Werk

Titel: An alternative description of incomplete sentences in Turkish and other agglutina...

Autor: Ido, Shinji

Ort: Wiesbaden

Jahr: 2002

PURL: https://resolver.sub.uni-goettingen.de/purl?666048797_0006|LOG_0029

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

An alternative description of incomplete sentences in Turkish and other agglutinative languages

Shinji Ido

Ido, Shinji 2002. An alternative description of incomplete sentences in Turkish and other agglutinative languages. *Turkic Languages* 6, 157-191

This paper analyses "incomplete sentences" in Turkish and other languages which utilise distinctively agglutinative components in their morphology. In the grammars of the languages dealt with in this paper, 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 and (Bukharan) Tajik, with some reference to Japanese. Since the system benefits from Set Theory, a potential computational application of the system is also discussed.

Shinji Ido, 2-10-23-607 Konan, Minato-ku, Tokyo, 108-0075 Japan. E-mail: shinji-ido@ybb.ne.jp

1. Introduction

This paper is the result of my attempt to explain some aspects of Turkish and several other 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 paper.

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, 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 tractible 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 paper 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 paper, namely to devise a system of analysis that is accessible to individuals whose areas of specialty are not within linguistics.

I see the main contribution of this paper 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.

All examples in the present paper 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.

The editorial plural is used throughout the remaining part of this paper, except when the use of it may cause confusion regarding the exact identity of its referent.[1]

[1] My use of 'we' rather than 'I' for self-reference has more to do with editorial reasons than with stylistic considerations. I 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.

2. Analysing information structure of agglutinative languages

The system of description of incomplete sentences that we propose in this paper has several features that distinguish it from the systems found in many grammars.

Firstly, our system has its basis in an analysis of Turkish, Japanese, and (Bukharan) Tajik. Concentrating on these languages, all of which have agglutination as a prominent component of their morphology, brings about an interesting consequence. It prompts us to approach incomplete sentences in a radically different way from that used by many grammarians and to explain them from a morphological point of view. This is the second feature of the system.

While existing grammars and linguistic theories typically explain incomplete sentences in terms of the lack of certain grammatical, syntactic, or lexical components (e.g. predicates, noun phrases, words), the system presented in this paper identifies non-occurrence of certain morphemes in these types of sentences. The feasibility of this morphological approach is in fact conditional on the first feature of the system, namely its concentration on languages with prominent agglutinative morphology. This is part of the reason why, as will be explained later, we restrict the applicability of the system to agglutinative languages.

For the purposes of the present work, the term "information structure" (hereafter referred to in its abbreviated form as IS) is used very loosely to refer to any line of thought that relates, directly or indirectly, sentence form and information. Accordingly, we call the analysis which this paper presents a system of information structure. The use of this term for our system may be misleading, considering its non-theoretical, practical nature (indeed, the basic concept of the morpheme-based system is adopted from the work of Greenfield, a psychologist) and also, because it does not attempt to contest current theories of information structure in linguistics. Wherever a distinction has to be made between our morpheme-based system and the linguistic theories of information structure, we will refer to the latter as "linguistic systems / analyses / theories" of information structure.

In the languages from which we take data, morphemes coincide with pieces of information to a considerable degree. This may not be surprising since morphemes are usually defined as the smallest meaningful units. However, as the existence of so-called portmanteau morphs or words like *ran* and *stood* (as opposed to *run* and *stand*, respectively) suggests, one-to-one correspondence between pieces of information / meaning and morphemes is not a salient feature in many languages, particularly in languages which have morphology characterised by its fusional (or replacive) nature.

On the other hand, in the morphology of agglutinative languages, this correspondence is relatively consistent. This enables one to establish direct links between morphemes and pieces of information, which, in turn, brings in observability of information in analysis of IS. The observability of distribution of pieces of information in sentences is the third of the distinctive features of the morpheme-based IS analysis.

In linguistic work on IS, pieces of information are typically "propositions" which can often be expressed only with clumsy long sentences. This is a natural consequence of linking syntactic phrases with pieces of information. On the other hand, in the present analysis where pieces of information are linked with morphemes, pieces of information are usually identical with the "meanings" of morphemes such as [negative], [plural], etc. As will be discussed later, this approach diminishes the descriptive ambiguity which seems almost inherent in IS analyses.

To sum up, the IS analysis proposed in this paper differs from most other IS analyses in the following aspects: It is based 1) on data taken from agglutinative languages and 2) on morphemes. As it will become clearer in the discussion in the next section, these features of the morpheme-based system of IS analysis are not independent from each other—the latter feature, which forms the backbone of the morpheme-based IS analysis, would not have arisen in the absence of the former.

We aim to make our exposition clear and conceptually tractible, allowing the reader to reason about the implications of the system. Accordingly, section four concentrates mainly on the major issue that is vital for the exposition of how the system works, following which it proposes a set of principles pertaining to the system. The section integrates the principles with the analysis of incomplete sentences.[2]

3. Remarks on the languages used in this paper

Before we proceed to the description of the languages, brief comments on each of them are in order.

Turkish is very often cited as a textbook example of agglutinative languages. Japanese too is generally considered to be an agglutinative language (see e.g. Katada 1995: 115). On the other hand, Tajik is usually not considered to be typically agglutinative although the Iranian languages to which Tajik belongs have some distinctively agglutinative components in their grammars and the strength of agglutinative character of Tajik may be comparable with that of Japanese. As for Bukharan Tajik, it is stronger in agglutination than literary Tajik is in some components of the grammar, while in some other components fusional morphology is more prominent than it is in their literary Tajik counterparts. A brief synchronic description of the agglutination of each of these languages is found in Ido (2001).

Restricting our source of data to Turkic languages would certainly have made it possible for us to set up principles applicable (exclusively) to the information structure in Turkic languages in general. However, we are aware of the possibility of a wider applicability of the principles which we will propose in the next section, which prompts us to include Japanese and (Bukharan) Tajik in our discussion. The inclusion of Japanese and (Bukharan) Tajik is inspired by several different reasons. Japanese is indispensable for our argument in the next section where we contrast

[2] The third chapter of Ido (2001) extends the framework to prominence in speech and discusses the relation between perceptive prominence and information structure.

Turkish and Tajik both of which utilise subject-verb agreement morphology with Japanese which lacks the agreement morphology altogether. (Bukharan) Tajik is of special usefulness for our argument regarding language-specificity in the operation of some of the information-structural principles which will also be proposed in the next section. These issues will be discussed at length in the following section.

4. Morpheme-based system of information structure analysis

4.1. Introduction

The main objective of the present paper is to propose a system of IS analysis which links information structure and morphology in Turkish and other agglutinative languages. The present section attempts to achieve this objective by presenting the most immediate consequence that linking of information structure with morphology would bring about, namely a direct connection between pieces of information and morphemes.

In our analysis of information structure, it will be not predicates or phrases, but morphemes that are linked to pieces of information. Consequently, our method of analysis will contrast sharply with those adopted in linguistics which see stretches of multiple morphemes as bearers / conveyers of information. It differs from many previous approaches also in not using the popular "old information"—"new information" dichotomy in explaining information structure of sentences. The system of analysis presented in this section sees pieces of information not as bearing characteristics of "new" or "old", but as entities which may be varied or remain constant.

4.2. Variability and values

Our system of analysis will be based essentially on "variability", a concept used by psychologist Greenfield (1982) in analysing children's language. She uses the notion of "variability" basically to account for children's choice of words at the stage of one- and two-word utterances, but we find the notion useful also in analysing adults' speech.

According to Greenfield, variability "exists to the extent there are alternatives in a given referential situation" and "the alternatives may exist across either space or time". She gives the following examples for variability (1982: 2):

> "For example, the mother might tell the child to take off a series of items of clothing, e.g. *hat, jacket*. Here action is constant, while object of action varies. Or, in the opposite case, she might tell the child to first put *on* a hat and then take it *off*. Now action is variable while object remains constant."

The "variable" and "constant" in the first example in this statement where the mother tells the child to take off items of clothing, then, may be schematised as follows:

Action:	Constant	taking off
Object of the action:	"Variable"	hat / jacket / etc.

Although Greenfield does not mention any other "constants" or "variables", the action is clearly not the only constant in this situation. Certainly, the performer of the action of taking off also remains constant, thus:

> Performer of the action: Constant child

So does the number of the performer of the action:

> Number of the performer: Constant one

So does, say, the tense in which the action takes place: [present]. Thus, a countless number of constants can be added to this list of constants. Accordingly, the constants and variables in this case may be shown as a chart as in the following:

Action:	Constant	taking off
Object of the action:	"Variable"	hat / jacket / etc.
Performer of the action:	Constant	child
Number of the performer:	Constant	one
Tense:	Constant	present
etc.	Constants	

Note that, in Greenfield's analysis, pieces of information are not given inherent "old" or "new" characteristics. For example, the information about the "performer of the action" maybe constant, but not "old" as many other theories may hold that it is.

In her analysis, there are pieces of information which remain constant before and after an utterance or linguistic expression. There are, on the other hand, pieces of information which are specified as a result of the utterance / linguistic expression. However, this difference between these two sorts of pieces of information does not derive from inherent "statuses" (such as "new" and "old") of the individual pieces of information. As Greenfield exemplifies by alternating "item of clothing" and "action" as "variables", the same piece of information, e.g. "hat" or "taking off", can belong to either of these two groups of pieces of information depending on what is "variable" in a given situation.

A more elaborated version of the idea of variability, which is apparently developed independently from Greenfield's, is found in Clark (1996: 296-297). He cites Wittgenstein's (1958: 3) example where a builder A and an assistant B communicate in what Wittgenstein calls "primitive language":

> "A is building with building-stones: there are blocks, pillars, slabs and beams. B has to pass the stones, and that in the order in which A needs them. For this purpose they use a language consisting of the words "block," "pillar," "slab," "beam." A calls them out; - B brings the stone which he has learnt to bring at such-and-such a call. [...]"

Clark calls the situation where Wittgenstein's "primitive language" takes place a "closed situation", "a social situation" which is "tightly circumscribed, fixed" and writes the following:

> "[...] A and B take it as common ground that there is one parameter (type of block) with four possible values (block, pillar, slab, beam), and A's utterance specifies the intended value."

Then, if A utters not "pillar" or "slab" or "block" but:

> "beam"

he specifies the value of the parameter, which corresponds to Greenfield's term "variable", "type of building-stone", whose possible values are [block / pillar / slab / beam], as "beam". If we symbolise the parameter with x, the utterance yields x = [beam].

That "type of building-stone" is the only parameter naturally entails that there are no other xs in this "closed situation". What Clark calls a "closed situation", then, can be interpreted as paucity of parameters. Fewness of parameters, in turn, seems to be interpretable as abundance of constants.

These infinite number of constants and a single parameter, then, can be schematised thus, with the specified values shown in the far right column:

Action:	Constant	passing
Object of the action:	Constant	building-stone
Number of the object:	Constant	one
Type of building-stone:	Parameter	block / pillar / slab / beam > beam
Performer of the action:	Constant	B
etc.	Constants	

Note that in the above table of parameter and constants,[3] only those that may directly concern our analysis are enlisted. By adding "etc" at the end of the above list, we

[3] Labels attached to constants and parameters in this paper admittedly lack rigour definitions. Thus, the parameter in Clark's example was "type of block", which is also Clark's own expression, and constants were "performer of the action", "number of object", etc. However, these labels secure intuitive understanding of what constants and parameters represent and we continue to use them in the following discussion. It has to be emphasised that constants and parameters are directly linked to morphological units, i.e. variables and non-variables. This link naturally presupposes that each parameter / constant is representative of (a) piece(s) of information to which morphemes correspond. Consequently, many of the labels of parameters / constants reflect the functions of the morphemes which they have links with. For example, the label of a parameter / constant whose value is specified by a passive morpheme may be repre-

allow for the existence of an enormous number of constants that make the joint activity by the interlocutors, i.e. discourse possible. The constants in the "et cetera" include, for example, "the language of communication" and "B is hearing-impaired" whose values are pre-specified as [English] and [negative], respectively. Without an immense number of constants, communication could easily be an impossibility. It should be also noted that the table is constantly updated every moment as the values of parameters are specified.

Constants, or rather, specified values for constants, then, seem to constitute something equivalent to "old information". This in turn means that the above analysis resolves "old information" into an immense number of (pre-)specified values.

Why, then, do we want to break down the single entity of "old information" into specified values? There is an advantage in having many sizeable (constant) pieces of information or specified values rather than a huge collective whole of "old information". This segmentation of "old information" into such values as [one], [beam], etc. is, in fact, instrumental in claiming a correspondence between information structure and morphology which we aim to present, because each one of pieces of information which we link to morphemes can be only as "big" as a piece of information encoded by a morpheme. That is, they are typically such pieces of information as [pl(ural)], [past], [neg(ative)], [beam], [you], etc.

Now, let us reexamine Clark's example above. In the morpheme-based system of IS analysis, the sentence / morpheme / beam / does not convey information such as "the type of block A wants B to pass to him is beam". /Beam/ occurs to specify the value [beam] for the parameter "type of building-stone". The morpheme encodes the piece of information [beam], but it is only instrumental in specifying the value for the parameter and does not have an inherent "old" or "new" status.

In Clark's example, the only parameter is "type of building-stone", whose value is specified as [beam] by A's utterance of the morpheme /beam/. Is this apparent correspondence between the specified value [beam] and A's choice of the only morpheme to be uttered, which is /beam/, a coincidence? We think there is a reason why specifically the occurrence of the morpheme /beam/ is called for here: the morpheme

sented as "passivity" or simply "passive". Similarly, parameters / constants may be labelled as "causality", "potentiality", "action performed" (or "sentence negation"), and "tense", in accordance with the types of morphemes which specifies their values.

A label may contain different types of information (semantic, grammatical, functional etc.) and will often be represented as a mixture of them. This is because morphemes are rarely only semantic or grammatical or functional and it is not our purpose here to be able to identify exactly which one of the types of information encoded by a given morpheme specifies (a) value(s) for a parameter or constant. It is unnecessary for the present discussion to classify the pieces of information which each morpheme encodes according to their types and make a list of them. More rigid formalisation of this labelling convention is desirable, but casual labelling suffices for the discussion here.

has to occur because only "beam" is needed to specify the value for the parameter "type of building-stone".

Accordingly, the first claim that we make regarding the relation between information structure and morphology will be the correspondence between values and morphemes.

4.3. Variables and non-variables

In this section, we will classify morphemes in sentences into two categories: morphemes that specify values for parameters and those that specify values for constants, which will be called variables and non-variables, respectively. However, before we go into discussion, the notion of "allosentences" which will be used extensively in the following discussion needs to be explained.

4.3.1. Allosentences

In explaining variables and non-variables, we will benefit from the notion of allosentence, a term introduced by Daneš (1966) and used extensively in Lambrecht (1994). Lambrecht uses the term in the sense of:

> "[...] semantically equivalent but formally and pragmatically divergent sentence pairs, such as active vs. passive, canonical vs. topicalized, canonical vs. clefted or dislocated, subject-accented vs. predicate-accented sentences, etc." (1994: 6)

He gives the following sentences as an example of a pair of allosentences (1994: 17):

> She likes GERMANS.
> It is GERMANS that she likes.

Put simply, Lambrecht's allosentences are "semantically equivalent but formally and pragmatically divergent surface manifestations of given propositions" (1994: 35). Lambrecht does not give a conclusive definition of "proposition", but since he says "we may refer to this sum of propositions [...] as the hearer's knowledge" (1994: 44), it is probably reasonable to think it as something like (a) piece(s) of information being conveyed by the speaker.

We adopt Lambrecht's definition of allosentence with a modification, which is in appearance minor, but in effect major. The modification is replacement of his "proposition" with "specification of values for parameters". Accordingly, our allosentences are "semantically equivalent but formally and pragmatically divergent surface manifestations of given specification of values for parameters". Allosentences in this paper, then, will mean sentences which, despite their morphological and syntactic differences, specify the same value for a given parameter. For example, 'Ali' and 'I met Ali' as two possible answers to the question 'Who did you meet?' are allosentences because they both specify the value of the parameter "object of meeting' as "Ali", despite their formal differences. Allosentences are explained in more detail in

the next subsection. In short, as long as the value they specify for the parameter is the same and the constants have the same values, they are a set of allosentences.

As a result of this modification, some sets of sentences Lambrecht identifies as sets of allosentences fall outside our version of category of allosentence or become unclear as to whether they are allosentences or not.

For example, we cannot tell with any certainty whether the "she likes Germans" pair cited above are allosentences or not, because what the parameter(s) is / are is not certain. If the parameter or parameters for which the sentences specify value(s) is / are unknown, naturally, whether they specify the same value(s) for the same parameter(s) is also unknown. As a result, we cannot be sure whether the sentences are allosentences or not.

Having introduced and defined the notion of allosentences, in the following subsection, we will explain what variables and non-variables are, using comparisons between allosentences.

4.3.2. Variables and non-variables

Clark's analysis reviewed before neatly explains the linguistic interaction in what he calls closed situations where there are few parameters. However, we think the notion of variability is applicable also to less "closed" situations. We will discuss below question-answer pairs in such less-closed situations. In doing so, we introduce and explain the notions of variables and non-variables with examples. The discussion will also familiarise us with the notions of parameters, constants, and allosentences.

Though these notions have been introduced to analyse the languages that are the subjects of this paper, in this preliminary section we use English sentences for ease of exposition. First, observe the following constructed question-answer pair:

(1) a *Who did you meet?*
b *Ali.*

Let us schematise the constants and parameter as well as their values as we did with the previously cited examples:

			Values
Action:	Constant		meeting
Performer of the action:	Constant		B
Object of the action:	Parameter	Jamila / Ali / Ayshamgul / etc. >	Ali
Tense:	Constant		past
etc.	Constants		

Now that the parameter and the value the answer specifies are clear, we can devise an allosentence of (b). As is explained in the preceding subsection, allosentences have to specify the same value for the parameter. In this case, then, an allosentence of (b) has to specify the value of the parameter "object of the action" as [Ali]. Accordingly, we can devise the following as (b)'s allosentence:

b' *I met Ali.*

This sentence is formally and probably also pragmatically divergent from (b). For instance, the "I met Ali" answer may induce the questioner to expect the answerer to go on to talk about someone who did not meet Ali, or his having met also Ali's wife, which expectation (b) does not arouse. Nevertheless, the sentence (b') specifies the same value "Ali" for the parameter "object of action" as (b) does. The sentences (b) and (b') are therefore allosentences by our definition.

Having devised a pair of allosentences, now we turn to "variables" and "non-variables". What are they and how do we classify morphemes in sentences into variables and non-variables? These will be explained in the next few paragraphs, using an English example.[4]

Variables are morphemes that specify values for parameters. As we will see below, allosentences prove to be instrumental in distinguishing variables from non-variables. First, let us compare the allosentences (b) and (b'):

a *Who did you meet?*
b **Ali.**
b' *I met* **Ali.**

The morpheme /Ali/, which is the value-specifier for the parameter in both (b) and (b'), is an example of what we call variables in this paper.

Note that both (b) and (b') contain the variable /Ali/, which is the only morpheme which the allosentences have in common—it appears as if the occurrence of the variable is essential in providing an answer to (a), regardless of the allosentences' formal and pragmatic properties. When we consider the fact that the single-word utterance we have seen in the preceding section also consists of one variable /beam/, there seems to be some sort of an information-structural constraint involved in the production of these sentences. The constraint will be discussed at length in the following section.

Non-variables, on the other hand, are morphemes that redundantly specify values for constants. Non-variables in the above example are, therefore, "I" and "met" in (b')

[4] Our use of an English example here is entirely expedient—it is certainly not ideal for explaining our system of IS analysis for agglutinative languages and induces some unwanted complexity in the explanation. However, the benefit of the reader's understanding of the notions of "variables" and "non-variables" which our use of an example in English may facilitate may outweigh the benefit of presenting a more appropriate example. Accordingly, we opt to use an English example, but only in this preliminary section.

which redundantly specify the values for the constants such as "performer of the action", "action", and "tense".[5]

To sum up, variables are morphemes that take part in the specification of the values of parameters. As for non-variables, they are morphemes other than variables. According to these definitions of variables and non-variables, not syntactic phrases or grammatical categories but morphemes are the units responsible for altering or retaining the information structure of sentences. This is to say that, in our system of IS analysis, the information structure of a sentence equals the distribution of variables and non-variables in the sentence.

By postulating the "morpho-informational" units of variables and non-variables, we have, in effect, asserted the existence of a direct connection between morphemes and pieces of information. Accordingly, our definition of morpheme-based IS for the present system is much simpler than most other definitions of IS:

Information structure:
Distribution of variables and non-variables in sentences.

There is a reason why in the present study which aims to provide a system of IS analysis specifically for agglutinative languages the morpheme-based approach to IS is preferred over the phrase-based approach. The preference has to do with the defining characteristic of agglutination, i.e. ready segmentability of words into morphemes.

The lexicon and inventory of morphemes of an agglutinative language, unlike those of an fusional language, consist predominantly of readily identifiable morphemes, whose forms are largely invariable. This feature of agglutinative languages is both what prompts us to propose a morpheme-based analysis of IS of sentences and a prerequisite for this proposal. Our assumption is that, if each morpheme can encode a piece of information, they can individually take part in specifying values for parameters and constants.

In sections 6, 7, and 8 where we analyse examples from agglutinative languages using the notions of variables and non-variables, this assumption that each morpheme can be a value-specifier for parameters (and also constants) will prove valid.

4.3.3. Summary

In this section, we have introduced some concepts which are essential for the morpheme-based system of IS analysis. In the next section, we will embark on analysis of example sentences from agglutinative languages, using these concepts of parameters, constants, allosentences, variables, and non-variables.

[5] We consider that the "met" morpheme (redundantly) specifies values for the constants "action" and "tense".

5. Rule 1

We will propose several rules based on the notions of variables and non-variables. Successful accounts for some linguistic phenomena by the rules, naturally, will support the feasibility of the postulation of the morpho-informational units. The rules are the following:

Rule 1: In informative communication in an agglutinative language, the following exist:
Rule 1-1: Variables.
Rule 1-2: Non-variables grammatically required to accompany variables.

Rule 1 lists two groups of morphemes which are essential for informative communication and hence have to appear in sentences. What, then, about morphemes which are neither variables nor their accompaniments? Our claim is that they are not essential in informative communication and hence may not appear even when their appearance is predicted by classical sentence grammar.

This rule, then, may be capable of accounting for the phenomenon which is variously called "ellipsis", "omission", etc., because these terms are used to refer to cases where many people think some elements in a sentence are "missing". While analyses that assume "ellipsis" and "omission" invariably contrast "elliptic" sentences with sentences in which no constituent is "missed" or "inaudible", our analysis involves no such contrast of "incomplete" and "complete" sentences, because rule 1 can yield only self-contained, i.e. "complete", sentences consisting only of morphemes that need to appear.

The difference between our analysis and those many others may be better explained by looking at how they differ in identifying self-contained sentences. Many previous analyses see sentences consisting of the most constituents they can possibly have as "complete". According to these analyses, sentences which have a less number of constituents than they can have in them have to be considered "incomplete" or to have "inaudible" constituents. In other words, this is a top-down approach to such sentences—in this approach, constituents in "complete" sentences are mapped onto actual sentences which may not contain all the constituents that appear in the assumed "complete" sentences. Hence Palmer's explanation that "horses" is a fragment of "they are horses".

On the other hand, the system of analysis employed here takes a bottom-up approach where morphemes appear as they are called for by rule 1. Sentences in our analysis are, then, assembled morphemes rather than (fragments of) sentences as the term is used in a large number of analyses. Naturally, in our analysis, there are no "complete" sentences to be mapped onto the assembled morphemes (= sentences).

If the rule accounts for "ellipsis" in a convincing way, the claimed feasibility of the rule, and hence also the validity of our postulation of variables and non-variables would be reinforced. In order to validate the notions of variables / non-variables, in

the remaining part of the this section we will examine the rule against examples from agglutinative languages.

The following discussion is divided into three sections. The first two sections (6 and 7) deal with sentences in each of which there is only one variable. The third section (8) analyses sentences with multiple variables.

6. Variable = free morpheme

"Elliptic" sentences have been analysed by a number of scholars (e.g. Kuno 1978, Boboeva 1978). However, their largely separate efforts have not established a unified way of analysing the phenomenon of "ellipsis" (or "omission" or "incompletion") in its entirety. We attempt to provide a new standpoint from which to approach the phenomenon as a whole, using variables and non-variables. It will be shown that rule 1 has an explanatory power for this phenomenon (at least in agglutinative languages).

Let us start our discussion with a simple Turkish example. The kind of "ellipsis" explanation mentioned above is adopted by the anonymous author of the Turkish grammar from which the following example is taken (Turska gramatika: 255, glosses and translations are mine):

(2) a *Ahmet ne zaman gel-miş?*
Ahmet what time come-IE
'When did Ahmet come?'

b *Dün.*
yesterday
'Yesterday.'

The grammar calls the answer *dün* 'yesterday' "incomplete (elliptic)" and explains that the answer represents the following:

b' *Ahmet dün gel-miş.*
Ahmet yesterday come-IE
'Ahmet came yesterday.'

Explanations similar to these ones are found also in Japanese works some of which prefer the term "omission" to "elliptic" or "incomplete" sentences (e.g. Kuno 1978: Chapter 1).

On the face of it, the explanation reviewed above may seem to account in sufficient detail the occurrence of the sentence *dün* 'yesterday'. However, this intuitively satisfying account gives us no clues as to why *dün*, and not, say, *Ahmet* or *gelmiş*, has to occur in (b). If (b) is an incomplete version of (b'), what makes only *dün* but not *Ahmet* or *gelmiş*, which are constituents in (b') like *dün* is, appear in (b)? Note that *dün* is the only morpheme the allosentences (b) and (b') have in common—does the appearance of only *dün* 'yesterday', for some reason, have an exclusive importance in answering the question?

In fact, if the answer consisting of every morpheme which (b') has in it but *dün* 'yesterday', the answer would be unacceptable despite its grammatical correctness, as can be seen below:

a 'When did Ahmet come?'
b"## *Ahmet gel-miş.*
Ahmet come-IE
'Ahmet came.'

The sentence (b") is grammatically correct, but is unacceptable as an answer to (a). What, then, causes the unacceptability of (b") as an answer to (a), which unacceptability becomes obvious even by looking at the English translations of (a) and (b")? As has been implied before, this is probably because the appearance of only *dün* 'yesterday' (and no other morphemes) has an exclusive importance in determining the acceptability of possible answers to (a). The reason why the appearance of *dün* alone is essential can be explained by rule 1-1—according to our analysis of the sentence, which will be explained in the following few paragraphs, the morpheme *dün* has to appear because it is a variable.

Let us analyse (b) and (b') to see the value(s) specified by them. Observe the following table of constants and parameter for the example above:

Action:	Constant	coming
Tense:	Constant	past
Mood:	Constant	indirect experience (inference)
Performer of the action:	Constant	Ahmet
When the action took place:	Parameter	yesterday / a week ago / etc. > yesterday
etc.	Constants	

Obviously, the sentences (b) and (b') specify the same value [yesterday] for the same parameter "when the action took place" although they are different in terms of linguistic expression. This makes (b) and (b') each other's allosentences. Which morpheme, then, is the variable, i.e. the morpheme that specifies the value? The variable is necessarily *dün* 'yesterday', because there are no morphemes but *dün* in (b). If *dün* was not a variable, (b), which consists only of *dün* would not be able to specify the value for the parameter.

Thus, the morpheme *dün* 'yesterday' is a variable. According to rule 1-1, variables are morphemes whose occurrence is required for informative communication. This in turn explains the acceptability of (b) and (b') as well as the unacceptability of (b"). The variable *dün* occurs in both (b) and (b'). On the other hand, (b") has all the morphemes which the acceptable (b') has except the variable *dün*. In other words, the morphemes aside from the variable may or may not appear in an answer to (a)—their occurrence does not contribute crucially to informative communication.

This suggests the exclusive importance which the occurrence of the variable has in determining the acceptability of sentences. Note that the variable appears in isola-

tion in (b)—the acceptability (and also naturalness) of (b) is evidence that rule 1-1 is in operation in the production of the mono-morphemic sentence.

Let us examine some other examples to see whether rule 1-1 works in other examples as it does in the example above. Gencan (1979: 141-142, the glosses and translations are mine) presents the following question-answer pair as an example where the predicate "drops":

(3) a *Kim* *uy-uyor?*
who sleep-pr.prog.3sg
'Who is sleeping?'

b *Turgut.*
Turgut
'Turgut is.'

Gencan claims that the answer *Turgut* is in fact *Turgut uyuyor* 'Turgut is sleeping' from which the predicate *uyuyor* 'sleeping' has been "dropped", thus:

a 'Who is sleeping?'
b **Turgut.**
b' **Turgut** *uyuyor.*

This explanation is along the same lines as the one given in the Turkish grammar previously cited. The deficiency of Gencan's explanation is in its not explaining why specifically *uyuyor* 'sleeping' is "dropped". Similarly, it does not explain why *Turgut* cannot be "dropped". Why is the following sentence from which *Turgut*, instead of *uyuyor*, is "dropped" not informative or even acceptable, despite its grammatical correctness?:

a 'Who is sleeping?'
b"## *Uyuyor.*
sleep-pr.prog.3sg
'(it / he / she) is sleeping'

Turgut's obligatory presence can be explained with rule 1-1. Let us repeat the process of analysis we have gone through when we analysed the previous example. The parameter and constants of Gencan's example are as follows:

Action:	Constant	sleeping
Tense:	Constant	present
Aspect:	Constant	progressive
Performer of the action:	Parameter	Ayşe / Ahmet / Turgut / etc. > Turgut
etc.	Constants	

The value for the parameter "performer of the action" is to be specified, for which, obviously, the morpheme *Turgut* is responsible. That is, *Turgut* is a variable, the obligatory presence of which is predicted by rule 1-1. On the other hand, the non-obligatory occurrence of any other morphemes (which are non-variables) is also predicted by rule 1-1. Thus, rule 1-1 provides an explanation for the reason why anything but *Turgut* cannot be absent in the answer in Gencan's example.

Note also that the single principle expressed as rule 1-1 accounts for the occurrence and non-occurrence of morphemes with different functional loads. Perhaps more importantly, rule 1-1 accounts for the non-occurrence of different syntactic units—in the *dün* example, non-variables which are present in (2b') but not in (2b) constitute an NP and VP, while in Gencan's example, non-variables in *Turgut uyuyor* make up a VP. Rule 1-1, then, is capable of handling different types of apparent non-occurrence of syntactic units without classifying them into pro, ellipsis, etc.

The examples cited above are both from Turkish, but the applicability of rule 1-1 is by no means limited to Turkish, as is clear in the following Tajik "incomplete sentence" (*jumlai nopurra*) taken from Niëzmuhammadov et al. (1955: 50, my glosses and translations) where the parameter is B's name, which is specified as [Pūlod][6] by the variable *Pūlod*:

(4) a *Nom-i tu chi-st?*
name-izf you what-cop
'What is your name?'

b *Pūlod.*
Pūlod
'Pūlod.'

We have seen cases where only variables have to appear, which, in turn, supports the feasibility of rule 1-1: "for informative communication, occurrence of variables is essential". The production of the sentences cited above which are explained to be "incomplete" are thus uniformly accounted for by rule 1-1.

In fact, the sentences cited above (*dün*, *Turgut*, *Pūlod*) share a property other than consisting exclusively of variables. In all of them, the variables are free morphemes. Rule 1 has proved to be capable of accounting for such sentences, but can it explain cases where variables are function or bound morphemes which do not appear in isolation, i.e. appear attached / accompanying other (free) morphemes? It can, and it is where rule 1-2 is in operation.

7. Variable = bound morpheme

In the previous section, we have looked at cases where variables are free morphemes. The restriction of examples to ones in which variables are always free morphemes

[6] As a common noun, *pūlod* means 'steel'.

allowed us to observe the operation of rule 1-1 in its simplest form. However, it also concealed the potential problem the rule has—occurrence of variables is called for by rule 1-1, but what happens if the variables are morphemes that cannot appear in isolation? If variables are bound morphemes such as affixes and suffixes, their occurrence which rule 1-1 predicts, would be unacceptable. For example, if, say, the prefix "un-" was a variable, the occurrence of the morpheme in isolation yields an unacceptable "sentence" / utterance: "un-". This is the problem attended to in this section where the problem is resolved by the application of rule 1-2.

First, observe the following constructed Turkish example:

(5) a *Ali-'yle tanış-tı-n mı?*
Ali-com meet-past-2sg[7] Q
'Did you meet Ali?'

b *Tanış-ma-dı-m.*
meet-neg-past-1sg
'I didn't'

What is the parameter in this example? The identity of the possible performer of the action of meeting is not—its value is constant: [B]. The action which may have been performed is not—the value is [meeting]. The tense in which the action is performed is not a piece of new information either—the value is [past]. The value specified by the answer is, then, the negativeness in regard to whether the action specified in the question is performed.

What part in the answer, then, is the variable? It is not the verb stem *tanış-* which specifies the action being carried out, because the value it specifies, i.e. [meeting] is constant. It is not *-dı* [past] either since the tense [past] also is a pre-specified value (see *-tı* [past] in (a)). The personal suffix *-m* [1sg] does not bear information yet to be specified either because the identity of the performer of meeting has already been specified in the question. We are left with only one candidate for the variable, namely the negative morpheme *-ma*.

Thus, whether the action is fulfilled, i.e. "action performed", is the parameter, the value of which is specified as [negative] by the morpheme *-ma* [neg]. The only variable in (b) is, then, *-ma*. However, in (b), other morphemes, namely *tanış-*, *-dı*, and *-m* are also present, despite their status as non-variables. This contradicts rule 1-1. Why are the non-variables present?

The answer is in the latter half of rule 1. According to rule 1-2, morphemes grammatically required to accompany variables have to appear. This explains why the non-variables *tanış-*, *-dı*, and *-m* are present in (b). They are there not to specify values, but to accompany the sole variable in (b), namely *-ma* [neg] and secure the

[7] Hereafter, the translation of the verb *tanış-* 'become acquainted' will be shown as 'meet' in glosses for simplicity's sake.

grammatical acceptability of (b) as an utterance in Turkish. The assumption that the postulation of rule 1-1 and rule 1-2 entails is that, if B could specify the value as "negative" without breaching the grammatical restrictions of Turkish, s/he would have.[8]

Let us observe a possible process of operation of rule 1-1 and 1-2 below. Rule 1-1 calls for the occurrence of variables. Accordingly, at least *-ma* has to occur:

-ma
neg

However, in Turkish one cannot say *-ma* in isolation, even if the speaker wanted to. Indeed, the morpheme *-ma* is the only variable, and any addition of a morpheme would be an addition of a non-variable which redundantly specifies a pre-specified value. Accordingly, though superfluous in terms of information or value-specification, the variable *-ma* needs to appear with some other morphemes to be grammatically acceptable. Above all, it needs a verb stem:

tanış-ma
meet-neg

This is grammatically correct, but means 'do not meet' in the imperative mood, which is not what the answerer means.[9] It is grammatically correct as it stands in isolation, but (grammatically) deficient as a sentence which is intended to specify the value for the parameter as [negative]. In order to avoid misapprehension, it needs more non-variables. Accordingly, in the following, *-dı* [past] is suffixed to the above:

tanış-ma-dı
meet-neg-past

This too is grammatically correct. However, this utterance does not fit the truth condition, because it means 's/he did not meet'. This is, then, another sentence whose grammar is incorrect for a sentence used to specify the value as [negative]. Accordingly, another morpheme is required for the value-specification by *-ma* to be executed without the deviation of the utterance from the truth:

[8] In fact, Turks often do breach grammatical rules in order to specify values only for parameters (and not to redundantly specify values for constants). See the section "pre-grammatical sentences" in Ido (2001).

[9] *Tanış-ma* may mean 'becoming acquainted' if the morpheme *-ma* is not a negative but a nominaliser. If it is a nominaliser, however, the location of the word accent shifts from *-nış-* to *-ma*.

tanış-ma-dı-m
meet-neg-past-1sg

With the addition of *-m* [1sg], the grammatical correctness of the sentence is secured. The variable *-ma* needs no more non-variables accompanying it. Consequently, this sequence of non-variables and a variable constitutes the answer to (a).

What the above analysis reveals is that only a fragment of (b) which rule 1-1 calls for, i.e. the [neg] morpheme, is important / indispensable in terms of information / value-specification. All the other morphemes are non-variables which redundantly specify values for constants. A point that is worth being pointed out in relation to this analysis is that the variable does not require the accompaniment of the subject pronoun *ben* 'I', whose occurrence would have yielded the following allosentence of (b):

a 'Did you meet Ali?'
b' *Ben tanış-ma-dı-m.*
I meet-neg-past-1sg
'I didn't'

This is because the performer of the action is a constant the value for which is already specified as [B] prior to the utterance of (b) and the variable *-ma* does not need *ben* to be grammatically acceptable.

Returning to the discussion of (b), according to rule 1-1, only *-ma* needs to appear for the value-specification of the parameter, but rule 1-2 requires the three non-variables *tanış-*, *-dı*, and *-m* to accompany the variable. This, then, allows the hypothesis that non-variables which need to be present in one language may not have to be present in another language with a different grammar. Taking (b) for example, while the [neg] morpheme requires the morphemes for [meet], [past], and [1sg] in Turkish, the grammar of another language may require different accompaniments for the variable or perhaps none at all. To see whether this is the case, let us look at Japanese translation of the Turkish *tanışmadım* example:

(6) a *Ari ni at-ta?*
Ali with meet-past
'Did you meet Ali?'

b *Aw-***ana**-*katta.*
meet-neg-past
'I didn't'

The parameter and constants for this example are the same as the Turkish ones. Naturally, as rule 1-1 predicts, appearance of the variable *-ana* [neg] is obligatory in the (b) above as the variable *-ma* [neg] is in the Turkish example. However, the Turkish and Japanese examples differ in the number of non-variables that accompany the variables. In the Turkish (b), the variable *-ma* has three accompanying non-variables,

whereas its Japanese counterpart has only two. The non-variable that is present in the Turkish example and absent in the Japanese example is the personal (agreement) suffix.

The answer (b) does not contain any personal suffixes, which Japanese does not have. Despite the lack of a personal suffix whose occurrence is grammatically required in the Turkish (b), the Japanese (b) is a perfectly natural translation of the Turkish (b). This is in accordance with our assumption that the personal suffix *-m* [1sg] in the Turkish (b) is present only to secure the grammatical correctness of (b).

This analysis of ours contrasts with the intuitively appealing analysis which considers that personal suffixes supply "the features of person and number of the referent of the Subject" (Napoli 1993: 85) and hence the subject can be "omitted". In our analysis, personal suffix appear only to secure grammatical correctness—they may be capable of providing "the features of person and number of the referent of the Subject", but at least in this example, the personal suffix does not appear for the sake of providing any information or specifying any value. (Perhaps this allows the assumption that personal suffixes may be of secondary importance in information structure and usually do not participate in value-specification for parameters. That is, personal suffixes may be constantly non-variables. Whether personal suffixes are an indispensable part of information structure or not is central to the discussion in the next section.)

Thus, the occurrence of the personal suffix has proved to be not essential in answering (a), i.e. in specifying the value [neg] for the parameter "action performed". On the other hand, the occurrence of the negative morpheme *-ana*, as a variable, is essential. What about, then, appearance of the morphemes other than the [neg] and [1sg] ones?

Comparing the Turkish and Japanese (b)s, we notice that morphemes specifying values for the constants "action" and "tense" as [meeting] and [past] (i.e. *tanış-*, *-dı*, *aw-*, *-katta*), respectively, are present in both of the Turkish (b) and Japanese (b). If they are non-variables whose appearance is determined by language-specific rules, why do they appear in both the Turkish and Japanese (b)s, as if they are variables whose appearance is called for by rule 1-1?

There are two possible reasons for this unexpected coincidence between Japanese and Turkish (b)s: 1) The rules that we set up are wrong; 2) Turkish and Japanese grammars happen to coincide in the rules concerning the sorts of morphemes by which a negative morpheme must be accompanied. We will claim in the next few paragraphs that the reason is the second. To validate our claim, we have to use an example from another language. Accordingly, we use the Bukharan Tajik translation of the above example:

(7) a *Ali kati šinos*[10] *šud-i-mi?*
Ali with acquaintance became-2sg-Q
'Did you meet Ali?'

b **Na**-*šud-am.*
neg-became-1sg
'(I) didn't.'

In this Bukharan Tajik example, the verbal compound *šinos šud-* 'become acquainted' does not have to appear in its entirety in (b). (Recall that Turkish *tanış-* is [become acquainted], though we gloss it as [meet] for simplicity's sake.) *šinos šudan* 'to become acquainted' is a verbal compound. Note that the answer could contain in it the first component of the verbal compound *šinos šud-*, namely *šinos*, which would yield *šinos našudam* ((b') in the following), an allosentence of (b):

b **Na**-*šud-am.*
neg-became-1sg
'I didn't.'

b' *Šinos* **na**-*šud-am.*
acquaintance neg-became-1sg
'(I) didn't.'

The variable in (b) is *na-* [neg]. This morpheme, like Turkish *-ma* [neg] and Japanese *-ana* [neg], cannot appear in isolation and calls for accompanying non-variables *šud-* [became] appears as rule 1-2 requires:

na-*šud*
neg-became

This is grammatically correct as it stands but attracts a third person singular interpretation, which does not conform with the truth. Therefore rule 1-2 calls for the presence of the [1sg] morpheme, i.e. *-am*:

na-*šud-am*
neg-became-1sg

However, the variable does not call for the appearance of another non-variable, i.e. the first half of the compound *šinos*, as its accompaniment. This is to say that in the above (b), there is no [meet / become acquainted] morpheme that specifies the value

[10] A Bukharan informant says the pronunciation of this word has two variations in Bukharan Tajik; either *šinos* or *šunos*. This is an excerpt from an e-mail from the informant: ... *yakta zanak kati šunos šudas* '(she) met a woman'.

for the "action", which has already been value-specified as "meeting". Thus, the Bukharan example in which no [meet] morpheme has to appear shows that the presence of Turkish *tanış-* in (5b) and Japanese *aw-* in (6b) is called for not by rule 1-1, but merely to secure the grammatical acceptability of the sentences.

Accordingly, we can now safely say that the validity of rule 1 is evident as far as the above examples are concerned. Significantly, the rule has accounted for the non-occurrence of subject pronouns without assuming any of the notions of "ellipsis" or "inaudible elements". This is particularly important because rule 1 manages to avoid the problematic argument that subject-verb agreement morphology somehow relates to the non-occurrence of subjects.

In the next section, we will examine cases where both the subject and agreement suffix must occur in a sentence. Such cases naturally pose another problem to the view which associates occurrence or non-occurrence of the subject with subject-verb agreement morphology. On the other hand, rule 1 proves to have explanatory power for such cases.

8. Multiple variables

In the preceding section, we have claimed that subject pronouns do not have to occur in Turkish, Japanese, and Bukharan Tajik examples because rule 1 does not require them to occur. This analysis of subjectless sentences contrasts sharply with Gencan's analysis in not associating the non-occurrence of subjects with agreement morphology. Our analysis reveals that subject pronouns do not have to occur unless rule 1 calls for their occurrence, whether or not there is agreement morphology.

This analysis entails that, if rule 1 calls for their occurrence, they have to occur regardless of whether there are agreement suffixes. According to our analysis, for example, if the subject pronoun is a variable in a sentence, it has to occur even in the presence of an agreement suffix attached to the verb. This type of obligatory co-occurrence of the subject pronoun and agreement suffix, which cannot be accounted for by Gencan's analysis, is the topic of this section.

Obligatory co-occurrence of a subject pronoun and its agreement suffix in Turkish has been analysed by Enç (1986: 195) from a semantic-pragmatic point of view. She claims that "the pronominal subject [as opposed to "null subjects"] signals topic change". An examination of her claim will serve as a good starting point to our discussion. Let us start with examining a Turkish example from Enç (1986: 205, the glosses are slightly modified by me):

(8) a *Herkes* *Ali-'yle* *tanış-tı* *mı?*
everybody Ali-com meet-past Q
'Did everybody meet Ali?'

b *Ben* *tanış-ma-dı-m.*
I meet-neg-past-1sg
'I didn't'

A peculiarity of the above example is the obligatory occurrence of the pronoun *ben* 'I' in (b). As Enç claims, the subject pronoun must be used in (b). If the pronoun is removed from (b), the sentence would be awkward as a response to (a):

b'# *Tanış-ma-dı-m.*
meet-neg-past-1sg
'I didn't'

Why does *ben* have to appear in (27b)? Clearly, the intuitively satisfying generalisation that "subjects can 'drop' when verbal inflection is rich enough to identify at least some features of the missing subject" (Franks 1995: 288-289)[11] does not work here as it did not in the preceding section, because in (b), the subject has to occur despite the occurrence of the agreement suffix. Enç, instead of looking for a syntactic solution to this problem, turns to a semantic notion which she calls (yes / no) contrast.

As the reason for the obligatory appearance of the pronoun in (b), Enç points to contrast which she claims (b) involves. Regarding the example, she writes (20a and 20b in the following comment signify (a) and (b), respectively, in the above example):

> "(20a) does not make an assertion, but implicates that the speaker was expecting everybody to have met Ali. (20b) [...] provides a counterexample" (Enç 1986: 205)

According to Enç, the response (b) involves provision of counterexample or "yes / no contrast" which she claims triggers the use of the subject pronoun *ben* in the example. As an explanation of "yes / no contrast" she cites a dialogue from Liberman and Sag (1974) and writes:

> "(23) a. All presidents are immortal.
> b. Well, Kennedy died.
>
> (23b), uttered with slightly rising intonation at the end, does not contrast Kennedy with anybody else. It is used to provide a counterexample to (23a), [...] I will call this 'yes / no contrast'." (Enç 1986: 206)

She presents some other Turkish examples where she thinks yes / no contrast is involved, one of which is shown below:

(9) a *Ekmeğ-imiz* *yok.*
bread-1PL:POSS non-existent
'We don't have any bread.'

[11] This sentence is taken from Franks' discussion on this matter and is not a representation of Franks' view.

b *Ben gid-ip al-ır-ım.*
I go GER buy AOR 1SG
'I'll go and buy some.'

She notes that the subject pronoun must be used in the (b) in this example. Certainly, (b) would be awkward without the *ben*, and the awkwardness could be ascribed to B's providing counterexamples as Enç claims to be the case. However, while Enç's theory is adequate for explaining the occurrence of the first person singular pronoun *ben* in this example, this example seems to be a partial manifestation of a rule pertaining to the entirety of the IS of Turkish as well as other agglutinative languages, namely rule 1.

We attempt to locate Enç's rule in a rule whose application is not limited to the analysis of the occurrence of *ben*, but open to agglutinative languages in general (of which, of course, Turkish is one). Firstly, we shall analyse the "meeting Ali" sentence, following which the other will be analysed. This separation of examples is not without a reason. Though they are both presented as different exemplifications of a single phenomenon in Enç's article, they are very different from each other in terms of information structure. One of the (b)s consists exclusively of variables, whereas the other has two variables each of which co-occurs with non-variables. The "meeting Ali" sentence, which is repeated here for convenience, is the latter.

Observe the following (the translation the verb *tanış-* in the gloss is changed from 'become acquainted' to 'meet' for simplicity's sake):

(10) a *Herkes Ali-'yle tanış-tı mı?*
everybody Ali-com meet-past Q
'Did everybody meet Ali?'

b *Ben tanış-ma-dı-m.*
I meet-neg-past-1sg
'I didn't'

Enç, in her analysis of the example, states that (10a) implies that the questioner "was expecting everybody to have met Ali" (Enç 1986: 205). In fact, in spite of her affirmation, the expectation does not become evident from the form of (a), since even without that expectation, the form can be the same, i.e. *herkes Ali'yle tanıştı mı?* However, we simply assume the existence of A's expectation which Enç claims to be there, instead of modifying her example sentence, since, after all, the assumption of the existence of such expectation with the utterance is highly feasible in, say, a classroom situation. (The same expectation will be assumed also for the example's Japanese and Bukharan Tajik interpretations appearing later.)

This, then, makes the utterance (a) resemble a soliloquy rather than a question, because A does not presume there is a parameter for which a value has to be specified. In other words, in the existence of the assumption of A's assumption that everyone met Ali, his uttering *herkes Ali'yle tanıştı mı?* in the above is an equivalent

of saying the affirmative sentence *herkes Ali'yle tanıştı* 'everybody met Ali'. Accordingly, A's constants can be charted thus:

Action:	Constant	Meeting
Tense:	Constant	Past
Performer of the action:	Constant	Everybody
Action performed:	Constant	Positive
Object of the action:	Constant	Ali
etc.	Constants	

B, however, does not comply with A's constants. For B, there are two parameters the values of which are yet to be specified. A thinks "everybody met Ali", whereas B thinks "B did not meet Ali". Accordingly, B's constants and parameters can be charted thus:

		Specification
Action:	Constant	Meeting
Tense:	Constant	Past
Performer of the action:	Parameter	B / Everybody / etc. > B
Action performed:	Parameter	Negative / Positive > Negative
Object of the action:	Constant	Ali
etc.	Constants	

Note that, unlike in the examples observed in preceding sections where there was only one parameter "action is performed", in this example, there are multiple parameters—In addition to the "action is performed" parameter, there is also the "performer of the action" parameter. B's intention is to specify the values for his parameters, because otherwise they would remain constant. A's constants and B's constants (and parameters) differ and a couple of A's pre-specified values need to be re-specified by B as follows:

	A's		B's	Value-specification:
Action:	Constant	✓	Constant	
Tense:	Constant	✓	Constant	
Action performed:	Constant	×	Parameter 1	Positive > Negative
Performer of the action:	Constant	×	Parameter 2	Everybody > B
Object of the action:	Constant	✓	Constant	
etc.	Constants	✓	Constants	

Having identified B's constants and parameters, we now must be able to tell which morphemes are the variables in (b). First, let us identify the variable which specifies the value of parameter 1 to [negative]. The variable which specifies the value for parameter 1, then, must be the only [negative] morpheme in (b), namely *-ma* shown in bold letters below:

*Ben tanış-***ma***-dı-m.*
I meet-neg-past-1sg
'I didn't'

What about, then, the variable for the other parameter, namely "performer of action"? The variable which specifies the value for parameter 2 has to be a morpheme that specifies the performer of meeting. There are two candidates for the variable for parameter 2 in (b). One is, of course, the first person singular pronoun, *ben* [I]. The other candidate is the first person singular suffix *-m* [1sg]. We are left with two morphemes with the potential of being variables. Both of *ben* [I] or *-m* [1sg] appear to be capable of specifying the identity of the performer of the action. Do *ben* and *-m* jointly specify the value, or is only one of them is value-specifier, the other one being a morpheme appearing merely to secure the grammatical acceptability of (b)?

The idea of them jointly specifying the value is unfeasible, because if they do, subjects and agreement suffixes must appear in pairs in all occasions. This is not the case since agreement suffixes can appear without subjects, as they do in the examples observed in the preceding sections. Does this mean, then, that the agreement suffix specifies the value by itself ? In other words, does *-m* [1sg] specify the value on its own, and hence is a variable?

Such value-specification by agreement suffixes seems to be found in some languages.[12] However, agreement suffixes in the Turkic languages do not appear to be capable of specifying values. If *-m* [1sg] was the variable for parameter 2, the sentence without *ben* [I], i.e. *tanışmadım*, has to be acceptable, because then all variables (the presence of which rule 1-1 calls for) are present and accompanied by the morphemes that secure the grammatical correctness of the sentence. However, this is not the case. As Enç notes, a (b) without *ben*, i.e. *tanışmadım*, is awkward as a response to (a), as shown below:

a *Herkes Ali-'yle tanış-tı mı?*
everybody Ali-com meet-past Q
'Did everybody meet Ali?'

b'# *Tanış-ma-dı-m.*
meet-neg-past-1sg
'(I) didn't'

Thus, the presence of *ben* is required, which in turn suggests that *ben* is a variable—if *-m* [1sg] was a variable which specified the value for the parameter "performer of

[12] For instance, judging from Bresnan and Mchombo's (1987: 745) comments, Chechewa, a Bantu language, seems to allow pronouns and agreement affixes to jointly specify values when they are both present, but allows only agreement affixes to specify values on their own. They take the latter case to be the pronominal subject interpretation of agreement affixes which are called S[ubject] M[arker]s in their paper.

action", rule 1 would not call for the appearance of *ben*. This is to say that the value-specification of the parameter is executed not with *-m* but with *ben*.

In other words, *-m* occurs only to accompany the variable *-ma* [neg], so that the utterance's conformity with the truth as well as its grammatical correctness are secured. Judging from the above discussion, agreement suffixes are, despite their often very obvious pronominal etymology in the Turkic languages, not value-specifiers.

The other piece of evidence which confirms that agreement suffixes are not variables in Turkish is obtained by modifying the above to construct an example where the only parameter in (b) is the identity of the performer of the action of meeting, thus:

(11) a *Kim Ali-'yle tanış-ma-dı?*
Who Ali-com meet-neg-past
'Who did not meet Ali?'

b **Ben.**
I
'I (didn't)'

If the agreement suffix *-m* was the variable and the pronoun *ben* was not, the answer in the immediately above example would have consisted of *-m* with morphemes it requires for grammatical correctness, making the presence of *ben* unnecessary:

(12) a *Kim Ali'yle tanışmadı?*
'Who did not meet Ali?'

b'## *Tanış-ma-dı-m.*
meet-neg-past-1sg
'(I) didn't'

However, (b') is clearly awkward as an answer to (a). This supports our analysis that variables are personal pronouns, not agreement suffixes.

The conclusion we can draw from the above discussion is that the morpheme *-m* is called for by rule 1-2 while the occurrence of *ben* is required by rule 1-1, thus:

Values specified by (b):	[I]		[neg]		
Variables (rule 1-1):	**b e n**		**-m a**		
Non-variables (rule 1-2):		*tanış-*		*-dı*	*-m*
Utterance:	**b e n**	*tanış-*	**-m a**	*-dı*	*-m*

Accordingly, the information structure of (b) can be represented as **ben** *tanış***ma***dım*, but not **ben** *tanışmadı***m**. Compare this with the IS of the allosentence in the example previously analysed, one of which is structurally identical with the this sentence:

(13) a *Ali'yle tanıştın mı?*
'Did you meet Ali?'

b *Tanış-***ma**-*dı-m.*
meet-neg-past-1sg
'(I) didn't'

b' *Ben tanış-***ma**-*dı-m.*
I meet-neg-past-1sg
'I didn't'

The latter of these allosentences is structurally identical with the response to *herkes Ali'yle tanıştı mı?* However, unlike that sentence where *ben* has to appear, *ben* in the (b') above does not have to appear. This difference can be ascribed to the difference between the information structures of the two sentences: the former is **ben** *tanış***ma***dım* which contrasts with the latter: *ben tanış***ma***dım.*

Values specified by (b):			[neg]		
Variables (rule 1-1):			**-ma**		
Non-variables (rule 1-2):		*tanış-*		*-dı*	*-m*
Non-variable (optional):	*ben*				
Utterance:	*ben*	*tanış-*	**-ma**	*-dı*	*-m*

Thus, rule 1 neatly explains obligatory and non-obligatory occurrences of morphemes in these Turkish examples.

Now, let us see whether rule 1 is identifiable in another of the languages discussed in this paper. Observe the following Bukharan Tajik counterpart of the Turkish "everyone met Ali" example:

(14) a *Hamma*[13] *Ali kati šinos šud(-mi?)*
'Everybody met Ali(?)'

b **Man** *šinos* **na**-*šud-am.*
I acquaintance neg-became-1sg
'I did not.'

The constants and parameters for this example is identical with those for its Turkish counterpart, that is, the parameters are "action is performed" and "performer of action", the values for which are [neg] and [I], respectively. Here again, despite the occurrence of the agreement suffix *-am* [1sg], the pronoun *man* [I] has to be present, as the awkwardness of the following sentences suggests:

[13] An informant provided not *hamma* 'all' (*hama* in literary Tajik and *hamma* in Uzbek) but *har kas* 'everybody'.

b'## *Šinos* *na-šud-am.*
acquaintance neg-became-1sg
'(I) did not.'

b"## *Na-šud-am.*
neg-became-1sg
'(I) did not.'

All five native Bukharan and one Samarkandi Tajik speakers rejected (b') claiming that it is strange or even utterly wrong and unanimously insisted on the obligatory appearance of *man* [I]. This is of course a judgment which would not take place if the agreement suffix *-am* [1sg] was a specifier of the value [I] for the parameter "performer of action". This suggests that the pronoun *man* is a variable. Accordingly, the IS of (b) can be tabulated thus:

Values specified by (b):	[I]		[neg]		
Variables (rule 1-1):	**man**		**na-**		
Non-variables (rule 1-2):				*šud-*	*-am*
Non-variables (optional):		*šinos*			
Utterance:	**man**	*šinos*	**na-**	*šud-*	*-am*

Thus, operation of rule 1 is clearly identifiable in this Bukharan Tajik example as it is in its Turkish example. Lastly, another piece of evidence verifying that agreement suffixes are not variables comes from a Japanese translation of the Turkish example.

Japanese is of great usefulness in examining whether the presence of the agreement suffix in the Turkish example is required by rule 1-1 or only to satisfy the requirement set by rule 1-2. If the former proves to be the case, the Japanese translation of Turkish (b) which lacks personal suffixes could not be acceptable whatsoever. This is not the case, as the following Japanese example shows:

(15) a 'Everybody met Ali'

b **Boku** (*wa*) *aw-***ana**-*katta*
I (top) meet-neg-past
'I didn't'

This example shows that the absence of the agreement suffix causes no deficiency in communication. On the other hand, the variable *boku* 'I', whose appearance is predicted by rule 1-1, must occur:

b'## *Aw-***ana**-*katta*
meet-neg-past
'(I) didn't'

Accordingly, the IS of Japanese (b) can be shown as a chart, thus:

Values specified by (b):	[I]			[neg]	
Variables (rule 1-1):	**boku**			**ana-**	
Non-variables (rule 1-2):			*aw-*		*-katta*
Non-variable (Optional):		(*wa*)			
Utterance:	**boku** (*wa*)		*aw-*	**ana-**	*-katta*

This chart, as well as those presented before suggest that agreement suffixes are of secondary importance in information structure at least in these languages. The awkwardness of (b)s without pronouns such as *ben* (Turkish) 'I' and *boku* (Japanese) 'I' also supports the view that the morpheme involved in specifying the value for the parameter "performer of action" is the first person singular pronoun in all of these languages. Taking into consideration this supporting evidence from Japanese and Bukharan Tajik, we now can safely acknowledge the analysis of the information structure of (5b) and (6b) presented at the end of the previous section.

The present analysis diverges from most other studies in ascribing the presence of the whole succession of morphemes of *tanışmadım* to the single morpheme *-ma* [neg]. This refutes the information-structural necessity of the appearance of *-m* [1sg] on which the existence of the "inaudible" *ben* [I] is hypothesised in many other works. The present analysis ascribes the appearances of *ben* and *-m* to two different rules. *Ben* is there because it specifies the value for the parameter, i.e. because it is a variable. On the other hand, *-m* is there only to accompany the other variable *-ma* so that the grammatical correctness of the utterance is secured. In sum, according to this analysis, 1) agreement suffixes are of secondary importance in information structure of the languages discussed in this paper, and 2) the appearance of the subject, whether "audible" or not, is not deducible from the appearance of a personal ending.

Before we proceed to the next section, let us briefly observe the other of Enç's examples cited before:

a *Ekmeğ-imiz yok.*
bread-1PL:POSS non-existent
'We don't have any bread.'

b *Ben gid-ip al-ır-ım.*
I go GER buy AOR 1SG
'I'll go and buy some.'

In this example, there are no pre-specified values among the values which morphemes in (b) specify. This lack of redundant specification makes all the morphemes in (b) variables. Accordingly, the production of (b) according to our analysis can be schematised as follows:

Values specified by (b):	[I]	[go]	[ger]	[buy]	[aor]	[1sg][14]
Variables (rule 1-1):	**ben**	**gid-ip**		**al**	**-ır**	**-ım**
Utterance:	**ben**	**gid-ip**		**al**	**-ır**	**-ım**

Unlike in some of the examples previously examined, (b) does not specify any value which has already been specified by (a). The sentence (b), then, consists only of variables as most of the examples cited in the section "variable = free morpheme" do.

To sum up, rule 1 has not only given an account for the obligatory occurrence of the first person singular pronoun *ben* in the first of Enç's examples (8b), but has also provided an account of why so-called "pro-drop" occurs freely in a language with no verb agreement morphology.

9. Unattended parameters

Let us attend to a complexity to which we did not attend in the analyses in the preceding sections, namely "unattended parameters". Unattended parameters are parameters whose values remain unspecified after an occurrence / utterance of a sentence.

There are parameters whose values remain unspecified in a dialogue and yet cause no deficiency to informative communication. Some examples of such unattended parameters are examined below. First, observe the following passage taken from a dialogue part of Muzaffer Buyrukçu's diary (1998: 112):

"Ne içiyoruz Buyrukçu?"	What shall / do we drink, Buyrukçu?
"Rakı", dedim.	'Rakı' I said.

Let us extract the dialogue and add glosses to it:

(16) a *Ne* *iç-iyor-uz* *Buyrukçu?*
what drink-pr.prog-1pl Buyrukçu
'What do we drink, Buyrukçu?'

b *Rakı*
rakı
'Rakı.'

The only morpheme appearing in the answer is *rakı* (a Turkish liquor). Is this utterance sufficient in specifying the values of parameters which are there in this dialogue? It certainly specifies the value for the parameter "object of the action (of drinking)". However, this is not the only parameter in this dialogue.

B does not specify the value for the parameter "identity of the prospective drinker". In this kind of situation, at any rate in Turkish, the identity would be open to at least the following two possibilities: "the answerer alone" and "the answerer and

[14] The verb *al-* means 'take' or 'get'. We adapt Enç's translation here.

the questioner". (This is despite the appearance of *-uz* [1pl] in the question.) Accordingly, both of the following sentences are eligible to be (b)'s allosentences:

b' *Biz rakı iç-iyor-uz*
we rakı drink-pr.prog-1pl
'We drink rakı.'

b" *Ben rakı iç-iyor-um*
I rakı drink-pr.prog-1sg
'I drink rakı.'

The identity of the prospective drinker or drinkers is, then, a parameter the value for which is unspecified by (b). Though it is a parameter, neither Buyrukçu nor his friend makes any attempt to specify the value of it. Consequently, the parameter is left unattended.

This case shows that not all values have to be specified. There are parameters that are neglected either intentionally or unintentionally by interlocutors. Since values for such parameters are not specified, there are no variables specifying values for them.

10. Summary and future directions

In this paper, we analysed a phenomenon commonly referred to as "ellipsis", following which unattended parameters were analysed. The production of all the examples involving these phenomena has been explained by two simple principles, namely rule 1-1 and rule 1-2. While we consider it erroneous to claim universality or wide cross-language applicability of principles elicited from a set of data taken from a very limited number of languages, we do think the applicability of rule 1 is extendable to agglutinative languages in general.

The "morpho-informational" units of variables and non-variables proved to be of high usefulness as analytic devices which can account for a number of seemingly mutually dissimilar linguistic expressions. In conclusion, the system of morpheme-based IS analysis that we presented in this paper has proved to have certain explanatory power for incomplete sentences in agglutinative languages.

As the reader has probably noticed, the system is very loosely based on the notion of "set", though our use of the notion is an informal one. Observe example (1) repeated here for convenience:

a Who did you meet?
b Ali.

Regarding this example, we stated that the object of meeting is a parameter whose value is not specified at the time of A's utterance. If we consider the "object" to be a set, which we symbolise with X_1 here, we can convert the statement into the following:

X_1 = {Jamila, Ali, Ayshamgul, ...}

or

$X_1 = \{x_1 \mid x_1$ is the object of $x_2\}$

where $x_2 \in X_2$ and X_2 is the finite set {meeting}. A part of the chart of constants and parameters that we presented with example 1, then, could be converted into the Cartesian product of sets, if we symbolise, say, the sets of "object", "action", and "tense" with X_1, X_2, and X_3, respectively:

$$X_1 \times X_2 \times X_3 = \{(x_1, x_2, x_3) \mid x_1 \in X_1, x_2 \in X_2, x_3 \in X_3\}$$

where

$X_1 = \{x_1 \mid x_1$ is the object of $x_2\}$
X_2 = {meeting}
X_3 = {past}

There are in fact more sets in the chart that need to be represented in the Cartesian product, such as $X_4 = \{x_4 \mid x_4$ is the performer of $x_2\}$ which equals {B} in example (1). "Etc." in the chart consists of a large number of sets. Accordingly, the entirety of the chart may be represented as:

$$X_1 \times ... \times X_n = \{(x_1, ..., x_n) \mid x_1 \in X_1, ..., x_n \in X_n\}$$

We can also generalise the element(s) which each set contains thus:

$X_1 = \{x_1 \mid x_1$ is the object of $x_2\}$
$X_2 = \{x_2 \mid x_2$ is action$\}$
$X_3 = \{x_3 \mid x_3$ is tense$\}$
$X_4 = \{x_4 \mid x_4$ is the performer of $x_2\}$

so that the Cartesian product could be used for other examples as well. The elements then may be converted into morphemes, though we did not work out exactly what functions we need for the mapping. We anticipate that, through rigid (mathematical) formalisation, the morpheme-based system could gain certain computability.

References

Boboeva, A. 1978. *Hozirgi o'zbek adabiy tilida to'liqsiz gaplar*. Toshkent: Fan.

Bresnan, Joan & Mchombo, Sam A. 1987. Topic, pronoun and agreement in Chichewa. *Language* 63-4, 741-782.

Buyrukçu, Muzaffer 1998. *Dillerinde dünya*. Ankara: T.C. Kültür Bakanlığı.

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.

Daneš, Frantisek 1966. A three-level approach to syntax. *Travaux linguistiques de Prague* 1, 225-240.

Enç, Mürvet 1986. Topic switching and pronominal subjects in Turkish. In: Slobin, Dan I. & Zimmer, Karl (eds.) 1986. *Studies in Turkish linguistics*. Amsterdam, Philadelphia: John Benjamins. 195-208.

Franks, Steven 1995. *Parameters of Slavic morphosyntax*. New York, Oxford: Oxford University Press.

Gencan, Tahir Nejat 1979. *Dilbilgisi*. Ankara: Ankara Üniversitesi basımevi.

Greenfield, Patricia Marks 1982. The role of perceived variability in the transition of language. *Journal of Child Language* 9, 1-12.

Ido, S. 2001. *Towards an alternative description of incomplete sentences in agglutinative languages*. [PhD dissertation, University of Sydney.]

Katada, Fusa 1995. On the nature of agglutination. In: *PAMAPLA* 19: *Papers from the nineteenth annual meeting of the Atlantic Provinces Linguistic Association*, Charlottetown, Prince Edward Island, November 10-11, 1995. 115-127.

Kuno, Susumu 1978. *Danwa no bunpō*. Tokyo: Taishukan.

Lambrecht, Knud 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Liberman, Mark & Sag, Ivan 1974. Prosodic form and discourse function. In: La Galy, Fox and Bruck (eds.) 1974. *Papers from the tenth regional meeting*. Chicago: Chicago Linguistic Society. 416-427.

Napoli, Donna Jo 1993. *Syntax: Theory and problems*. New York, Oxford: Oxford University Press.

Niëzmuhammadov, B & Niëzī, Š. & Buzurgzoda, L. 1955. *Grammatikai zaboni tojikī*, 1: *Fonetika va morfologiya baroi maktabhoi haftsola va miëna*. Stalinobod: Našriëti davlatii tojikiston.

Wittgenstein, Ludwig 1958. *Philosophical investigations*. New York: Macmillan.

Glosses

1sg	first person singular
2sg	second person singular
3sg	third person singular
com	commitative
cop	copula
IE	indirect experience
izf	izafet
neg	negative
pr.prog	present progressive
Q	interrogative
top	topic marker