Monsters, agreement, and anaphora: evidence from Tamil[*]

Sandhya Sundaresan

University of Tromsø(CASTL)/Stuttgart

1. The puzzle

The focus of this presentation is structures like (1) in Tamil:

(1) raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ jey-pp-een/*aan-nnu] so-nn-aan
raman-NOM *taan*-NOM$_i$ win-FUT-1SG/*3MSG-that say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

The structure in (1) contains a speech predicate *sonnaan* (told.3MSG) which embeds a (finite) clausal complement. This CP complement has a subject *taan*: *taan* is traditionally described as a nominative, 3SG (optionally) long-distance (or SE- (Reinhart and Reuland 1993)) anaphor (Annamalai 1999).[1] *taan* is obligatorily coreferent with the matrix (3MSG) subject and attitude-holder, *raman* and is, crucially, interpreted obligatorily *de se* with respect to its DP antecedent. Crucially, the embedded indicative verb *poo-n-een* (go-PST-1SG) doesn't seem to agree with either its clausemate embedded subject *taan*. Also, this 1SG agreement only surfaces when the clausemate subject is *taan* (as shown in (1)) or *naan* (I) – showing that it is the clausemate subject that is responsible for the surface verbal agreement.

[*]I thank Pavel Caha, Gennaro Chierchia, Alessandra Giorgi, Angelika Kratzer, and Peter Svenonius for taking the time to meet with me and discuss aspects of this paper – and Ken Safir for additionally sharing generously of unpublished work and curious binding data. I also thank the audiences at the Local Pronouns Workshop in Nijmegen, GLOW-in-Asia in Beijing, Peculiar Binding Configurations Workshop in Stuttgart, NELS 41 in Philadelphia, and the Grammar of Attitudes Workshop at DGFS in Göttingen, where versions of this paper were presented, for their feedback, and my Tamil native-speaker informants for confirming the existence of monsters in Tamil. I am especially grateful to Thomas McFadden and Gillian Ramchand for extensive meetings and discussions of the fledgling and final forms of this paper. Any errors are mine alone.

[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 clictic *-ee* which attaches to the focussed DP. In the course of this paper, I argue that *taan* is simultaneously anaphoric and indexical in nature. Since this is far from obvious at the moment, I simply gloss "*taan*" as *taan* throughout this paper.

It has been proposed in Woolford (1999) that the lack of embedded subject-verb agreement in (1) is a grammatical conspiracy to avoid a violation of the Anaphor Agreement Effect/AAE – the idea, going back to Rizzi (1990), that anaphors are crosslinguistically disbarred from agreement(-triggering) positions. Problematically for this hypothesis, however, when the same subordinate clause in (1) is embedded under a non-speech predicate, like *kaɳḍupiḍi* (find out) or a perception verb like *paar* (see) "real" agreement is not only possible, it is obligatorily required.

(2) raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ jey-pp-aan/*een-nnu] paar-tt-aan
raman-NOM$_i$ *taan*-NOM$_i$ win-FUT-3MSG/*1SG-that see-PST-3MSG
"Raman$_i$ saw [$_{CP}$ that he$_{\{i,*j\}}$ would win.]"

Crucially, furthermore, this 3MSG agreement on the embedded verb may only surface with a masculine, 3rd-person clausemate subject, thus cannot be dismissed as default agreement. Based on these reasons, I propose, contra Woolford (1999) that structures in Tamil like (1) are not a function of the Anaphor Agreement Effect and, in fact, stand as direct counterexamples to the AAE. But then, what causes the apparent lack of agreement between the embedded verb and anaphoric *taan* in (1)? Conversely, why is subject-verb agreement necessitated in (2)?

2. A first inkling: similarities to indexical shift

The matrix predicate in (1), *so-nn-aan* (say-PST-3MSG) is a speech-predicate; its clausal complement thus denotes a speech-report. The matrix predicate in (2) *paar-tt-aan* (see-PST-3MSG) is a non-reportive predicate and its clausal complement does not constitute a speech-report. Now consider the shifted indexical structure from Zazaki below (Anand and Nevins 2004):

(3) hɛsen-i$_j$ (mɨ$_k$-ra) va kɛ ɛz$_{j/k}$ dɛwletia
Hesen-OBL I-OBL-TO said that I rich.be-PRES
"Hesen said that {I am, Hesen is} rich."

In (3), the indexical pronoun *ɛz* (I) can refer to the author of the utterance context or, anomalously, to that of the speech-report *hɛsen*. Thus, the embedded indexical pronoun is referentially ambiguous: a shifted indexical or Kaplanian monster (Schlenker 1999, von Stechow 2002, Anand 2006). Significantly, shifted indexicality typically only occurs in contexts that are under the scope of speech-predicates.

This provides us with an initial clue as to what might be going on in structures like (1) in Tamil. Let us review the similarities between the two constructions. Just as in (3), the 1st-person marking on the embedded verb in (1) is shifted to refer to the attitude-holder (*Raman*), not to the speaker. This 1st-person marking, furthermore, appears to be induced under conditions that are identical to those that induce attested cases of shifted indexicality, as in (3) – specifically, under conditions of embedding by a reportive/speech predicate. In Tamil, other speech-predicates besides *soll* in (1) can induce "monstrous agreement" on the embedded verbs of their clausal complements : *kattu* (shout/scream), *arivi* (inform), *pukaarsey* (advertise), *oppukkoḷ* (admit). At the same time, attitude-predicates like *nene*

(think/believe) don't seem to induce such shift: thus, this phenomenon appears restricted to speech-predicates alone. This is also considered characteristic of indexical shift (Anand 2006).

3. Confirming the suspicion: eliminating the usual suspects

Before embarking on an investigation of shifted indexicality in structures like (1), it is important to ascertain that what we're dealing with is actually an instance of indexical-shift. In this spirit, I discuss two of the usual suspects (see among others Schlenker To appear, for more discussion) – full-on quotatives and partial quotatives – candidates that may emulate the effects of monsters without actually being underlyingly monstrous.

3.1 Against a quotative analysis

Quotations may emulate shifted-indexicality effects because they "form a closed domain with respect to syntactic and semantic operators" (p. 81, Anand 2006). This grammatical opacity property of quotatives itself yields some useful diagnostics for testing the existence of quotatives. Thus, for instance, *wh*-extraction out of a quoted domain is expected to be ungrammatical. Similarly, an NPI inside a quotative may not be licensed by an operator outside the quote. I apply both the *wh*-extraction (4) and NPI (5) tests below:

(4) krishnan$_i$ yaar-ai$_1$ [$_{CP}$ taan$_{\{i,*j\}}$ t$_1$ paar-tt-een-nnu] so-nn-aan
krishnan who-ACC *taan* see-1SG-that say-PST-3MSG
"Who(m)$_1$ did Krishnan say [$_{CP}$ that he$_{\{i,*j\}}$ saw t$_1$]?"

(5) raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ oru tapp(u)-um senjeen-nnu] ottukka-le
raman *taan* one mistake-even made-that admit-NEG
"Raman$_i$ didn't admit that he$_{\{i,*j\}}$ made any mistake."

Both structures are fully grammatical, showing conclusively that the embedded CP in these structures is not a full clausal quotative. For the sake of completeness, I present the minimal pairs to (5) below – the result of replacing embedded subject *taan* with the "rigid" 1SG indexical *naan* (I). The structures below are both ungrammatical, indicating that the embedded CPs here are full-on quotatives. When such operations are not performed, the sentences are, of course, fully acceptable:

(6) * krishnan yaar-ai$_1$ [$_{CP}$ naan t$_1$ paar-tt-een-nnu] so-nn-aan
krishnan who-ACC I see-1SG-that say-PST-3MSG
"*Who(m)$_1$ did Krishnan say "I saw t$_1$?"

(7) * raman [$_{CP}$ naan oru tapp(u)-um se-nj-een-nnu] ottukka-le
raman I one mistake-even made-1SG-that admit-NEG
"*Raman didn't say "I made any mistake."

It is, nevertheless, still possible that only a sub-part of the embedded CP in (1) (specifically perhaps the part with the anomalous agreement) is actually quoted. Unlike full on quotatives of the kind discussed above, partial quotatives *can* influence the truth-conditional

semantics of the larger context they occur in. However, in theory, any string, including a nonsense string, could be quoted. Thus, if the embedded CP in (1) *did* involve a partial quotative, we wouldn't expect the 1SG embedded verb-agreement to be restricted to clauses embedded under reportive predicates. This conclusively shows that the embedded CP in (1) does not involve a partially quoted string either.

3.2 An important difference from (classical) indexical shifting

The discussion above indicates that structures like (1) involve bonafide indexical in the embedded clause. At the same time, (1) shows one important difference from standard structures involving indexical shift. In the "classic" indexical structure in Zazaki (1), the morphosyntactic agreement mismatch, between the shifted indexical and the attitude-holder DP that it denotes, is cross-clausal. In (1), on the other hand, it appears as if there is a person-feature mismatch within the embedded clause itself.[2] How? Because the embedded subject DP *taan* may only take 3rd-person antecedents and is, as such, traditionally treated as a 3SG anaphor. Its clausemate verb is, however, marked 1SG. To make matters even worse, and most problematically, this verbal suffix *-een* is an agreement marker, thus must have been triggered by *taan*! Independent of what ϕ-features, if any (Heinat 2006, Kratzer 2009), anaphoric *taan* may possess – how is it that the (superficially identical) instances of *taan* in (1) and (2) trigger differing agreement-markings (1SG vs. 3MSG, respectively) on their clausemate verbs? Is it simply the case that, despite appearances, there are two underlyingly different DPs both of which are spelled out syncretically as *taan*, or is there a more complex and potentially more interesting explanation?

In the following sections, I consider three predominant analyses of indexical shift in the literature – those in Anand (2006), von Stechow (2002), and Schlenker (2003) – and demonstrate that the Tamil agreement patterns presented above cannot be (non-stipulatively) dealt with under these analyses. I then proceed to outline a novel approach to indexical shift – one that will argue that it is the same element *taan* that surfaces in both (1) and (2) and which explains the peculiar agreement facts in both structures in a natural and elegant manner.

4. Analysis of indexical shift in Anand (2006)

Consider again the case from Zazaki (3), repeated as (8) below:

(8) hɛsen-i$_j$ (mɨ$_k$-ra) va kɛ ɛz$_{j/k}$ dɛwletia
Hesen-OBL I-OBL-TO said that I rich-be-PRES
"Hesen said that {I am, Hesen is} rich."

One way to understand the different readings obtained above might be to claim the following. In the case of the conventional, unshifted reading, the AUTHOR function of the indexical ranges over the utterance context whereas the shifted reading obtains when the

[2] Rahul Balusu (p.c.) informs me that the pattern in (1) also obtains in Telugu, another Dravidian language.

same AUTHOR function ranges over the context introduced by the speech predicate. The central intuition behind Anand (2006)'s approach is that the difference between the shifted and unshifted readings above is not the denotation of the indexical *ɛz* ('I'), which remains constant (at $\lambda c.Author(c)$) – but the value of the context that the AUTHOR function ranges over. In his account, shifted indexicality is thus taken to be the result of context shifting due to context-overwriting, argued to be implemented by parametrized operators introduced by the speech predicate and presumably located in C. Anand proposes that the operator responsible for context-overwriting in Zazaki is: $OP_{\forall}$; $[\![OP_{\forall}\alpha]\!]^{c,i} = [\![\alpha]\!]^{j,i}$ where j = $<Auth(i), Addr(i), Time(i), World(i)>$. The shifted reading for (3) is then derived as follows:[3]

[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]]$^{i',i',g}$ (READING B: SHIFTED)

4.1 Problems with extending Anand (2006)'s analysis to Tamil

The essence of Anand (2006)'s analysis of Malayalam *taan* is that it is underlyingly a 1SG indexical – thus, ultimately, a *shifted* indexical since it refers to the speech-report author and not the utterance-context speaker. It is tempting to apply his context-overwriting approach to the Tamil data in (1) and (2). But there are three main problems with such an analysis.

The first major problem for a context-overwriting approach is that Tamil as well as Malayalam *taan* only take 3rd-person antecedents (9 vs. 10):

(9) raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ jey-pp-een-nnu] sol-r-aan
raman *taan* win.FUT-1SG-that say-PRES-3MSG
Raman$_i$ says [$_{CP}$ that he$_{\{i,*j\}}$ will win.]

(10) * naan$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ jey-pp-een-nnu] sol-r-een
I *taan* win-FUT-1SG-that say-PRES-1SG
I say that [$_{CP}$ I will go.] (Intended)

This antecedence restriction is unexpected if Malayalam and Tamil *taan* are indeed 1SG indexicals, shifted or otherwise. Under Anand (2006)'s context-overwriting approach, the solution to this dilemma would be to either claim that the *taan* in (1) is underlyingly different from that in (2) or to propose that a speech-predicate like SAY in Tamil (and Malayalam) may introduce a context-shifting operator only if it is itself specified as 3SG. But the latter option, while workable, is, as Anand himself concedes, very stipulative. The second major problem under a context-overwriting approach is the "regular" embedded verbal agreement in structures like (2) (a problem that, interestingly, Anand (2006) doesn't address for Malayalam, again because it lacks verbal agreement-marking) In Anand's system, we are thus essentially forced to claim that *taan* in structures like (2) is underlyingly different from

[3] **Unshifted Reading:** [Hesen said [$_{CP}$ I$_{sandhya}$ am rich.]]
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).

that in (1). E.g. that it is not an indexical, but a logophor/SE-anaphor, which inherits the ϕ-features of its antecedent and transmits these via Agree to the embedded verb (Kratzer 2009)). A potentially serious empirical challenge to this type of analysis is that structures like 2 also seem to be interpreted as *obligatorily de se* – there is no clear independent reason to assume that the syntactico-semantics of *taan* in (2) is different from that in (1). Finally, a third problem is that, on top of the obligatorily shifting *taan*, Tamil has a purely "English-like" 1SG indexical *naan* (see also (7) and (6) above):

(11) raman [$_{CP}$ naan jey-pp-een-nnu] so-nn-aan
raman I win-FUT-1SG-that 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, it is unclear how the morphophonological distinction between *naan* and *taan* is to be captured. Of course, one could postulate different spell-out rules for the unshifted indexical *naan* and the shifted indexical *taan*, as follows:[4]

UNSHIFTED: [AUTHOR (c)]$^{c,i,g} \rightarrow$ SPELL-OUT: *naan*
SHIFTED: [OP_{auth} AUTHOR (c)]^{c,i,g} = [AUTHOR (i)]$^{i,i,g} \rightarrow$ SPELL-OUT: *taan*

While possible, this requires the context-shifting operator to already be present in Narrow Syntax (in order to affect both the semantics at LF and Spell-Out rules at PF). The syntacticization of context-shifting/overwriting, however, still needs to be precisely formalized in Anand's system which deals predominantly with the semantic aspects of indexical shift. This is thus perhaps more a problem of omission, in contrast to the first two problems outlined above which, I believe, pose a more serious challenge to Anand's analysis.

4.2 Variable binding and LF feature-deletion: von Stechow (2002)

Descriptively, shifted indexical pronouns show an inherent split along the LF/PF domains: semantically, they denote 3rd person individuals (the attitude-holder); morpho(syntactically), however, they are marked 1st person. von Stechow (2002) proposes that shifted indexicals are variable bound by attitude verbs (which, he proposes, quantify over contextual tuples: ⟨Author, (Addressee), Time, World, Location⟩). Furthermore, he proposes, the ϕ-features of a DP that is variable bound are deleted at LF. Such a model neatly accounts for the person-feature split noted in shifted indexical DPs, as follows. The indexical pronoun is born with a [+*Author*] feature – this is the feature that it bears in the Narrow Syntax. At PF, this 1st-person feature is spelled out. At LF, the indexical DP is variable bound by the quantificational attitude verb that c-commands it – this ensures that the indexical denotes the author of the speech-report. Crucially, by assumption, when the embedded indexical pronoun is thus variable bound, its [+*Author*] feature is deleted (thus preventing a presupposition failure/clash). This is formalized below:

[4] I extend Anand (2006)'s proposal, that the context-shifting operator in Malayalam is OP_{auth} (an operator which shifts only the Author coordinate of the context-tuple), to Tamil which behaves very similarly.

SENTENCE: Raman$_i$ says that he$_i$ will win. (*Obligatory De se* reading: Raman says, "I will win".)
[Raman 3rd] λi [say 3rd] $\lambda \langle x_j$3rd, z_k, w_i, $t_i \rangle$ [x_j + ~~author~~] will win.

What happens when we apply this analysis to the Tamil data in (1)/(2) above? First of all, the 3rd-person antecedence restriction of the embedded DP *taan* in both (1) and (2) would have to be reframed as a restriction on the quantificational attitude-verbs binding them. Specifically, we would have to stipulate that only quantificational attitude-verbs specified 3rd-person may delete and thus variable-bind [+author] elements. More problematically, under this model, there is no way to distinguish between non-shifting *naan* (11) and shifting *taan* ((1)/(2)), assuming that both are [+author] elements in the syntax, again because variable binding is held to occur only at LF. Finally, assuming that *taan* is the same underlying element, specifically an [+author] indexical pronoun, in both (1) and (2), it is unclear how to derive the distinction between agreement patterns (1st vs. 3rd) on their corresponding clausemate verbs. Thus, we end up with essentially the same problems that we encountered with Anand (2006)'s context-overwriting analysis.

5. Pronoun-centric model: Schlenker (2003)

Schlenker (2003) (and subsequent) attempts to derive indexical shift as a function of the syntactico-semantics of the indexicals themselves. Simply put, some indexical pronouns, like those in English, are *inherently* specified as being rigid/unshifting. Thus, English $[\![I]\!]^{c,g}$ = $[\![x_i + author * (x_i)]\!]^{c,g}$ = s(x_i) $\wedge$ δ(g(x_i) is Author(c*)) (for c* = $c_{utterance}$). Other indexicals, like those in Amharic and Zazaki, are underspecified as to whether they are always shifting or rigid – in other words, they can optionally shift or not shift. Thus, Amharic $[\![I]\!]^{c,g}$ = $[\![x_i + author(x_i, c_j)]\!]^{c,g}$ = s(x_i) $\wedge$ δ(g(x_i) is Author(c_j)). Still others, like the classic logophoric pronouns in e.g. Gokana and other languages (and presumably also controlled PRO (Landau 2004)), are semantically specified to always shift. Thus, logophoric $[\![I]\!]^{c,g}$ = $[\![x_i + author(x_i, c_j)]\!]^{c,g}$ = s(x_i) $\wedge$ δ($\forall$ $c_j \neq$ c* $\rightarrow$ g(x_i) 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* (11) and shifting *taan* in Tamil (1)/(2). Specifically, *naan* would have the denotation of an English-like "I" whereas *taan* would have the denotation of a logophoric "I" – as formalized above.[5] I take this to be an advantage of Schlenker's system over both Anand's and von Stechow's. Nevertheless, the other two problems mentioned above remain for Schlenker's analysis as well. That is, it is unclear how to derive the distinction between monstrous agreement in (1) and "regular" agreement in (2) given that both are presumably triggered by the same underlying indexical pro-form *taan*. The 3rd person antecedence restriction of *taan* in (1)/(2) is also problematic.

To sum up, although the details of the three analyses portrayed above described above vary, they all face the same empirical issues with regards to the Tamil data. Why is this? The proposals of Anand (2006), von Stechow (2002) and Schlenker (2003) are

[5]Of course, this still needs to be *syntactically* formalized to be able to effect Spell-Out rules at PF in a Y-modular architecture.

all purely *semantic* in nature. As such, they fall short of of dealing with a primarily morphosyntactic phenomenon like subject-verb agree(ment) and additionally make some wrong predictions.

6. An alternative analysis: access to both contexts

The proposal outlined here motivates indexical shift within the Narrow Syntax itself. Specifically, I detail (the beginnings of) an analysis which implements indexical shift as a form of Agree which, in turn, yields the correct input conditions for further operations at LF and PF.

6.1 Preliminary assumptions and claims

Integral to the analysis developed here is the idea that (at least some) contextual information is represented in the syntax. This enables a syntactic implementation of indexical shift without compromising the central role of the context. The syntacticization of (some) contextual information is not an entirely new idea: it has been proposed in various forms by Bianchi (2003), Sigurðsson (2004), among others. Evidence for the claim that (at least some) contextual information is present early in the derivation, directly in the syntactic structure, comes from the observation that contextual information appears to be able to affect LF (e.g. obligatory *de se* readings) as well as PF (e.g. special complementizer forms to introduce logophors in many languages like Tuburi (Sells 1987) and the well-known morphophonological properties of controlled PRO) as well as syntax-internal intervention/ "blocking" effects due to indexicals, such as those found in long-distance anaphora in Chinese, Italian, Icelandic and many other languages (Giorgi 2006).

The analysis that I develop here to deal with the Tamil patterns in (1)/(2), rests on the following specific assumptions. First, I assume that a context is a speech situation (Schlenker To appear), and is enriched with information pertaining to the following parameters: $\langle Author, Addressee, Time, World, Location \rangle$ (Anand 2006). Featurally, it is represented as a tuple containing the parameters above as coordinates. Second, I assume that information with regards to these coordinates is *featurally* represented in the syntax – specifically in the cartographic C-layer. Such an idea is not novel and has been proposed in various forms by many researchers: Cinque (1999) proposes syntactic projections to encode certain pragmatic information, such as an Evaluative Mood, Evidential Mood, and Speech Act Mood in the left periphery, Sigurðsson (2004) and Bianchi (2003) explore the idea of the syntacticization of certain pragmatic features in C, Speas (2004) adapts the functional hierarchy proposed in Cinque (1999) to argue for a syntactic instantiation of contextual features to capture the representation and distribution of logophors and evidentials, and more recently, Giorgi (2010) motivates and develops a syntactic account of contextual features in the cartographic C and T layers. I notate the features pertaining to the utterance context as Context_{utt} and to those of the embedding context as Context_{emb}. Finally, and crucially, I also assume that both Context_{utt} and Context_{emb} can be simultaneously present in the derivation (an idea used to derive Double Access Readings (DAR) for tense (Giorgi, 2010)/mood (Schlenker, 2004). In the absence of specific co-occurrence restrictions for

these features, this is, in fact, the null hypothesis. In such cases of dual contextual representation, I follow Giorgi (2010) in assuming that Context_{emb} is located below Context_{utt} in the embedded C-layer. I also propose that Context_{emb} is "updated" with information about Context_{utt} (perhaps it is marked as [same] vs. [different]. Presumably, this is the result of an Agree operation between Context_{emb} and Context_{utt}. Given that an embedded CP cannot exist "by itself", precisely by virtue of its being embedded, and must be syntactically anchored to a higher clause, it makes sense to view Context_{emb} (the context containing information about the embedded CP) as the Probe in this Agree relationship. The details of this need to be further formalized, of course – but interestingly, there is a precedent to this claim as well: Iatridou (2000) proposes a similar categorization for modal anchoring, one that Ramchand (2011) extends and adapts.

6.2 A descriptive observation about Tamil (1/2 and the formalization thereof

The starting *empirical* point for this analysis is the following descriptive observation with regards to the structures in both (1) and (2). The embedded CP in (1)/(2) is interpreted as an eventuality that is from the perspective of an individual who is simultaneously the Author of Context_{emb}, but Other(/3rd) from that of Context_{utt}. I.e. CP_{emb} in both structures embodies the *first-personal account of a 3rd-personal narrator*. The logophoric element *taan* that is the matrix subject of the TP complement of CP_{emb} in (1)/(2) morphosyntactically instantiates the dual nature of this individual. I formally model this property as follows. The $\text{DP}_{embedded}$ in (1)/(2), is featurally specified in the syntax as both Author and Other(/"3rd") – each evaluated against a different context.

I further claim that, in (2), Context_{utt} and Context_{emb} are both present in the embedded C-layer, whereas in (1), only the Context_{emb} is: this will be shown to be instrumental in driving the distinct agreement patterns on the respective embedded verbs in (1) vs. (2). But first I show that there is some promising independent evidence that the embedded CP in examples like (1) is structurally smaller than those in (2) – this could be made to automatically follow from Claim (ii) above.[6] Clauses like (1) seem to also be capable of being represented as gerundivals:

(12) raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ poo-v-adaaga] so-nn-aan
raman *taan* go-FUT-GER say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would go]"

In contrast, structures like (2) seem to resist this:

(13) */?? raman$_i$ taan$_{\{i,*j\}}$ poo-v-adaaga paar-tt-aan
raman$_i$ *taan*$_i$ go-FUT-GER see-PST-3MSG

[6]Notice, incidentally, that such a claim appears to be at odds with prior observations (see among others Speas 2004, Cristafaro 2003) that the complements of speech predicates are structurally larger than those of factive predicates like *know* and even larger than complements of direct-perception predicates like *see*. Nevertheless, the gerundival data presented for Tamil in (12)/(13) is quite robust and potentially also mirrored in parallel structures in English ("John spoke of going" vs. "*John discovered/saw of going") – suggesting that it is not a quirk of Tamil. It is unclear at this juncture how this tension is to be resolved – the investigation of this is part of my ongoing dissertation research.

"Raman$_i$ saw [$_{CP}$ that he$_{\{i,*j\}}$ would go] " (Intended)

6.3 Deriving the patterns in (1) and(2) in the syntax

I propose that the embedded DP which is ultimately spelled out as *taan* in (1) and (2 has the following feature-specification in the syntax:

Valued privative features: [Singular, Masculine].
Unvalued features: [Author: __], [Other: __].

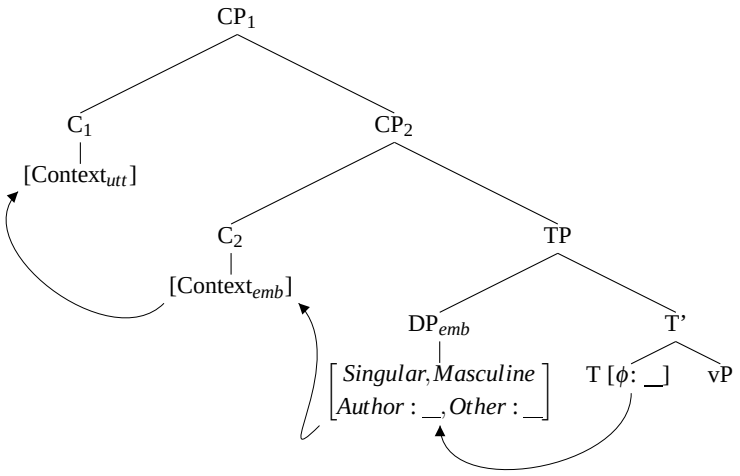
The [*Author*] and [*Other*] attributes take contextual features as values. This is the syntactic instantiation and extension of the idea that, at LF, $[\![Author]\!]^{c,i,g} = \lambda$c. Author(c) (Anand 2006). Since [*Author*] and [*Addressee*] are treated as unvalued feature-attributes on the embedded DP in (1) and (2), the embedded DP is a Probe in the syntax to get these attributes valued by contextual features. Below, I present detailed derivations for the paradigms in (1) and (2).

Step-by-Step Derivation for (2)

As per my assumption, both Context$_{utt}$ and Context$_{emb}$ are featurally present in the embedded C-layer – the latter being structurally lower than the former. This is thus a structure that yields a "double access" reading. Context$_{emb}$ is "updated" with the information that its Author/Addressee coordinates are different from those of Context$_{utt}$ in the manner briefly described above. In the next step, DP$_{embedded}$ upward-probes and reaches the C node containing the feature Context$_{emb}$ first; its unvalued feature, [Author: __] gets valued as: [Author: Context$_{emb}$]. Finally, since Context$_{emb}$ is enriched with information about Context$_{utt}$, its unvalued feature, [Other: __], is valued as [Other: Context$_{utt}$]. DP$_{embedded}$ values clausemate T as: [Author: Context$_{emb}$, Other: Context$_{utt}$, Singular, Masculine]. Under a feature-sharing approach such as that in Pesetsky and Torrego (2007), the unvalued ϕ-features on T and the unvalued Author and Other features on DP$_{embedded}$ in [Spec, TP] could form a link and both be simultaneously valued by Context$_{emb}$. This derivation is diagrammatically represented in the simplified structure in (14) below:[7]

[7]Arrows indicate Agree, not movement. As the arrows show, I have indicated upward probing throughout. The upward probing relationship between DP$_{emb}$ and Context$_{emb}$ on the one hand, and that between Context$_{emb}$ and Context$_{utt}$, on the other, instantiates the core intuition that DP$_{emb}$ and Context$_{emb}$ are the respective Probes in these two Agree relationships. However, nothing crucial hinges on the upward probing arrow between embedded T and DP$_{emb}$: if we assume that DP$_{emb}$ is still in [Spec, vP], the agreement would be downward from T to DP$_{emb}$. I have represented the feature-specification on T merely as [ϕ: __], as it is not yet clear if T is specified with the same feature attributes as DP$_{emb}$. Finally, I have ignored Case features for now, for reasons of simplicity.

(14)

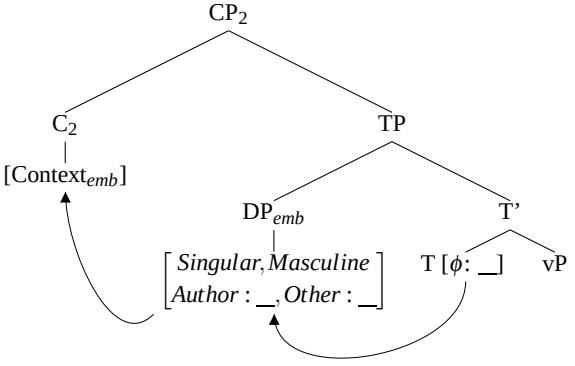


Step-by-Step Derivation for (1)

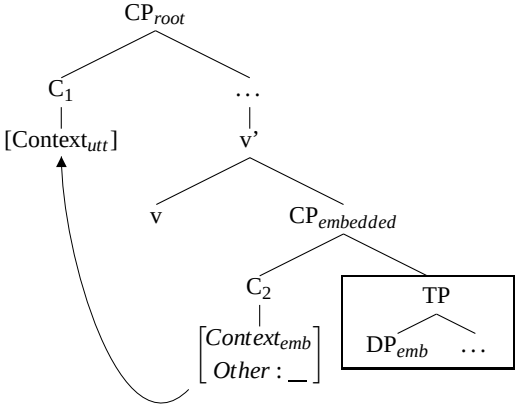
In the Tamil structure in (1), only Context_{emb} is featurally present in embedded C, as per my claim. As such, Context_{emb} is not "updated" with featural-information for Context_{utt} in the lower phase. $\text{DP}_{embedded}$ upward-probes and reaches the C node containing the [Context_{emb}] feature; its [Author] feature gets valued as: [Author: Context_{emb}]. [Other: __] on $\text{DP}_{embedded}$ remains unvalued for now – it presumably gets passed up to Context_{emb} in the left-edge. At the higher phase, it gets valued as [Other: Context_{utt}], by Context_{utt} in root C. $\text{DP}_{embedded}$ values embedded T as: [Author: Context_{emb}, Singular, Masculine]. The [Other] feature doesn't get copied to T because it is unvalued. In (13) below, I present the tree structures for the two stages of the derivation in (1). In Stage 2 of the derivation (15b), the embedded TP is boxed in to illustrate that it is invisible to the matrix Phase, due to the PIC:

(15)

a. EMBEDDED PHASE: STAGE 1



b. MATRIX PHASE: STAGE 2



There are of course many technical details that need to be further developed and formalized: what triggers percolation of the unvalued [Other: __] feature from DP_{emb} to Context_{emb}, thereby preventing a crash? How does the contextual comparison between Context_{emb} and Context_{utt} in the syntax work and lead to the valuation of the [Other] attribute on DP_{emb}? The investigation of these issues is part of the ongoing research.

Table 1: SPELL-OUT RULES IN A LATE INSERTION MODEL

AGR	[Author, sg]	⟷	*-een*
	[Other, masc, sg]	⟷	*-aan*
PRONOUN	[Author: Context$_{utt}$, sg]	⟷	*naan*
	[Author: Context$_{emb}$, sg]	⟷	*taan*
	[Other: Context$_{utt}$, masc, sg]	⟷	*avan*

6.4 Exploiting Late Insertion spell-out rules

Here I outline a series of Late Insertion (Embick and Noyer 2007) Spell-Out Rules which take the completed derivations in the syntactic module, detailed above, as input values and yield the appropriate morphophonological paradigms in (1) and (2) at PF. Table (1) shows the following. *taan* syncretically spells out [Author: Context$_{emb}$, Other: Context$_{utt}$] in (2) and [Author: Context$_{emb}$] in (1) ([Other: __] is presumably not "seen"). Distinct AGR underspecification rules yield the 1SG *-een* on the embedded verb in (1) and 3MSG *-aan* on the embedded verb in (2).

6.5 The derivation of *naan* and 3rd-person antecedence of *taan*

It was noticed earlier, that in addition to *taan*, Tamil has a rigid, unshifting indexical *naan* which may only refer to the author of the utterance context. One of the three main problems with prior analyses of indexical shift, particularly those of Anand (2006) and von Stechow (2002) (and to a lesser extent, that in Schlenker (2003)) was their inability to properly derive the distinction between *naan* and *taan* in Tamil. Here, I briefly describe the mechanism by which this distinction is motivated in the syntax, within the current analysis. According to Table (1) above, *naan* spells out a DP syntactically specified [Author: Context$_{utt}$, Singular]. However, I assume that the DP that ultimately spells out *naan* is born with its [Author] attribute *already valued*. Within the Narrow Syntax, this has the consequence that the DP doesn't probe for this feature and, instead, refers to Context$_{utt}$ *deictically* – this is the syntactic instantiation of a Kaplanian "rigid" indexical. In languages, like Amharic and Zazaki (3), where the indexical DP is shifting and non-rigid, the 1st-person indexical would not be born with the [Author] attribute valued, as with *naan* above.

There is one final problem that needs to be addressed which was also mentioned as an issue that could not be non-stipulatively dealt with by the prior indexical shift analyses. This is the fact that *taan* may only take 3rd-person antecedents. I repeat the relevant facts below in (16) and (17):

(16) raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ poo-r-een-nnu] sol-r-aan
raman *taan* go-PRES-1SG-that say-PRES-3MSG
Raman$_i$ says [$_{CP}$ that he$_{\{i,*j\}}$ will go.]

(17) * naan$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ poo-r-een-nnu] sol-r-een
I *taan* go-PRES-1SG-that say-PRES-1SG

I say that [$_{CP}$ I will go.] (Intended)

Here, I briefly demonstrate how the current analysis captures this restriction. Assume, as a first case scenario, that the matrix DP is specified with an unvalued [*Author*] attribute: [*Author* : _]. At the root CP, only the utterance context is present by default (the exception is cases of free indirect discourse). Thus, this will by default get valued as [Author: Context$_{utt}$]. This directly contradicts [Other: Context$_{utt}$] on DP$_{emb}$. However, if, as I have assumed for the unshifting indexical *naan*, the matrix DP is born with a prevalued [*Author* : *Context*$_{utt}$] feature, the same result as in (i) will obtain. Both DPs will get spelled out as *naan*. The same logic can be extended to [*Addressee*] features. What if the matrix DP has the same feature-specification as DP$_{emb}$? I.e. if DP$_{matrix}$ is also specified as: [*Author* : _] and [*Other* : _]? Again, since there is only one contextual feature, Context$_{utt}$ at the root (by default), only one of these attributes will be capable of being valued. As such, the presence of the other unvalued feature will cause a crash. This is why, for the most part, *taan* cannot occur as the matrix subject. Interestingly, the exception to this is the case of free indirect discourse – *taan* is capable of being a matrix subject in such cases. Finally, for the sake of completeness, it makes sense to explore the following scenario and ask whether a DP can be *born* with the feature: [*Author* : *Context*$_{emb}$]. Technically, this pairing should be possible. However, while there might be no problems in Narrow Syntax with such a featurally-specified DP, this would nevertheless be ill-formed because there will be no embedding context to evaluate [*Author* : *Context*$_{emb}$] against since, with the exception of free indirect discourse, only the utterance context is present at the root CP.

7. Conclusion

The analysis presented above deals with the agreement patterns in the embedded CP in structures like (1) and (2) by means of tweaking the following independent parameters: the ϕ-features that the embedded DP is born with and those that it ends up with due to Agree; the types (utterance vs. embedding) and number (both Context$_{utt}$ and Context$_{emb}$, or just the latter) of contexts that are syntactically represented, thus locally accessible, in the embedded C-layer; distinct patterns of underspecification on the exponents for pronouns and verbal agreement in the Spell-Out component. A crucial component of this analysis is that it motivates indexical shift within the Narrow Syntax itself. As such, it is able to account for LF as well as PF properties of indexical shift in a Y-modular architecture.

References

Anand, Pranav. 2006. De de se. Doctoral Dissertation, MIT.

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

Annamalai, E. 1999. Lexical anaphors and pronouns in Tamil. In *Lexical and anaphors and pronouns in selected South Asian languages: a principled typology*, ed. Lust, Wali, Gair, and Subbarao, 169–216. Mouton de Gruyter.

Bianchi, Valentina. 2003. On finiteness as logophoric anchoring. In *Tense and Point of View*, ed. Guéron and Tasmovski, 213–246. Nanterre: Paris.

Cinque, Guglielmo. 1999. *Adverbs and functional heads*. OUP.
Cristafaro, Sonia. 2003. *Subordination*. Oxford: OUP.
Embick, David, and Rolf Noyer. 2007. Distributed Morphology and the syntax/morphology interface. In *The Oxford handbook of linguistic interfaces*, ed. Ramchand and Reiss. OUP.
Giorgi, Alessandra. 2006. From temporal anchoring to long distance anaphors. *NLLT* .
Giorgi, Alessandra. 2010. *Towards a syntax of indexicality*. Oxford: Oxford University Press.
Heinat, Fredrik. 2006. Probing phrases, pronouns, and binding. *The Department of Eglish: Working Papers in English linguistics* VI. Lund University.
Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31:231–2.
Kratzer, Angelika. 2009. Making a pronoun: fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40:187–237.
Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22:811–777.
Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture: syntactic derivation an dinterpretation*, ed. Karimi, Samiian, and Wilkins, 262–294. Amsterdam: John Benjamins.
Ramchand, Gillian. 2011. Indexical vs. anaphoric modals. Talk given at the Tense, Aspect, and Modality workshop, CASTL, Tromsø.
Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.
Rizzi, Luigi. 1990. On the anaphor-agreement effect. *Rivista di Linguistica* 2:27–42.
Schlenker, Philippe. 1999. Propositional attitudes and indexicality: a cross-categorial approach. Doctoral Dissertation, MIT.
Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and Philosophy* 26:29–120. Kluwer Academic Publishers.
Schlenker, Philippe. To appear. Indexicality and *de se* reports. In *Handbook of Semantics*, ed. von Heusinger, Maienborn, and Portner. Mouton de Gruyter.
Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18:445–479.
Sigurðsson, Halldór Ármann. 2004. The syntax of person, tense, and speech features. *Italian Journal of Linguistics* 16:219–251. Ed. Bianchi and Safir.
Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114:255–276.
von Stechow, Arnim. 2002. Feature deletion under semantic binding: tense, person, and mood under verbal quantifiers. In *Proceedings of NELS 33*.
Woolford, Ellen. 1999. More on the anaphor agreement effect. *Linguistic Inquiry* 30:257–287.

CASTL, Universitet i Tromsø
9037 Tromsø
Norway

sandhya.sundaresan@ uit.no