Tamil allocutive agreement and the syntax of the SpeechAct Phrase

GLOW 41, Research Institute for Linguistics, MTA, Budapest, April 12th, 2018

Thomas McFadden
Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin
mcfadden@leibniz-zas.de

1 Introduction

In many colloquial varieties of Tamil (Dravidian; South Asia), one commonly comes across utterances of the following kind:[1]

(1) Naan ʤaaŋgiri vaaŋg-in-een-ŋgæ.
I Jangri buy-PST-1SG.SBJ-ALLOC
'I bought Jangri.'[2]

Aside from the good news it brings, (1) is of interest because it contains two different types of agreement stacked on top of each other.

1. *-een* marks quite normal agreement with with the 1sg subject.
2. *-ŋgæ* marks something far less common: so-called **allocutive agreement**.

☞ Rather than cross-referencing properties of one of the arguments of the main predicate, allocuative agreement provides information about the addressee.

☞ The addition of *-ŋgæ* specifically indicates a plural addressee or a singular one who the speaker uses polite forms of address with.

☞ If the addressee is a single familiar person, the suffix is simply lacking, as in (2).

(2) Naan ʤaaŋgiri vaaŋg-in-een.
I Jangri buy-PST-1SG.SBJ
'I bought Jangri.'

Allocutive agreement (henceforth AllAgr) has been identified in a handful of languages and is characterized by the following properties (see Antonov, 2015, for an initial typological overview):

- It marks properties (gender, politeness...) of the addressee of the current speech context.
- It is crucially not limited to cases where the addressee is an argument of the local predicate.
- It involves the use of grammaticalized morphological markers in the verbal or clausal inflectional system, thus is distinct from special vocative forms like English *ma'am* or *sir*.

[1]Thanks to my informants: Jegan Murugesan, Champa Sundaresan, Subramania Sundaresan and Sandhya Sundaresan.

[2]Jangri is a delicious flower-shaped sweet made of deep-fried lentil batter soaked in sugar syrup.

The most extensively discussed example of AllAgr comes from Basque (Bonaparte, 1862, Oyharçabal, 1993, Alcázar and Saltarelli, 2014).

- Here, the use of AllAgr depends, in dialect-specific ways, on politeness and the number of the addressee, with the specific form reflecting the gender of the addressee.
- The Souletin Basque examples in (3) from Antonov (2015) 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)

There are some additional interesting properties of Basque AllAgr that should be noted here.

☞ AllAgr is banned when there is a second person argument, which will be coindexed with the appropriate (ergative, absolutive or dative) 2nd person argument agreement.[3]

☞ When the conditions for it are met, AllAgr is obligatory.

☞ AllAgr is restricted to *root* declaratives.

Miyagawa (2017) has argued that the kind of politeness marking found in Japanese examples like (4) should also be analyzed as a type of AllAgr.

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

b. Watasi-wa piza-o tabe-ru.
I-TOP pizza-ACC eat-PRS
'I will eat pizza.' (colloquial)

- Here again, the marker gives information about the addressee, and it is a clearly grammaticalized part of the verbal inflectional system.
- What makes the case here a bit less obvious is that Japanese doesn't have straightforward argument agreement for more familiar ϕ features like person, number and gender.

This talk will give a detailed consideration of Tamil data, which present some interesting novel properties, with a view to what AllAgr can tell us about the representation of the speech act.

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

- The only prior work on the phenomenon I am aware of is Amritavalli (1991), which reports the central facts, including a number of insightful observations.
- I will report a number of additional facts, and will focus in particular on the interesting patterns that arise when AllAgr interacts with question-formation and embedding.
- With the former there are some rather surprising ordering and doubling facts that can hopefully help us tease apart alternative analyses.
- With the latter we are presented with questions for our understanding of embedded root phenomena and also find a cool interaction with indexical shift.

2 Background on Tamil

First some basic descriptive and sociolinguistic information that matters for AllAgr:

- Tamil shows marked diglossia, with an extraordinarily conservative written standard.
- AllAgr is very much a phenomenon of the colloquial language, and the form of the data presented will reflect this fact.
- Spoken Tamil is also characterized by extensive regional, social and communal dialect variation, and the possibility and prevalence of AllAgr is highly sensitive to this.
- Fortunately, my primary informant speaks a dialect (a variety of Kongu Tamil from Pollachi, Coimbatore district) that makes heavy use of AllAgr, and he has robust intuitions about the phenomenon.[4]

Now a brief primer on relevant aspects of Tamil morphosyntax:

- Tamil is a highly inflected language with a strongly agglutinative character, is strictly head-final and almost exclusively suffixing.
- Finite verbs can be marked for transitivity, aspect, voice, mood, negation, tense and agreement, but mood, negation and agreement are essentially in complementary distribution (see Amritavalli and Jayaseelan, 2005, Sundaresan and McFadden, 2017).
- Verbal agreement targets the highest nominative argument and reflects person and number, plus gender in the 3rd person and politeness in the 2nd and 3rd persons.
- Plural forms of pronouns and agreement are used in the 2nd person to indicate politeness.

Table 1 shows the regular agreement paradigm with an example of the simple present tense and imperative forms of *oo*ɖ̯*ŭ* 'run' (the *-r-* before agreement marking present tense).

☞ Note that agreement follows all aspect, tense and voice markers, as we can see in (5), an example of a moderately complex, fully inflected finite verb.

[4]My other informants speak the Iyer dialect of central and northeast Tamil Nadu, which does not show AllAgr, but they also have command of the common colloquial dialect, where AllAgr is found. The data reported in the main text generally reflect my Kongu Tamil speaker's intuitions, though I will report a few cases where those of my other informants differ.

Table 1: Regular verb agreement in Tamil

	SG	PL
1	oodŭ̱-r-een	oodŭ̱-r-oom
2	oodŭ̱-r-e	oodŭ̱-r-iiŋgæ
3F	oodŭ̱-r-aa	oodŭ̱-r-aaŋgæ
3M	oodŭ̱-r-aan	oodŭ̱-r-aaŋgæ
3POL	oodŭ̱-r-aarŭ	oodŭ̱-r-aaŋgæ
3N	oodŭ̱-dŭ	oodŭ̱-dŭ
IMP	oodŭ̱	oodŭ̱-ŋgæ

(5) Kausalya padi̱-ččŭ-kiṯṯŭ-ru-nd-aa
Kausalya study-ASP-PROG-BE-PST-3F.SG
'Kausalya was studying.'

☞ The finite verb, terminated by agreement, is typically the final element in a root declarative clause, but it can be followed by further suffixes in the C domain, e.g. the complementizer *-nnŭ* in (6a) or the polar question particle *-aa* in (6b):[5]

(6) a. Venkaṭ [Kausalya padi̱-ččŭ-kiṯṯŭ-ru-nd-aa]-nnŭ so-nn-aan
Venkat [Kausalya study-ASP-PROG-BE-PST-3F.SG]-C say-PST-3M.SG
'Venkat said that Kausalya was studying.'

b. Kausalya padi̱-ččŭ-kiṯṯŭ-ru-nd-aaḷ-aa?
Kausalya study-ASP-PROG-BE-PST-3F.SG-Q
'Was Kausalya studying?'

3 The form and position of Tamil AllAgr

Now the we have some background, let's work out the basic morphophonology of the Tamil allocutive marker. I repeat (1) from above as a basis for the discussion:

(7) Naan ʤaaŋgiri vaaŋg-in-een-ŋgæ.
I Jangri buy-PST-1SG.SBJ-ALLOC
'I bought Jangri.'

- The allocutive suffix is *-ŋgæ*. This is actually a fairly general *plural* marker in the language.
- It is the final component of all 2nd and 3rd person (non-neuter) plural agreement markers, and it attaches to the verb root to form (2nd) plural imperatives (see Table 1).
- Furthermore, it is used as the plural marker in a number of nominal categories, both nouns (*maram* 'tree', *maraŋgæ* 'trees') and pronouns (*nii* 'you.SG', *niiŋgæ* 'you.PL').
- (7) 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.

[5]The -ḷ before the question particle in (6b) is part of the underlying form of the agreement suffix, but only surfaces when a vowel-initial suffix immediately follows in 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 *-ŋgæḷ* before vowel-initial subjects. As we will see, this includes AllAgr marker.

- When the verb has a modal or negative suffix rather than agreement, *-ŋgæ* still follows at the very end of the verb form, as in (8), based on data from Amritavalli (1991).

(8) a. koɻandæ ippaɖi sejjæ-kkuuɖaadŭ-ŋgæ
child like this do-must not-ALLOC
'The child should not act in such a way.'

b. Venkaʈ varæ-læ-ŋgæ
Venkat come-NEG-ALLOC
'Venkat isn't coming.'

- The AllAgr marker can also co-occur with unambiguous vocatives, strictly adjacent to the verb, with the vocative obligatorily coming outside (typically extraposed). This confirms that *-ŋgæ* itself cannot be a vocative:

(9) a. naan va-r-een-ŋgæ saar
I come-PRS-1SG-ALLOC sir
'I'll take my leave, sir.'

b. * naan va-r-een saar ŋgæ
I come-PRS-1SG sir ALLOC

- The marker can also appear in clauses with no verb, like (10a), and even in fragmentary or elliptical utterances that are smaller than clauses, as in (10b) and (10c):

(10) a. naan aaʈʈookkaaran-ŋgæ
I automan-ALLOC
'I am an auto rickshaw driver.'

b. indæ pajjan-ŋgæ
this boy-ALLOC
'this boy' (e.g. as answer to 'Who's next?')

c. illi-ŋgæ
no-ALLOC
'No' (as answer to polar question)

☞ The basic generalization is that the marker attaches to whatever is final in the clause or sub-clausal utterance (excluding extraposed material).

☞ It is clearly a bound form. It never appears alone or after a pause, always being attached to a preceding word. Indeed, the sequence *ŋg-* is not licit word-initially in Tamil.

➪ We can thus conclude that it is a grammaticalized marker, presumably of agreement.

4 The distribution of Tamil AllAgr

Let us now turn to the conditions under which AllAgr does and does not appear. The central determining factor is the identity of the addressee and their relationship with the speaker.

- There is only one allocutive suffix in the language, so the number of distinctions that can be made is minimal.

- Simply put, allocutive *-ŋgæ* is found whenever *niiŋgæ* would be the appropriate 2nd person pronoun, i.e. when the addressee is plural or polite singular.
- Thus an utterance like (11) would be appropriate when addressed to a group of friends or to an adult stranger, but not to an individual friend.[6]

(11) enæ-kkŭ teri-læ-ŋgæ
me-DAT know-NEG-ALLOC
'I don't know'

An important question is what happens with allocutive agreement when the 2nd person *is* an argument of the main predicate.

☞ Recall that in Basque, AllAgr is blocked in such circumstances.

☞ Tamil shows a somewhat mixed behavior. When a 2nd person subject triggers regular argument agreement on the verb, AllAgr is ruled out, as shown by the examples in (12):[7]

(12) a. * eppaḍi iru-kk-iiŋgæ-ŋgæ?
how be-PRS-2PL-ALLOC
'How are you?'
b. * niiŋgæ rombaa smart-aa iru-kk-iiŋgæ-ŋgæ
you.PL very smar-PRED be-PRS-2PL-ALLOC
'You're very smart.'

☞ But when a 2nd person argument doesn't trigger argument agreement, AllAgr is just fine.

☞ (13a) shows this with a 2nd person accusative direct object, and (13b) shows the same with a quirky dative subject, neither of which can trigger argument agreement.

☞ (13c) has a 2nd person nominative subject, but the main predicate is in a participial form which doesn't host argument agreement. In all of these examples, AllAgr is possible.

(13) a. naan ongaḷ-æ paḍatt-læ paa-tt-een-ŋgæ
I you.PL.OBL-ACC film-LOC see-PST-1SG-ALLOC
'I saw you in a film.'
b. ongaḷ-ŭkkŭ coffee veeɳum-aa-ŋgæ?
you.PL-OBL-DAT coffee want-Q-ALLOC
'Do you want coffee?'
c. niiŋgæ saapṭ-aačč-aa-ŋgæ?
you.PL eat-RES-Q-ALLOC
'Have you eaten?'

Now let us consider the further conditions on the appearance of AllAgr, once we've restricted our attention to speech contexts with the right kind of addressee.

[6] For my Central Iyer speakers, the allocutive suffix is only used to reflect politeness, not plural, i.e. for them (11) could not be used with a group of friends.

[7] At least one of my Central Iyer speakers accepts examples like these where AllAgr appears on top of 2nd person subject agreement.

- We've seen that it *can* appear in root declaratives and various fragmentary utterances.
- Furthermore, unlike in at least some dialects of Basque, it can appear in root interrogatives. (13b) and (13c) above show polar questions, and (14) a *wh*-question use:

(14) evḷavŭ aag-um-ŋgæ?
how much become-MOD-ALLOC
'How much will it come to?' (i.e. 'How much does it cost?')

- And it turns out that it can also appear in certain embedded environments. We'll look at relevant data in detail in Section 6 below.

A final, quite crucial point is that, when there is no (overtly expressed) 2nd person argument, AllAgr is actually obligatory, at least for my primary Kongu Tamil informant.

☞ I.e. when one would use *niiŋgæ* with the addressee, only (15a) is possible. Leaving off the *-ŋgæ* signals non-politeness, and thus (15b) is ill-formed in such a discourse context.

(15) a. rombaa thanks-ŋgæ
very thanks-ALLOC
'Thanks a lot'
b. * rombaa thanks (to a polite or plural addressee)
very thanks

☞ This is strong evidence that the use of *-ŋgæ* as AllAgr is fully grammaticalized.

➪ From all of this we can securely conclude that the *-ŋgæ* suffix in (relevant dialects of) Tamil is indeed an instance of fully grammaticalized allocutive agreement.

5 Ordering and doubling

Now that we have the basics of Tamil AllAgr, we can turn to some interesting points of detail that will have important consequences for any analysis.

As we have seen, Tamil is perfectly happy to have *-ŋgæ* on a root *wh-* or polar-interrogative.

- Indeed, it is extremely common on tags and other short interrogative utterances marked by the polar question particle *-aa*.
- What is odd is how *-ŋgæ* is ordered relative to the particle. Consider two examples that involve this combination:

(16) a. niiŋgæ saapṭ-aačč-aa-ŋgæ?
you.PL eat-RES-Q-ALLOC
'Have you eaten?'
b. niiŋgæ saapṭ-aačču-ŋgæḷ-aa?
you.PL eat-RES-ALLOC-Q
'Have you eaten?'

☞ They represent a minimal pair, differing only in the order of the AllAgr marker and the question particle.

☞ In (16a), the AllAgr suffix comes at the end, outside of the question particle, while in (16b) it comes before it.[8] In other words, *both* orderings of the two suffixes are possible.

More examples show that this ordering alternation is fairly general:[9]

(17) a. illij-aa-ŋgæ? / illi-ŋgæɭ-aa?
no-Q-ALLOC / no-ALLOC-Q
various uses, e.g. 'Isn't it?', 'No?', tag question

b. appaɖij-aa-ŋgæ? / appaɖi-ŋgæɭ-aa?
like.that-Q-ALLOC / like.that-ALLOC-Q
'Oh really?', 'Is that so?'

c. koɻandæ ippaɖi sejji-laam-aa-ŋgæ? / koɻandæ ippaɖi sejji-laam-ŋgæɭ-aa?
child like.this do-SBJV-Q-ALLOC / child like.this do-SBJV-ALLOC-Q
'Is it right for the child to do this?'

- Especially with the fragment utterances, the order with the AllAgr preceding the question particle is the prefered one. But both are entirely possible under the right circumstances.

- This variation in the order of the affixes is surprising, and is not generally found elsewhere in the inflectional morphology of the language.

But it gets even more interesting. In the cases where both orders are avaiable, it is actually possible to find the the allocutive suffix *doubled* on either side of the Q particle:

(18) a. appaɖi-ŋgæɭ-aa-ŋgæ?
like.that-ALLOC-Q-ALLOC
'Oh really?'

b. niiŋgæ saapʈ-aaccŭ-ŋgæɭ-aa-ŋgæ?
you.PL eat-RES-ALLOC-Q-ALLOC
'Have you eaten?'

c. ongaɭ-ukkŭ coffee veɳum-ŋgæɭ-aa-ŋgæ?
you.PL-DAT coffee want-ALLOC-Q-ALLOC
'Would you like coffee?'

☞ It should be noted that, at least for my primary Kongu Tamil informant, such structures are not particularly marked, nor do they correspond to elevated or exaggerated politeness.

☞ Such doubling is quite unexpected. Again, I am aware of no other bit of grammaticalized morphology in the language that behaves this way.

[8] The other minor differences we see are the result of regular morphophonology.
[9] The first version of (17c) is from Amritavalli (1991).

6 Embedded AllAgr

What, then, about embedded environments? Again, AllAgr has been reported to be blocked or at least heavily restricted under embedding in other languages.

☞ Perhaps not surprisingly, AllAgr is possible in direct speech, where the *-ŋgæ* suffix is understood as part of what is quoted:

(19) Raman avar-ʈʈæ "saap-ʈ-aaccŭ-ŋgæɭ-aa?" so-nn-aan.
raman him.POLITE-LOC eat-ASP-RES-ALLOC-Q say-PST-3SG.M
'Raman said to him "Have you eaten?".

☞ What is more interesting is that it can also be found in indirect speech, embedded under at least some attitude predicates:[10]

(20) Maya [avæ pooʈʈi-le ʤejkkæ-poo-r-aaɭ-ŋgæ-nnŭ] so-nn-aa
Maya [she contest-LOC win-go-PRS-3SF-ALLOC-COMP] say-PST-3SF
'Maya said that she would win the contest.'

➪ This suggests that AllAgr constitutes an embedded root phenomenon.[11]

I am currently investigating how it might fit into the typology of strong vs. weak root phenomena (see Frey and Meinunger, 2017, and references there).[12]

7 The representation of the speech act

AllAgr in general is of great theoretical interest, in particular because it seems to involve a clear interaction between the morphosyntax and information about the speech act.

☞ It is clear that some information about the utterance context — the identity of the author and the addressee, the time and location — is relevant for semantics and pragmatics, e.g. for the interpretation of indexicals.

☞ AllAgr provides evidence that this information is represented in the syntax as well.

[10] My data on this point are still very preliminary. I don't yet have a clear picture of exactly which predicates allow AllAgr in their complements, but the outlines are reminiscent of the classic bridge verbs that allow classic embedded root phenomena.

[11] Note, however, that Baker and Alok (2017) report AllAgr to actually be possible in a far wider range of embedding types in Magahi, so there is apparently some cross-linguistic variability here that will need to be investigated.

[12] Potentially surprising in this regard is the fact that AllAgr *is* possible in certain temporal adverbial clauses, as in (i), which we might expect to be CACs and thus to exclude both strong and weak root phenomena:

i. [naan viʈʈ-ukkŭ poo-ji-ʈʈŭ-ŋgæ], call paɳɖ-r-een-ŋgæ
I house-DAT go-PTCP-COMPL-ALLOC call do-PRS-1SG.SBJ-ALLOC
'When I get home, I'll call.'

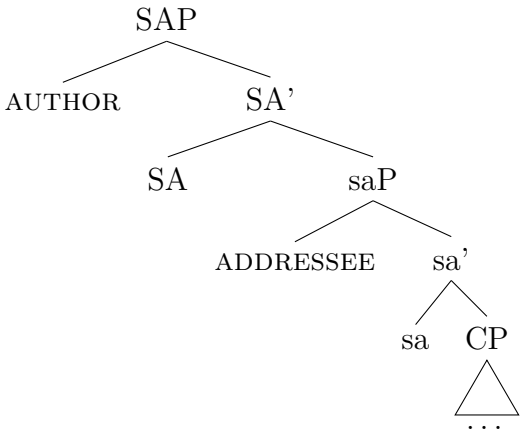
Note, however, that the temporal clause in (i) is built, not with a complementizer explicitly marking a temporal relation like 'when', but with the temporal use of a completive participial structure (a bit reminiscent of a Latin ablative absolute). It is plausible that such clauses are more loosely connected to the main clause than 'when' clauses would be, as also suggested by a clear prosodic break between them — perhaps they are non-integrated adverbials, where embedded root phenomena would be unsurprising.

☞ It is morphosyntactic agreement, reflecting grammatical properties of the addressee, thus under standard views of agreement there must be a representation of the addressee that it is agreeing *with*.

A number of authors have thus used AllAgr to argue for a literal syntactic representation of the discourse context, including the speech-act participants (Hill, 2007, Haegeman and Hill, 2011, Miyagawa, 2012, Haegeman and Miyagawa, 2016, Miyagawa, 2017, Baker and Alok, 2017):

- This work adopts a version of the neo-performative hypothesis, building on Speas and Tenny (2003)'s reinterpretation of Ross (1970)'s idea that the speech-act participants are represented by normal syntactic material that happens (generally) not to be pronounced.
- The idea is that the author and addressee are represented by (silent pronominal) elements in a (potentially internally complex) Speech Act Phrase (SAP) in the left periphery, as in (21), based essentially on trees given by Miyagawa (2017), following Hill (2007):

(21)



➪ AllAgr then constitutes syntactic agreement with the ADDRESSEE in Spec-saP.

☞ Among other things, this accounts for the status of AllAgr as a root phenomenon, given the assumption that SAP & saP are only projected in root clauses (see especially Miyagawa, 2012, for discussion).

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.

- Speas and Tenny (2003) collect a whole series of phenomena from the previous literature that seem to require reference to a syntactic representation of the speaker and hearer.
- Haegeman and Hill (2011) use the SAP to analyze verbal particles in Romanian and West Flemish which "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].
- Sundaresan (2012) argues that the possibility of projecting a SpeechActP in the complement of certain attitude predicates is crucial for deriving indexical shift.
- Zu (2015) uses the SAP to analyze 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, as it also includes data from Tamil, in particular the pattern of indexical shift she calls 'monstrous agreement':

(22) Maya$_i$ [$_{CP}$ taan$_{i,*j}$ pooʈʈi-læ ʤejkkæ-poo-r-een-nnŭ] so-nn-aa
Maya ANAPH contest-LOC win-go-PRS-1SG-COMP say-PST-3FSG
'Maya$_i$ said that she$_i$ would win the contest'

- Sentences like (22) have a matrix speech verb, which embeds a clause where the subject — realized as an anaphor — is co-referent with the the matrix subject.
- What is interesting is that the agreement on the embedded verb in cases like this can be 1SG, but this refers then not to the actual speaker of the utterance, but to the author of the speech act described by the matrix speech verb.

Sundaresan (2012) analyzes this as involving indexical shift:

☞ The 1SG agreement in the embedded clause is not co-indexed with the speaker of the utterance, but with Maya, in her capacity as the speaker of the matrix speech predicate.

☞ The analysis makes crucial use of an SAP in the embedded clause, containing a representation of the speech act associated with matrix 'say'.

➪ We thus have independent evidence for the syntactic representation of information about speech-act participants in the language.

Now, if both monstrous agreement and AllAgr crucially involve the presence of an SAP, we expect them to interact.[13]

☞ Indeed they do, and with fascinating interpretive results. Since monstrous agreement occurs in the complement of speech predicates, we have utterances involving two different speech acts, so there are two different addressees that AllAgr could be tellling us about.

☞ Consider the context in (23), where it would be appropriate for me to utter either (23b) or (23a). Both are perfectly grammatical and have more or less the same assertive content.[14]

(23) Maya has told Lila that she (Maya) is going to win a contest. I (Tom) witnessed this and want to report it to Kausalya, who wasn't there.

a. Maya$_i$ Lila-ʈʈæ [taan$_{i,*j}$ pooʈʈi-le ʤejkkæ-poo-r-een-ŋgæ-nnŭ] so-nn-aa.
Maya Lila-LOC [ANAPH contest-LOC win-go-PRS-1S-ALLOC-COMP] say-PST-3SF
'Maya$_i$ told Lila that she$_i$ would win the contest.' (Maya being polite to Lila)

b. Maya$_i$ Lila-ʈʈæ [avæ$_{i,j}$ pooʈʈi-le ʤejkkæ-poo-r-aa[-ŋgæ-nnŭ] so-nn-aa.
Maya Lila-LOC [she contest-LOC win-go-PRS-3SF-ALLOC-COMP] say-PST-3SF
'Maya told Lila that she would win the contest.' (Tom being polite to Kausalya)

- (23a) shows 1SG monstrous agreement in the embedded clause, hence indexical shift.

[13] Baker and Alok (2017) give a first look at intriguing data on the interaction between AllAgr and indexical shift in Magahi, which appear to be similar to what we will see here for Tamil.

[14] The main difference is that in (23a), the embedded subject is obligatorily co-referent with the matrix subject *Maya*, whereas in (23b) it *can* co-refer with *Maya*, but need not.

- Here, AllAgr characterizes the reported speech act. I.e. I am reporting that Maya made a polite utterance to Lila.
- On the other hand, (23b) does *not* involve indexical shift, as the embedded verb bears unexceptional 3SF subject agreement.
- In this case, the AllAgr characterizes the utterance speech act. I.e. it indicates that I, Tom, am being polite to Kausalya.

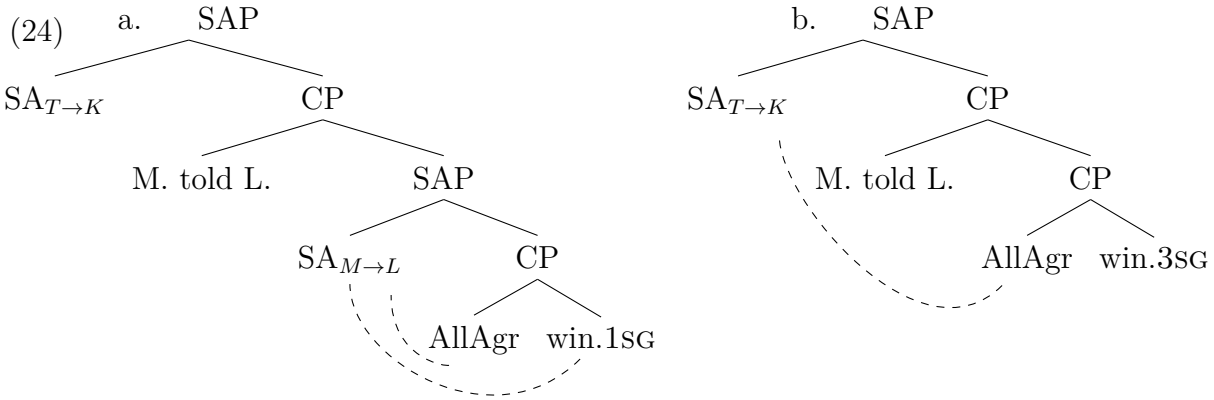
If both indexical shift and AllAgr depend on the same syntactic representation of speech acts, i.e. a SAP, we can make perfect sense of this, as shown schematically in (24):

☞ The monstrous agreement in a sentence like (23a) diagnoses the presence of a dedicated SAP for the embedded clause, shown in (24a)

☞ The lack thereof in sentences like (23b) tells us that the only SAP is the one anchoring the root clause to the utterance context, as in (24b).

☞ In both cases, AllAgr then simply targets the closest SAP and reflects information about the Addressee represented there.

(24)



8 The realization of AllAgr

Once we've decided that AllAgr reflects properties of the representation of the addressee in the left periphery, a bunch of open questions remain:

? How does the SAP interact with other elements in the left periphery, in particular the material relevant for forming interrogatives?

? How does the overt AllAgr morphology actually relate to that syntactic representation?

? What's the story with the ordering variation and doubling facts?

One simple analysis would be that *-ŋgæ* directly realizes the ADDRESSEE in the SAP:

☞ In other words, the Tamil phenomenon wouldn't really be allocutive *agreement*, but the direct spell out of (some r features of) the otherwise silent ADDRESSEE.

But this runs into several problems, e.g. with the facts about the ordering of the *-ŋgæ* suffix relative to the polar question particle *-aa*, repeated here in (25).[15]

[15] Another problem is presented by examples like (23b), where *-ŋgæ* reflects information about the root speech act but is clearly realized internal to an embedded clause.

(25) a. illij-aa-ŋgæ? illi-ŋgæɭ-aa?
no-Q-ALLOC no-ALLOC-Q
various uses, e.g. 'Isn't it?', 'No?', tag question

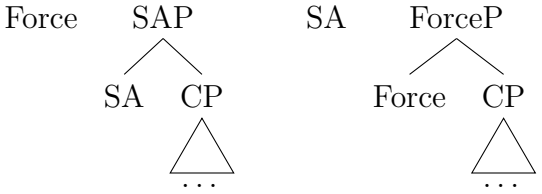
b. appaɖij-aa-ŋgæ? appaɖi-ŋgæɭ-aa?
like.that-Q-ALLOC like.that-ALLOC-Q
'Oh really?', 'Is that so?'

c. koɻandæ ippaɖi sejji-laam-aa-ŋgæ? koɻandæ ippaɖi sejji-laam-ŋgæɭ-aa?
child like.this do-SBJV-Q-ALLOC child like.this do-SBJV-ALLOC-Q
'Is it right for the child to do this?'

Which of these orders should we actually expect, if *-ŋgæ* is realizing something in SAP?

- Both the question particle and the SAP are somewhere in the C domain, and that has been employed to explain why AllAgr is incompatible with questions in Basque — AllAgr and question particles compete for a single slot and thus are in complementary distribution.
- But since at least Rizzi (1997) we are generally willing to recognize a richer structure in the C domain, 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 for concreteness, though this may be an oversimplification — in addition to the SAP head(s).
- So the question is which of the orders in (26) do we we expect?

(26) a. ForceP b. SAP



It seems to me that, semantically speaking, only (26b) 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 speech act.

☞ It provides the background against which a question is asked, including who is asking and answering, and the order in (26b) seems to best reflect this.

The order in (26a), on the other, 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 (27), this would imply a meaning along the lines of 'Is it the case that the identity of the speaker and the addressee is such that the speaker uses the polite form of address with the addressee and furthermore that the addressee has eaten?'.

(27) niiŋgæ saapʈ-aačč̆ŭ-ŋgæɭ-aa?
you.PL eat-RES-ALLOC-Q
'Have you eaten?'

☞ This kind of reading is not available, and to my knowledge it is never possible for a normal question to scope over the representation of the speech act participants in this way.

So what does all of this mean 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.
- So if we assume that allocutive *-ŋgæ* realizes something in SAP, and the question particle *-aa* realizes Force, then we should get the order in (28):

(28) niiŋgæ saapʈ-aačč-aa-ŋgæ?
you.PL eat-RES-Q-ALLOC
'Have you eaten?'

- Indeed we do, but of course the whole point here is that we also get the reverse order in (27). So this simplistic analysis won't work.

Now, we could conclude that this is simply a morphological quirk:

- We could say that the syntax corresponds to (28) with the expected ordering of AllAgr outside of Q, but then there is a post-syntactic process that optionally flips their order.

Again though, this has problems:

☞ It doesn't have a good way of dealing with doubling. The fact that the AllAgr marker can simultaneously show up before and after the Q particle at least suggests that these are two distinct syntactic positions rather than two orderings of a single element.

☞ It leads us to expect that there should be no meaning difference attached to the two orderings. While the details remain a bit murky, this seems to be incorrect.

The data here are tricky, and I'm still working with my main informant to nail things down, but to a first approximation, the two orders seem to differ in how the question is biased:

(29) a. illij-aa-ŋgæ? vs. illi-ŋgæɭ-aa?
no-Q-ALLOC no-ALLOC-Q
'It's not, is it?' 'Isn't it?'

b. appaɖij-aa-ŋgæ? vs. appaɖi-ŋgæɭ-aa?
like.that-Q-ALLOC like.that-ALLOC-Q
'So it is the case?' 'Is that the case?'

c. Naan ʤej-čč-een-aa-ŋgæ? vs. Naan ʤej-čč-een-ŋgæɭ-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.[16]
- I hesitate to draw any firm analytical conclusions from this, but what it suggests is that the difference here involves something more substantial than just post-syntactic morphology.[17]

So I'd like to suggest something (slightly) more interesting. First, for the ALLOC-Q order:

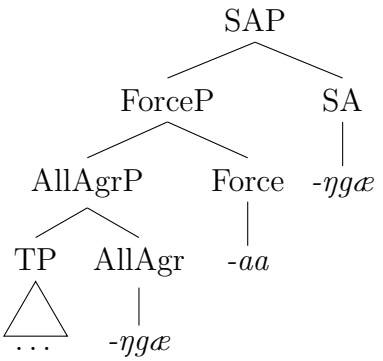
- We need to insist that *-ŋgæ* really is agreement, i.e. not a direct realization of something in SAP, but rather the features of something in SAP reflected elsewhere in the structure.
- This lets us put the question particle in Force, below SAP, getting the broad semantics right, and then AllAgr realizes an even lower head agreeing with the addressee in SAP.
- Baker and Alok (2017) propose for Magahi that AllAgr is in T itself (or FIn), because it is adjacent to and sometimes portmanteau'd with subject agreement, and is found in only those clauses that have subject agreement.
- In Tamil, however, even the AllAgr inside Q must be in a head distinct from and higher than T — it always comes outside argument agreement and can easily appear in clauses and fragments where argument agreement is impossible. I'll simply call the head AllAgr.

For the Q-ALLOC order, something else must be going on:

- The *-ŋgæ* suffix must be realizing a higher position in the left periphery, and as far as I have found, there is nothing that appears even higher.
- We can thus put it as high as SAP itself, so that perhaps *this* position for *-ŋgæ* isn't agreement, but more directly spells out (a part of) the representation of the addressee.
- At present I don't have enough evidence to take a strong position on this, so I will leave the details open, simply placing the suffix in the SA domain.

The following structure shows how this might look, indicating both positions for AllAgr:

(30)



☞ This approach has the clear advantantage that it provides two distinct structural positions for the allocutive suffix, and thus can automatically handle the doubling data.

☞ It does not immediately offer any insight into the subtly different readings available with the two orders, but it at least provides structural distinctions that an account could make use of, once the facts are better understood.

[16]The data remain murky in cases of doubling, but the interpretation *seems* to go together with the Q-ALLOC order in being biased towards confirmation.

[17]The odd thing is that the change in the *question* semantics suggests that it is the Q particle occurring in two different places, but the doubling facts show that it is rather the AllAgr suffix that does so.

References

Alcázar, Asier, and 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., and K.A. Jayaseelan. 2005. Finiteness and negation in Dravidian. In *The Oxford Handbook of Comparative Syntax*, ed. Guglielmo Cinque and Richard S. Kayne, 178–220. Oxford: OUP.

Antonov, Anton. 2015. Verbal allocutivity in a crosslinguistic perspective. *Linguistic Typology* 19:55–85.

Baker, Mark, and Deepak Alok. 2017. Allocutive agreement and indexical shift in Magahi. Handout from MIT Colloquium Talk, October 27th.

Bonaparte, Louis-Lucien. 1862. *Langue basque et langues finnoises*. London: Strangeways & Walden.

Frey, Werner, and André Meinunger. 2017. On strong and weak root phenomena. Workshop 'Speech Acts', Leibniz-ZAS.

Haegeman, Liliane, and Virginia Hill. 2011. The syntacticization of discourse. Ms., Ghent University and University of New Brunswick-SJ.

Haegeman, Liliane, and Shigeru Miyagawa. 2016. The syntax of participants. Talk presented at GLOW 39, Workshop on Representation of Speech Acts.

Hill, Virgina. 2007. Vocatives and the pragmatics-syntax interface. *Lingua* 117:2077–2105.

Miyagawa, Shigeru. 2012. Agreements that occur mainly in main clauses. In *Main clause phenomena: New horizons*, ed. Lobke Albrecht, Liliane Haegeman, and Rachel Nye, 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 *Generative studies in Basque linguistics*, ed. José Ignacio Hualde and Jon Ortiz de Urbina. Amsterdam: Benjamins.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.

Ross, John Robert. 1970. On declarative sentences. In *Readings in English transformational grammar*, ed. Roderick Jacobs and Peter Rosenbaum, 222–277. Waltham, Mass.: Ginn.

Speas, Peggy, and Carol Tenny. 2003. Configurational properties of point of view roles. In *Asymmetry in grammar*, ed. Anna Maria DiSciullo, 315–344. Amsterdam: John Benjamins.

Sundaresan, Sandhya. 2012. Context and (co)reference in the syntax and its interfaces. Doctoral Dissertation, University of Tromsø (CASTL)/Universität Stutgart.

Sundaresan, Sandhya, and Thomas McFadden. 2017. The articulated *v* layer: evidence from Tamil. In *The verbal domain*, ed. Roberta D'Alessandro, Irene Franco, and Ángel Gallego, 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.