Linguistics Initiative

ISSN 2775-3719

Volume 1, Number 1, (2021)

<https://doi.org/10.53696/27753719.19>

Published by *Pusat Studi Bahasa dan Publikasi Ilmiah*
(Center for Language and Scholarly Publishing Studies)

Copyright © 2021 The author(s)

Article License: CC BY-NC 4.0

Restricted and Unrestricted Syntactic Distribution of *ar* in Bangla

Samrat Halder ⁽¹⁾

⁽¹⁾ The English and Foreign Languages University, India |
samratphdlandp18@efluniversity.ac.in

Abstract

This approach examines the syntactic distribution of *ar* in Bangla language. Apparently, it seems to occur almost every plausible position across a clause. It may occur clause initially, internally and finally. In some extreme cases, it may occur with each word in a clause. This surface observation does not specify what its exact syntactic distribution is. So, its across clause distributional phenomenon is controversially disputed. In this paper, I shall argue that *ar* can be restrictedly and unrestrictedly distributed in their respective contexts, thus its apparent multiple syntactic position is a misconception. However, each syntactic distribution either restricted or unrestricted producing a distinct semantic interpretation has specific significance depending on its particular contextual circumstances. This investigation aims at enunciating how every plausible interpretation of *ar* can be determined based on its restricted and unrestricted syntactic distribution only which in turn reflects its forming of a constituent with the elements following it in a clause. It also shows that the distribution of *ar* in Bangla is both syntactically and semantically determined.

Keywords

Bangla *ar*, restricted and unrestricted distribution, coordinator and adverb, mysterious *ar*

Introduction

Bangla (or Bengali) is one of the widely spoken South Asian languages that typically follows the SOV or Subject Object Verb word order. As a head final language, it manifests multiple characterizations such as scrambling, pro-drop, gender neutrality, person-tense-inflectional agreement and so on. In addition to that, Bangla uses a wide range of functional category elements among which many have been identified as ‘discourse particles’ by [Bayer et al. \(2014\)](#). This is because of their direct correspondences with discourse information. These particles such as *to*, *ar*, *naki*, *ba*, *bujhi*, *na*, *ki* etc. ([Bayer et al. 2014](#)) may change and direct or redirect the course of discourse utterances. Therefore, one might identify *ar* as one of the discourse particles which naturally leads to the anticipation of a discourse function. In this paper, the concerned *ar* has shown to have association with morphologically ‘functional’ category elements status but its function in a clause is extremely crucial and rarely addressed in linguistic literature. But its decisive role in both spoken and written forms of Bangla language may not be a mere linguistic coincidence.

ar in this language Bangla, is typically understood as a “*coordinator*” as noted in [Bayer & Dasgupta \(2016\)](#) followed by [Karmakar et al. \(2017\)](#) including one of my own research paper presentations in ICOLSI-42, 2020 conference. [Karmakar et al. \(2017\)](#) and in my work, it has been argued that *ar* also functions as an adverb besides its being a polysemous item. Now, one can simply assume that these two mentioned functions; coordinator and adverbial are determined by its respective syntactic distribution theoretically, as simple as that. So, why there would be any need to show the syntactic distribution that happens restrictedly and unrestrictedly? I shall get back to this question in the discussion section, bear with me until then. Let me first introduce you with the *ar*.

As a coordinator, *ar* corresponds to the conjunction ‘and’ and as adverb it usually corresponds to ‘again’ and ‘anymore’ simultaneously. However, a wide range of multiple adverbial interpretations can be contextually derivable. To see its two majorly decisive functions in communication, consider the following examples;

1. Sam ar Dan aam kha-chhe
 Sam.Nom.3 and Dan.Nom.3 mango eating
 Sam **and** Dan are eating mangos
2. Saul ar Delhi ja-be -na
 Saul.Nom.3 again delhi.Loc go.Fut.3 not
 Saul will not go to Delhi **again**
3. Paul ar aam kha-i -na
 Paul.Nom.3 anymore mango eats not
 Paul does not eat mango **anymore**

All the three instances *ar* above in (1-3), at a surface observation might seem to occupy the exact same position that is right after the subject noun. What is to look at more carefully in this small set of data is that despite its apparent occurrence at the exact same position each occurrence means something distinct from the other. So, we see the

conjunction ‘and’ (1), adverb ‘again’ (2) and another adverb ‘anymore’ in (3) respectively. Therefore, one should be skeptical about either its apparent syntactic occurrence or its open plausibility of having multiple interpretations; its “polysemous nature” (Karmakar et al., 2017). Since, *ar* as a polysemous item has already been addressed in earlier works by Karmakar et al. (2017) followed by my own work presented in ICOLSI-42, 2020 conference, I would not repeat the same here, rather, in this paper I shall argue for a transparent picture of how every plausible interpretation; coordination, adverbial (and determinative as we shall see soon) of *ar* can be determined for which I assume its restricted and unrestricted syntactic distribution is solely responsible.

Only one or two familiar available research works on *ar* make/s it intriguing to reconsider the existing research from a different perspective. The identification of this linguistic expression as a “coordinator” in Bayer & Dasgupta (2016) followed by some basic work in Karmakar et al. (2017) have been the preliminary base of this present investigation. Later on, a conference paper in ICOLSI-42, 2020 carried out by myself can be considered as the backbone based on which this present study is built up.

However, distributing *ar* in this way, namely restrictedly and unrestrictedly follows from the framework originally proposed in Crnić (2019; part 1 and part 2) followed by Lin et al. (2021) among many others. Interested readers should refer to their original works cited in the reference section below for more details. Following their framework investigating *ar* has some empirical consequence, as we shall eventually see that it would be convincing to carry out the present study following their proposed framework. As most of the examples will show a negative licensing context, one might consider *ar* as an NPI or negative polarity item. However, at this position me and my supervisor agree to term it a species of NPI until we finalize how exactly that sounds okay. Until then, calling it an NPI seems justifiable. Seeing it from an NPI perspective *ar* shows significant similarities with previous works by Crnić (2019) & Lin et al. (2021) especially.

I have articulated above that *ar* is a crucial linguistic element among many other identified functional elements or noted discourse particles in (Bayer et al. 2014). It also behaves differently from ‘*naki*’ as argued in Bhadra (2017). What makes this study interesting and isolates *ar* from discourse function is that unlike other particles it is free to roam around across the clause and land anywhere. That is why it was mentioned above that the apparent appearance in every plausible position, with each word is ‘a misconception’. However, it is true that its occurrence across the clause does not affect its semantic interpretation, in this case it must an adverbial one. Therefore, the claim made in (Bhadra, 2017) for any other particles to occur clause initially is banned is contradictory with *ar*’s.

Using ‘*to*’ at the clause initial position to meet the contextual demand may be found in spoken Hindi as argued in (Montaut, 2015). Similarly, in Bangla ‘*to*’ might contextually be used clause initially (Bayer et al. 2014). But it would be unquestionable, I assume, to show the clause initial occurrence of *ar* that makes it significantly distinct from other function category elements. It shows that *ar* may not be treated as a discourse particle in this paper, but it would argue how its various meanings can be sensitive to its respective syntactic distribution. This generalization, then, suggests, all the different semantic

interpretation that one might get from *ar* can be derived from its syntactic distribution which I propose happens in two ways; restricted and unrestricted.

I am anticipating some relevant questions that might occur while treating *ar* in this way, which would be linguistically appropriate. But treating it appropriately, covering every aspect may not be possible in a single paper. Thus, this paper has the following limitations. I would not address how does it qualify for a discourse particle, neither would I address how many adverbial interpretations it might provide us, but a hint can be given; I have noted 14 different adverbial meanings of *ar* that can be derived from different contextual uses. These multiple adverbial interpretations both correspond to its NPI and PPI (positive polarity item) status. Even more striking phenomenon, on which I am currently working, concerns its canonical syntactic position in clause. Therefore, this paper could be considered as an initiative to address the issue of investigating its canonical position. To do this, it is necessary to differentiate the functions it posits before us and how does this happen.

Method

This novel approach is an empirically data-based study. Following the framework, analyzing the data, doing a comparative study with other languages show that somewhat phonologically identical linguistic element can be found across languages, even though I have considered only three in this study; Hindi, Odia and Bhojpuri. These languages are spoken in geographically approximated regions across India. Thus, different language convergence might influence the phonological status of the *ar* like elements found in the noted languages. Besides that, the data from Bangla language that have been provided here are expected to have some similarities with earlier works in [Karmakar et al. \(2017\)](#) including mine own ([Halder 2020](#)). To avoid plagiarism, I have used different contexts and other relevant linguistic elements.

Every set of data provided in this study have been collected and cross checked with Bangla, Hindi, Bhojpuri and Odia native speakers across India. After the data were gathered together, I had consulted with many different native speakers for accuracy judgements. Initially, the informants were provided with the English sentences and were asked to translate into their respective languages. This test was expected to result if it is the exact same item that means differently when occurred in different syntactic position or there exists some different linguistic elements for each job. The result was impressive and met the hypothesis; phonologically and functionally identical item serves the purposes. The data were share with the informants using social media platforms such as WhatsApp and they were then interviewed using phone calls, Zoom meetings, Skype calls and G-Meet.

During my stay at village-home at Murshidabad district, West Bengal, India, I have observed the actual conversation people were engaging in. A little input of *ar* resulted significant semantic differences. This stage of experiment shown that the natives know contextually appropriate use of *ar* especially. Their way of acquiring this particular linguistic item enables them to use it distinguishably such as when *ar* needs to introduce a new topic as a topic changer, or when to add something with previously given

information as an additive particle, or when to change the discourse of an ongoing conversation as a discourse particle and so on. Therefore, the expectation of treating *ar* properly covering every aspect is really hard to carry out in a single paper. It proves to be a much harder job to provide an accurate explanation when a five-year old child uses *ar* or answers just like adults to an *ar*-contained context, however, not entirely. At my home, I had conducted this experiment on my two five-years old nephews, their response to my *ar*-contained utterances were impressive. This stage also shows that the natives acquire this item more systematically. Therefore, the limited observations that this investigation is aiming at concerns if these meaning multiplicity of *ar* is caused by its different syntactic distribution. This observation leads to the next question that deals with what are its different syntactic distribution to which I propose two; restricted and unrestricted.

The paper has been organized in the following ways. It starts with an abstract followed by the keywords important for this paper. A short introduction of *ar* has been given in section 1 followed by the methodology in section 2. The methodology section shows how I collected the data, consulted with native speakers and how different stages of data collection has led to formulation of hypotheses. In section 3 I have specified the research problems addressed in this paper. This section is intended to isolate the present investigation from other relevant aspects for the limitation's sake. In section 4, I have provided an elaborate discussion with data from three other different languages other than Bangla, namely Hindi, Bhojpuri and Odia. This discussion shows that cross linguistic it is plausible to find out the same, however, phonologically distinct, linguistic element with identical purpose. In section 5 I present the findings and results from the discussion section followed by a conclusion of the discussion in section 6. This paper then ends with a list of works I have cited while preparing this paper.

Research Problems

Lack of existing works on *ar* in particular, in this language Bangla makes it difficult to address an issue properly. I assume that this specific linguistic item is much powerful and extremely crucial than one could think of. This assumption has the following consequences; there is no significant study on its syntactic distribution in earlier works, neither it has been received much attention from an appropriate linguistic perspective. Thus, choosing the present topic has some risks. The risks contain its restricted and unrestricted syntactic distribution that I am aiming to enunciate in this paper has no prior syntactic judgement. Therefore, conducting an experiment purely on empirical data basis might lack some theoretically important issue or there might be a need to develop a new theory that would explain the syntactic phenomenon of *ar*, both are equally possible.

As I have articulated above, this paper is an initiative based on empirical data to show the multiple manifestations that *ar* characterizes in communication can be determined based on its syntactic distribution. And it has also mentioned before that this distribution happens restrictedly and unrestrictedly. So, this paper aims at addressing the following research questions.

So, the first question in this paper concerns why the syntactic distribution of *ar* should be restricted and unrestricted?

The second question concerns what triggers the aforementioned distributions?

The third question concerns if there is a canonical syntactic position of *ar* in a clause?

Addressing the third research question is beyond the scope of this paper. It is also hard to tell what its canonical syntactic position would be. However, the former two questions such as why *ar* has to be restrictedly and unrestrictedly distributed and what is the motivation behind this could be covered in this paper.

Discussion

The investigation begins with three different sentences containing the same linguistic expression *ar*, apparently at the exact same position (right after the subject noun) yet they mean different from the other. At this point two natural questions arise; first, either the apparent syntactic position that is right after the subject is misleading or two, as a polysemous item *ar* shows multiple associated meanings sitting exactly at the same position. Since, the second approach has been articulated in some earlier works by [Karmakar et al. \(2017\)](#) including one of my own research papers, we are left with the first approach that is the apparent syntactic position of *ar* should be questioned.

As a *coordinator* ([Bayer & Dasgupta 2016](#)) *ar* conjoins two or more similar constituents together [Carnie \(2012\)](#), hence, obeys the principle of coordination. This phenomenon is noted in (1) above. As an adverb, it modifies the verb phrase (VP) in (2) and (3). Although, two different adverbial meanings have been noted in respective examples depending on the action types (habitual or recent) and the tense inflection marker with the verb. We shall see in this later section that *ar* also functions as a determiner 'more' where it takes a DP as its complement. Therefore, we have three different interpretations we can derive from *ar* based on its syntactic position. Let us see each one separately.

In (1) above two NPs were conjoined together which in turn forms a larger (Complex) Noun Phrase ([Carnie 2012](#)) out of which the movement of *ar* would lead to ungrammaticality. It then follows from this NP conjugation (1) that as a conjunction *ar* conjoins two similar constituents together as argued in [Carnie \(2012\)](#), we can call it the principle of coordination following which we see that two VPs (4), two DPs (5) and two TPs (6) can be conjoined together. Consider the following as examples;

4. Danny [[semantics por-che] **ar** [gan gai-che]]
 Danny.Nom.3 semantics studying and song singing
 Danny is reading semantics and singing a song
5. Tabil-e [[du-to pen] **ar** [tin-te khata]] ache
 Table.Loc two.CL pen and three.CL copybook have
 There are two pens and three copybooks on the table
6. [[Sam ranna kor-che] **ar** [Paul TV dekh-che]]

Sam.Nom.3 cook doing and Paul.Nom TV seeing
Sam is cooking a dish and Paul is watching TV

It is because as a conjunction *ar* is connecting two similar constituents together, it functions like a glue that establishes the relation between the two conjuncts it is conjoining together, thus its alternation elsewhere results ungrammaticality;

7. * (**ar**) Sam __ Dan (**ar**) aam (**ar**) kha-chhe (**ar**)
8. * (**ar**) Dan (**ar**) semantics porche __ gan (**ar**) gai-che (**ar**)
9. * (**ar**) Tabil-e (**ar**) duto pen __ tinte (**ar**) khata (**ar**) ache (**ar**)
- 10.* (**ar**) Sam (**ar**) ranna korche Paul TV (**ar**) dekhche (**ar**)

On a similar account, Hindi *aur* (11) Bhojpuri *auri* (12) and Odia *aa* (13) respectively, shown in bold face, have their coordinative function when they occur in between two similar constituents. Apparently, a phonological identity among such particles can be established which might have been influenced by language convergence among geographically related languages. Consider the following instances;

11. [[Sam] aur [Danny]] aam khaa rahe hain (Hindi)
Sam.Nom.3 and Danny.Nom.3 mango eat. Prog.Pl. is
Sam and Danny are eating mangoes
12. [[Sam] auri [Danny]] aam khaat baatan (Bhojpuri)
Sam.Nom.3 and Danny.Nom.3 mango eat Prog.Pl
Sam and Danny are eating mangoes
13. [[Sam] aau [Danny]] amba khau-chanti (Odia)
Sam.Nom.3 and Danny.Nom.3 mango eat.Pres.Prg
Sam and Danny are eating mangoes

Since, the respective conjainers (11-13) are establishing a connection between two NPs, its conjunctive function behaves like *ar* in Bangla. What is interesting to notice is that the above evidence suggests that the coordinative functions by respective linguistic particles can only be determined after the distribution of such particles have taken place, not a priori. This generalization gives birth to a question that says what would be the syntactic identity of respective particles noted in four different languages? If we are to determine their syntactic category after their distribution have taken place and not before that, then the question is what category do they belong to? This question resembles the first question formulated in our research problem section above. It will eventually be clearer that there are absolute reasons to see *ar* in this way; after its distribution took place. So, the *ar* that appears in between similar constituents that are conjoined together by it is identified as a conjunction. Since, a conjunction outside of its conjuncts would lead to ungrammaticality, we can say that in this way distributing *ar* is syntactically restricted.

As a VP modifying adverb *ar* is seen to have occupied a clause internal position (2-3). One should remember in this case that *ar* is free to occur anywhere across the clause yet its adverbial meaning can be preserved. This across clause occurrence of *ar* could be the

result of scrambling found in this language, thus, the example (2) can still have the same meaning even if *ar* occurs with each word.

14.¹**ar** Saul Delhi ja-be -na
again Saul.Nom.3 delhi.Loc go.Fut not
Saul will not go to Delhi again

15.Saul Delhi **ar** ja-be na

16.Saul Delhi ja-be **ar** na

17.??Saul Delhi ja-be na **ar**

This phenomenon of *ar* is much powerful than any other linguistic elements in this language Bangla. To my knowledge there is no such particles that can cooccur with literally every word in a clause, of course, if we are not concerned with respective constituents that might be followed by some repeated linguistic items. Thus, these multiple occurrences of *ar* (14-17) might be misleading, because no matter where it occurs its adverbial meaning seems to be preserved (14-17). Thus, one might be curious to identify its origin in the clause, as I am currently working on this project. So, we see that unlike coordinative *ar*, the adverbial *ar* does not need to be restrictedly distributed, so we can call the adverbial *ar* is unrestrictedly distributed. As soon as it preserves the adverbial meaning by its elsewhere occurrence calling it unrestrictedly distributed may not be controversial. However, there is room for debate.

A noticeable phenomenon of adverbial *ar* is that it can be interpreted with two different adverbials namely 'again' and 'anymore' as noted in (2) and (3) respectively. Besides these two adverbials in separate clauses, the adverbial *ar* can be ambiguous between the two in a single clause as shown in (18) below.

18.Saul **ar** Delhi ja-be na
Saul.Nom.3 AR delhi.Loc go.Fut not
Reading 1: Saul will not go to Delhi **again**
Reading 2: Saul will not go to Delhi **anymore**

It would be a theoretical escape to argue that depending on the contextual information one can distinguish between the two adverbial meanings shown in (18). In such cases, one might argue that native Bangla speaker-hearers possess a mutually shared CG knowledge (Krifka 2006; 2015) that enables them to choose one particular (and in most cases the intended one) adverbial meaning rejecting another. So, it seems that even if (18) is ambiguous between two adverbial meanings, using CG knowledge the native speakers can make an educated guess and could avoid falling into the ambiguous trap. Thus, it sounds reasonable to use the contextual information to disambiguate (18) based on the speaker's CG knowledge. So, if context dependency would have been the right

¹ With *ar* at the clause initial position is ambiguous in a different sense other than the two possible adverbials. One can read (14) as 'And, Saul will not go to Delhi' which can be done by a pause after pronouncing *ar*. Or one can read it as 'It is unlikely for Saul to do to Delhi again' which roughly means 'Saul will not do to Delhi again'. So, if one is taking a pause after *ar* it may mean the conjunction 'and' to which something is following, or it could be the adverbial 'again'.

path to make a distinction between two noted adverbial interpretations of *ar* then why context sensitivity does not work in the following cases;

19. Sam ar aam kha-i -na²
 Sam.Nom.3 anymore mango eat.Pres.3 not
 Sam does not eat mango **anymore/*again**
20. Sam ar aam kha-be -na
 Sam.Nom.3 again/anymore mango eat.Fut not
 Reading 1: Sam will not eat mango again
 Reading 2: Sam will not eat mango **anymore/again**

While *ar* in (19) unambiguously means the adverb 'anymore' and not 'again', in (20) it could ambiguously mean both. The possibility of perceiving *ar* ambiguously arises because of the tense inflection marker that the V takes, in this case, the present tense marker "*~i*". This certainly does not mean any present tense marker could give rise to an ambiguity. If this were to be the case then we would have an ambiguous sentence like (19) without an *ar*. So, this argument suggests that the relation *ar* forms with the VP, in some sense may create an ambiguity just the way they did in (19). However, just like (2) *ar* in (3) can also occur with each word in the clause, thus we can get the same adverbial 'anymore' meaning across its clausal positions as can be seen in the following (21-24).

21. **ar** Paul aam kha-i na
22. Paul aam **ar** kha-i na
23. Paul aam kha-i **ar** na
24. Paul aam kha-i na **ar**

Note that (21) is ambiguous in the same way it was in (14) above. Similar to Bangla adverbial *ar* Hindi, Bhojpuri and Odia languages use the same lexical item that gave a conjunction meaning also gives adverbial meanings thus, we see their adverbial interpretations in (25-27) below.

25. Paul **aur** aam nahi khata (Hindi)
Paul.Nom.3 more mango not eats
Paul doesn't eat mango **anymore**
26. Paul **auri** aam naahi khaalaa (Bhojpuri)
Paul.Nom.3 now mango not eat
Paul doesn't eat mango **anymore**
27. Paul **aau** amba khau-ni (Odia)
Paul.Nom.3 anymore mango eat not
Paul doesn't eat mango **anymore**

² One can argue that instead of 'na' using 'ni' might give adverbial 'again' meaning which is true but the with 'ni' the adverbial would mean something like 'ever again' as in "*Sam ar aam kha-i ni*" which roughly means either 'Since then Sam has not eaten mango again' or 'Sam has not eaten mango ever again', which I assume entails the 'anymore' meaning. However, the debate is open to other judgements.

Their (*aru*, *auri* and *aau*) alternations in respective languages noted in (25-27) would not change the adverbial meanings they express. Therefore, we might consider the adverbial *ar* in Bangla and *aru*, *auri* and *aau* in Hindi, Bhojpuri and Odia languages respectively have been unrestrictedly distributed, they can occur clause initially, internally and finally without affecting their adverbial interpretations.

Apart from the conjunctive and adverbial meanings of *ar*, we can also see it functioning as a determiner. In this case, as a determiner *ar* takes a DP as its complement across the noted languages.

28. Sam [ar [du-to aam]] nie bari phir-e gelo (Bangla)
Sam.Nom.3 more two mango take house.Loc return went
Sam went back home with two more mangoes

29. Sam [aur [do aam]] lekar ghar gaya (Hindi)
Sam.Nom.3 more two mango take house.Loc went
Sam returned home with two more mangoes

30. Sam [auri [dui-tho aam]] leke ghar-e gail (Bhojpuri)
Sam.Nom.3 more two.CL mango take house went
Sam went home with two more mangoes

31. Sam [aaau [dui-ta amba]] neiki ghara-ku gala (Odia)
Sam.Nom.3 more two.CL mango take house.Loc went
Sam went home with two more mangoes

As a determiner, *ar* in Bangla (28), *aur* in Hindi (29), *auri* in Bhojpuri (30) and *aaau* in Odia (31) takes a DP as its complement. With the following DP *ar* forms a constituent, a larger DP constituent. And there is no way *ar* can be extracted to some other places from the larger DP constituent, hence, the following would be ungrammatical (33), neither can any part of this larger DP constituent can move to some other position (32).

32. * Sam duto ar aam niye bari phire gelo

33. * ar Sam __ duto aam (ar) niye (ar) bari (ar) phire (ar) gelo

However, the DP constituent can be moved to clause initial position for focus reasons as shown in (34). In (34) the larger DP to which focus sensitivity has been applied can occur clause initially '*ar duto aam*'.

34. [ar duto aam] niye Sam bari phire gelo
more two.CL mango take sam.Nom.3 house return went
Sam went home with two more mangoes

The possible unmarked order from which (34) has been derived is shown in (35) where the moved DP occurs post verbally. Notice that there exists another verb "*niye*" which meaning 'take' also moves with the focus sensitive DP. This could mean that it is not only the DP to which the focus sensitivity has been applied rather it is the smaller embedded clause to which the focus sensitivity has been applied, thus the movement of this smaller clause in (34) is okay.

35. Sam bari phire gelo **ar** duto aam niye
 Sam.Nom.3 house return went more two mango take
 Sam went home with two more mangoes

Following (14) and (21) one might argue that the clause initial *ar* in (34) should also be ambiguous; it may mean both a conjunction ‘and’ and the determiner which is true. Reading (34) as in (36) opens this possibility.

36. **ar**..., [duto aam] niye Sam bari phire gelo
 and two mango take Sam.Nom.3 house.Loc return went
 and, Sam went home with two more mangoes

So, if the clause initial *ar* is meant to be a conjunction ‘and’ then does not contradict with the principle of coordination we have argued earlier? This question triggers two primitive questions; either clause initial is an absolute ambiguous item that gives rise to ‘and’ and ‘adverbial’ or ‘determiner’ meaning, or it might function as a discourse particle as argued in [Bayer et al. \(2014\)](#). Using one’s mutually shared CG knowledge s/he can distinguish between the possibilities. But it is likely that when *ar* is unrestrictedly distributed, hence, occurring clause initially without a DP complement, it functions as an adverb. And when it takes a DP complement it is likely to have a determinative meaning. As a topic introducer or to function as a discourse particle it must be followed by a pause as in (36). Keeping aside the discourse function for the moment, if we see its determinative function, we see its distribution is restricted.

So, the discussion on restricted and unrestricted syntactic distribution of *ar* in Bangla language up to this point gave us three types of *ars* for consideration. The first type corresponds to conjunction ‘and’, the second to adverbs ‘again’ and ‘anymore’ and the third to the determiner ‘more’. Recall that we had considered the coordinator *ar* as restricted, the adverbial *ar* as unrestricted but we are yet to determine the third type. In the third type, the determiner *ar* also has restricted distribution.

The evidence we have seen so far, *ar* corresponds to syntactic and semantic distribution simultaneously. This means from earlier discussion it is plausible to anticipate the meaning of *ar* based on its syntactic distribution. If it occurs in between two similar constituents then it is likely to be a coordinator one, thus it is restrictedly distributed. On a different account, when *ar* takes a DP as its complement it functions as a determiner, thus once again its distribution is restricted. But when it occurs preverbally or clause initially or finally it functions as an adverb. And since, it can freely occur across the clause, we might consider this distribution as unrestricted. This generalization on *ar* shows direct syntactic and semantic correlation, until we encounter with more crucial instances of *ar*. which I call the ‘mysterious *ar*’, where its syntactic distribution is fixed and nonobligatory but its semantic interpretation is unlikely to be determined. Consider the following examples;

37. .. ey chol-e elam **ar** ki
 this walk came AR what
 I/we Came just like that (to surprise you)

38. ... pelam **ar** koi go
got AR where GO
Didn't get yet (even after trying so hard)
39. ... tay bol-chi-lam **ar** ki
So saying AR what
So, I was saying (in case you would listen to me)

These utterances are colloquially used in response to show surprise (37), out of desperation (38) and being a victim of ignorance (39). However, contextually the interpretations may vary across their uses. What is striking in these examples is that the *ar* is in no way optional, its syntactic distribution is fixed just like its restricted distribution as a coordinator and as a determiner. But, in the given utterances they hardly mean anything yet its presence is strongly required or else we ran into ungrammatical constructions (40-42).

- 40.*... ey chol-e elam __ ki
41.*... pelam __ koi go
42.*... tay bol-chi-l-am __ ki

I am assuming some conventional correspondence for the *ar* + *ki* (or some other WH-words) sequence in such utterances. My point is, in such occurrences the presence of *ar* is obligatorily required even if their semantic correspondence is hard to determine. At the same time the *ar* has its position fixed, it is unmoved to some other positions, hence, we can conclude its distribution has been restrictedly happened.

Findings

This investigation enunciates the different meanings of *ar* and *ar* like elements found in other noted languages like Hindi, Bhojpuri and Odia respectively, that can be determined based on their respective syntactic distribution. The discussion suggests that such *ar* like particles to their very extent are polysemous. But this polysemous nature arises because of their contextual uses and the elements they are taking as their complements. Just because *ar* is a polysemous item it does not necessarily follow that it has to be distributed across the clause. As we have seen that the apparent elsewhere occurrence of adverbial *ar* is a misconception as long as it preserves the adverbial meaning.

This generalization might seem to miss out the focus sensitivity effect due to which the movements of other elements in a clause took place, which is true, for the space and time I have intentionally avoided introducing the focus sensitivity effect that triggers the movement of other elements pre-*ar* but in relevant portion I have mentioned it. With limited data set a comparative study with other languages Hindi, Bhojpuri and Odia with Bangla language shows that the use of same particle in different positions can mean different. One can conduct a typological study to find out more relevant similarities and dissimilarities among the *ar*-like particles whose semantic function can be anticipated based on its syntactic position.

We have seen four different *ars* in Bangla; first one corresponds to conjunction ‘and’ hence, we call it a coordinator *ar* following Bayer & Dasgupta (2016), second one corresponds to adverb/s ‘again’ and ‘anymore’ (omitting other possible adverbs) hence, we call it adverbial *ar*, third one corresponds to determiner ‘more’ hence, we call it a determiner *ar*, the fourth one that corresponds to the ‘mysterious *ar*’ is semantically hard to determiner at this moment, so let us call it the ‘mysterious *ar*’. All these four types of *ars* have something in common; their syntactic distribution is either restricted or unrestricted. Therefore, the apparent four types of *ar* can be syntactically reduced to two types; restricted and unrestricted.

Conclusion

So, the purpose of distributing *ar* in this way is to minimize the many different positions in which it is likely to occur for many different purposes. Similar constituents internally, clause internally, DP internally and in between some specific utterances where it is semantically null would be hard to deal with, so if we can determine its specific occurrences by reducing the many into few then we can also anticipate its semantic functions, to some extent. Of course, the tough question that we did not see it discussed here is the origin of *ar* in a clause. Notice that we may not go far with contextual dependency, we must acknowledge the linguistic factors that are responsible for making an utterance ambiguous as in the discussion we have seen some. The linguistic requirements are what triggering the *ar* to be distributed in two different ways. In case of conjunction, to establish a connection between two similar constituents, for adverbial as a VP modifier and for a determiner a DP modifier *ar* has been distributed. This seems very difficult to address the third question we asked in research problem section that concerns with determining its canonical position in a clause. I am assuming that even if we are to determine its canonical position in a clause it must go through the respective restricted and unrestricted syntactic distribution.

Acknowledgments

N/A.

References

- Bayer, J., Dasgupta, P., Mukhopadhyay, S., & Ghosh, R. (2014). Functional Structure and the Bangla Discourse Particle “to”. SALA 30, Hyderabad, 06-08-2014
URL:https://ling.sprachwiss.uni-konstanz.de/pages/StructureUtterance/web/Events_files/Bayer_Dasgupta_MukhopadhyayGhosh_SALA.pdf
- Bayer, J. & Dasgupta, P. (2016). Emphatic Topicalization and the Structure of the Left Periphery: Evidence from German and Bangla. *Syntax A Journal of Theoretical Experimental and Interdisciplinary Research*, 19(4).
<https://doi.org/10.1111/synt.12125>
- Bhadra, D. (2017). Evidentiality and Questions: Bangla at the Interfaces, Rutgers University PhD Dissertation. Lingbuzz/003695

- Carnie, A. (2013). *Syntax: A Generative Introduction*, Third Edition. Wiley-Blackwell, A John Wiley & Sons, Ltd, Publication.
- Crnič, L. (2019). Any: Logic, likelihood, and contexts (Pt.1). *Research Article. Lang Linguist Compass*. 2019;13:e,12354. <https://doi.org/10.1111/Inc3.12354>
- Crnič, L. (2019). Any: Logic, likelihood, and contexts (Pt.2). *Language and Linguistics Compass*. 2019;13:e,12353. <https://doi.org/10.1111/Inc3.12353>
- Halder, S. (2020). Polysemous 'ar' in Bangla. Conference paper presented on 42nd *International Conference of Linguistic Society of India (ICOLSI-42)* held from 10th -12th December, 2020. Organized by Department of English, GLA University, Mathura, India
- Karmakar, S., Ghosh, S. S., & Banerjee, A. (2017). Functional Significance of Contextual Distribution: Discourse Particles ar in Bangla. *Acta Linguistica Asiatica*, 7(1). <https://doi.org/10.4312/ala.7.1.23-39>
- Krifka, M. (2006). Basic Notions of Information Structure. Humboldt Universität zu Berlin and Zentrum für Allgemeine Sprachwissenschaft, Berlin. Interdisciplinary Studies on Information Structure 06 (2006):000-000. Fréy, Fanselow and Krifka (eds.) https://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka_InformationStructure.pdf
- Krifka, M. (2015). Quantification and Information Structure. Humboldt- Universität Berlin & Zentrum für Allgemeine Sprachwissenschaft (ZAS) Berlin. https://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka_QuantificationInformationstructure.pdf
- Lin, J., Weerman, F., & Zeijlstra, H. (2021). Distributionally restricted items Acquisition and beyond. *Nat Lang Linguist Theory*. <https://doi.org/10.1007/s11049-021-09502-5>
- Montaut, A. (2015). The discourse particle to and word ordering in Hindi: From grammar to discourse. In M. Fernandez-Vest & R. Van Valin (Ed.), *Information Structuring of Spoken Language from a Cross-linguistic Perspective* (pp. 263-282). Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110368758-013>

Author's Biography

Samrat Halder, PhD Linguistics and Phonetics, Department of Linguistics and Contemporary English, School of Language Sciences, The English and Foreign Languages University, Hyderabad 500 007, India | samratphdlandp18@efluniversity.ac.in