Towards an Alternative Description of
Incomplete Sentences in Agglutinative Languages

S. Ido

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

2001

University of Sydney
I declare that this thesis is all my own work. I have acknowledged in formal citation the sources of any reference I have made to the work of others.

____________________________
Shinji Ido

____________________________
Date
Title:

Towards an Alternative Description of Incomplete Sentences in Agglutinative Languages

Abstract:

This thesis analyses ‘incomplete sentences’ in languages which utilise distinctively agglutinative components in their morphology.

In the grammars of the languages dealt with in this thesis, there are certain types of sentences which are variously referred to as ‘elliptical sentences’ (Turkish eksiltili cümleler), ‘incomplete sentences’ (Uzbek to‘liqsiz gaplar), ‘cut-off sentences’ (Turkish kesik cümleler), etc., for which the grammarians provide elaborated semantic and syntactic analyses.

The current work attempts to present an alternative approach for the analysis of such sentences. The distribution of morphemes in incomplete sentences is examined closely, based on which a system of analysis that can handle a variety of incomplete sentences in an integrated manner is proposed from a morphological point of view. It aims to aid grammarians as well as researchers in area studies by providing a simple description of incomplete sentences in agglutinative languages.

The linguistic data are taken from Turkish, Uzbek, and Japanese, with special reference to (Bukharan) Tajik.
## CONTENTS

Acknowledgment v  
Preface vii  
Notes on the Text ix

1 INTRODUCTION 1  
1.1 Introduction 1  
1.1.1 Information Structure and Agglutinative Languages 3  
1.1.2 Morpheme-based Analysis and Observability of Information Structure 5  
1.1.3 Definition of Information Structure for This Thesis 11  
1.1.4 Summary 12  
1.2 Agglutination 12  
1.2.1 Agglutinative Morphology 12  
1.2.2 Agglutination and Fusion 17  
1.2.3 Remarks on the Languages Used in This Thesis 22  
1.2.3.1 Turkish 24  
1.2.3.2 Uzbek 26  
1.2.3.2.1 General Remarks 26  
1.2.3.2.2 Orthography 27  
1.2.3.3 Bukharan Tajik 29  
1.2.3.3.1 Tajik 29  
1.2.3.3.2 Agglutination in Tajik 30  
1.2.3.3.3 Tajik Orthography 32  
1.2.3.3.4 Bukharan Tajik 33  
1.2.3.3.4.1 Tajik and Bukharan Tajik 33  
1.2.3.3.4.2 Characteristics 34  
1.2.3.3.4.3 Agglutination in Bukharan Tajik 36  
1.2.3.4 Japanese 37  
1.2.3.5 Summary 39  
1.3 Summary of the Chapter 39

2 MORPHEME-BASED SYSTEM OF INFORMATION STRUCTURE ANALYSIS 41  
2.1 Introduction 41  
2.2 Preliminary Notes 41  
2.3 ‘Old/New Information’: A Review of previous works 42  
2.3.1 Old/New Information 42
ACKNOWLEDGEMENTS

To two of my best friends, Leyla Alioğlu and Sarvar Rahmatullaev, my deep gratitude.

I would like to express sincere thanks to the following friends (who were also patient informants) and people without whose kindness and support this study could not have been completed:


I am very grateful to Hugh Clarke, William Foley, Nerida Jarkey, Aet Lees, Christopher Manning, Sumru Özsoy, and Jane Simpson for their helpful comments on proposals and drafts of this thesis and for their guidance and encouragement. I also thank Hamid Ne'matov for hours of pleasant discussion on information structure in Uzbek.

The present thesis has greatly benefited from Tajik and Uzbek books, articles, and microfilm which the staff of the Interlibrary Loans Section of Fisher Library spent much time and effort to obtain. My thanks go also to my Tajik teacher Qahramon
Tüxsanov at Bukhara State University who gave me books which proved to be essential in writing the present thesis.

Finally, I would like to thank the Australian Commonwealth Department of Employment, Education, Training, and Youth Affairs (DEETYA) and the University of Sydney. DEETYA and the University of Sydney helped me by awarding me the Overseas Postgraduate Research Scholarship and International Postgraduate Award, respectively.
This thesis is the result of my attempt to explain some aspects of agglutinative languages in a simple and psychologically convincing manner.

The present research started with studying certain types of sentences which are variably referred to as ‘elliptical sentences’ (Turkish eksiltili cümleler), ‘incomplete sentences’ (Uzbek to’liqsiz gaplar, Tajik jumlahoi nopurra), ‘cut-off sentences’ (Turkish kesik cümleler), etc. in grammars of the languages dealt with in this thesis.

Such sentences, which I call collectively ‘incomplete sentences’ here, are analysed by the authors of the grammars into a number of disparate phenomena such as pro-drop and null subjects, case marker drop, deletion, etc., for each of which an elaborated syntactic or semantic analysis existed.

After a study of existing analyses of such sentences, analyses that make little reference to one another, I began to think of the possibility of an integrated account for these apparently diverse phenomena. This led me to a somewhat unconventional approach to incomplete sentences – I started to examine them from a morphological point of view.

It seemed to me that such phenomena which have been analysed in terms of syntax or semantics could be given simpler accounts if the distribution of morphemes in sentences in which the phenomena take place was examined closely. For any analysis, simplicity is a positive attribute – I therefore embarked on setting up a simple system of analysis that can handle a variety of phenomena yielding incomplete sentences in an integrated manner.
As the reader will immediately notice, simplicity and general accessibility are the key concepts in this system. Particular attention was paid to make the system conceptually tractable and independent of theoretical complications. The parsimonious nature of the system is obvious from the fact that a very limited number of principles have been posited. For example, all the linguistic data that this thesis contains are analysed in terms of the operation of only two principles.

The importance attached to simplicity is due partly to the fact that the accounts derived from the simple system seemed to comply with the psychological reality of many of my informants as well as myself. It is also due to another purpose of this thesis, namely to devise a system of analysis that is accessible to individuals whose areas of specialty are not within linguistics, e.g. researchers in area studies (of whom I am one), language teachers, etc.

I see the main contribution of this thesis as being not so much in the area of theoretical linguistics, in spite of some reference to works on pragmatics in the literature review and the employment of the term ‘information structure’, but more within the tradition of descriptive linguistics.

The system which has emerged from my analysis of the data was not constructed to fit any particular established theoretical framework. However, it benefits a great deal from previous linguistic work on information structure and I believe that it will be as much of interest to general linguists as it will be to teachers and students of Japanese, Turkish and other agglutinative languages.
NOTES ON THE TEXT

The editorial plural is used throughout the text, except when the use of it may cause confusion regarding the exact identity of its referent. My use of ‘we’ rather than ‘I’ for self-reference has more to do with editorial reasons than with stylistic considerations.¹

Explanations of the transliteration systems for Tajik and Bukharan Tajik used in this thesis are found in the subsection 1.2.3.3. and appendix 2.

All examples in the present thesis are taken from my field notes and e-mails from native speakers, unless indicated otherwise. A large majority of Bukharan Tajik examples are taken from the language used by young Bukharans in their twenties who have had no formal education in standard (literary) Tajik.

¹I had written a fair amount of text before I attempted to convert the editorial plural to the first person singular, which conversion proved to be difficult because the text contains citations as well as English translations of sentences in which first person plural pronouns and agreement morphology are used.