change in Tamil due to the formation of new words: B. Padmanabha Pillai; Deriving nouns in Dakkhini: An experimental study: B. Lakshmi Bai and Najmus Sahr; Word formation in Panjabi: Derivation and compounding: Sukhvinder Singh; The possible and the impossible in Bengali word formation: Some problems in nominalization: Udaya Narayana Singh and Suchita Singh; Word formation in Ashtadhyayi: H.S. Ananthanaryana; Sanskrit word formation and the Samartha theory of Panini: Siniruddha Dash; The treatment of word formation in traditional Sanskrit grammar: G. Rangarajan; Word formation in an Indian variant of English: Priya Hosali.

OSMANIA PAPERS IN LINGUISTICS
Volume 24, 1998

CONTENTS

Printed at: Apollo Graphics, Tarnaka, Secunderabad.
Ph : 7015195.