Topicalization in Pashto

Abstract
The present study is concerned with the phenomenon of topicalization in Pashto. It analyzed the information status of the referents in a sentence and their grammatical functions. The study is another contribution after the work of Roberts (2000) and Rahman (2014). The study was descriptive in nature and it found out the relationship between subject and topic in a sentence. The grammatical function of the subject and the topic were found out coincidentally occurring with the topic of the sentence. Both the subject and topic were found out commonly appearing at the initial position of the sentence. If the topic was any other constituent than the subject, then topicalization happened. The topicalized element is moved to the second or initial position of the sentence. Word order is determined by the information status of the referents where the topic comes first followed by the focused constituents, if not overruled by intonational prominence.

Key Words: Grammatical Function, Information Status, Pashto, Topicalization.

Introduction
Different complex meanings (like who did what by whom etc.) are organized through relations between words conveyed in the form of different grammatical strategies like case marking (Rahman & Bukhari, 2014), verbal agreement (Rahman, 2014), intonation (Trijp, 2017) and word order (Trijp, 2017; Rahman, Riaz & Iqbal, 2017). Most of these strategies are used to make communicative success sure with minimum cognitive resources utilization allowing the hearer to correctly interpret the utterances without observing the events (meaning g therein) themselves and meaningfully decoding the information structure packages. Word order, for example, in English, complemented with a preposition is among the new conventions in English (Trijp, 2017).

Pashto being an SVO language, is one of the most frequent canonical words order pattern languages, along with SVO word order for other languages (Tomlin, 1986; Comrie, 1989; Dryer, 2008). In both these types of word order languages (SVO and SOV), the sentence starts with the universally preferred subject. This makes the subjects syntactically prominent and frequently found at an initial position of the sentence (Levelt, 1989; Levin & Rappaport Hovav, 2005). But the subject does not always come first in a sentence. English, for example, with rather strict word order, in some constructions allows the subject after the finite verb. The following examples from Vogels (2009) illustrate the subject after the finite verb.

1) a) There is something fishy going on here.
   b) Down the stairs came John.

In the first example, the subject follows the finite verb, while in the second example, locative inversion of down the stairs shifts the subject (John) after the finite verb, which is a stylistic variation (operation). So, constituents, other than the subject (objects or adjuncts) may also be found at the initial position of the sentence. This syntactic prominent initial position is mostly because of taking into account the informational status of the referents (or of other constituents) in the sentence. The old information, commonly expressed in earlier sentences or understood from the discourse context, is
expressed early, in a new sentence, followed by new information in the sentence. In other words, what comes first in a sentence is older than what comes later (Haiman, 1983; De Swart and De Hoop, 2000).

Aims and Objectives
The main objectives of the present study are:
- To find out the relation between the information status of referents and grammatical functions
- To analyze the topicalization of referents

Information Status of the Referents in Sentence
The structure of the clause is influenced by two aspects of discourse-pragmatics. Both of these aspects are concerned with either the pragmatic relations or with the pragmatic properties of the referents. The first aspect of discourse-pragmatics concerns the ‘focus structure’ or distribution of information in terms of the pragmatic relations among the referents based on Lambrecht (1994). The second aspect of discourse-pragmatics includes the properties of the referents involving the forms chosen to encode the activation status in the mind of the interlocutor (Rahman, Riaz & Iqbal, 2017). Lambrecht (1994) proposes that information is organized across languages in ‘focus types’. In an utterance, the referring expression can have two primary information statuses, i.e., ‘Topic’ and ‘Focus’.

The Subject in a Sentence
There are two at least two core roles in a language around which grammar of that language resolves. If a theory has the grammatical relations of subject and object (direct object), these terms are clearly defined within that theory. Out of these two core roles, the most syntactically privileged one is the subject (Farrell, 2005), which is mostly the topic of the sentence, at least in accusative languages (including English) or in languages having accusative type constructions (including Pashto). Pashto having both accusative and ergative type constructions (Rahman, 2014), allows omission of both subject and direct object (omission of subject in non-past and omission of the object in past). Tegay and Robson (1996) claim this to happen under predictable circumstances and Rahman (2014) analyzes them in terms of ergative and accusative type constructions. Subject, being the most prominent constituent syntactically comes at the initial position of the sentence. In both SVO and SOV types languages (or constructions), the word order follows the prominence scale of the functions where the subject is more prominent than the object and oblique complement and adjunct. So, the higher on a prominence scale happens at the initial position of the sentence.

Topic in a Sentence
The topic in the sentence is what the speaker intends to increase the addressee’s knowledge about, what he gets the addressee to act concerning it, or what the information is requested about (Trijp, 2017). The topic is synonymous with ‘presupposed’ or ‘given’ part of the utterance. The given information is that part of the sentence that is being presented by the speakers as readily available to the hearers’ memory against what is being asserted about the topic. It is one of the salient elements in the proposition. Commonly, the subject of the clause, being the presupposed information is the topic. In English, the sentence usually initiates with the topic and the predicate (focus) appears at the end. The decision about which part of the sentence makes the topic and which part makes the focus depends on the information needed to be conveyed. The following English sentence is a good example of a topic in English.

My Uncle will Leave Tomorrow.
The topic in a sentence (2) is ‘my uncle’. We are introduced to the topic ‘my uncle’ but we do not know what he does and therefore we need to be told about him further. Such needed information about the topic is called the focus. In most of the cases, the topic is definite (Givón, 1979) and so, is typically
referential, convening old information in the context of discourse or in the real world related to the shared knowledge between the speaker and hear (Skopeteas & Verhoeven, 2009) making sure that the new information follows the old one (Haiman, 1983). That is why definiteness makes the constituent semantic and discourse prominent (Farrel, 2005). On the other hand, the indefiniteness conveys new information (Vallduví, 1990) and the indefinite NPs are not the good topics being not familiar and so, cannot be referential but existential (Farkas & de Swart, 2006). The definiteness and old information are just two of the factors making a constituent as the topic of a sentence (Farrell, 2005). Other factors making a constituent as topic include humanness, grammatical category of person (first person more suitable to be the topic than the second which in turn is more suitable than the third person) in the sentence and possessor and possessed relation, i.e. the possessor being a good topic than the possessed (Givón, 1976). In sum, factors like concreteness, animacy, definiteness, and agentivity contribute to topicality and constituents having these properties are good topics occupying the initial position compared to those having few of these properties and not preferred in the prominent or initial position (Vogels, 2009). Different writers have used different terminologies for the topic. The following ones are some of these terminologies for the topic.

Table 1. Showing different Terminologies for topic Adapted from Hirst (1981)

<table>
<thead>
<tr>
<th>The Boy</th>
<th>Is Riding the Horse</th>
<th>Used by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic</td>
<td>Comment</td>
<td>Sgall et al (1973)</td>
</tr>
<tr>
<td>Theme</td>
<td>Rheme</td>
<td>Halliday (1967)</td>
</tr>
<tr>
<td>Old</td>
<td>New</td>
<td>Chafe (1970)</td>
</tr>
<tr>
<td>Given</td>
<td>New</td>
<td>Haviland and Clark (1974), Clark and Haviland (1977), and Allerton (1978)</td>
</tr>
<tr>
<td>Logical subject</td>
<td>Logical object</td>
<td>Chomsky (1965)</td>
</tr>
<tr>
<td>Psychological subject</td>
<td>Psychological predicate</td>
<td>Hornby (1972)</td>
</tr>
</tbody>
</table>

From a generative perspective, the topic is one of the functional projections that move to the left periphery of the TP node (Rizzi, 1997). Topic from this perspective is an accusative object which can be separated from the rest of the conditional clause but not if it is dative. This crossing of certain syntactic nodes for the topic is only possible under certain semantic and syntactic conditions like it conveys known/old information and it doesn’t belong to focus constituents (Alkala, 2014; Benineca & Poletto, 2004). This makes the topic suitable to be studied under a pragmatic/semantic approach and so, should be studied in context. This allows a kind of room for violation of certain syntactic constraints like island constraints, for example (Alkala, 2014). The present study focuses on the analysis of these pragmatic/discourse and semantic conditions responsible for the topicalization of constituents in Pashto.

Topic in a Pashto Sentence

Topic in Pashto is about the discourse-pragmatic information status of the referents in the structure of the clause. Pashto has unmarked SOV word order like other South Asian Languages, and other word orders are derived from this unmarked order for topicalization, etc. (Tegey & Robson, 1996; Roberts, 2000; Rahman, 2014). The following Pashto examples show the subjects are the topic of the following clauses.

3) a) Taa kitab lwasto 2.SG.ERG book.SG.ABS read.SG.PST.IMP.ABS You were reading a book.
   b) Mahmood kar kai Mahmood.NOM work.M.ACC do.M.PRES.NOM Mahmud is working.
   c) Spi me pisho khog kra Dog.SG.ERG 1.SG.CL cat.SG.ABS hurt do.PST.ABS My dog hurt the cat.

One of the easiest ways to recognize the topic in the Pashto clause is its recoverability in the form of the missing argument. The ‘presupposed’ or ‘given’ part of the utterance in some languages is so frequently used like their use in Pashto that these languages omit the topic of the sentence from the
clause structure. We have different slots for different referents in the clause structure. The first slot is commonly filled by the topic of the clause (in languages having rigid syntax) which is not necessarily the subject of the clause in terms of the functions of the NPS. The pro-drop of the subject and direct object and even indirect object indicates the topic in Pashto sentences. The following examples (the a examples show full arguments and b examples show the pro-drop of the subject, direct object and indirect object respectively (Rahman, 2014).

4) a) Ahmad kitab lwali 3.SG.NOM book.ACC read.PRES.IMP.3.SG.NOM Ahmad is reading a book.
   b) _________kitab lwali Book.ACC read.PRES.IMP.3.SG.NOM He is a book.
   c) hagha kitab lwasto 3.SG.ERG book.ABS read.PST.IMP.ABS He was reading a book.
   d) hagha __________ lwasto 3.SG.ERG read.PST.IMP.ABS He was reading a book.
   f) taa________ kitab WAR kro 2.ERG.SG book.ABS.VC give.PST.PF
   You gave him/her/them a book.
   g) ta ________ kitab WAR kay 2.NOM.SG book.ACC. VC give.PST.IMP
   You are giving him/her/them a book.

The above examples show that the subject in present (4b) and direct object (4d) in the past can be omitted as they are realized by the verbal suffixes in the present and past respectively. Similarly, the indirect object can be omitted in both present (4f) and past (4g) as the verbal clitic realizes the dative argument. The verbal agreement suffixes are obligatory which may co-occur with the overt arguments and may result in pro-drop of subject and object as well. The non-past has the pro-drop of the subject with the verbal agreement suffixes showing agreement with the subject, while past has the pro-drop of the object with verbal agreement suffixes showing agreement with the object. The verbal agreement suffixes always show agreement with direct case (nominative, absolutive) which may co-occur with full NPs (Rahman & Bukhari, 2014). Generally, the entities having a high degree of individuation and discourse saliency (Rahman, 2014; Greenberg, 1974; Given, 1979; Belloro, 2007) are the topics of the sentence. The arguments, in Pashto sentences, are linked according to the topic and focus positions of the arguments in the language. The topical position is the first position in the sentence while the focal position is the position immediately before the verb.

**Topic and Subject**

Prototypically, the topic is the subject in a sentence. The subject commonly is the agent of the clause, i.e. the most active participant for an action. The agent is a semantic role most commonly functioning as a subject in a clause. Other dependents, too can function as a subject (Farrell, 2005), but in an active voice clause, the prototypical subject is the agent. Similarly, the prototypical direct object has the semantic role of the patient, i.e. the primary target of the action of the agent (subject). Other dependents, too can function as a direct object like the subject but in an active voice clause, the prototypical object is the patient. From this perspective, it follows from the above discussion that what counts is the speaker’s conception, i.e., the speaker’s conception (subjective point of view) of the situation described in a sentence not the listener’s conception from an objective point of view. That is why even dative cases can function as subjects (macro role of the subject is more important than the case system in Pashto) and that is why the most important macro role occupied the initial position in clauses (Farrell, 2005). In English, the sentence starts with a subject (a grammatical function), in line with the strong syntactic preference to start sentences with subjects which are the position for a topic as well, resulting in the coincidence of subject and topic in a single clause. From this perspective, the subject is often the topic and the topic is often the subject. The conditions for the Pashto clause are
almost the same if the topicalization of any other constituent does not happen. The sentence starts with
the subject which is the most common, the topic of the sentence.

**Topicalization and Dislocation**
Topicalization is the movement of a certain element to the front (left edge) of the sentences. This
movement is the movement of elements other than the subject, as prototypically, the subject comes
first in every sentence (Speyer, 2005). When an NP is separated from the rest of the sentence by a
marked fall in intonation and a boundary pause, it is called dislocation which is an extreme form of
topicalization. The dislocated constituent may be located at the beginning or end of the sentence. The
DP moved the left is called left dislocation the DP moved to the right is called right dislocation (Jaeggli,
1986; Holton, 1997). This difference between right and dislocation is not always clear. The only
difference between the topicalized and dislocated constituent is the degree of the length of the pause.
The greater length of the pause and deeper intonation fall are correlated with dislocation (Holton,
1997). Dislocation is often a prosodic (in some cases syntactic) phenomenon.

**Topicalization in Pashto**
Topicalization in Pashto is the promotion of any other constituent than the topic (syntactically, the
subject) to the topic position (which is syntactically the subject position in the sentence). Subject,
commonly, being more prominent than other constituents of the sentence occupy the initial position in
a sentence, which can only be occupied by other constituents if they are topicalized. As the topic is a
semantic/pragmatic prominence, any of the constituents in a sentence can be the topic of the sentence.
Pashto has an OSV order derived from the unmarked and rigid SOV order if the object is scrambled for
topicalization, etc. The following examples show the OSV order derived from the unmarked SOV order.
The OSV order is commonly used in the past tense and this is not a preferred order in the present tense
(Roberts, 2000). The following examples from Rahman (2014) show the topical objects in Pashto.

5) a) Kitab maa lwastalay day Book.M.ABS 1.SG.ERG read.PST.PF.ABS be.PRES.ABS
   I have read the book.
   b) Ta sari wahaley ye 2.SG.ABS man.SG.ERG beat.PST.PF.ABS be.PRES.ABS
   The man has beaten you.
   c) Pisho me spi khog kra Cat.ABS 1.SG.CL dog.SG.ERG hurt do.PST.ABS
   The dog injured my cat.

   The above examples show the objects in the subject position (because of the topicalization of the
   object). It is apparent from the above examples that topics are moved (if the topic is not the subject
   of the sentence) to the initial position of the sentence, as a result occupying the subject position. The
   object then occurs in a topical position and precedes the subject. This marked order in the past is
   because of the disambiguation in the sentence (Rahman, 2014; Tegey, 1977; Babrakzai, 1999; Roberts,
   2000). If the subject precedes or follows the object, it does not create ambiguity in terms of the
   agreement because the intended interpretation is easy without any ambiguity to identify the subject and
   object through the resultant agreement markers on the verb. The subjects and objects are determined
   on the bases of the agreement markers on the verb in the past (with object) and non-past (with the
   subject). In other words, in Pashto, the topic position can be filled by the object in the Pashto clause if
   we have topicalization. This topicalization resembles the passive constructions in English where the
   object moves to the topic position (subject position) of the clause. Similarly, even indirect objects are
   correlated with topical participants (Weissenrieder, 1995). If the temporal or spatial settings of the
   predicate are referred to in a sentence, we can have the topicalization of adjuncts describing a sentence
   in a temporal and spatial location (Crasborn et al., 2009).

   In the case of the topicalization of a constituent, the usual focus position (being part of the
   predicate/comment) is occupied by the other constituent. For example, if a direct object is moved to
   the initial topical position of the clause, the subject becomes part of the focus. This distinction between
   topic and focus is easy if the referents have different semantic properties, for example, if one of the
referents is inanimate and the other is animate and the verb requires an animate subject and so, to identify the topic becomes easy. But if both the referents are either inanimate or animate, ambiguity arises in the form of who did something to whom. That is why the following examples, clarify these points.


The above examples clearly show that when one of the referents is animate (6a), the animate functions as a subject being the topic even if there a small pause or comma after it (6b) because the action of pushing can never be performed by book and it requires an animate subject like Ali. Similarly, if the same object is topicalized (6c), the topic is now not the subject but the object at the initial position. The ambiguity arises when both the referents are animate (6d), which could be solved to some extent by the use of a shorter pause or comma after the subject preceding the topic. The topic in the form of a topicalized constituent is easily recognized in sentences having clitics instead of full NPs (Rahman, 2014). The grammatical properties of referents like a number, person and gender, particularly, are important ones and cannot be ignored (omitted) and if the full NP is not available, the clitic shows these features at least. The NP with which agreement happens could be omitted because its presence is shown by the agreement markers on the verb. The grammatical features of the other referent are represented by clitic at least. The following examples show the use of clitics for these features.

7) b) kitab ye guzar ko SG.OBJ.ABS.INA 3.CL.ERG.SUB.ANI throw do.PST.ABS He threw the book.
e) ______________ kitab guzarai 3.M.SG.ANI ook.M.SG.INA throw.PRES.NOM.ANI He is throwing it/He throws it.

The above examples show that when topicalization of the object happens (7a), the other referent can appear as clitic instead of full NP. The topic itself can be omitted (7b), being understood from the context of discourse or the recently mentioned in the previous sentences. Similarly, when the book is not topic (7c and 7d), it cannot be omitted but appear either as full NP (7d) or clitic (7c). The choice between the clitic and the agreement markers is clear from the fact that if the verb agrees with a particular NP, that NP can be omitted and if the verb does not agree with an NP in a transitive construction, the non-agreed NP is at least represented by clitic. So, agreement markers/clitic/pro help us in recognizing the topic in a Pashto sentence.

If we have more than two arguments (referents) in a Pashto clause, most commonly the indirect object precedes the direct object like we find in English. The only difference between an indirect object in English and Pashto is, in English, the direct object could be in form of preposition phrase or in the form of NPS along with shifting or because of shifting the positions for objects. In Pashto, the indirect object always appears in the form of postpositional phrases with preceding or following the direct object. The indirect object, in Pashto, precedes the direct object and if the direct object is topicalized, it can precede the indirect object. The direct object if it precedes the indirect objects, shows it being in the topical position. The following examples, show three arguments verb. Any of the three arguments, being topical can be omitted in the sentence.
Conclusion
The present study analyzed the information status of the referents in a sentence and their grammatical functions. The study found the relationship between subject and topic in a sentence. The grammatical function of the subject coincidentally co-occurs with the topical information structural content of the sentence. Both the subject and topic commonly appear at the initial position of the clause. If the topic is any other constituent than the subject (the initial position of the sentence), then topicalization happens. This topicalization is different from dislocation, which is represented from a long pause after the first constituent in the sentence. If the topicalized element functions as an object, it is moved to the initial position in montransitive constructions having two arguments verb. On the other hand, in ditransitive constructions, having three arguments, the topicalized direct object, is moved to the second position, from its real third position after the indirect object. The indirect object, on the other hand, if topicalized is moved to the initial position of the sentence. The topicalized element is easily recognized from the use of clitic instead of full NP if the verb does not agree with the topic. The agreement in the language takes into account the grammatical function in collaboration with the tense. The agreement markers take into account the grammatical function and case marking irrespective of the word order of the constituent, while word order is determined by the information status of the referents where the topic comes first followed by the focused constituents if not overruled by intonational consideration.
References


