

The morphosyntax of allocutive agreement in Tamil

Thomas McFadden

Leibniz-Zentrum Allgemeine Sprachwissenschaft

August 15th, 2017

Abstract

In this paper I will examine allocutive agreement in Tamil, a phenomenon in which an agreement suffix attached to the verb or other clause-final element indexes features not of any argument, but of the addressee of the utterance. I will report in detail on the morphophonology, syntactic distribution and discourse use, supplementing the basic facts reported by Amritavalli (1991) with several additional crucial details, and will compare the Tamil data with what has been reported for other languages, especially Basque and Japanese. I will then discuss the consequences of Tamil allocutive agreement for the theoretical treatment of how discourse information interacts with the morphosyntax, leading to a preliminary analysis of the patterns I find. We will see that the Tamil data provide interesting insights into the structural representation of the addressee and into how allocutive agreement is derived, in particular from how the relevant suffix is ordered relative to other verbal material.

1 Background

1.1 Introduction by example

In many colloquial varieties of Tamil (Dravidian; South Asia), one commonly comes across utterances of the following kind:

- (1) Naan džangiri vaang-in-een-ŋgæ.
I Jangri buy-PST-1SG.SBJ-ALLOC
'I bought Jangri.'¹

Aside from the good news it brings to the hearer, (1) is of interest because it contains two different types of agreement stacked on top of each other. First, there is the suffix *-een*, which marks the unremarkable agreement of a finite verb with

¹ Jangri is a delicious sweet made by deep-frying a type of lentil batter in flower shapes and then soaking them in sugar syrup.

the subject that is found in a significant portion of the languages of the world. Second, there is the suffix *-ŋgæ*, glossed here as *ALLOC*, which marks a rather different kind of agreement that is far less widely attested. Specifically, rather than cross-referencing properties of one of the arguments of the verb, it provides information about the addressee, specifically that this sentence is addressed to either a group or an individual with whom the speaker would use the polite form of address. If addressed instead to a single person with whom the speaker would use the familiar form, this suffix would simply be lacking, as in (2):

- (2) Naan *ɟaangiri vaang-in-een*.
I Jangri buy-PST-1SG.SBJ
'I bought Jangri.'

As we will see directly, similar types of agreement with the addressee have been described for a number of other languages. It is most famously found in Basque (Oyharçabal 1993), where it is referred to as 'allocutive agreement', a term I will adopt here. Allocutive agreement is of considerable interest, both for the theory and typology of agreement systems, and for what it can tell us about the grammatical representation of speech acts and speech act participants. As such, looking into the precise distribution of and constraints on this kind of agreement will hopefully shed some light on currently ongoing discussion and controversy over the extent to which certain phenomena normally associated with discourse and semantics may actually have a morphological and syntactic side.

In this paper, I will present newly collected, detailed data on allocutive agreement in Tamil and compare it with what has been reported for other languages, including especially Basque and Japanese. We will see that the Tamil facts resemble those from the other languages in the broad strokes, but that there are a number of interesting points of detail where Tamil differs in ways that are potentially quite relevant for our theoretical understanding. I will discuss the data before the background of a recent body of work on the syntactic side of the representation of speech acts. Of particular interest here will be how allocutive agreement interacts with the phenomenon of monstrous agreement in the language discovered by Sundaresan (2012), where speech act participants also play a crucial role. I will conclude the paper with a preliminary analysis of the patterns and some discussion of how it can serve as a starting point for future investigations.

1.2 On allocutive agreement

Allocutive agreement (henceforth abbreviated as AllAgr), while far from common, has been identified in a number of languages from a wide selection of families (see Antonov 2015: for an initial typological overview). Adapting Antonov (2015) slightly, we can identify something as AllAgr if it has the following properties. First, it marks properties of the addressee of the current speech context, i.e. it provides information about the gender, number or politeness status of the person or persons to whom the utterance is directed. Second, it is not limited to cases where the addressee is an argument of the local predicate, so it is to be clearly distinguished from all forms of subject or object agreement, even instances where the 2nd person might behave in a special way. Third, it involves the use of grammaticalized morphological markers in the verbal or clausal inflectional system. This is meant to exclude e.g. special vocative forms like ‘madam’, ‘sir’ or ‘captain’ which may serve similar functions but are not grammaticalized in the same way and show different morpho-syntactic behaviors than true AllAgr.

As already noted, the classic example of AllAgr comes from Basque, for which the term was first introduced by Bonaparte (1862). In Basque, the use of AllAgr depends, in ways that vary across dialects, on the politeness relationship between the speaker and addressee as well as the number of the addressee, with the form reflecting the gender of the addressee (see also Oyharçabal 1993; Alcázar & Saltarelli 2014: ch. 5). In Standard Basque e.g., the agreement only crops up when the speaker and addressee would use the highly familiar form of an address, and then only when the addressee is singular and is not additionally an argument of the verb. The examples in (3) (from Antonov 2015), which all have the same basic meaning, but differ in the properties of the addressee, illustrate the phenomenon:

- (3) a. etʃe-a banu
 house-ALL 1.SG.go
 ‘I am going to the house.’
- b. etʃe-a banu-k
 house-ALL 1.SG.go-ALLOC:M
 ‘I am going to the house.’ (familiar male addressee)
- c. etʃe-a banu-n
 house-ALL 1.SG.go-ALLOC:F
 ‘I am going to the house.’ (familiar female addressee)

- d. etʃe-a banu-sy
house-ALL 1.SG.GO-ALLOC:RSP
'I am going to the house.' (respected addressee)

(3a) gives the baseline, where verb only shows 1.SG agreement with the subject. The remaining examples add allocutive suffixes to this verb form cross-referencing the addressee, in (3b) a familiar male addressee, in (3c) a familiar female, and in (3d) an addressee with whom the speaker would use the polite form.² These suffixes indicate information about the addressee independent of it being an argument. Furthermore, they are fully grammaticalized verbal inflection forms, appearing in the normal position for agreement in the language and involving (nearly) the same forms as those used to agree with a 2nd familiar ergative argument (see Antonov 2015: p. 66f. for discussion of the forms). What we have here thus clearly meets our criteria for AllAgr.

There are some additional interesting properties of Basque AllAgr that should be noted here. First, the appearance of the allocutive suffixes is not actually independent of the addressee being an argument, but rather requires that it is not. If the addressee *is* one of the arguments, it will be coindexed with the appropriate (ergative, absolutive or dative) 2nd person argument agreement, and AllAgr will not appear.³ Second, in contexts where the conditions for it are met, AllAgr is obligatory, i.e. we are dealing with a fully grammaticalized system, not some optional marking of familiarity or respect. Third, AllAgr is generally restricted to root clauses and, at least in many dialects, is not possible in questions.

Miyagawa (2017) has argued that Japanese politeness marking should also be analyzed as a type of AllAgr. Japanese has a range of constructions and markers belonging to its system of 'honorifics', which encode various types of social relationships between the speech act participants and different nominal arguments in a given clause. These include lexical choices and verbal affixes that reflect honorification toward the subject or the object which will not be of direct concern to us here, because they crucially involve arguments. There is also, however, verbal marking used to indicate politeness or honorification from the speaker toward the addressee, as in (4), from Miyagawa (2017):

² This last form is only in use in certain dialects. In others, including the standard, the form with no AllAgr in (3a) would be used with a formal addressee.

³ This plausibly reduces to the fact that Basque independently blocks a single referent from being coindexed with multiple agreements (e.g. in reflexives), and argument agreement is obligatory (Antonov 2015).

- (4) a. Watasi-wa pizza-o tabe-mas-u.
I-TOP pizza-ACC eat-ALLOC-PRS
'I will eat pizza.' (formal)
- b. Watasi-wa pizza-o tabe-ru.
I-TOP pizza-ACC eat-PRS
'I will eat pizza.' (colloquial)

Here again, the marker is clearly giving information about the addressee, independent of what the arguments of the verb are, and furthermore it is a clearly grammaticalized part of the verbal inflectional system, appearing as a suffix on the verb, inside of a tense suffix.

What makes the case here a bit trickier than the Basque one is that Japanese doesn't have straightforward argument agreement, never seeming to indicate standard ϕ -featural information (i.e. person, number or gender) about subjects or objects. Indeed, Japanese is usually regarded as an agreementless language. However, as noted above, the language *does* indicate honorification towards the subject or object in certain constructions, at least some of which have been argued to involve a type of agreement (see e.g. Boeckx & Niinuma 2004). Furthermore, Miyagawa (2017) argues that the lack of prototypical agreement elsewhere in the language should not at all dissuade us from recognizing the politeness marking as a type of agreement. If we assume that there is a universal set of grammatical features, which are overtly manifested in all languages (his principle of Strong Uniformity, from Miyagawa 2010), then the lack of ϕ -agreement on T actually leads us to *expect* ϕ -agreement on C, i.e. something like AllAgr. It just happens to be the case that the type of ϕ -features overtly manifested in the language have to do with honorification rather than person, number or gender.

AllAgr patterns have also been reported for Pumé (isolate; Venezuela), Nambikware (isolate; Brazil), Mandan (Siouan; North America) and Beja (Cushitic; Northeast Africa), as summarized by Antonov (2015). Beyond the criteria for identifying AllAgr we have already discussed, Antonov notes several points about the typology of the phenomenon. First, languages differ in what information about the addressee they encode, with gender and varying types of familiarity or politeness being perhaps most common, and number being rather less common. Indeed, in his sample it seems to be found only in Basque, and even here it is limited in most dialects to the fact that AllAgr only appears when the addressee is singular.⁴ Second, languages also differ in how AllAgr interacts with the ar-

⁴ That is, one can infer a partial singular/plural contrast from the fact that allocutive marking is always lacking when the addressee is plural, though the lack of marking may have other

gument status of the addressee. As noted above, it is ruled out in Basque when one of the arguments is 2nd person, but this restriction does not seem to apply in any of the other languages surveyed. Finally, while there is some variation in the distribution of AllAgr across clause types, there are clear generalizations to be made. The core environment, where AllAgr is found in all of the languages considered, is root declarative clauses. There is then a fair amount of variation across the languages in whether it is also found in other types of root clauses, i.e. interrogatives, exclamatives and imperatives. Basque, for example, excludes it in all of these, Beja allows it in all of them, and Japanese allows it in interrogatives and exclamatives, but not imperatives. Finally, in all of these languages, AllAgr is heavily restricted or entirely ruled out in embedded clauses. The details about embedding have been, as far as I am aware, most carefully examined for Japanese, and it is perhaps no surprise then that it is here that some embedded environments have been reported to allow the phenomenon. I will return to consider the issue once I have presented the relevant data from Tamil.

From a theoretical perspective, AllAgr is highly intriguing because it seems to involve a clear role for information about the speech act in the morphosyntax. It is clear that the identity of the author and addressee of an utterance, as well as its time and location, play a role in the semantic and pragmatic interpretation, and thus must be encoded somehow in the discourse context. This is necessary, among other things, for the appropriate interpretation of so-called indexical items, like 1st and 2nd person pronouns and expressions like ‘here’ and ‘now’. What is less obvious is whether we need to assume that a representation of such information is accessible in the syntax. AllAgr potential offers evidence that we do. One could argue that, if it really is an instance of morphosyntactic agreement, then there must be some representation of the addressee in the syntax for it to be agreeing *with*. As we will discuss in section 3.1, this has led to the idea in much recent work that AllAgr targets syntactic representation of the discourse context, a representation that has been argued to be necessary in independent grounds for example by Speas & Tenny (2003), ultimately going back to the ideas of Ross (1970).

The only prior work on AllAgr Tamil that I am aware of is Amritavalli (1991). That unfortunately all too brief festschrift contribution reports the central data, including a number of insightful observations, and compares the Tamil facts with those in the closely related languages Kannada and Telugu. However, it does not have the space to explore the data in detail, and there are a number of important

causes as well. There are, however, apparently some dialects that allow explicit marking of a plural addressee. See Antonov (2015) for brief discussion and references.

points that it does not touch upon. The current paper will attempt to fill this gap by providing a careful and extensive description of the empirical situation with respect to AllAgr in Tamil. We will see that it displays a number of properties in the language that are of theoretical interest, some of which distinguish it crucially from what has been reported for other languages.

1.3 Some relevant properties of Tamil

Tamil is a Southern Dravidian language, spoken by approximately 70 million people, primarily in southern India and Sri Lanka, as well as a significant diaspora, e.g. in Malaysia, Singapore, Mauritius and South Africa. As the second classical language of India after Sanskrit, it has a written tradition going back over two thousand years. For better or worse, the written standard is extremely conservative, approximating a rather archaic variety of the language, which differs significantly in all aspects of grammar and lexicon from contemporary spoken varieties. This leads to a marked diglossia, such that there is even a spoken version of the literary language used e.g. for newscasts and political speeches. While there is arguably a contemporary standard version of the spoken language, used e.g. in films and television talk shows (see e.g. Schiffman 1999; Asher & Annamalai 2002), there is no generally agreed-upon written form for this variety, nor is there any standard romanization. The AllAgr that is of interest here is very much a phenomenon of the colloquial language, not the written standard, and the form of the data presented will reflect this fact. I adopt essentially the transliteration used by Sundaresan (2012), which attempts to reflect the phonology of the standard colloquial variety, without going into too much phonetic detail.

Contemporary Tamil is also characterized by extensive dialectal variation, reflecting communities defined by geography as well as socio-economic and religious factors. Here as well, the variation is relevant for our considerations of AllAgr. For one thing, the phenomenon is largely restricted to non-Brahmin dialects, plausibly related to the fact that Brahmin dialects use different forms for 2nd plural and polite agreement than the one that is involved in AllAgr. I conducted sessions with three speakers of Brahmin Tamil from Chennai and Tiruchirappalli, in which I was able to confirm that they make little or no use of AllAgr in their native dialect.⁵ Furthermore, while the basic AllAgr patterns are found

⁵ However, they generally use a hybrid of Brahmin and non-Brahmin colloquial Tamil when speaking with non-Brahmins. In such circumstances they do make use of AllAgr, though to a more limited extent. Being regularly exposed to non-Brahmin varieties in the media and daily interactions, they also have passive command of its use by other speakers. I found in my sessions that they have clear intuitions about core uses of AllAgr, but are less certain about

in a wide array of colloquial varieties, there is some geographic variation in the frequency with which it is used. My primary informant is from Pollachi, in Coimbatore district, which is reputed to be an area that makes particularly heavy use of AllAgr. This has the advantage that he has quite robust intuitions about the phenomenon. I also collected preliminary data from two speakers of Singapore Tamil, which suggest that the phenomenon is more restricted there and also subject to different constraints regarding the ordering of affixes, as will be briefly mentioned below. Given all of these factors, I have chosen to focus here on the patterns found in the speech of my informant from Pollachi, as he was able to provide the most extensive and consistent data on AllAgr. Thus unless otherwise indicated, the examples provided here come from my sessions with him. Note that I have made no attempt to systematically investigate the dialectal distribution of the phenomenon, but am simply registering here that relevant differences do exist. Speakers of other dialects of the language should thus not be expected to agree with all of the judgments reported.

Regarding the research methodology, with my informants from Pollachi, Chennai and Tiruchirapalli, I used a questionnaire of pre-constructed sentences, combined with elicitation based on translation of English examples for some of the more complex structures. The questionnaire was based on a combination of my own prior observations of the phenomenon in naturally occurring speech, data from Amritavalli (1991) and additional sentences constructed based on patterns reported for other languages in the literature. With the two informants from Singapore, I collected basic judgment data on the core patterns in brief, informal interviews.

Tamil is a highly inflecting language with a strongly agglutinative character, though it shows some fusional tendencies, and is almost exclusively suffixing. Syntactically speaking, it is SOV and indeed quite generally head-final, allows pro-drop of all arguments and has long-distance anaphors. The language has a nominative-accusative case system with differential object marking, and distinguishes a total of six or seven cases marked by suffixes on nouns and pronouns in addition to the unmarked nominative. The pronominal system includes an inclusive/exclusive distinction in the first person and a local/distal distinction in the third person forms. Politeness is indicated by plural forms in the 2nd person and by distinct pronominal forms in the 3rd person (historically related to older plural forms), and occasionally also with plural marking on nouns. Three genders are distinguished – masculine, feminine and neuter – corresponding essentially to the notional status of the referent, and play a role in both the pronominal and

points of detail.

verbal agreement systems. The language has a range of complex predication constructions, and its verbs display an impressive array of participial and nominalized forms, many of which can head particular types of non-finite clauses, and often include aspectual marking. Finite verbs can be marked for transitivity, aspect, passive and middle voice, mood, negation, tense and agreement (Sundaresan & McFadden 2017). There are, however, interesting restrictions on co-occurrence, as e.g. mood, negation and agreement are essentially in complementary distribution (Amritavalli & Jayaseelan 2005).

Let us focus then on agreement. Standard verbal agreement targets the highest nominative argument in the clause, which is typically the subject, but may also be an object if the subject is marked with a (quirky) dative or locative case, as in (5) from Baker (2015).⁶

- (5) En-ækkü andæ poŋŋü teeve-ppaḍ-r-aa
 I-DAT that girl.NOM need-suffer-PRS-3FS
 ‘I need the girl’

The form of the agreement reflects person and number, as well as gender in the 3rd person and politeness in the 2nd and 3rd persons. The inclusive/exclusive distinction is *not* reflected by agreement, there being a single 1st plural form used for both. Table 1 shows the regular agreement paradigm with an example of the simple present tense and imperative forms of *oodü* ‘run’.⁷ The agreement

Table 1: Regular verb agreement in Tamil

	SG	PL
1	oodü-r-eeŋ	oodü-r-oom
2	oodü-r-æ	oodü-r-iingæ
3F	oodü-r-aa	oodü-r-aangæ
3M	oodü-r-aan	oodü-r-aangæ
3POL	oodü-r-aarü	oodü-r-aangæ
3N	oodü-dü	oodü-dü
IMP	oodü	oodü-ŋgæ

⁶ Transliteration and formatting have been modified to fit the system used elsewhere in the paper.

⁷ The *-r-* suffix found before the agreement suffixes outside of the 3rd neuter forms marks the present tense. Tense marking generally interacts in odd ways with neuter agreement.

suffix follows all aspect, tense and voice markers. We can see an example of a moderately complex, fully inflected finite verb in (6), where the combination of an aspectual marker followed by the suffix *kittũ* and a form of ‘be’ forms a progressive, to which tense and agreement are further suffixed.

- (6) Kausalya paḍi-ččũ-kiṭṭũ-ru-nd-aa
Kausalya study-ASP-PROG-BE-PST-3F.SG
‘Kausalya was studying.’

The finite verb, terminated by the agreement suffix, is typically the final element in a root declarative clause (aside from extraposed material), but it can be followed by further suffixes that we might expect to be in the C domain, e.g. the complementizer *-nnũ* as in (7a) or the polarity question particle *-aa* as in (7b):

- (7) a. Venkaṭ [Kausalya paḍi-ččũ-kiṭṭũ-ru-nd-aa]-nnũ so-nn-aan
Venkaṭ [Kausalya study-ASP-PROG-BE-PST-3F.SG]-COMP say-PST-3M.SG
‘Venkaṭ said that Kausalya was studying.’
b. Kausalya paḍi-ččũ-kiṭṭũ-ru-nd-aa]-aa?
Kausalya study-ASP-PROG-BE-PST-3F.SG-Q
‘Was Kausalya studying?’

The *-l* that suddenly appears before the question particle in (7b) is part of the underlying form of the agreement suffix, which is deleted in coda position, but surfaces when a vowel-initial suffix immediately follows within the same word. This is a common phenomenon in the morphophonology of Tamil, and crops up also in the various plural agreement forms ending in *-ŋgæ*, which surface as *-ŋgæl-* before vowel-initial subjects. As we will see, this includes AllAgr marker.

2 The core data

In this section I will present the empirical details on Tamil AllAgr, again based on a combination of what was already reported by Amritavalli (1991) and what I have collected in the work with my informants. We will see that it meets all criteria to be considered genuine allocutive agreement, but that it also shows interesting details in its behavior that distinguish it from what has been reported for Basque, Japanese and other languages.

2.1 The morphophonology of the suffix

A central part of the argument that the Tamil phenomenon of interest here really is a type of agreement, rather than e.g. a specialized vocative (along the lines of English *sir/ma'am* or certain uses of *guys*) or a speech act particle (like those discussed for Romanian and West Flemish by Haegeman & Hill (2011)), comes from the form and position of the actual marker. I repeat here the basic example from (1) above as a basis for the discussion:

- (8) Naan *ɕaangiri vaang-in-een-ŋgæ*.
 I Jangri buy-PST-1SG.SBJ-ALLOC
 'I bought Jangri.'

As we see, the shape of the allocutive suffix is *-ŋgæ*, which turns out to be a rather general plural marker throughout the language. Looking back at Table 1, we see that it is the final component of all of the 2nd and 3rd person plural agreement markers (setting aside the 3rd neuter marker, which simply doesn't distinguish plural from singular), and it is the marker attached to the verb root to form (2nd) plural imperatives. Furthermore, it, or the related form *-gæ*, is used as the plural marker in a number of nominal categories, as shown by the examples in Table 2. In the (rather common) case that a noun or pronoun stem

Table 2: Number marking in Tamil nominals

	SG	PL
1EXCL	naan	naaŋgæ
2	nii	niingæ
3M	avan	avaŋgæ
'girl'	poŋŋũ	poŋŋũgæ
'tree'	maram	marangæ

ends in a nasal, it is impossible to tell whether the plural suffix is *-gæ* or *-ŋgæ*.⁸ But even with vowel-final nouns, there is a fair amount of variation between the two. Interestingly enough, the two main instances where the plural ending is unambiguously *-ŋgæ* are both in the 2nd person, in the 2nd person pronoun itself (since 2nd singular *nii* ends in a vowel) and in the plural imperative suffix,

⁸ Note that in all of these cases, whether with *-gæ* or *-ŋgæ*, there is an underlying final *-l*, which surfaces when any vowel-initial suffix follows.

which is *-ngæ* regardless of what the verb root ends in. Thus what we can say to summarize all of this is that *-ngæ* is a plural ending which always occurs in the 2nd person and variably occurs elsewhere. Note again that the 2nd plural forms are also used for politeness with singular addressees.

Example (8) also demonstrates that the allocutive marker attaches to the clause-final verb, after all of the other inflectional suffixes that might precede it, including tense, aspect, voice and argument agreement. This is also true when the verb has a modal or negative suffix rather than agreement. Again, the allocutive marker follows at the very end of the verb form, as demonstrated by the sentences in (9), based on examples from Amritavalli (1991).

- (9) a. *ko.ɽandæ ippaɽi sejjæ-kkuuɽaadü-ngæ*
child like this do-must not-ALLOC
'The child should not act in such a way.'
- b. *Venkaɽ varæ-læ-ngæ*
Venkat come-NEG-ALLOC
'Venkat didn't come.'

As we will see in more detail below, the marker can also appear in clauses without a verb, like (10a), and even in fragmentary or elliptical utterances that aren't even clauses, as in (10b) and (10c):

- (10) a. *naan aaɽtookkaaran-ngæ*
I automan-ALLOC
'I am an auto rickshaw driver.'
- b. *indæ payyan-ngæ*
this boy-ALLOC
'this boy' (e.g. as answer to 'Who's next?')
- c. *illæ-ngæ*
no-ALLOC
'No' (as answer to polar question)

It is important to note that the AllAgr marker can also co-occur with unambiguous vocatives. It occurs strictly attached to the verb, with the vocative obligatorily coming outside, which confirms that *-ngæ* itself cannot be a vocative:

- (11) a. *naan va-r-een-ngæ saar*
I come-PRS-1SG.SBJ-ALLOC sir
'I'll take my leave, sir.'

discussed by Antonov (2015) allow it. Tamil shows a somewhat mixed behavior, which is quite instructive. When the *subject* is 2nd person and it triggers regular argument agreement on the verb, AllAgr is ruled out, as shown by the examples in (13):

- (13) a. *eppaḍi iru-kk-iingæ-ṅṅæ?
how be-PRS-2PL-ALLOC
'How are you?'
- b. *niingæ rombaa smart-aa iru-kk-iingæ-ṅṅæ
you.PL very smar-PRED be-PRS-2PL-ALLOC
'You're very smart.'

The question is whether the problem here is the fact that the addressee is an argument at all, or that it is specifically the subject, or simply that it triggers agreement. This is resolved by the examples in (14), all of which involve 2nd person arguments combined with AllAgr:

- (14) a. naan onga|-æ paḍatt-læ paa-tt-een-ṅṅæ
I you.PL.OBL-ACC film-LOC see-PST-1SG.SBJ-ALLOC
'I saw you in a film.'
- b. onga|-ūkkū coffee veṅṅum-aa-ṅṅæ?
you.PL-OBL-DAT coffee want-Q-ALLOC
'Do you want coffee?'
- c. niingæ saap-ṭ-aach-aa-ṅṅæ?
you.PL eat-ASP-RES-Q-ALLOC
'Have you eaten?'
- d. *niingæ saap-ṭ-iingæ|-aa-ṅṅæ?
you.PL eat-PST-2PL-Q-ALLOC
'Did you eat?'

(14a) shows that AllAgr is perfectly fine with a 2nd person direct object, and (14b) shows the same with a quirky dative subject. Datives never trigger agreement in the language, so here the AllAgr is the only agreement with the addressee. We see something similar in (14c), where the main predicate of the clause is in a resultative participial form which doesn't host argument agreement. AllAgr only fails in cases like (14d), where there is a (finite, non-negative, non-modal, non-participial) verb form capable of bearing argument agreement, with a 2nd person subject in the nominative case, which thus triggers that agreement. At

least descriptively then, it seems that double expression of agreement with the addressee – both argument agreement and AllAgr – is ruled out.

Now let us consider the further conditions on the appearance of AllAgr, once we've restricted our attention to utterance contexts with the right kind of addressee and no 2nd person argument agreement. We've already seen that AllAgr *can* appear in root declaratives and various fragmentary utterances. Furthermore, unlike in at least some dialects of Basque, it can appear in root interrogatives. (14c) above shows a polar question, and (15) demonstrates its use in a *wh*-question:

- (15) ev|avũ aag-um-ŋgæ?
 how much become-FUT-ALLOC
 ‘How much will it come to?’ (i.e. ‘How much does it cost?’)

One crucial point in all of this is that, when its conditions are met, AllAgr is actually obligatory, at least for my Pollachi informant. I.e. when one would use *niingæ* with the addressee, only something like (16a) is possible. Leaving off the *-ŋgæ* signals non-politeness, and thus (16b) is ill-formed in such a discourse context.

- (16) a. rombaa thanks-ŋgæ
 very thanks-ALLOC
 ‘Thanks a lot’
 b. *rombaa thanks (to a polite addressee)
 very thanks

This is strong evidence that the use of *-ŋgæ* is fully grammaticalized as allocutive agreement.

Next, we must consider embedded environments, where AllAgr has been reported to be blocked or at least heavily restricted in other languages. Interestingly enough, Tamil seems to be more permissive here, though there are some complications in the judgments. As a starting point, examples like (17) are grammatical. Note that we have the *-ŋgæ* suffix showing up outside of the argument agreement *-aa[-*, and inside of the complementizer *-nnũ*, in a clause that is the complement of a verb meaning ‘say’.

- (17) Maya [avæ poot[ti-le dzejkkæ-poo-r-aa[-ŋgæ-nnũ] so-nn-aa
 Maya [she contest-LOC win-go-PRS-3SF-COMP] say-PST-3SF
 ‘Maya said that she would win the contest.’

Exactly which conditions must be satisfied to allow such embedded AllAgr is still under investigation, as the empirical situation and the relevant intuitions quickly get rather tricky. One recurring issue is that, when the AllAgr suffix immediately follows the argument agreement, as in (17), there are some processing difficulties that arise, in particular the tendency to interpret the two suffixes as a single plural argument agreement suffix. It thus takes some care to ensure that judgments of ungrammaticality do not reflect an unintended parse.

A more interesting complication with embedding, especially under attitude predicates, is that there is often some flexibility or ambiguity as to whether the embedded clause is interpreted for various purposes relative to the utterance speech act or relative to the speech act or attitude expressed by the matrix attitude predicate. This is relevant of course for AllAgr, because it expresses information about the addressee and potentially also the relationship between that addressee and the author (i.e. whether the latter would use the familiar or formal form of address with the former). Consider in this light example (17), assuming that the entire sentence has been uttered by Tom to Venkat, and that it is reporting on Maya saying the equivalent of ‘I’m going to win’ to Kausalya. We can reasonably ask now whether the embedded AllAgr reflects Tom showing respect to Venkat (the utterance speech act) or Maya showing respect to Kausalya (the embedded speech act). In this case, my informant reports that it can only reflect respect being shown by Tom in the utterance speech act toward Venkat. There are other cases where intuitions are more uncertain, and what we observe is clearly affected by things like the type of the matrix attitude predicate and the plausibility of the various scenarios in a given context. These concerns make it especially difficult to determine the constraints on when AllAgr is possible in embedding contexts, which readings are available, and what a speaker’s rejection of a particular example should be attributed to. For now I will thus simply report that AllAgr is possible in some complement clauses, but probably not in all, and leave a more complete investigation of the facts for future work.

There is, however, one point on which I will already say more, because the judgments here seem to be relatively clear, and the pattern is extremely interesting and will be highly relevant for the theoretical treatment of AllAgr. As background, consider the pattern of indexical shift investigated in detail by Sundaresan (2012) in detail which she calls ‘monstrous agreement’:

- (18) Maya_i [taan_{i,*j} pooʈʈi-læ dʒejkkæ-poo-r-eeen-nnũ] so-nn-aa
 Maya ANAPH contest-LOC win-go-PRS-1S-COMP say-PST-3FS.SBJ
 ‘Maya_i said that she_i would win the contest’

Sentences like (18) have a matrix speech verb embedding a clause where the subject, expressed as an anaphor, is co-referent with the matrix subject. What is interesting is that the argument agreement on the embedded verb in cases like this can be 1SG. But this indicates not the actual speaker of the utterance (i.e. not Tom in the example we discussed above), but Maya, the author of the speech act described by the matrix speech predicate. This should make it clear that the representation of speech act participants will be highly relevant for the derivation and interpretation of such sentences, and indeed, as we will see in section 3, Sundaresan (2012) analyzes monstrous agreement in terms of indexical shift, where the syntactic representation of the embedded speech act plays a crucial role.

Now, if both monstrous agreement and AllAgr imply the involvement of information about speech act participants in the morphosyntax, we might expect interesting things to happen if we can manage to get them to co-occur. Fortunately, we can, and the results do not disappoint. Consider (19):

- (19) *Maya_i* [taan_{i,*j} poot[ti-le dʒejkkæ-poo-r-eeŋgæ-nnũ] so-nn-aa
Maya [ANAPH contest-LOC win-go-PRS-1S-COMP] say-PST-3SF
 ‘*Maya_i* said that she_i would win the contest.’

The combination of the two interesting types of agreement does indeed seem to be possible, as we have the monstrous 1s suffix *-een* immediately followed by the allocutive *-ŋgæ*. The reading is similar to what we saw in (17), but with two important differences. First, whereas in (17) the embedded subject *avæ* could be either coreferent with the matrix subject *Maya* or not, here the coreference is obligatory, as *taan* is a long-distance anaphor, and *Maya* is the only appropriate binder around.¹⁰ Second, in this case the politeness implied by the AllAgr is reported as having been shown by Maya to the addressee of the embedded speech act (in our scenario above this would be Kausalya). In other words, when we have AllAgr in an embedded clause that also contains monstrous agreement, it seems to make reference to the embedded speech act. In the absence of monstrous agreement, reference is apparently made instead to the utterance speech act, even when the suffix shows up in the embedded clause as we saw in (17). These facts will need to be investigated in more exhaustive detail in future work, but the preliminary picture they paint is highly suggestive about the kind of analysis that we should

¹⁰ Actually, monstrous agreement structures place very specific restrictions, such that the long-distance anaphor in subject position must be bound by the attitude holder of the immediately embedding predicate. Thus, while in other contexts *taan* has some more flexibility in the choice of its antecedent, in (19) it would be quite fixed to *Maya* even if some additional potential perspective holder were present elsewhere in the sentence (see Sundaresan 2012: for details).

pursue for AllAgr, as we will see in (3).

Moving away from complement clauses, Tamil also allows AllAgr in some other types of embedding. For example, it can be found in certain adverbial clauses, like the temporal adjunct built on a completive participle in (20):

- (20) [naan vitt-ukkū poo-ji-[[t̃-ŋgæ], call paŋd-r-eeŋ-ŋgæ
I house-DAT go-PTCP-COMPL-ALLOC call do-PRS-1SG.SBJ-ALLOC
'When I get home, I'll call.'

Perhaps relatedly, it is perfectly fine on a very common kind of hanging topic construction, built by following the topic itself with a participial form of the verb meaning 'come', as we seen in (21):

- (21) [naan vandū-ŋgæ], naa|ækki coimbattuur-ukkū
I come-PTCP-ALLOC, tomorrow Coimbatore-DAT
poo-v-eeŋ-ŋgæ
go-FUT-1SG.SBJ-ALLOC
'As for me, I'm going to Coimbatore tomorrow.'

Notice incidentally that ALLOC marking actually appears twice in (20) and (21) – once on the embedded part and once on the root clause. This doubling seems to be optional.

2.3 Affix ordering and doubling

Now we get to the data that make Tamil AllAgr stand out and present some interesting puzzles. Recall that in Basque, AllAgr is ruled out in questions in addition to embedded clauses. It has been proposed that this is because AllAgr realizes the C position and is thus in competition with question particles and with the complementizers found in embedding. Japanese shows that this can't be a general property of AllAgr, since it does allow the marking in questions, as we see in (22) from Miyagawa (2017):

- (22) Dare-ga ki-mas-u ka?
who-NOM come-ALLOC-PRS Q
'Who will come?'

Note then that the AllAgr marker *-mas* appears below not just the question particle *ka*, but also the tense suffix *-u*. This leads Miyagawa (2017) to argue that,

while AllAgr involves the C domain, its morphological realization in Japanese is lower in the structure, near T.

Consider now what happens in similar cases in Tamil. We have seen above that Tamil is also perfectly happy to have its AllAgr marker *-ɲgæ* appear on a *wh*- or polar-interrogative. Indeed, it is actually quite common on short fragment and tag-question-like utterances, which are marked by the polar question particle *-aa*. What is potentially odd is how *-ɲgæ* is ordered relative to this question particle. I repeat here two of the examples above that involve this combination:

- (23) a. *niɲgæ saap-t-aačč-aa-ɲgæ?*
 you.PL eat-ASP-RES-Q-ALLOC
 ‘Have you eaten?’
 b. *saap-t-aaččü-ɲgæ|-aa?*
 eat-ASP-RES-ALLOC-Q
 ‘Have you eaten?’

(23a) and (23b) are essentially a minimal pair (the pro-drop in (23b) is irrelevant for current concerns), differing only in the order of the AllAgr marker and the question particle. In (23a), the AllAgr suffix comes at the end, outside of the question particle, while in (23b) it comes before it.¹¹ In other words, both orders of the two suffixes are possible. More examples show that this ordering alternation (again accompanied by predictable morphophonological effects) is fairly general:¹²

- (24) a. *illij-aa-ɲgæ? ~ illi-ɲgæ|-aa?*
 no-Q-ALLOC no-ALLOC-Q
 various uses, e.g. ‘Isn’t it?’, ‘No?’, tag question
 b. *appadij-aa-ɲgæ? ~ appadi-ɲgæ|-aa?*
 like.that-Q-ALLOC like.that-ALLOC-Q
 ‘Oh really?’, ‘Is that so?’

¹¹ The other minor differences we see are the result of regular morphophonology. In (23a), the final *ü* of the resultative suffix is deleted before a vowel-initial suffix, and the final *l* of the allocutive suffix is deleted in coda position. In (23b), both surface, the conditions for their deletion not being met.

¹² The first version of (24c) is from Amritavalli (1991). She did not discuss the other order, with ALLOC before Q, in her paper.

- c. ko.ɔandæ ippaɖi seji-laam-aa-ŋgæ? ~ ko.ɔandæ ippaɖi
 child like.this do-SBJV-Q-ALLOC child like.this
 seji-laam-ŋgæ]-aa?
 do-SBJV-ALLOC-Q
 ‘Is it right for the child to do this?’
- d. Naan ɖej-čč-eeen-aa-ŋgæ? ~ Naan ɖej-čč-eeen-ŋgæ]-aa?
 I win-PST-1SG.SBJ-Q-ALLOC I win-PST-1SG.SBJ-ALLOC-Q
 ‘Did I win?’

Especially with the fragment utterances, the order with the AllAgr preceding the question particle is the preferred one, but both are entirely possible under the right circumstances.¹³ This variation in the order of the affixes is surprising, and is not generally found in the inflectional morphology of the language. That is, the various temporal, aspectual, voice-related and other suffixes that can appear on verb forms are rigidly ordered relative to each other, and two given suffixes generally cannot have their order reversed (for extended discussion on this point, see Sundaresan & McFadden 2017).

Indeed, it gets even more interesting. In the cases where the AllAgr suffix can appear either before or after the Q particle, it is actually possible for it to be doubled, appearing simultaneously in both positions:

- (25) a. appaɖi-ŋgæ]-aa-ŋgæ?
 like.that-ALLOC-Q-ALLOC
 ‘Oh really?’
- b. niŋgæ saapt-aaččü-ŋgæ]-aa-ŋgæ?
 you.PL eat-RES-ALLOC-Q-ALLOC
 ‘Have you eaten?’
- c. onga]-ukkü coffee veŋum-ŋgæ]-aa-ŋgæ?
 you.PL-DAT coffee want-ALLOC-Q-ALLOC
 ‘Would you like coffee?’

It should be noted that, at least for my primary informant from Pollachi, such structures are not particularly marked, nor do they correspond to elevated or exaggerated politeness. Doubling of this kind is quite unexpected, and again, I

¹³ This actually seems to be a point of dialectal variation. The preference described in the main text holds for all of my informants from India, but my Singapore informants reported that they would only use the AllAgr-Q order, and that they associate the other order with speakers from India.

am aware of no other piece of grammaticalized morphology in the language that behaves this way.

An obvious question to ask then is whether the different ordering and doubling possibilities are associated with interpretive differences. We might expect, e.g., that the two orders would correspond somehow to distinct scope readings of some kind. The data here are tricky on this point, and it will ultimately require more careful empirical work, ideally with multiple informants who speak relevant dialects. However, to a first approximation, the two orders seem to differ in how the question is biased. Consider the following pairs, with the distinct translations offered by my informant:

- (26) a. *illij-aa-ŋgæ?* vs. *illi-ŋgæ[-aa?*
 no-Q-ALLOC no-ALLOC-Q
 ‘It’s not, is it?’ ‘Isn’t it?’
- b. *appaɟij-aa-ŋgæ?* vs. *appaɟi-ŋgæ[-aa?*
 like.that-Q-ALLOC like.that-ALLOC-Q
 ‘So, it is the case?’ ‘Is that the case?’
- c. *Naan ɟej-čč-eeen-aa-ŋgæ?* vs. *Naan ɟej-čč-eeen-ŋgæ[-aa?*
 I win-PST-1SG.SBJ-Q-ALLOC I win-PST-1SG.SBJ-ALLOC-Q
 ‘I won, didn’t I?’ ‘Did I win?’

The Q-ALLOC order seems to be biased towards confirmation, whereas the ALLOC-Q order seems unbiased, a genuine request for information. So in the first variant in (26b) with Q-ALLOC order, the speaker is expecting that the answer will be ‘yes’, and is just asking for confirmation — e.g. in order to get the addressee to admit something or just to be absolutely sure of something. But in the second variant with ALLOC-Q order, the speaker genuinely doesn’t know what the answer will be, and is asking in order to find out.

We can ask then what happens to the interpretation in cases of doubling. Here the judgments are subtle, and not all of the examples I checked seem to behave the same, but in one case where my informant did have a clear intuition, the doubling pattern goes together with the Q-ALLOC order in being biased towards confirmation. Consider the following minimal triplet of examples repeated from above, now with information added about the bias on the question:

- (27) a. *niŋgæ saap-t-aačč-aa-ŋgæ?*
 you.PL eat-ASP-RES-Q-ALLOC
 ‘Have you eaten?’ (speaker expects that addressee has)

- b. niingæ saap-t-aaččü-ŋgæ]-aa?
you.PL eat-ASP-RES-ALLOC-Q
'Have you eaten?' (speaker doesn't know)
- c. niingæ saapt-aaččü-ŋgæ]-aa-ŋgæ?
you.PL eat-RES-ALLOC-Q-ALLOC
'Have you eaten?' (speaker expects that addressee has)

It should be stressed at this point that the description of the interpretations here, especially in the doubling case, is highly preliminary. Note for example that it is a bit difficult to lock down exactly what the bias is. In the pair in (28) below (again based on examples from above, but now with additional interpretive information), it is still the ALLOC-Q order that comes with a bias, and the Q-ALLOC order that is neutral, but the direction of the bias is difficult to pin down. My informant reports that this might be how you ask someone as part of a routine, where you can anticipate what the will answer be based on your familiarity with their coffee drinking habits. But it does not seem to be restricted to either an expectation that they will say yes or an expectation that they will say no. The Q-ALLOC order in (28b) again shows no bias and is a genuine request for information, while the doubling in (28c) again patterns with the ALLOC-Q in showing a difficult to pin down bias.

- (28) a. onga]-ükkü coffee veenŋum-aa-ŋgæ?
you.PL-OBL-DAT coffee want-Q-ALLOC
'Do you want coffee?' (some bias, direction not clear)
- b. onga]-ükkü coffee veenŋum-ŋgæ]-aa?
you.PL-OBL-DAT coffee want-ALLOC-Q
'Do you want coffee?' (no bias)
- c. onga]-ükkü coffee veenŋum-ŋgæ]-aa?
you.PL-OBL-DAT coffee want-ALLOC-Q
'Do you want coffee?' (some bias, direction not clear)

Wide-reaching generalizations about the affects of ALLOC-affix ordering beyond the specific examples discussed are thus not yet warranted at this time.

3 Towards an account

3.1 Theoretical preliminaries

The phenomenon of AllAgr is clearly of great theoretical interest. As already noted in section 1.2, some basic information about each utterance – including the identity of the author and the addressee, as well as the time, location and other similar parameters – is obviously relevant for semantic and pragmatic interpretation, and so it is uncontroversial that such information must be encoded in some way in the representation of the discourse context. The question is whether such information is already represented in some form in the syntax. For indexical pronouns, for example, it is clear that we must assume a morphosyntactic status for person features, as they play a role in various agreement and binding operations, but it is entirely plausible that their relationship with the actual author and addressee of a given utterance is only established in the semantics. That is, a feature like [2] or [+participant, –author] would be treated no differently in the morphosyntax than a feature like [pl], and the association with the utterance context and in particular the identity of the discourse participants would only play a role in the interpretive component when reference is determined.

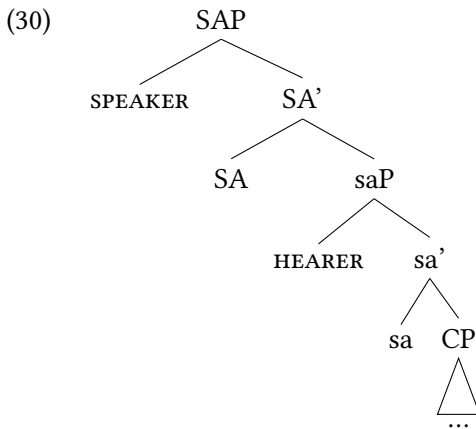
However, it is more difficult to imagine how an approach like this would work in the case of AllAgr, for two reasons. First, AllAgr encodes information about addressees beyond just their role in the discourse context, including their gender, number or status relative to the speaker, depending on the specific language. This seems to imply sensitivity to the actual identity of the addressee of a given utterance, as opposed to the simple fact that there *is* an addressee for every typical utterance. Second, while indexical pronouns appear as syntactic arguments or adjuncts, and clearly contribute to the asserted meaning of the sentences where they are found, AllAgr is again orthogonal to whether the addressee plays a role as an argument or adjunct, and clearly does not contribute anything to the assertion. That is, a Basque sentence like (29), repeated from above, does *not* assert something like ‘The speaker is going to the house, and the addressee is a female individual familiar to the speaker’.

- (29) etʃe-a banu-n
house-ALL 1.SG.go-ALLOC:F
‘I am going to the house.’ (familiar female addressee)

Rather, it asserts something more like ‘The speaker is going to the house’, and comes with something like a presupposition, such that it can only be uttered felicitously when the addressee is a female individual familiar with the speaker.

Thus we have morphosyntactic agreement, which by the nature of what ‘agreement’ means must be *with* something, but that something is not an argument or adjunct that we would normally expect to be part of the syntactic representation.

This has led a number of recent authors to conclude that AllAgr provides evidence for a literal syntactic representation of the discourse context, including information about the speech-act participants (Haegeman & Hill 2011; Miyagawa 2012; Zu 2015; Haegeman & Miyagawa 2016; Miyagawa 2017). This work generally adopts a version of what is sometimes called the neo-performative hypothesis of Speas & Tenny (2003). This is in turn a reinterpretation of Ross (1970)’s idea that the speech-act participants are represented not just in the semantics and pragmatics, but also in the syntax, by quite normal syntactic material that happens (generally) not to be pronounced. Setting the details aside for the moment, the basic neo-performative proposal is that the left periphery of the syntactic representation of the relevant clause types includes a (potentially internally complex) Speech Act Phrase (SAP).¹⁴ The *SPEAKER* and *HEARER* (or *AUTHOR* and *ADDRESSEE*) of the speech act are then directly represented by (silent, essentially pronominal) elements introduced in specifiers of this functional structure, as in (30), based essentially on the trees given by Miyagawa (2017):



For Miyagawa, AllAgr then represents straightforward agreement with the *HEARER* in Spec-saP. Since the *HEARER* is actually represented syntactically, its various ϕ -features will be available and thus can be reflected in the form of the

¹⁴ For Speas & Tenny (2003), who adopt assumptions about phrase structure from Larson (1988); Hale & Keyser (1993) etc., this SAP has a shell structure including two head positions, two specifier positions and a complement. Haegeman & Hill (2011) and Miyagawa (2012; 2017) update this by splitting it up into two phrases, SAP for the speaker and saP for the addressee.

allocutive suffixes on the verb. The usual assumption is that the probe for AllAgr is located somewhere in the C domain. This helps to explain why the agreement targets something in the speech act domain, whereas classic subject agreement in T targets something lower down in the argument domain. It also provides a way to approach various restrictions on the occurrence of AllAgr in things like interrogatives in certain languages, under the idea that a question particle competes with the AllAgr probe to realize C. Finally, the impossibility of AllAgr in (most) embedding contexts can be attributed to the fact that SAP and saP are only projected in root clauses. Miyagawa thus argues in detail (see especially Miyagawa 2012), that AllAgr is a ‘root phenomenon’ in the sense of (Emonds 1970).

In fact, these attempts to analyze AllAgr fit into a broader trend of arguing for an expanded left periphery containing a syntactic representation of the speech act and its participants. Sundaresan (2012) argues that the (limited) possibility of projecting a SpeechActP in the complement of certain attitude predicates (primarily speech predicates) is responsible for the phenomenon of indexical shift. Haegeman & Hill (2011) make crucial use of SAP in their analysis of a series of verbal particles in Romanian and West Flemish, which serve to “signal the speaker’s attitude or his/her commitment towards the content of the utterance and/or of his relation towards the interlocutor” [p. 9]. Zu (2015) uses the SAP to analyze speaker-related allocutive agreement in Jingpo and conjunct marking in Newari, which relates the subject of a clause either to a preceding subject or to the speech act participants. Sundaresan’s work here is especially relevant because, as discussed above, it investigates the monstrous agreement pattern in Tamil. Her analysis makes crucial use of an SAP in the embedded clause, containing a representation of the speech act associated with matrix ‘say’, which then plays an important role in the determination of argument agreement in the embedded clause. The anaphoric subject *taan* is unable to trigger agreement, thus agreement is instead with a coreferent 1st person form, the interpretation of which is ‘shifted’ by the embedded SAP to be relative to the author of the matrix speech predicate rather than the author of the utterance context. This thus provides evidence, independent of AllAgr, for the syntactic representation of information about speech-act participants in Tamil.

3.2 The proposal

I will assume to begin with that the work mentioned in the previous section is on the right track. In particular, there is a syntactic representation of the speech act in the left periphery of the clause which includes information about the speech

act participants, and the phenomenon of AllAgr is a case of the addressee playing an active role in the morphosyntax. But this still leaves a number of interesting issues open. First, where exactly is the SAP located and how does it interact with other elements in the left periphery, in particular the material relevant for forming interrogatives? Second, how does the overt AllAgr morphology actually relate to that syntactic representation? Third, what is behind the facts we observed in section 2.3 involving variation in the ordering and even doubling of the allocutive suffix?

The simplest analysis would be that *-ŋgæ* directly realizes the HEARER in the SAP. In other words, the Tamil phenomenon wouldn't really be allocutive *agreement* per se, but rather a direct spell out of (at least one of the ϕ -features of) the otherwise silent HEARER in Spec-saP in Miyagawa's tree in (30). This has some initial plausibility given the basic facts about where the suffix shows up, again on the basis of our first example:

- (31) Naan dʒaŋgiri vaŋg-in-een-ŋgæ.
 I Jangri buy-PST-1SG.SBJ-ALLOC
 'I bought Jangri.'

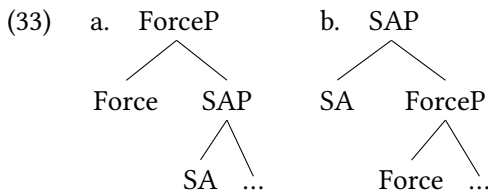
As already noted, the *-ŋgæ* suffix generally comes all the way at the end of the sentence, attached after all of the other inflectional suffixes on the clause-final verb. Given that Tamil is a strictly head-final language, this is where we expect something near the top of the functional sequence, like something in the SAP, to show up.

There are some issues with this prosal, however. First, the fact that it shows up as a suffix on the clause-final verbal material actually suggests that it realizes a head in the functional sequence, not a specifier, i.e. not the actual representation of the addressee. Second, it runs into trouble with the facts about the variable ordering of the *-ŋgæ* suffix relative to the polar question particle *-aa*, a subset of which are repeated in (32).

- (32) a. illij-aa-ŋgæ? illi-ŋgæ[-aa?
 no-Q-ALLOC no-ALLOC-Q
 various uses, e.g. 'Isn't it?', 'No?', tag question
 b. appadij-aa-ŋgæ? appadi-ŋgæ[-aa?
 like.that-Q-ALLOC like.that-ALLOC-Q
 'Oh really?', 'Is that so?'

We should ask at this point which of the two orders we should actually expect, if *-ŋgæ* is realizing something in SAP. Both the question particle and the SAP

should be somewhere in the C domain, i.e. in the left periphery of the clause higher than T, and it is that idea that has been employed to explain why AllAgr is incompatible with questions in Basque. But since at least Rizzi (1997) we are generally willing to recognize a richer structure in this region of the clause, involving a series of (more or less strictly ordered) heads. We can assume that there is a head responsible for indicating whether a clause is interrogative, declarative etc. — let’s adopt Rizzi’s Force, though this may be an oversimplification — in addition to the SAP heads. The trees in (33) give the two obvious logical options for how these two heads could be ordered with respect to each other above the rest of the clause, and we must decide which is more likely to be correct.¹⁵



It seems to me that, semantically speaking, only (33b) is plausible. If SAP really introduces the representation of the speech act participants and related information, then it is setting the stage for the entire sentence. It provides the background against which a question is asked, including who is asking and answering, and the order in (33b) seems to best reflect this. The order in (33a), on the other hand, suggests a very odd and inappropriate semantics. It would seem to imply that the contents of the SAP are part of what the question in Force is being asked about. In very rough terms the information about the speech act participants would be under the scope of the polar question. For an utterance like (34), this would imply a meaning along the lines of ‘Is it the case that the identity of the speaker and the hearer is such that the speaker uses the polite form of address with the hearer and furthermore that the hearer has eaten?’. This kind of reading is certainly not available, and to my knowledge it is simply never possible for a normal question to scope over the representation of the speech act participants in this way, in Tamil or any other language.

¹⁵ There is of course a third option, where we would say that both hierarchical orders are available, and this is what is responsible for the variation in the linear order of the suffixes. However, this would run into semantic difficulties, and would go against the spirit of the more or less cartographic approach being adopted here, at least under the assumption that we’re dealing with the same SAP and the same ForceP in both structures.

- (34) niingæ saapt-aaččü-ŋgæ[-aa?
you.PL eat-RES-ALLOC-Q
'Have you eaten?'

Let us consider then what all of this means for affix orders. If SAP comes above Force, and both are realized by overt affixes, then by the Mirror Principle, SAP should come after Force in linear order. So if we assume that allocutive *-ŋgæ* realizes something in SAP (we'll say more below on what exactly that something might be), and the question particle *-aa* realizes Force, then we should get the order in (35):

- (35) niingæ saapt-aačč-aa-ŋgæ?
you.PL eat-RES-Q-ALLOC
'Have you eaten?'

Indeed we do get this order, but of course the whole point is that we *also* get the reverse order shown in (34). Now, we could conclude that this is simply a morphological quirk. We could say that the syntax really corresponds to something like (35) with the expected ordering of AllAgr outside of the question particle, but that there is then a post-syntactic process that optionally flips their order. Again however, this has some problems. For one thing, it doesn't have a good way of dealing with doubling, i.e. the fact that the AllAgr marker can simultaneously show up in both positions relative to the question particle. For another, if the two orderings are identical in the syntax, being differentiated only in the morphological portion on the PF branch the derivation, it predicts that there should be no meaning difference between them, under the standard assumption that PF doesn't feed into LF. But this is incorrect — as we saw in section 2.3, the Q-ALLOC order seems to be biased towards a particular answer, while the ALLOC-Q seems to be more neutral, an open request for information. I hesitate to draw any firm analytical conclusions from this, again because the empirical situation remains somewhat murky, and because I have not carried out an investigation of the syntax and semantics of bias in polarity questions in the language. However, what it suggests at the minimum is that the ordering difference involves something more substantial than just post-syntactic morphology.

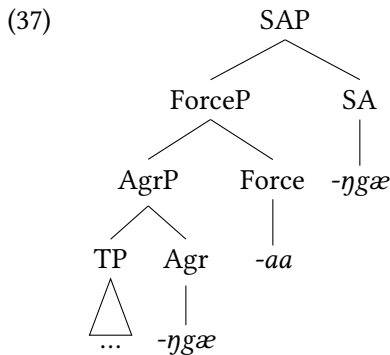
Therefore, I'd like to propose something (slightly) more interesting. First, for the ALLOC-Q order we need to reaffirm the idea that *-ŋgæ* really is agreement. That is, it does not realize anything in SAP directly, but rather the features of something in SAP being reflected elsewhere in the structure. This lets us put the question particle in Force, below SAP, getting the broad semantics right, with

AllAgr realizing an even lower head that agrees with the addressee in SAP. This has to be distinct from the head that realizes argument agreement. This is because we get both types of agreement simultaneously, even when both are below the question particle, as in the second variant in (24d), repeated in (36).

- (36) Naan džej-čč-*een-ηgæ*|-aa?
 I win-PST-1SG.SBJ-ALLOC-Q
 ‘Did I win?’

Given the ordering facts, though, the locus of low AllAgr should be very close to, and just a bit higher than, that of argument agreement, which itself is just above T.

For the order where *-ηgæ* shows up after the question particle, something different must be going on. Again by the Mirror Principle, it must be realizing a higher position in the left periphery, which as far as I have found only comes below the (rather high) complementizer *-nnū*. We can thus potentially place it even as high as SAP itself. This opens up the possibility that *this* instance of *-ηgæ* isn’t agreement in the T region, but is more directly spelling out something relating to the representation of the addressee, but given the logic discussed above, it must still be a head rather than e.g. a pronominal directly representing the addressee in Spec-saP. Perhaps the simplest assumption is that it is the sa head itself, which agrees in ϕ -features with the representation of the addressee in its specifier, but at present I am not in a position to defend any specific proposal. The following structure gives an idea of how this might look:¹⁶



This approach has the clear advantage that it provides two distinct structural positions for the allocutive suffix, and thus will allow a natural account of

¹⁶ One should not take the label ‘Agr’ too seriously, and of course this leaves open how exactly the higher *-ηgæ* relates to the actual representation of the addressee.

the doubling data. It does not explain the subtly different readings available with the two orders, but it does at least allow an account to be formulated, once the facts are better understood, because the two positions for allocutive suffixes have different sources. The idea is basically as follows. The baseline is that an Agree relationship is established between the representation of the addressee and the Agr head above T, which itself has no semantic consequences. An additional relationship can be established with a higher head in the SAP region, but if this is done, it has the semantic consequence of introducing a bias with respect to a polar question. The results of this higher Agree relation are always pronounced when they obtain. The lower one is usually also pronounced, but can be optionally left off when the higher one is pronounced. When only the lower is pronounced, we get the ALLOC-Q order and no bias, because the higher Agree operation has not occurred. Anytime the higher is pronounced, we get the bias, because this is derived by the higher Agree operation. But in this case pronunciation of the lower Agr is optional, so we get both doubling and the Q-ALLOC order, both with the semantics of bias.

We can also take some first steps towards an analysis of the embedding facts and the interaction of AllAgr with monstrous agreement. As noted above, Sundaresan (2012) analyzes monstrous agreement in terms that require an SAP in the embedded clause which encodes information about the embedded speech act rather than the matrix one. Among other things, the contexts in which monstrous agreement is possible are constrained by the conditions on when such an embedded SAP can be projected, and the optionality of monstrous agreement in certain speech complements amounts to optionality in whether the SAP will be present. Now, under the proposal being considered here, at least the lower position of AllAgr is not itself in the SAP region, thus we do not necessarily predict that AllAgr will only be available in embedded clauses that contain their own SAP – a lower AllAgr above T could at least potentially Agree with some higher representation of an addressee.¹⁷ We do, however, have clear expectations about how the presence of an embedded SAP should be relevant for the *interpretation* of AllAgr. If there is an embedded SAP, we expect on minimality groups that an embedded *-ŋgæ* suffix must be Agreeing with that, and not with the matrix SAP associated with the utterance speech act. If monstrous agreement implies an embedded SAP, then we predict that an accompanying embedded AllAgr will

¹⁷ This of course raises important and complicate questions of mechanics. In particular, how would such an Agree relation be constrained to probe specifically for the ϕ -features of addressee representations, and what are the locality implications? Consideration of such issues must wait until we have a better grasp on the empirical situation.

reflect properties of the addressee of the embedded speech act (and her relationship with the embedded author), not of the addressee of the utterance speech act. This is precisely what we observed in the contrast between (17) and (19) above, repeated in (38):

- (38) a. Maya [avæ poot[ti-le ɕejkkæ-poo-r-aa[-ɲgæ-nnũ] so-nn-aa
 Maya [she contest-LOC win-go-PRS-3SF-COMP] say-PST-3SF
 ‘Maya said that she would win the contest.’ (speaker being polite)
- b. Maya_i [taan_{i,*j} poot[ti-le ɕejkkæ-poo-r-ee-ɲgæ-nnũ]
 Maya [ANAPH contest-LOC win-go-PRS-1S-COMP]
 so-nn-aa
 say-PST-3SF
 ‘Maya_i said that she_i would win the contest.’ (Maya being polite)

In (38a) without monstrous agreement, we can presume that there is no embedded SAP, thus the embedded AllAgr is interpreted relative to the utterance speech act, and so it is the utterance speaker, Tom in our scenario from above, who is showing politeness to the addressee, Venkat. In (38b), on the other hand, monstrous agreement establishes the presence of an embedded SAP, relative to which the embedded AllAgr must be interpreted. Thus it is the author of the embedded speech act Maya who is showing politeness towards her addressee, Kausalya.

An interesting final point to note is that in both AllAgr agreement positions, the only ϕ -feature we see realized is plural number (which can either be interpreted as actual plural number or polite address). We don’t see gender or person. Again, the Tamil ALLOC marker is the general plural suffix *-ɲgæ*, not e.g. the 2nd plural argument agreement suffix *-iɲgæ*. While I have nothing to say at this point about why gender is missing,¹⁸ the lack of person is explicable and quite interesting. Consider that, unlike gender, person seems to never be marked in AllAgr in the various languages for which the phenomenon has been reported. One could imagine that this is due to a sort avoidance of redundancy – the addressee is always 2nd person, so there is no actual contrast to be achieved by explicitly marking the person. But I think this is actually mistaken. In fact, the grammatical representation of the addressee (syntactically in the left periphery or also in the semantics) is *not* 2nd person, and indeed it does not make sense

¹⁸ Tamil *does* distinguish three genders in verbal argument agreement, so it certainly has the formal material to mark such distinctions in AllAgr. And while masculine and feminine are not generally distinct in the plural, there is nothing in principle that could stop the marking of the gender of familiar singular addressees, as Basque does.

to attribute any person value to it. Rather, this representation identifies a particular (singular or plural) individual as the addressee of the speech act, much as an accompanying representation identifies another individual as the author of the same speech act. Neither of these individuals properly has a person value. Rather, the person values of all other individuals in the scope of the speech act will be determined or interpreted relative to these two individuals. Any nominal element bearing 2nd person will be interpreted as coreferent with the representation of the addressee, and any element bearing 3rd person will be distinct from both author and addressee. The representations of author and addressee are the anchor points for the interpretation of person on other nominals, and thus it makes no sense to attribute person to these representations themselves. Number and gender, on the other hand, are properties of individuals independent of their role in the current speech act, and thus it is reasonable that these features might be present on the representations in the SAP. It is thus entirely reasonable that Tamil AllAgr reflects the number of the addressee, but not her person.

Abbreviations

Table (3) lists the abbreviations used in the paper that are not found in the Leipzig Glossing Rules.

Table 3: Abbreviations not in the Leipzig Glossing Rules

ALLOC	allocutive marker
ANAPH	anaphor
ASP	aspect
RSP	respected addressee

Acknowledgements

I am grateful first of all to my informants, Jegan Murugesan, Champa Sundaresan, Subramania Sundaresan and Sandhya Sundaresan, for patiently exploring this particular detail of their language with me. Thanks also to Sandhya and Jegan, as well as the audience at TripleA 4 in Göteborg, for very helpful discussion of both empirical and theoretical issues. Finally, I am indebted to Rahul Balusu for not

only directing my attention to, but also sending me a copy of, the very important (but not so easy to obtain) Amritavalli (1991).

References

- Alcázar, Asier & Mario Saltarelli. 2014. *The syntax of imperatives*. Cambridge: Cambridge University Press.
- Amritavalli, R. 1991. Addressee markers on verbs: form of agreement? In *Linguistics at large: papers in general and applied linguistics dedicated to prof. Shivendra K. Verma*, 38–51. Hyderabad: Booklinks.
- Amritavalli, R. & K.A. Jayaseelan. 2005. Finiteness and negation in Dravidian. In Guglielmo Cinque & Richard S. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*, 178–220. Oxford: OUP.
- Antonov, Anton. 2015. Verbal allocutivity in a crosslinguistic perspective. *Linguistic Typology* 19. 55–85.
- Asher, R.E. & E. Annamalai. 2002. *Colloquial Tamil*. London: Routledge.
- Baker, Mark. 2015. *Case: its principles and its parameters*. Cambridge: CUP.
- Boeckx, Cedric & Fumikazu Niinuma. 2004. Conditions on agreement in Japanese. *Natural Language and Linguistic Theory* 22. 453–480.
- Bonaparte, Louis-Lucien. 1862. *Langue basque et langues finnoises*. London: Strange-ways & Walden.
- Emonds, Joseph. 1970. *Root and structure-preserving transformations*. MIT PhD thesis.
- Haegeman, Liliane & Virginia Hill. 2011. The syntacticization of discourse. Ms., Ghent University and University of New Brunswick-SJ.
- Haegeman, Liliane & Shigeru Miyagawa. 2016. The syntax of participants. Talk presented at GLOW 39, Workshop on Representation of Speech Acts.
- Hale, Ken & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Ken Hale & Samuel Jay Keyser (eds.), *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19. 335–391.
- Miyagawa, Shigeru. 2010. *Why Agree? why Move?* Cambridge, Mass.: MIT Press.
- Miyagawa, Shigeru. 2012. Agreements that occur mainly in main clauses. In Lobke Albrecht, Liliane Haegeman & Rachel Nye (eds.), *Main clause phenomena: new horizons*, 79–112. Amsterdam: John Benjamins.
- Miyagawa, Shigeru. 2017. *Agreement Beyond Phi*. Cambridge, Mass.: MIT Press.

- Oyharçabal, Beñat. 1993. Verb agreement with non arguments: on allocutive agreement. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *Generative studies in basque linguistics*. Amsterdam: Benjamins.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer Academic Publishers.
- Ross, John Robert. 1970. On declarative sentences. In Roderick Jacobs & Peter Rosenbaum (eds.), *Readings in English transformational grammar*, 222–277. Waltham, Mass.: Ginn.
- Schiffman, Harold. 1999. *A reference grammar of spoken Tamil*. Cambridge: Cambridge University Press.
- Speas, Peggy & Carol Tenny. 2003. Configurational properties of point of view roles. In Anna Maria DiSciullo (ed.), *Asymmetry in grammar*, 315–344. Amsterdam: John Benjamins.
- Sundaresan, Sandhya. 2012. *Context and (co)reference in the syntax and its interfaces*. University of Tromsø (CASTL)/Universität Stuttgart PhD thesis.
- Sundaresan, Sandhya & Thomas McFadden. 2017. The articulated *v* layer: evidence from Tamil. In Roberta D'Alessandro, Irene Franco & Ángel Gallego (eds.), *The verbal domain*, 153–178. Oxford: Oxford University Press.
- Zu, Vera. 2015. A two-tiered theory of the discourse. In *Proceedings of the poster session of WCCFL 33* (Simon Fraser University Working Papers in Linguistics, vol. 5), 151–160.