A plea for syntax and a return to first principles: monstrous agreement in Tamil*

Sandhya Sundaresan
Universities of Tromsø/Stuttgart

Abstract The paper focuses on an interesting form of (person) indexical shift in the Dravidian language Tamil which surfaces as 1SG agreement marking in a clause embedded under a speech predicate. I show that this agreement is an instance of indexical shift and label it "monstrous agreement". However, I demonstrate that its full range of empirical properties cannot be adequately explained by the major analyses of indexical shift in the literature. The bulk of these, I argue, in addition to being predominantly semantic in spirit, and thus ill-equipped to deal with a morphosyntactic phenomenon like agreement, also involve two core misconceptions regarding indexicality vs. logophoricity on the one hand and speech vs. attitude predicates on the other. I propose that these core assumptions be strongly re-evaluated from first principles and that syntactic and typological clues on the subject be paid more heed. I propose a new analysis of the Tamil paradigms which derives indexical shift within an enriched grammatical model involving contextual features instantiated in a structurally articulated cartographic left periphery.

Keywords: Monsters, indexicals, anaphors, Tamil, contextual and ϕ-features, Agree(ment), cartography, propositional complements.

1 The phenomenon: monstrous agreement

Consider the structure in (1) from the Dravidian language Tamil below:

(1)   Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ jey-pp-een-nnu]   so-nn-aan.
      Raman[NOM.SG]   SE[NOM.SG] win-FUT-1SG-COMP say-PST-3MSG

   "Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]."

   (1) contains a matrix speech predicate *sonnaan* (told.3MSG) which embeds a finite clausal complement. The subject of this embedded sentential complement is

---

taan[1] which is treated in the literature as a long-distance (i.e. SE-) anaphor that takes animate 3SG antecedents (Annamalai 1999). In (1), *taan* is obligatorily coreferent with the matrix (3rd person) subject *Raman* and is interpreted obligatorily *de se*.

The really curious property of (1) is the 1SG agreement marking on the embedded verb *jeyppeen*. Tamil is a subject-agreement language: as such, verbal agreement in Tamil typically reflects the ϕ-features of its clausemate subject. But in (1) the clausemate subject for the embedded T head is *taan* whose own ϕ-features, given its referential anaphoricity, may, indeed, be entirely absent (Kratzer 2009; Reuland 2011). So where does the 1SG agreement marking on the embedded verb come from? To make matters even more complicated, this 1SG agreement marking on the embedded verb (as well as *taan* itself) refer(s) not to the utterance speaker but to the matrix 3MSG subject *Raman* which denotes the speaker of the matrix speech event. This is the central puzzle.

1.1   Against a default agreement analysis

One plausible response to the agreement facts in (1) might be to claim that the 1SG agreement on the embedded verb actually exemplifies default agreement. However, default agreement in Tamil typically obtains in constructions involving experiencer subjects with dative case and is invariably marked 3NSG.[2] Also, the 1SG agreement in (1) isn't a frozen morphological form: it faithfully reflects the number feature of the matrix subject – see (2) below:

(2)   Pasaŋ-gaḷ$_i$ [$_{CP}$ taan-gaḷ$_{\{i,*j\}}$ jey-pp-oom-unnu]   so-nn-aan-gaḷ.
      boy[NOM.PL]   SE[NOM.PL$_{\{i,*j\}}$] win-FUT-1PL-COMP say-PST-3MPL

      "The boys$_i$ said [$_{CP}$ that they$_{\{i,*j\}}$ would win]."

Thus, a default agreement analysis doesn't seem viable for the embedded agreement marking in (1).

---

1 Crucially, *taan* does indeed seem to be an embedded subject in such stuctures and not, for instance, a focus adjunct reflexive on the matrix subject DP. Emphasis of the form "Raman himself ..." in Tamil is created by means of a focus clitic *-ee* which attaches to the focussed DP.

2 This is shown in (i) below:

i. Raman-ukku   koovam   va-nd-adu.
   Raman-DAT.SG anger[NOM.SG] come-PST-3NSG

   "Raman became angry."
   Lit. "Anger came to Raman."

1.2 Against a quotative analysis of 1SG agreement

Quotations may emulate shifted-indexicality effects because they "form a closed domain with respect to syntactic and semantic operators" (Anand 2006: 81). This grammatical opacity property of quotatives itself yields some useful diagnostics for testing the existence of quotatives:[3] for instance, an NPI inside a quotative may not be licensed by an operator outside the quote.

However (3), which involves the licensing of an embedded NPI by a matrix NEG operator, is fully grammatical showing that the embedded CP is not a full clausal quotative:[4]

(3) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ oru tapp=um
     Raman[NOM.SG] SE[NOM.SG] one mistake=even
     se-nj-een-nnu] ottukka-le.
     make-PST-1SG-COMP admit-NEG

"Raman$_i$ didn't admit [$_{CP}$ that he$_{\{i,*j\}}$ made any mistake]."

Finally, we can also rule out a partial quotative analysis for the structure in (1). A partially quoted string should be possible under any type of predicate. Thus, if the embedded CP in (1) did involve a partial quotative, we wouldn't expect the kind of systematic syntactico-semantic restriction of the phenomenon to the complements of propositional speech predicates.

1.3 Not a Dravidian quirk: crosslinguistic support for monstrous agreement

1SG agreement such as that in Tamil (1) cannot be dismissed as a Dravidian quirk either. Culy (1994); Curnow (2002) present data from Donna Sɔ, a Niger-Congo

---

3 In English, quoted elements may never be embedded under a complementizer, so the concomitant presence/absence of an overt complementizer can be used to tease quotatives and non-quotatives apart. But languages like Tamil and Japanese embed quotatives and non-quotatives alike under the same complementizer – thus, different diagnostics are needed.

4 For the sake of completeness, I present the minimal pair to (3) in (i) below – the embedded subject *taan* is replaced with the 1SG *naan* (I); this sentence is ungrammatical. (ii) shows that when the NEG operator locally binds the NPI, the sentence becomes fully acceptable:

i. * Raman$_i$ [$_{CP}$ naan$_{\{utt-speaker,*i\}}$ oru tapp=um se-nj-een-nnu] ottukka-le.
   Raman[NOM.SG] I[NOM.SG] one mistake=even make-PST-1SG-COMP admit-NEG

   "*Raman$_i$ didn't admit ["I$_i$ made any mistake."]"
   "Raman$_i$ didn't admit that I$_{utt-speaker}$ made any mistake"

ii. Raman$_i$ [$_{CP}$ naan$_i$ oru tapp=um seyya-le-nnu] so-nn-aan.
   Raman I one mistake=even make-NEG-COMP admit-PST-3MSG

   "Raman said "I didn't make any mistake"."

language, which has a structure that resembles that in Tamil (1) to a startling degree. The relevant paradigm is reproduced below:

(4) Oumar$_i$ [$_{CP}$ inyemɛ$_{\{i,*j\}}$ jɛmbɔ paza bolum] miñ tagi.
      Oumar SE[SUBJ] sack.DF drop left.1SG 1SG.OBJ informed

   "Oumar$_i$ told me [$_{CP}$ that he$_{\{i,*j\}}$ had left without the sack.]"

Curnow further argues that the embedded complement in such sentences does not constitute a direct speech report and states that such 1SG agreement, while quite rare, is evidenced in half a dozen other African languages. Further research must be undertaken to ascertain the full range of empirical properties evidenced by such structures – but I take the existence of sentences like (4) as evidence that such agreement is not unique to Tamil or to the Dravidian languages. It is indeed possible that further research on this topic will uncover other languages with similar monstrous agreement structures.[5]

   Now that we have seen what types of phenomena the structure in (1) does not represent, we can move on to the types of phenomena that it does represent. Below, I articulate two wellformedness conditions that appear to be necessary for the 1SG agreement seen on the embedded verb in (1).

1.4   Prerequisite 1: embedding under a propositional speech predicate

The first prerequisite for the "mismatched" 1SG agreement in (1) to obtain is that the superordinate verb be a speech predicate.[6] Thus, in (5) below, where the embedded CP is selected by a direct perception and attitude predicate, respectively, the agreement marking on the embedded verb straightforwardly reflects the ϕ-features of the matrix subject:

(5) Seetha$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ jey-pp-aa[-nnu]   nene-tt-aaḷ.
      Seetha[NOM.SG]   SE[NOM.SG] win-FUT-3FSG-COMP nene-PST-FSG

   "Seetha$_i$ thought [$_{CP}$ that she$_{\{i,*j\}}$ would win.]"

Equivalent patterns obtain under other propositional verbs such as epistemic SEE, factive KNOW, counter-factive PRETEND, and emotive FEAR. Speech predicates

---

5 A plausible reason for the typological rareness of monstrous agreement phenomena might be traced to the uncommonness of anaphors in agreement-triggering position (Rizzi 1990; Woolford 1999). As I will argue later, the presence of such an anaphor is a necessary condition for monstrous agreement.

6 The grammaticality judgments provided here correspond to my own native dialect of Tamil. An online survey that I've conducted among various speakers, however, suggests that, for some dialects, monstrous agreement can obtain under the scope of propositional attitude predicates like *nene* (THINK) in (5) as well. However, it is clear that speech predicates are special even for such dialects, in the sense that: if monstrous agreement obtains under an attitude predicate, it invariably also obtains under a speech predicate.

such as that in (1) are, thus, special in that they are capable of triggering this special 1SG agreement on the embedded verb in their scope.[7]

1.5 Prerequisite 2: the clausemate embedded subject is anaphoric

The second prerequisite is that the clausemate embedded subject be an anaphor. If the clausemate subject is any kind of non-anaphoric element, the agreement in that clause straightforwardly reflects the ϕ-features of this subject – this is illustrated below:

(6) Seetha   [$_{CP}$ ni:ŋ-gaɭ   jey-tt-i:ŋgaɭ-unnu]   so-nn-aaɭ.
     Seetha[NOM.SG] [you[NOM.PL]$_i$ win-PST-2PL-COMP] say-PST-3FSG

     "Seetha said [$_{CP}$ that you$_{pl}$ won]."

The 3rd person antecedence restriction of *ta(a)n*, noted earlier, in turn places the corollary restriction that the peculiar 1SG embedded agreement in structures like (1) only obtains when the superordinate subject is 3rd-person.

1.6 Enter indexical shift: 1SG agreement in (1) = "monstrous agreement"

In English, indexicals like *I*, *here* and *now* are always interpreted relative to the context of utterance. Thus, once the context of utterance is known, the denotations of these terms are fixed – an intuition that Kaplan (1989) captured by claiming that indexicals refer "directly" and that their denotation cannot be manipulated by intensional operators. Under Kaplan's model, *I*, *here* and *now* have the "rigid" semantic denotations given below:

$$[\![I]\!]^{c,i} = \lambda c\lambda i.Author(c)$$
$$[\![now]\!]^{c,i} = \lambda c\lambda i.Time(c)$$
$$[\![here]\!]^{c,i} = \lambda c\lambda i.Location(c)$$

However, typological evidence from hitherto less-studied languages, like Amharic (Schlenker 2003) and Indo-Iranian Zazaki (Anand & Nevins 2004), suggests that the denotations of indexicals in these and other languages can be manipulated by intensional operators. The Zazaki example from Anand & Nevins (2004) serves to illustrate this point:

(7) Hɛsen-i$_j$ (mɨ$_k$-ra)   va [$_{CP}$ kɛ   ɛz$_{j/k}$ dɛwletia].
     Hɛsen-OBL I-OBL.TO said   that I   rich.be.PRES

---

[7] Manner variants of SAY such as *kattu* (SCREAM), *muɳumuɳu* (MUMBLE), *monagu* (GROAN), and *polambu* (MOAN) as well as *arivi* (ANNOUNCE) appear to take clausal complements with 1SG verbal agreement.

"Hesen said (to me) [$_{CP}$ that {I am, Hesen is} rich.]"

In (7), the indexical pronoun εz (I) can refer to the speaker of the utterance context, just like in English; but it can crucially also refer to the speaker of the saying event denoted by the matrix verb, namely *Hɛsen*. Indexicals like εz, which are capable of altering their reference under certain intensional predicates, are called "shifted indexicals" or "monsters" (the Kaplanian moniker for such elements).

Returning to the Tamil data, the fact that 1SG embedded agreement as in (1) obtains under the scope of propositional speech predicates immediately suggests that it is an instantiation of indexical shift. First, indexical shift has also been reported to obtain predominantly under the scope of speech predicates (Anand 2006; Schlenker 2003; von Stechow 2002; Speas 1999; Shklovsky & Sudo To Appear). Second, the agreement in (1), just like the indexical in (7), is 1st-person, and refers to the speaker of the matrix speech context rather than to the speaker of the utterance context. I will thus take it as uncontroversial that the 1SG agreement in Tamil (1) involves an instantiation of indexical shift and will henceforth refer to it as "monstrous agreement".

At the same time, there is an important difference between the more "standard" cases of indexical shift, as in Zazaki (7) above, and Tamil monstrous agreement. The indexical shift in Zazaki involves a full-on DP whereas the Tamil case involves agreement marking on a functional head. This is a non-trivial distinction: agreement marking, unlike the ϕ-features on a DP, is not intrinsic to the category that hosts it but must be inherited, via some sort of feature-transmission or feature-copying mechanism like Agree (Chomsky 2001). Structures like (1) thus raise some additional questions. For instance, where is the 1st-person marking on the embedded predicate inherited from? What is the syntactico-semantic contribution of the matrix *say*-predicate, and what is the syntactico-semantic contribution of the clausemate anaphoric DP *taan* in effecting this agreement?

In the following sections, I discuss three major recent analyses of indexical shift: the context-overwriting analysis of Anand (2006), the pronoun-centric analysis of Schlenker (2003), and the LF feature-deletion and quantificational binding analysis of von Stechow (2002). In the process, I show that none of these analyses can satisfactorily account for the full range of empirical properties seen in (1) because they are not syntactic enough to deal with a primarily morphosyntactic phenomenon like agreement and because they involve some core misconceptions about the relative status of anaphors and indexicals and of speech vs. intensional predicates.

2   Analysis of indexical shift in Anand (2006)

Anand (2006) proposes that the difference between the shifted and unshifted readings above is not the denotation of the indexical *I*, which remains constant (at

$\lambda c.Author(c)$), but the value of the context that the AUTHOR function ranges over. In Anand's acccount, shifted indexicality is the result of *context shifting* due to context-overwriting by an operator introduced by an intensional predicate. Anand proposes that the type of context-overwriting operator is parametrized: in some languages, the operator overwrites every contextual coordinate in its scope whereas in others, only a proper subset of contextual coordinates is overwritten. This correctly yields the typological differences in shiftable indexicals observed crosslinguistically. In Zazaki, all types of indexicals may be shifted. Thus, the Zazaki operator functions as follows:

$$[\![OP_\forall\alpha]\!]^{c,i} = [\![\alpha]\!]^{j,i} \text{ where } j = <Auth(i), Addr(i), Time(i), World(i)>.$$

The shifted reading in (7) is thus derived as follows:[8]

[Hesen said [$_{CP}$ $OP_\forall$ I am rich]]^{c,i,g} = [Hesen said [$_{CP}$ I am rich]]$^{i',i',g}$
= 1 iff, [Hesen said [$_{CP}$ AUTHOR(i') am rich]]$^{i',i',g}$
= [Hesen said [$_{CP}$ I$_{hesen}$ am rich]]$^{i',i',g}$

2.1 Anand's analysis of Dravidian *taan*

Can Anand's context-overwriting analysis for shifted indexicals be exploited to shed light on the monstrous agreement phenomenon in Tamil? Interestingly, Anand discusses the anaphoric pronoun *taan* in Malayalam, a closely related Dravidian language, and claims that *taan* is not an anaphor but a shifted indexical based on two diagnostics. First, he shows that Malayalam *taan* receives an obligatory *de se* interpretation with respect to its antecedent. This interpretive property, for Anand, means that *taan* is one of two sorts of elements with dedicated *de se* LFs: an anaphor or an indexical. Anand then applies a diagnostic, termed the De Re Blocking Effect, to distinguish between these two possibilities for *taan*. The De Re Blocking Effect is a well-formedness condition on the interaction between *de se* anaphors and *de re* elements in the syntax which states that a(n obligatorily) *de se* anaphor may not be c-commanded by a *de re* element.[9] Thus, anaphors are expected to be sensitive to the De Re Blocking Effect whereas indexicals are not. In the Malayalam sentence in (8) below (Anand 2006: 155) (formatting mine), the embedded DP *tannai* which is

---

8 **Unshifted Reading A:** [Hesen said [$_{CP}$ I$_{sandhya}$ am rich.]] (UNSHIFTED)

There is no context overwriting; the indexical *I* is interpreted against the utterance context (perhaps introduced at root C) and is thus mapped onto the utterance speaker = me, Sandhya (in this utterance context).

9 The interested reader is referred to Chapter 1, Section 1.4 of Anand (2006) for detailed argumentation and motivation of this generalization.

de se is c-commanded by the *de re* element *awanooɖə* in direct violation of the De Re Blocking Effect, but the resulting sentence is grammatical:

(8) [[Mary   tan-nai$_{\{i,j\}}$ sneehikkuŋŋu eŋŋə] Bill$_j$ awan-ooɖə$_i$ paraññu
     Mary[NOM.SG] SE-ACC.SG loves   COMP Bill   him-to   said
     eŋŋə] John$_i$ wicaariccu.
     COMP John$_i$ thought

     "John$_i$ thought [that Bill$_j$ told him$_i$ [that Mary loves him$_{\{i,j\}}$]]"

Based on this, Anand proposes that *taan* is not an anaphor but an (obligatorily shifted) indexical and attempts to derive its properties via his context-overwriting approach. Unfortunately, there are some non-trivial problems with this, as I discuss below. These issues are compounded for Tamil which differs from Malayalam in one very important respect. Malayalam, unlike Tamil, lacks all (verbal) agreement marking and thus exhibits no morphological agreement differences such as those we have seen for Tamil.[10]

2.2 Problem I: Antecedent-restriction to 3rd person DPs

The first major problem for a context-overwriting approach is that Tamil as well as Malayalam *taan* only take 3rd person antecedents (9). However, such a restriction is entirely unexpected if, as Anand claims, these elements really are 1st-person indexicals underlyingly:

(9) * Naan/Nii   [$_{CP}$ taan   poo-r-een-nnu]
      I[NOM.SG]/You[NOM.SG]   SE[NOM.SG] go-PRES-1SG-COMP
      sol-r-een/sol-r-aai
      say-PRES-1SG/say-PRES-2SG

      "I say, 'I will go.'"
      "You say, 'I will go.'" (Intended)

Under Anand (2006)'s context-overwriting approach, this antecedence restriction would have to be captured by claiming that a speech predicate like SAY in Tamil (and Malayalam) may introduce a context-shifting operator only if it is itself specified as 3rd person. But such a condition, while workable, is, as Anand himself concedes, very stipulative and perhaps even unexpected: why, after all, should the (putatively) semantic context-overwriting operator be sensitive to the morphosyntactic features of the predicate that introduces it?

---

10 The other major Dravidian languages are Telugu and Kannada, both of which do have overt agreement marking, and I am informed (by Rahul Balusu, p.c.) that Telugu exhibits monstrous agreement in structures analogous to that in Tamil (1).

2.3 Problem II: The 3rd-person agreement paradigm in Tamil (5)

The second major problem under a context-overwriting approach is how to account for 3rd-person embedded verbal agreement as in (5). Interestingly, this is a problem that Anand (2006) doesn't deal with for Malayalam because it has no verbal agreement, as noted above. If we were to extend Anand's analysis to the Tamil data, we would, essentially, be forced to derive these differences as a function of featural distinctions on *taan* in (5) vs. (1) – such as, for instance, by stipulating (in the absence of any independent evidence) that *taan* is an anaphor in the former but an indexical in the latter. This is clearly also not an attractive result.

2.4 Problem III: *naan* vs. *taan*

Tamil has a purely "English-like" 1SG indexical *naan*:

(10)   Raman [$_{CP}$ naan jey-pp-een-nnu]   so-nn-aan.
       Raman   I   win-FUT-1SG-COMP say-PST-3MSG

   "Raman said [$_{CP}$ that I$_{utt-speaker}$ will win]." (Indirect speech)
   "Raman said, 'I will win'." (Quotative/Direct speech)

If indexical shift obtains due to context-overwriting at LF, as Anand proposes, and *taan* and *naan* are both indexicals, it is unclear how the morphophonological distinction between the two is to be captured. The problem, it should be noted, is not insurmountable: one could, for instance, propose that a 1SG indexical that is introduced by a context-overwriting operator is spelled out as *taan* whereas one that is not is spelled out as *naan*. However, the formulation of such rules requires an articulated feature structure for indexicals and contextual operators in the syntax, which is lacking in Anand's analysis.

3   Problems with the pronoun-centric model (Schlenker 2003)

Schlenker (2003) attempts to derive indexical shift as a function of the syntactico-semantic properties of the indexicals themselves. Simply put, some indexical pronouns, like those in English, are inherently specified as being unshifting with respect to the utterance context; others, like those in Amharic and Zazaki, are underspecified as to whether they can shift; still others, such as logophors in e.g. Gokana (Hyman & Comrie 1981) and other languages, are semantically specified to always shift relative to the utterance context. The formal denotations of these indexicals are, thus, as follows:[11]

---

11 c* is a free variable that refers to the context of utterance c; δ represents the presuppositions introduced by the features; and a pro-form is represented as a free variable whose reference is constrained by semantic presuppositions (Heim & Kratzer 1998).

$$[\![I_{Eng}]\!]^{c,g} = [\![x_i + Author * (x_i)]\!]^{c,g} = s(x_i) \wedge \delta(g(x_i) \text{ is Author}(c))$$
$$[\![I_{Amh}]\!]^{c,g} = [\![x_i + Author(x_i, c_j)]\!]^{c,g} = s(x_i) \wedge \delta(g(x_i) \text{ is Author}(c_j))$$
$$[\![I_{Log}]\!]^{c,g} = [\![x_i + Author(x_i, c_j)]\!]^{c,g} = s(x_i) \wedge c_j \neq c^*: \delta(g(x_i) \text{ is Author}(c_j))$$

Since much of the action is localized to the denotations of the indexicals themselves, Schlenker's model would be able to capture the distinction between unshifting *naan* (10) and shifting *taan* in Tamil (1) and (5). Specifically, *naan* would have the denotation of an English-like "I" whereas *taan* would have the denotation of a logophoric "I" – as formalized above. I take this to be an advantage of Schlenker's system over Anand's. Nevertheless, the other two problems mentioned for Anand remain an issue for Schlenker as well. That is, it is unclear how to derive the distinction between monstrous agreement in (1) and 3rd-person agreement in (5) given that both are presumably triggered by the same underlying indexical pro-form *taan*. It is also unclear how to derive the 3rd person antecedence restriction of *taan* in both (1) and (5) if *taan* is, indeed, underlyingly a 1st-person indexical.

3.1   Why do the semantic analyses fail?

The fact that Anand's and Schlenker's analyses come up against the same empirical issues with regards to the Tamil data is not an accident. Below, I argue that these problems can be traced back to three core issues. The first issue has to do with assumptions regarding the relative syntactico-semantic status of indexicals and anaphors/logophors. The second problem is that neither analysis presents a clear model of the morphosyntactic issues associated with indexical shift. The third problem has to with the fact that speech predicates are not syntactico-semantically privileged relative to other intensional predicates with respect to their role in inducing indexical shift.

4   Housecleaning issue I: indexicals vs. anaphors

As we've already seen, Schlenker treats all anaphors/logophors as indexicals that are specified to obligatorily shift. Anand assumes that there is a fundamental difference between indexicals and anaphors but, as discussed above, treats (Malayalam) *taan* as a shifted indexical for independent reasons. The end result is that both Anand and Schlenker end up treating *taan* as an obligatorily shifted indexical, rather than an anaphor. It is this assumption, I propose, that is at the root of many of the problems that these analyses face in accounting for the Tamil data.

If we assumed instead that *taan* is actually an anaphor rather than an indexical, two core problems, namely the 3rd person antecedence restriction and the *naan* vs. *taan* distinction, would disappear straightaway. Anaphors typically take 3rd person

antecedents (Reinhart & Reuland 1993); thus, the 3rd person antecedence restriction of *taan*, far from being anomalous, would be characteristic. The *naan* vs. *taan* problem would also be obviated. Under the new taxonomy, *naan* would be the only 1st-person indexical pronoun in Tamil and it would be a rigid, Kaplanian indexical; as an anaphor, *taan* simply wouldn't compete for its place.

Crucially, claiming that *taan* is an anaphor in Tamil is not merely a strategic ploy to account for the facts at hand. Strong empirical evidence both from thorough crosslinguistic typological (Sells 1987) and syntactic (Huang & Liu 2001; Speas & Tenny 2003; Bianchi 2003) studies shows that anaphors occur in a wider range of syntactico-semantic environments than do shifted indexicals. According to Sells (1987), anaphors may indicate the SOURCE of information, the internal SELF whose mind is being reported or the physical PIVOT or viewpoint of an individual. As such, anaphoric elements may occur under the scope of propositional attitude (BE-LIEVE), factive (KNOW), epistemic (SEE), motion (COME) as well as propositional speech predicates (SAY). In contrast, shifted indexical pronouns occur in much more restricted environments: in languages where this phenomenon is morphophonologically distinguished, they have been shown to occur predominantly under the scope of speech predicates. The distribution of Tamil *taan* thus fits in perfectly with that of logophors crosslinguistically and less with that of indexicals, shifted or otherwise. Finally, it is worth noting that there is a more familiar distinction between anaphors and indexicals. This is the fact that the latter, but crucially not the former, may occur as the subjects of matrix clauses (Baker 2008; Delfitto & Fiorin 2011): *taan* resembles an anaphor in this respect as well, as it typically cannot occur as a matrix subject.[12]

Based on such pervasive differences, I thus submit that *taan* should be treated as an anaphor and not an indexical. Therefore, both the Schlenkerian idea that all anaphors are obligatorily shifted indexicals and Anand's argument that the De Re Blocking effect is both a necessary and sufficient diagnostic to evaluate the distinction between indexicals and anaphors should be subjected to further scrutiny.

5   Housecleaning issue II: Agree(ment) matters

If *taan* is an anaphor in both (1) and (5), which element is the indexical? The obvious candidate would be the agreement marking on the embedded verb in (1) since this alone alters its surface form under the scope of a speech predicate like *sonnaan* (SAY.3MSG) in (1). However, as we've seen, agreement is not inherent to the categorial head (T) on which it appears but is inherited from another element in its local syntactic domain. In Tamil, agreement on T is typically triggered by

---

12 The well-known exception to this is the case of free indirect discourse where *taan* may occur as the matrix subject. But this is a property that is common to logophors crosslinguistically (Sells 1987).

a clausemate subject; thus, the obvious source for the monstrous agreement in (1) would be the clausemate subject *taan* itself. Claiming that *taan* itself is an anaphor and not an indexical, therefore, is equivalent to severing this direct connection between *taan* and the agreement on its clausemate verb.

But why should agreement behave differently just in those cases where an anaphoric subject is involved? There is a potential answer to this question: anaphors crosslinguistically have been known to either not trigger agreement at all or trigger special anaphoric agreement forms, a property that has been termed the Anaphor Agreement Effect (AAE) (Rizzi 1990; Woolford 1999). Indeed, Woolford (1999) proposes that the monstrous agreement in sentences like (1) represents a grammatical conspiracy to avoid a violation of the AAE. This proposal, though interesting, is not sufficient to explain the Tamil paradigms: specifically, it cannot explain why the same apparent element *taan* should trigger monstrous agreement in (1) but a different agreement in (5). Nevertheless, the idea that anaphors have a special relationship with agreement has extensive crosslinguistic support and is thus worth taking seriously. I will outline a preliminary analysis toward the end of this paper that proposes that the syntactic feature content – or more specifically, the lack of the relevant syntactic feature content – of *taan* prevents it from valuing the ϕ-features (specifically the person feature) of an upward-probing clausemate T head, with the result that this T gets its features valued higher up in the clausal left periphery.

6 Housecleaning issue III: attitude vs. speech predicates

The final issue is that there isn't enough of a principled distinction made between speech- and other attitude predicates. The logic behind this (often tacit) conflation is a semantic one: attitude verbs like BELIEVE, as well as speech verbs like SAY, have been observed to be capable of quantifying over worlds or over an enriched intensional index involving worlds, times, and locations (Lewis 1980). Schlenker (2003) proposes that these intensional indices are ontologically on a par with contextual indices of the Kaplanian type, thus blurring the distinction between predicates that introduce contextual features and those that introduce intensional ones. Thus, analyses involving indexicality and indexical shift, which are thought to involve these same contextual variables, treat speech- and attitude predicates on a par. Here, I do not dispute the notion that speech predicates and attitude predicates are semantically similar in both being able to introduce rich propositions. However, I will argue, primarily from the typological and syntactic perspectives, that speech predicates alone may introduce contextual variables in their scope and that there is a relationship of structural containment (rather than identity) in the representation of speech- and (other) attitude predicates in the syntax.

Strong empirical evidence from the syntactic and typological literature indicates

that propositional speech predicates are underlyingly special. Cristofaro (2005) points out that there is a strong implicational hierarchy in the size of complement clauses that are selected by propositional predicates which corresponds to the lexico-conceptual semantics of these verbs. Interestingly, this same hierarchy has been independently argued to serve as an implicational hierarchy for logophoric predicates (Culy 1994; Speas & Tenny 2003):

(11) COMPLEMENT SIZE AND LOGOPHORIC PREDICATE HIERARCHY:
      Speech » Thought » Knowledge » Direct perception
      Implication for logophors: If a language allows a logophor in the complement of one type of predicate, then it also allows a logophor in the complement of predicates higher in the hierarchy.
      Implication for complement size: Speech complements are structurally the largest; Direct perception complements are structurally the smallest.

Hooper & Thompson (1973), classifying predicates according to somewhat different criteria, demonstrate that certain root-transformations (such as VP-preposing and topicalization) can apply to embedded clauses with a degree of ease that corresponds to the hierarchy above: thus, root-transformations apply most easily to speech predicates, and least easily to direct perception predicates. Wiklund, Bentzen, Hrafnbjargarson & Hróarsdóttir (2009) use the possibility of embedded V2 in Scandinavian languages to confirm the ordering above. Finally, Cinque (1999) presents data from the relative ordering and interpretive scope of adverbs and verbal affixes to demonstrate that these also respect the same hierarchy. This type of evidence yields the Cinque Hierarchy below:

(12) **Cinque Hierarchy:** Speech Act » Evaluative » Evidential » Epistemological.

Cinque's hierarchy essentially rephrases Cristofaro's predicate ranking into a hierarchy involving functional heads selected by the respective predicates. Thus, a speech predicate selects a complement of the largest size (containing the Speech Act head and all the ones below it) whereas an epistemological one selects the smallest one.

6.1 Relevance to indexical shift

How is all this relevant to the case of indexical shift? Recall that indexical shift is predominantly induced under the scope of speech predicates. I reproduce an updated version of the mini typological table for indexical shift from Anand (2006: 110).

It is immediately apparent from Table (1) that all these languages allow indexical shift under propositional speech predicates. Slave, the lone language which also permits indexical shift under WANT, nevertheless obeys the implicational hierarchy: it also allows indexical shift under a predicate that is higher in the hierarchy, like

LANGUAGE	VERB	CLASS DESCRIPTION
Amharic, Aghem	SAY	optionally shifts 1st/2nd person indexicals
Navajo	SAY	optionally shifts 1st/2nd person indexicals
Slave	TELL	optionally shifts 1st/2nd person indexicals
	WANT	optionally shifts 1st person indexicals
	SAY	obligatorily shifts 1st person indexicals
Zazaki	SAY	optionally shifts all 1st person indexicals
Tamil, Telugu	SAY	obligatorily shifts 1st-person verbal agreement
Donna Sɔ (?)	SAY	obligatorily shifts 1st-person verbal agreement
English	ALL	no indexical shift

Table 1   Mini-typology of indexical shift across languages.

SAY and TELL. If we take seriously the idea that speech predicates are special and different from attitude predicates and all other propositional predicates (and we should) the fact that indexical shift is always effected under propositional speech predicates would not be an accident.

Before moving on to an explication of my own analysis, I first briefly discuss another prominent, and slightly more syntactic, analysis of indexical shift, namely that of von Stechow (2002).

7   Feature-deletion and verbal-quantifier binding (von Stechow 2002)

von Stechow (2002) proposes that attitude predicates are "universal quantifiers which take an egocentric proposition [a proposition that contains contextual variables pertaining to the person, tense, and mood] as their complement". Indexical shift, in this model, obtains when an indexical pronoun is bound by such an attitude-verbal quantifier.[13] Von Stechow crucially claims that the features of a variable are deleted at LF when they are bound: such feature-deletion is, in fact, a precondition on variable binding. The difference between languages that manifest (person) indexical shift, like Amharic or Tamil, and those that don't, like English, is a function of whether the features on the indexical may be deleted by the c-commanding verbal quantifier or not (respectively).

For languages like English, von Stechow claims that feature-deletion obtains just in case the features of the verbal quantifier match those of the indexical in value.

---

13 This is an extension of von Stechow's analysis for bound-variable indexicals, – like "my" in the sentence: "Only I finished drinking **my** beer". This sentence, perhaps more patently, involves the variable-binding of an indexical by a c-commanding quantifier, namely "only I".

This ensures that an attitude predicate that is marked 3rd person will never delete the features of an (1st/2nd) indexical pronoun in its scope and will, by extension, never bind it. As such, an indexical pronoun in English will never "shift" its reference to a 3rd person attitude-holder that c-commands it but will always refer deictically, as desired. In languages that do evidence indexical shift, such as Amharic, the rules governing feature-deletion are different. In such languages, the "verbal quantifiers delete ([at] LF) the feature **1st** of the person variable they bind, regardless of what their person checkee is." Thus, in a sentence like "John said$_{3sg}$ [$_{CP}$ that I am hungry]" in Amharic, the features of the embedded indexical pronoun "I" could be deleted and then bound by the 3SG features of the attitude predicate checkee. Crucially, this feature deletion will happen at LF, so for the purposes of morphophonology, the pronoun will still surface as "I". The formal binding operations are reproduced below (von Stechow 2002):[14]

(13)   Target sentence: "John says I am a hero." (Amharic)

a. John^{3rd} says*^{3rd} λ<x^{1st} …>…x$^{\text{1st}}$ am*$^{\text{1st}}$ a hero.
   "John says 'I am a hero.'" (**Shifted reading**)

b. John^{3rd} says*^{3rd} λ<x^{1st} …>…y^{1st} am*$^{\text{1st}}$ a hero.
   "John says 'You are a hero.'" (**Unshifted reading**)

What about logophors? Von Stechow claims that logophoric pronouns are variables with Case- and ϕ-features and additionally proposes (though without much independent argumentation) that they possess an inherent feature LOG (presumably to differentiate them from 3rd person deictic pronouns). He tentatively also proposes that they be assigned a feature -1ST to account for the fact that they do not occur in the 1st person.

Von Stechow's system does seem better equipped to handle the issues manifested by the Tamil data in (1) and (5). For instance, deriving the distinction between *naan* and *taan* is fairly unproblematic in von Stechow's system: the former would be an English-like, rigid indexical whereas the latter would be treated as a logophoric pronoun.[15] In von Stechow's model, this means that *taan* would have the features {3RD, NOM, SG, LOG} whereas *naan* would be marked as {1ST, NOM, SG}. Crucially, both would thus be treated as underlyingly different elements in the morphosyntax. The featural properties of *taan* would also automatically account for the 3rd person antecedence requirement on this element – a property that was genuinely difficult to

---

14 A quick note on the notation adopted here. For a privative feature F, [F] = the interpretable counterpart of F; [*F] = the uninterpretable/inherited feature F.

15 *taan* fulfills von Stechow's description of logophoric elements: it is a pronoun that is associated with an obligatory *de se* interpretation, occurs in the scope of attitude predicates, and takes a 3rd person antecedent. Thus, it seems reasonable to think that, all else being equal, *taan* would be treated as a logophor in von Stechow's system.

account for in both Anand's and Schlenker's analyses where *taan* would have been treated as an indexical, not a logophor.

7.1 Problematic: Monstrous vs. 3rd-person agreement in Tamil

The central problem that the von Stechow model faces is the difference between 3rd-person and monstrous agreement patterns in (5) and (1), respectively. Von Stechow distinguishes between interpretable and uninterpretable features and assumes a standard Minimalist model of local feature-checking where agreement marking on a T head in a subject-agreement language is straightforwardly inherited from a clausemate subject. Recall that logophoric *taan* possesses the interpretable features {3RD, LOG, NOM, SG} in both (1) and (5) in this analysis. As such, in both structures, *taan* would transmit the features {3RD, LOG, NOM, SG} to its clausemate T head in the Narrow Syntax. The latter would thus possess the uninterpretable features: {*3RD, *LOG, *NOM, *SG} in the syntax – and these features would be straightforwardly spelled out at PF. As such, the agreement marking on the embedded verb in both (5) and (1) is expected to be the same – in blatant contradiction of the actual facts.[16] Thus, although unlike Schlenker and Anand, von Stechow does provide a basic model of feature-structure and feature-checking, he doesn't go far enough. A related problem for von Stechow is that, in line with both Anand and Schlenker, he doesn't distinguish between the types of propositions introduced by speech, attitude, and other types of predicates. As such, under his model, an attitude predicate like that in (5) and the speech predicate in (1) are both expected to introduce egocentric propositions of the same type. As such, the fact that monstrous agreement obtains under a speech predicate in (1) and not under the intensional-predicate in (5) must be reduced to mere accident.

8 The beginnings of an alternative model: cartography meets contextual features

I take seriously the idea that speech predicates are underlyingly different from other types of propositional predicates. Specifically, I assume that the complements of speech predicates crosslinguistically are structurally larger than those selected by other propositional predicates, in line with the Cinque hierarchy discussed above. I will also assume, following proposals by Speas & Tenny (2003); Baker (2008) and others that the Kaplanian contextual coordinates of <*Speaker, Addressee, Time, World, (Location)*> are present in [Spec, CP] of all root clauses and some embedded ones and play a crucial role in determining person, tense and mood features in the

---

16 Note that the feature-deletion rule covariable-binding should not affect this in any way since it is held to only happen at LF.

syntactic structure in their scope. Putting these assumptions together, I propose the following generalization:

(14) CONTEXTUAL FEATURE GENERALIZATION (CFG):

   a. For any language, only those clausal complements that contain a Speech-Act Phrase are large enough to host their own contextual features in the syntax.

   b. These contextual features are hosted in the specifier of the Speech-Act Phrase (defined in the sense of Cinque (1999).

I additionally put forth the following formal claims:

   I. Informally, an anaphor is looking to represent the mental/physical POV of some sentient individual (Sells 1987; Koopman & Sportiche 1989; Speas 2004). Formally, an anaphor is an element with an unvalued POV feature; this feature takes a reference index (or ID feature (Hicks 2009, Adger and Ramchand 2005) of a sentient individual as its value.

   II. One of the specifiers in the clausal functional sequence hosts a valued POV feature which is responsible for checking the unvalued POV feature on an anaphor in its c-command domain. This specifier is located crucially below the Speech-Act Phrase in the Cinque hierarchy.[17]

   III. Tamil *taan* is an anaphor, not an indexical. Given that that it can be bound even under direct perception/evidential predicates, I propose that the valued POV feature it agrees with is hosted in [Spec, EvidP].

   IV. Anaphors are incapable of triggering ϕ-agreement, presumably because they are themselves ϕ-deficient: i.e., the Anaphor Agreement Effect (AAE) holds in some form.

8.1 Deriving the monstrous agreement structure in (1)

Consider again the Tamil monstrous agreement structure in (1) repeated as (15) below:

(15) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ jey-pp-een-nnu]   so-nn-aan.
      Raman[NOM.SG]   SE[NOM.SG] win-FUT-1SG-COMP say-PST-3MSG

      "Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]."

---

17 For now, I follow Speas (2004) in assuming that a POV head may, in theory, be hosted within any phrase in the functional sequence. The choice of phrase is parametrized and results in different distributions for anaphors crosslinguistically, as reported by Culy (1994).

This sentence can now be derived as follows. The SAY-class verb in (15) selects a maximally large complement, i.e. Speech-ActP which hosts contextual features pertaining to the matrix speech event. The ϕ-features of the matrix subject in (15) are transmitted to the left periphery (via feature-transmission, in a manner to be formalized). The 3MSG person feature of the matrix subject *Raman* is re-evaluated under the new context: as Raman is the speaker of the matrix speech event, it is reset as 1st-person. The embedded T head has unvalued ϕ-features but cannot get these valued by its clausemate anaphoric subject *taan* (due to the AAE). As such, T keeps probing upward until its person feature gets valued (as 1st-person) by the contextual person feature in [Spec, SpeechAct-P]. The reference index *Raman* is separately feature-transmitted to [Spec, EvidP]: this in turn checks the unvalued POV feature of *taan* as [POV: *Raman*]. The result is that the embedded verb is valued as 1SG and refers to the matrix speaker *Raman*; *taan* also refers to the matrix speaker *Raman* by virtue of representing its mental point-of-view.

8.2 Deriving the 3rd-person agreement paradigm in (5)

Consider the structure in Tamil (5), repeated below:

(16) Seetha$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ jey-pp-aa[-nnu]   nene-tt-aaḷ.
       Seetha[NOM.SG]   SE[NOM.SG] win-FUT-3FSG-COMP nene-PST-FSG

  "Seetha$_i$ thought [$_{CP}$ that she$_{\{i,*j\}}$ would win.]"

The crucial difference between (16) and (15) is the choice of matrix predicate in each; this, in turn, brings about the difference in agreement paradigms. The matrix verb in (16), being an evaluative predicate, selects an EvalP which is structurally smaller than a SpeechAct-P. By (14) above, contextual features may only be hosted on [Spec, SpeechAct-P]; as such, the embedded EvalP in (16) is not anchored to a new speech context.

The 3FSG ϕ-features of the matrix subject in (16) are featurally transmitted to the highest projection in the left periphery, just like with (15). However, in this case, the leftmost position is [Spec, EvalP]. Thus, the person features of the matrix subject *Seetha* are not context-shifted or "reset" (as with (15)) but are straightforwardly evaluated against the context of utterance, as 3FSG. The embedded Agr probes upward, again cannot get its ϕ-features valued by the anaphoric taan (due to the AAE), and keeps probing until it reaches EvalP in the left periphery. The inherited ϕ-features on [Spec, EvalP] value the embedded Agr with the ϕ-features of the matrix subject, namely as 3FSG. Meanwhile, the reference ID of the matrix subject *Seetha* is feature-transmitted to [Spec, EvidP] and values *taan* as [POV: *Seetha*]. The relevant agreement and binding paradigms are thus accounted for.

References

Anand, Pranav. 2006. *De de se*: Massachusetts Institute of Technology dissertation.

Anand, Pranav & Andrew Nevins. 2004. Shifty operators in changing contexts. In *Proceedings of SALT 14*, .

Annamalai, E. 1999. Lexical anaphors and pronouns in Tamil. In Barbara C. Lust, Kashi Wali, James W. Gair & K.V. Subbarao (eds.), *Lexical and anaphors and pronouns in selected South Asian languages: a principled typology*, 169–216. Mouton de Gruyter.

Baker, Mark C. 2008. *The syntax of agreement and concord*. United Kingdom: Cambridge University Press.

Bianchi, Valentina. 2003. On finiteness as logophoric anchoring. In Jacqueline Guéron & L. Tasmovski (eds.), *Temps et point de vue/Tense and Point of View*, 213–246. Nanterre: Université Paris X.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, Cambridge, Mass.: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective* Oxford Studies in Comparative Syntax. New York: Oxford University Press.

Cristofaro, Sonia. 2005. *Subordination*. Oxford: OUP.

Culy, Christopher. 1994. Aspects of logophoric marking. *Linguistics* 32(6). 1055–1094.

Curnow, Timothy Jowan. 2002. Verbal logophoricity in African languages. In *Proceedings of the 2002 Conference of the Australian Linguistic Society*, .

Delfitto, Denis & Gaetano Fiorin. 2011. Person features and pronominal anaphora. *Linguistic Inquiry* 42(2). 193–224.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell Publishers.

Hooper, Joan & Sandra Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4. 465–497.

Huang, James & Luther Liu. 2001. Logophoricity, attitudes, and *ziji* at the interface. In Peter Cole, Gabriella Hermon & James Huang (eds.), *Long-distance reflexives*, 141–195. New York: Academic Press.

Hyman, Larry M. & Bernard Comrie. 1981. Logophoric reference in Gokana. *Journal of African Languages and Linguistics* 3(1). 19–37.

Kaplan, David. 1989. Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In John Perry, Joseph Almog & Howard Wettstein (eds.), *Themes from Kaplan*, 481–563. Oxford: Oxford University Press.

Koopman, Hilda & Dominique Sportiche. 1989. Pronouns, logical variables, and

logophoricity in Abe. *Linguistic Inquiry* 20(4). 555–588.

Kratzer, Angelika. 2009. Making a pronoun: fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2). 187–237.

Lewis, David. 1980. Index, context, and content. In Stig Kanger & Sven Öhman (eds.), *Philosophy and Grammar*, 79–100. Dordrecht: Reidel.

Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24. 657–720.

Reuland, Eric. 2011. *Anaphora and language design*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1990. On the anaphor-agreement effect. *Rivista di Linguistica* 2. 27–42.

Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and Philosophy* 26. 29–120. Kluwer Academic Publishers.

Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18(3). 445–479.

Shklovsky, Kirill & Yasutada Sudo. To Appear. Indexical shifting in Uyghur and the syntax of monsters. In *Proceedings of NELS 40*, .

Speas, Margaret. 1999. Person and point of view in Navajo. In *Proceedings of the west coast conference on formal linguistics 18 (wccfl 18)*, .

Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114(3). 255–276.

Speas, Peggy & Carol Tenny. 2003. Configurational properties of point of view roles. In Anna Maria Di Sciullo (ed.), *Asymmetry in grammar*, vol. I Syntax and Semantics, 315–372. Amsterdam: John Benjamins.

von Stechow, Arnim. 2002. Feature deletion under semantic binding: tense, person, and mood under verbal quantifiers. In *Proceedings of NELS 33*, .

Wiklund, Anna-Lena, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson & Þorbjörg Hróarsdóttir. 2009. On the distribution and illocution of V2 in Scandinavian *that*-clauses. *Lingua* 119. 1914–1938.

Woolford, Ellen. 1999. More on the anaphor agreement effect. *Linguistic Inquiry* 30(2). 257–287.

Sandhya Sundaresan
CASTL, University of Tromsø
9037 Tromsø, Norway
sandhya.sundaresan@uit.no