Allocutive marking and the theory of agreement

SyntaxLab, University of Cambridge, November 5th, 2019

Thomas McFadden

Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), 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 initial informants, Jegan Murugesan, Champa Sundaresan, Subramania Sundaresan and Sandhya Sundaresan, as well as to the twelve speakers in the areas surrounding Pollachi and Thenur who shared their varieties with me during fieldwork in August 2018.

[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 (26) 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-n
          house-ALL 1.SG.go-ALLOC:F

          'I am going to the house.' (familiar female addressee)

       c. 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, focusing 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.

---

[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 that argument agreement is obligatory.

- With the latter we are presented with questions for our understanding of embedded root phenomena and also find a cool interaction with indexical shift.

- Based on all of this I will propose an analysis that, if correct, has consequences for the representation of the speech act and the workings of agreement.

2 Background on Tamil

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

- Tamil shows marked diglossia, and AllAgr is a phenomenon of the colloquial language.

- The occurrence and behavior of AllAgr are also dependent on dialect. Fortunately, my primary informant and several of my fieldwork subjects speak dialects (from Pollachi and surrounding villages) that make heavy and systematic use of AllAgr.[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 *ooḍŭ* 'run' (the *-r-* before agreement marking present tense).

Table 1: Regular verb agreement in Tamil

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

---

[4]My other informants speak either the Iyer dialect of central and northeast Tamil Nadu, which does not natively show AllAgr at all (though they have intuitions about the phenomenon via their command of the common colloquial dialect), or dialects from villages in Perambalur district, north of Tiruchirapalli, which do have AllAgr, but far less frequently and systematically. The data reported in the main text mostly reflect my Kongu Tamil speakers' intuitions, though I will report a few cases where those of my other informants differ.

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

(5) Kausalya paḍi-ččŭ-kiṭṭŭ-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 paḍi-ččŭ-kiṭṭŭ-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 paḍi-ččŭ-kiṭṭŭ-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 vaaŋ-in-een-ŋæ.
     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', *mara-ŋgæ* 'trees') and pronouns (*nii* 'you.SG', *nii-ŋgæ* 'you.PL').

- (7) also demonstrates that the allocutive marker attaches to the clause-final verb, after all of the other inflectional suffixes, including tense, aspect, voice and argument agreement.

- 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. koṯandæ ippaḍi   sejjæ-kkuuḍaadŭ-ŋgæ
         child   like this do-must not-ALLOC
         'The child should not act in such a way.'
     b. Venkaṭ varæ-læ-ŋgæ
         Venkat come-NEG-ALLOC
         'Venkat isn't coming.'

---

[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. Several suffixes show this alternation, including the AllAgr marker itself.

- 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.

- 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'

---

[6]This is subject to dialectal variation. For most of my speakers from outside the Pollachi area, 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.

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 strongly degraded:[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 veeṇum-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.

• 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.

---

[7]A few of my speakers accept examples like these where AllAgr appears on top of 2nd person subject agreement, but even they clearly prefer to leave it off in such cases.

A final, quite crucial point is that, when there is no (overtly expressed) 2nd person argument, AllAgr is obligatory for my Pollachi informants.[8]

☞ 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

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

     b. niiŋgæ saapʈ-aaččŭ-ŋgæɭ-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.[9] In other words, *both* orderings of the two suffixes are possible.

More examples show that this ordering alternation is fairly general:[10]

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

---

[8]For speakers of the other dialects, this is not the case, i.e. AllAgr is generally optional and can be left off in examples like (15, especially if some other marker of respect (e.g. a special vocative like 'sir') is present.

[9]The other minor differences we see are the result of regular morphophonology.

[10]The first version of (17c) is from Amritavalli (1991).

    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.[11] 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-aaččŭ-ŋgæɭ-aa-ŋgæ?
       you.PL eat-RES-ALLOC-Q-ALLOC
       'Have you eaten?'

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

☞ It should be noted that, at least for my Pollachi informants, 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.

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. Tamil is different.

☞ AllAgr is possible in clauses embedded under at least **some** attitude predicates:[12]

(19)   Maya [avæ pooʈʈi-le   ɖʑejkkæ-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]Here again we find dialectal variation. While Pollachi speakers are reasonably happy with both orders, Iyer and Perambalur speakers clearly prefer AllAgr-Q in most contexts, and two speakers of Singaporean Tamil I interviewed only have Q-AllAgr and associate AllAgr-Q with speakers from India.

[12]My data on exactly which predicates allow AllAgr in their complements are still preliminary, but the outlines are reminiscent of the classic bridge verbs that allow classic embedded root phenomena.
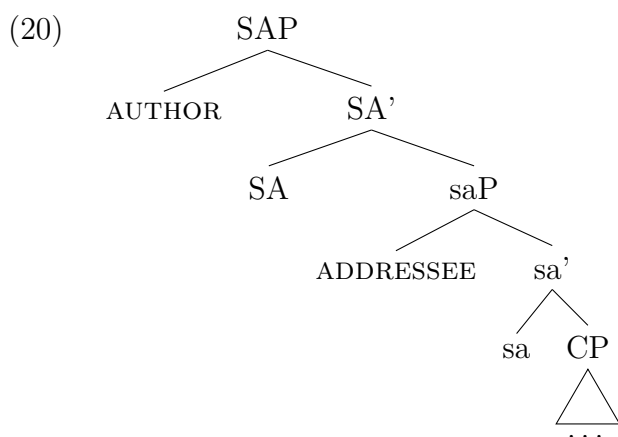
7   The representation of the speech act

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 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.

☞ It is morphosyntactic agreement, reflecting properties of the addressee, thus 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, 2013, 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 (20), based essentially on trees given by Miyagawa (2017), following Hill (2007):

(20)

```
                    SAP
                   /   \
            AUTHOR      SA'
                       /   \
                     SA     saP
                           /   \
                 ADDRESSEE      sa'
                              /   \
                            sa     CP
                                   /\
                                  /__\
                                  ...
```

⇨ 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 (2013) 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.

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':

(21) 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 (21) 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.

If both monstrous agreement and AllAgr crucially involve an SAP, they should 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 (22), where it would be appropriate for me to utter either (22b) or (22a). Both are perfectly grammatical and have more or less the same assertive content.[14]

(22) 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)

---

[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 (22a), the embedded subject is obligatorily co-referent with the matrix subject *Maya*, whereas in (22b) it *can* co-refer with *Maya*, but need not.

    b. Maya$_i$ Lila-ʈʈæ [avæ$_{i,j}$ pooʈʈi-le   ʤejkkæ-poo-r-aaɭ-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)

- (22a) shows 1SG monstrous agreement in the embedded clause, hence indexical shift. Here, AllAgr characterizes the reported speech act. I.e. I am reporting that Maya made a polite utterance to Lila.

- On the other hand, (22b) 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 (23):

☞ The monstrous agreement in a sentence like (22a) diagnoses the presence of a dedicated SAP for the embedded clause, shown in (23a)

☞ The lack thereof in sentences like (22b) tells us that the only SAP is the one anchoring the root clause to the utterance context, as in (23b).

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

(23)
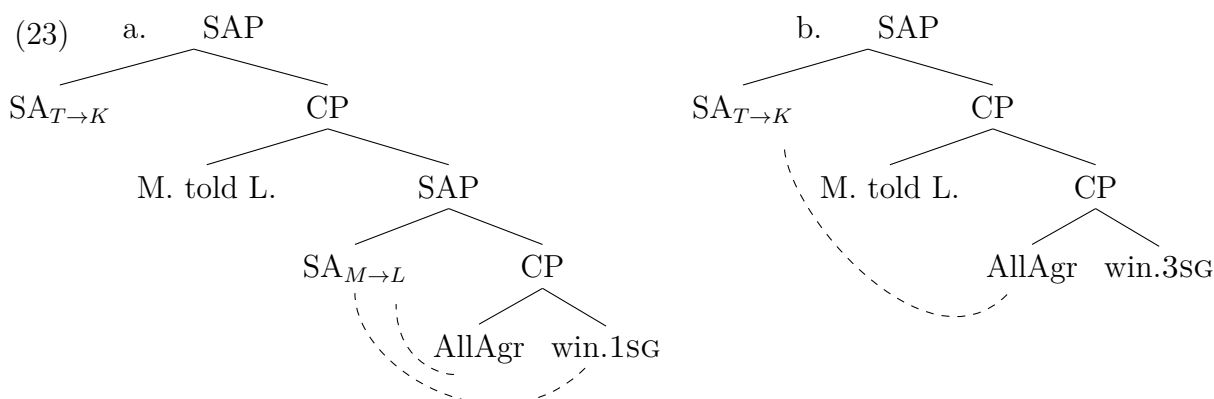
8   Integrating AllAgr into a theory of agreement

Now that we have a basic idea about AllAgr, as involving something in the C domain agreeing with something in a SAP, let's see how it might fit into a general, restrictive theory of agreement.

- To do that, we need above all to integrate it with other types of agreement that involve the clause periphery or C-domain.

- For this I'll present the basics of some ongoing joint work with Sandhya (McFadden and Sundaresan, 2018).

We can identify at least 3 different kinds of C-domain agreement:[15]

1. Downward complementizer agreement (DCA); C agrees with embedded subject; see (24) from West Flemish:

---

[15]Switch reference might belong here too…

(24) K peinzen da-n   die   studenten nen buot gekocht ee-n.
      I think   that-3PL those students a   boat bought have-3PL
      'I think that those students have bought a boat.'

2. Upward complementizer agreement (UCA); C agrees with matrix subject; see (25) from
   Lubuksu:

   (25) ba-ba-ndu ba-bol-el-a   Alfredi ba-li   a-kha-khil-e
         2-2-people 2-said-AP-FV 1Alfred 2-that 1-FUT-conquer
         'The people told Alfred that he will win.'

3. Allocutive Agreement (AllAgr); peripheral affix agrees with addressee; see (26) from
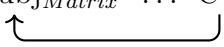   Basque:

   (26) Pettek   lan   egin   din.
         Peter.ERG work.ABS do.PRF 3.S.ABS-2.S.C.FM.ALLOC-3.S.ERG
         'Peter worked.' *Utterance to a close female friend*

The theoretical challenge that arises now is that the simplest theories of these three phenomena
are mutually incompatible.

- For DCA, we could assume (along the lines of e.g. Carstens, 2003, van Koppen, 2005)
  that C^0 bears unvalued ϕ; it probes **downward** and Agrees with the embedded subject,
  which it minimally c-commands.

- For AllAgr, we could assume (with Miyagawa, 2017, McFadden, To appear) a C^0 with
  unvalued ϕ for Japanese and Tamil, but there probing **upward** for the representation of
  the addressee in the left periphery.

- For UCA we would need unvalued ϕ to Agree upward (unlike DCA), but with the matrix
  subject rather than the addressee (unlike AllAgr).

This presents a problem if we want a principled account that covers all three:

- We need to deal with the variation in upward vs. downward, ideally without resorting to
  parametrization of directionality of Agree a là Baker (2008).

- And we need a story about why the representation of the addressee is a potential agree-
  ment target in cases where we get AllAgr, but not in those where we get UCA.

(27) **DCA** ... Subj_{Matrix} ... C ... $\text{Subj}_{Embedded}$ ...

(28) **UCA** ... Subj_{Matrix} ... C ... $\text{Subj}_{Embedded}$ ...

(29) **AllAgr** ... (Subj_{Matrix}) ... C ... $\text{Subj}_{Embedded}$ ...
               Addressee

Observation I: embedded vs. root clauses
A key difference between AllAgr and both types of CA is that, while AllAgr is a root phemo-
menon, CA is restricted to embedded clauses.

- This is not (just) about the presence of an overt complementizer — in Bavarian, CA can attach to a *wh*-phrase in Spec-CP in an embedded clause with no overt C (Bayer, 1984):

  (30) Du sollst   song [[an wäichan Schuah]-st du   wui-st]
        you should say [[the which   shoe]-2SG you want-2SG]
        'You should say which one of the shoes you want.'

- And when embedded clauses have root syntax, like the embedded V2 in Frisian (32), they seem to actually disallow CA (de Haan, 2001):

  (31) Heit sei **dat-st**   do soks net leauwe **moa-st.**
        dad said that-2P.SG you such not believe must-2P.SG
  (32) Heit sei **dat(*-st)** do **moa-st**   soks net leauwe.
        dad said that-2P.SG you must-2P.SG such not believe
        'Dad said that you should not believe such things.'

Observation 2: UCA vs. DCA is about more than directionality

- There is evidence that the 'complementizer' in UCA is different from — and structurally higher than — that in DCA.

1. Unlike DCA, UCA can apparently have a semantic effect. Agreeing *a-li* in Lubukusu (33) is only possible if the speaker considers the reported information reliable (Diercks, 2013), otherwise non-agreeing *bali* appears.

   (33) Mosesi a-lom-ile a-li Sammy k-eb-ile   chi-rupia.
         1Moses 1-say-PRF 1-C 1Sammy 1-steal-PST 10-money
         'Moses has said that Sammy stole the money.'

2. In some Bantu languages, the complementizer (with UCA) can appear overtly in the matrix clause, replacing the matrix verb, as in the Kipsigis example (34) from (Diercks, van Koppen, and Putnam, 2017)

   (34) kɔ-lɛ-ndʒin   Kiproono ko-∅-ɾuuja tuɣa amut.
         3-C-2SG.OBJ Kiproono PST-3-sleep cows yesterday
         'Kiproono told you that the cows slept yesterday.'

3. Indeed, at least some of the relevant complementizers seem to be grammaticalized from verbs meaning 'say', i.e. at least historically they belong more to the matrix than the embedded clause.

Proposal I: different probe locations plus (phase) locality
The differences between UCA and DCA derive from a difference in the height of the ϕ-probe in the C domain (building on Carstens, 2016, Diercks et al., 2017).

- ☞ **West Germanic complementizers** and their accompanying ϕ probes are relatively low in the left periphery, something like Rizzi (1997)'s Fin^0.

- Following Carstens (2016), we propose that this is **below** the CP phase boundary, meaning they can probe material in TP, yielding (downward) Agree with the embedded subject.

- The relevant head is also present in all types of finite embedded clause, hence the presence of CA doesn't diagnose any special left-peripheral structure and so isn't associated with any particular semantics.

☞ **Bantu complementizers** and their accompanying ϕ probes are relatively high in the left periphery, something like Rizzi (1997)'s Force^0.

- Again following Carstens (2016), this is **above** the CP phase boundary, meaning they cannot Agree downward with the embedded subject. Instead, they must find an appropriate goal in the structure above them.

- Something beyond the assumption of simple upward Agree is needed here, since they target the matrix subject, not necessarily the closest higher DP. See Carstens (2016), Diercks et al. (2017) for discussion.

- The presence of this Bantu CA diagnoses the projection of an elaborated left periphery, with concomitant intepretive consequences — e.g. the presence of Krifka (2017)'s CommitmentP, explaining the effects in (33).

Proposal II: root clauses, SAP and minimality

The differences between AllAgr and CA derive from the structural representation of the Speech Act characterizing Root clauses.

☞ **Root clauses** are distinguished by the presence of SpeechActP at the top, including representations of Author and Addressee.

☞ **AllAgr ϕ-probes** are even higher in the left periphery than Bantu CA probes.

- This again makes them non-local to any ϕ-bearing material in the clause below, forcing them to find something higher to Agree with.

- It also implies the presence of the extended left-peripheral structure that is only available in root clauses, explaining their restricted distribution.

- The SAP, with its syntactic representation of the Addressee, provides them with something local to Agree with, pre-empting Agree with anything higher and yielding AllAgr.

Here's how it looks:

(35) **DCA**
Subj_{Matrix} $[_{phase}\ [\phi{:}\ ]\ \text{Subj}_{Embedded}\ ]$

(36) **UCA**
Subj_{Matrix} $[\phi{:}\ ]\ [_{phase}\quad \text{Subj}_{Embedded}\ ]$

(37) **AllAgr**
$(\text{Subj}_{Matrix})\ [_{SAP}\ \text{Addr}\ [\phi{:}\ ]\quad [_{phase}\quad \text{Subj}_{Embedded}\ ]\ ]$

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. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.

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

Bayer, Josef. 1984. COMP in Bavarian syntax. *The Linguistic Review* 3:209–274.

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

Carstens, Vicki. 2016. Delayed valuation: A reanalysis of goal features, "upward" complementizer agreement, and the mechanics of case. *Syntax* 19:1–42.

Carstens, Vicky. 2003. Rethinking complementizer agreement: Agree with a Case-checked goal. *Linguistic Inquiry* 34:393–412.

Diercks, Michael. 2013. Indirect Agree in Lubukusu complementizer agreement. *Natural Language and Linguistic Theory* 31:357–407.

Diercks, Michael, Marjo van Koppen, and Michael Putnam. 2017. Agree probes down: Anaphoric feature valuation and phase reference. Ms. Pomona College, Utrecht University and Penn State University.

de Haan, Germen. 2001. More is going on upstairs than downstairs: embedded root phenomena in West Frisian. *Journal of Comparative Germanic Linguistics* 4:3–38.

Haegeman, Liliane, and Virginia Hill. 2013. The syntacticization of discourse. In *Syntax and its limits*, ed. Raffaella Folli, Christina Sevdali, and Robert Truswell, 370–390. Oxford: Oxford University Press.

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.

van Koppen, Marjo. 2005. One probe, two goals: Aspects of agreement in Dutch dialects. Doctoral Dissertation, Leiden University.

Krifka, Manfred. 2017. Assertions and judgements, epistemics, and evidentials. Presented at Workshop 'Speech Acts', ZAS.

McFadden, Thomas. To appear. The morphosyntax of allocutive agreement in Tamil. In *Agree to Agree: Agreement in the Minimalist Programme*, ed. Katharina Hartmann, Johannes Mursell, and Peter Smith. Language Science Press.

McFadden, Thomas, and Sandhya Sundaresan. 2018. Unorthodox C-domain agreements: Complementizer agreement vs. allocutive agreement. Presented at CGSW 33, Göttingen.

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.