SOUTH ASIAN LANGUAGE REVIEW
VOL. XV. No.2. June 2005.

Dative Subjects in Tamil: A Computational Analysis

Umarani Pappuswamy
University of Pittsburgh

Abstract. This paper describes the dative constructions in Tamil and discusses their theoretical and computational implications within the Lexical Functional Grammar (LFG) framework. Not all dative nominals can be subjects. The main aim of the paper is to distinguish the dative nominals that behave like surface subjects from those that do not. My approach is to recognize the case marking through simple morphological rules and assign grammatical functions based on the lexical semantic and syntactic properties of the dative nominals with respect to various subject-hood tests. I also describe the role of unification and morphological blocking in assigning grammatical functions to the dative nominals.

1. Introduction

One of the most interesting issues of grammatical functions, perhaps the problematic, is SUBJECT. Several grammatical theories have addressed this issue in a number of ways. Two approaches have been taken so far to the nature of 'subject'. For Government and Binding (GB) and its antecedents, and for Head-driven Phrase Structure Grammar (HPSG), subject can be defined in terms of syntactic properties. For Lexical Functional Grammar (LFG), subject (SUBJ) is an argument function, specifically the most prominent argument function on the relational hierarchy. Traditionally in LFG, grammatical functions have been identified by functional descriptions stating relations such as (↑SUBJ) = ↓ and (↑CASE) =NOM. Nordlinger (1998) proposes constructive functions for case markers, such as (SUBJ ↑), in order to account for various properties of case morphology in Wambaya (an Australian language, non-configurational in nature). Nordlinger (1997) pointed out that 'Case markers, in addition to carrying a regular case feature may carry a designator specifying information about the GF to which they belong to'.

ni: (↑CASE) = ERG, (SUBJ ↑)

The above is a case specification for Wambaya. A nominal with an ergative case marker as above will serve as the SUBJ. Thus, syntactic relations are constructed directly from the case morphology.

My approach is slightly different from Nordlinger's in that I use morphological rules to recognize the case markers of all the NPs present in the sentence and then assign Grammatical Functions (GF) based on some lexical semantic and syntactic conditions. For example, the specification for a Dative case marker in Tamil is as follows:

ukku: (↑CASE) = DAT, (SUBJ ↑) V (OBJ2 ↑

It contains two designators indicating that the Dative NP could either be a subject or an object. The Tamil morphology does not allow us to construct the grammatical function directly as in the case of Wambaya. Hence, it is necessary to examine this phenomenon in depth in order to distinguish the dative nominals that behave like surface subjects from those that do not.

This paper is organized as follows: section 2 provides a brief overview of case marking and grammatical function assignment in Tamil; section 3 describes the types of Tamil dative constructions and provides evidence that dative nominals behave like nominative subjects with respect to various syntactic phenomenon related to subject-hood tests.; section 4 describes the role of morphological blocking and unification in recognizing the dative nominals and section 5 presents the conclusions.

2. Case Marking in Tamil[1]

2.1 Case in LFG

Unlike Chomskyan frameworks, "case" within the LFG framework is not invoked to account for the distribution of lexically filled NPs. The distribution of arguments that may be subject to grammatical or anaphoric control is handled by the theory of control, in terms of grammatical functions rather than syntactic positions (Neidle 1994). Words can carry the same kinds of functional information as syntactic phrases; words and phrases are alternative means of encoding the same syntactic relations (Bresnan 1996). That 'morphology competes with syntax' (Bresnan 1996) is true of many nonconfigurational languages. I will provide evidence from Tamil to support this claim.

Tamil is an agglutinating language in which morphs are 'stuck on' in a sequence as suffixes (Lehmann, 1989:8). Words are made up of lexical roots, or stems followed by inflectional suffixes) which mark categories such as case, number, person, number, gender, tense, etc. As the focus of this paper is on the subjects, let us look into the Tamil Case system. There are seven cases: (Arden 1942) the nominative (first case), accusative (second case), instrumental (third), dative (fourth), ablative (fifth), genitive (sixth), and locative (seventh). All the cases are marked except the nominative. The accusative is not marked for inanimate nouns. These markers (or suffixes) are added to the oblique stem of the respective noun. On the other hand, the plural suffix is added only to the

noun stem and not to an oblique stem. The case forms for the noun *maram* 'tree' are given below:

Table 1 Case Forms of *maram*

Case	***Marker***	**maram *'tree'***
Nominative	*Ø*	maram
Accusative	ai *or Ø*	marattai
Instrumental	aal	marattaal
Dative	ukku	marattukku
Ablative	iliruntu	maratiliruntu
Genitive	uTaiya, in	marattin
Locative	il	marattil

I adopt Lehmann (1989: 14-17)'s classification of nouns[2] for our analysis to build the Morphological Analyzer (MORPHA) of Tamil. Simple 'sandhi' rules related to the noun stems are incorporated into the LFG package developed by Andrews (1990)[3].

2.2 Assigning Grammatical Functions

Grammatical Functions (GF) in LFG denote surface grammatical relationships and not the underlying, logical relationships commonly represented in transformational deep structures (Kaplan and Bresnan, 1982). GFs are normally assigned to predicate arguments. It should be noted that the same GF cannot be assigned to different predicate-arguments and that different GFs cannot be assigned to the same predicate argument. This is called as the Principle of Function Arugments assignments (Kaplan and Bresnan, 1982:163). Assignment of GFs to predicates varies across languages. It may be encoded by the position of a constituent structure (as in English) or by means of morphological marking (as in Tamil). This is what Bresnan (1982) calls 'configurational' and nonconfigurational' encoding of GFs across languages.

I follow Umarani (1996) to assign GFs to Tamil NPs. Fig.1 shows that subject is marked either by the nominative or the dative case. In this paper, we will focus only on the dative subjects. The morphological component related to nouns includes the morphological specifications for CASE and NUMBER.

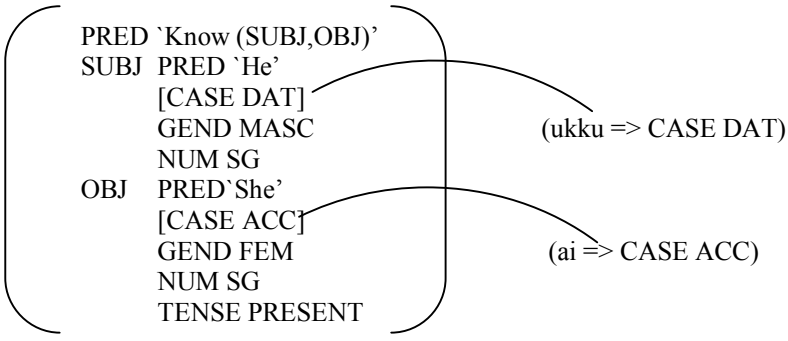
For a sentence like (1), the parser built in Umarani (1996) does a lexical check-up and finds *avan* 'He' and as it is inflected the noun morphological parser is called upon. The parser also recognizes *-ukku* as the DATive case marker.

(1) *avan-ukku* *avaL-a.it* *teriyum*
He-dat she-acc know-fu+3sn
'He knows her'

FORM	**GF**
PRED [NP] X	SUBJ
PRED [NP NP] X	*α = SUBJ*
α β	*if CASE = NOM/DAT*
	else α = OBJ
	β = SUBJ
	if CASE = NOM/DAT
	else β =OBJ
PRED [NP NP NP] X	
α β γ	
if NP's CASE = NOM	GF = SUBJ
if NPs CASE = DAT	GF = DATOBJ/0BJ2
the other NPs	GF = OBJ
(NP = α/β/γ & X = Adjunct in all the 3 cases)	

Fig. 1 Grammatical Function Assignment in Tamil

Similarly the other elements are also recognized in the lexicon and by the morphological parser. This information is sent to the GF module where the NPs get their corresponding GFs in the given context. Out of the two NPs in (1), the NP containing the DAT marker is realized as the SUBJ. The f-structure for (1) would be as follows:



The attribute-value pairs indicating the CASE are the results of unification of the information found in the morphological component. All the other attribute-value pairs come from the lexicon. The rules of GF assignment are

applied at this stage and then a particular NP gets its GF value. Thus, lexical, structural and morphological information are integrated and unified within the f-structure. The validity of an f-structure is ensured by a number of well-formedness conditions (Principles of uniqueness, completeness and coherency as stated in Kaplan and Bresnan 1982).

It is not as simple as the above when it comes to complex constructions that contain dative nominals. Not all dative nominals exhibit the properties of subjecthood normally associated with grammatical subjects. Those that fail to behave like subjects (at least in some respects) could be I-nominals[4] (in the sense of Moore and Perlmutter 2000).

The next section discusses the types of Dative constructions and their various types of syntactic environments.

3. Types of Dative constructions

There are two types of Dative constructions in Tamil:

1. Dative -Accusative pattern
2. Dative-Nominative pattern

3.1 Dative-Accusative pattern

This type of constructions allow a small set of verbs like verbs of knowing, verbs of liking, verbs denoting physical and mental states, modal verbs etc. For example[5]:

teri: V, Subclass KNOW, PRED 'Know (SUBJ,OBJ)'.
piTi: V, Subclas PSYCH, PRED 'Like (SUBJ,OBJ)'.

All these verbs are morphologically defective i.e. they are not infected for person, number and gender. They all carry 3-person singular neuter suffix as their default value. They do not trigger agreement with the dative NP or with the accusative NP. Some examples are given below:

(2) *avan-ukku avaL-ai.k kaNTaal-ee ericcal*
He-dat she-acc see-cond-emp irritation
'He gets irritated by the very sight of her'

The dative nominal behaves like a surface subject in (2) and it controls the reflexive in (3).

(3) *raamu.v-ukku* $_i$ *avaL-ai.p/tan.n-ai*$_i$ *piTi-kk-um*
Ram-dat she-acc/self-acc like-fu+3sn
'Ram likes her / himself

It is an experiencer in (4).

(4) *appa.v-ukku un.n-ai.ppatti kavalai.n-aa enakku puvanaa- patti*
father-dat you-acc- about worry-cli I-dat *Bhuvana-about*
kavalai puvana.v-ooTa freNTu cari.y-illee-nu enakku.t tooNu-tu
worry Bhuvana-gen friend correct-not-comp I-dat appear-pst
'Father is worried about you and I am worried about Bhuvana.
It appears to me that Bhuvana's friend is not good'.

It allows Equi-NP deletion as in (5).

(5) *enakku unnaip patti-taan kavalai vasant enakku*
I-dat you-acc about-emp worry Vasanth I-dat
maTTum-illai unkappaa.v-ukku-m kavalai
only-not yourfather-dat-also worry
'I am worried only about you. Vasanth, not only me, your father is also worried (about you).

Verbs of knowledge like *teri* show a different type of behaviour with dative nominals. In a complex construction such as (6), the antecedent of the reflexive is the dative nominal of the higher clause. The reflexive dative co-refers to the subject argument of the *teri* verb. But the reflexive object *tannai* in (8) does not co-refer to the subject argument of 'teri'. Here, the lexical pronoun *avanai* co -refers to r*agu.* This is true of some nominative-accusative constructions too[6].

(6) *tanakku$_i$ avan-ai paa-kka.p piTikkum-nu umaa.v-ukku$_i$*
self-dat he-acc see-inf like-be-pst-3sn-that uma-dat
naNR-aaka teri.y-um
well-adv knows
'Uma knows very well that self would like to see him'.

(7) *unakk-enna teriyu-maa kaar-aip.patti?*
you-dat-what know-mother car-acc about
'Mother, what do you know about car?'

(8) *enakku, avanai/*tannai$_i$ piTikkum enRu ragu.v-ukku$_i$ teriyum*
I-dat he-acc/*self-acc like-fu+3sn that Raghu-dat know-fu+3sn
'Raghu knows that I like him/*myself.

It controls PRO as in (9).

(9) ([PRO] *pooToo.v-ai* *paarttatum* *ravi-kku* *azukai* *vantatu*
picture-acc see-CP Ravi-dat cry came
'Having seen the photograph Ravi cried'
(Lit: To Ravi cry came having seen the photograph).

It can be observed that the Dative nominals in the DAT-ACC constructions given above (except (6) and (8)) satisfy most of the properties of the notion 'subject'. So these could be classified as SUBJ by the parser. The ones in (6) and (8) show the differences mainly due to the presence of personal pronouns in the matrix clause. Example (6) contains reflexive dative where as (8) contains a first person dative.

3.2 Dative-Nominative pattern

This type of dative constructions 'involve verbs of existence/ possession, verbs expressing need/obligation and a few other verbs like *teri* 'know', *keel* 'hear', *peecu* 'speak' (when occurring with nouns *kaN* 'eyes', *kaatu* 'ear' and *vaay* 'mouth') and the verb *vaa* 'come' when occurring with nouns denoting 'mental or emotional experience' (Lehmann 1989). I further extend his analysis and propose that the dative nominal can also occur with some of the psych verbs such as *puri* 'understand' and a small subset of verbs related to physical and biological processes such as *paci* 'be_hungry', *vali* 'pain' etc. I also look into the various syntactic contexts in which dative nominal can occur. Not all dative nominals exhibit the properties of subject-hood normally associated with grammatical subjects. Those that fail to behave like subjects (at least in some respects) could be I-nominals (in the sense of Moore and Perlmutter 2000).

With verbs of existence/possession

(10) *amma.v-ukku* *ippa* *enta* *kavalai.y-um illai*
mother-dat now which worry-emp not
'Mother has no worries at all now'.

(11) *enakku* *enga* *viTT-iliruntu* *etirppu* *iru-nt-atu*
I-dat our house-abl opposition *be-pst-3sn*
'I had oppositions from my home'

(12) *enakku* *romba* *naaL-aa* *oru* *TavuT*
I-dat long day-adv a doubt
'I have a doubt for a long time'.

(13) *iTuppu* *vazi-kku* *pala* *kaaraNang-kaL* *uNTu*
hip pain-dat many reason-pl be
'There could be many reasons for hip pain'

It should be noted that though the nominative NP is in plural form, it does not trigger agreement on the existential verb '*uNTu*' in (13). This is unlike Telugu where there is overt agreement between the singular / plural theme and the existential verb (see examples (7) and (8) of Subbarao 2001:121).

(14) *pinu.v-ukku* *talai-kkuL* *kuRukuRu.v-enRu* *iru-nt-atu*
Pinu-dat *head-dat+loc* *something-running* *be-pst-3sn*
'Pinu felt something strange inside his head'

(14) is very interesting: the dative marker appears along with the locative. There are double datives: one the experiencer and another is the goal. The experiencer dative is a subject in this case.

The experiencer dative when occurring with existential verb '*iru*' be and the psychological predicate 'bayam' fear behaves like a subject as in (15a) and (15b).

(15a) *saamb-ukku* *inta* *tiruppam* *bayam-aaka* *iru-nt-atu*
Sambu-dat this turning fear-adv be-pst-3sn
'Sambu was afraid of this change in the events'.

The existential verbs allow reduplication of infinitive markers in temporal adverbs with a dative nominal (an experiencer) as in (15b).

(15b) *neeram-aaka* *aaka* *saambukku* *bayam-aaka* *iru-nt-atu*
time-become-inf become-inf Sambu-dat fear-adv be-pst-3sn
'As time passed by, Sambu was afraid'.

Though most of the verbs of existence/possession (except in 13) trigger agreement with the nominative NP, the nominative NPs cannot serve as the subject because the dative nominal is an experiencer and all experiencers are subjects.

With modal verbs

The modal auxiliaries *kuuTaatu, kiTaiyaatu*, and the auxiliaries indicating necessity such as *teevai*, *veeNTum* trigger a dative subject as in (16) - (20).

(16) *raamu-vukku* *ittanai* *akambavam* *irukkak-kuuTaatu*
Ram-dat this much arrogance be-inf-neg
'Ram should not have this much arrogance'

(17) *namma amma.k-kaL-ukku veera velai.y-ee kiTai.y-aatu*
our mother-pl--dat other work-em exist-neg-3sn
'Our mothers don't have any other work'

(18a) attakaiya vaatiyaar-ukku paNam teevai.p-paTT-atu
that-kind teacher-dat money need-pass-3sn
'That kind of teacher needed money'

The dative nominal in (18a) can be expanded by one or more reflexive possessors as exemplified in (18b) and (18c).

(18b) *attakaiya vaatiyaar-ukku*$_i$ *tannooTa*$_i$ *makaL(in)*
*th*at-kind teacher-dat self daughter-poss
kalyaaNatt-ukku paNam teevai.p-paTT-atu
marriage-for money need-pass-3sn
'That kind of teacher needed money for self's daughter's marriage'

(18c) *attakaiya vaatiyaar-ukku*$_i$ *tannooTa*$_i$ *makaL-ukku kalyaaNam*
that-kind teacher-dat self-poss daughter-dat marriage
paNNarat-ukku paNam teevai.p-paTT-atu
do-inf-dat money need-pass-3sn
'That kind of teacher needed money to get self's daughter married'

When there are three datives as in (18c), the dative nominal that controls the reflexive becomes the subject: the one that immediately follows the possessive reflexive nominal is the indirect object and the other dative is the goal.

(19) *cinimaa.v-ukku teevai.y-ellaam oor eLuttaaLanaLLa*
cinema-dat necessary-all a writer-not
'To cinema (one) doesn't need a writer'

(20) *avan-ukku putu peenaa veeNTum*
he-dat new pen-nom want-3sn
'He wants a new pen'

It can be observed in all the above constructions that there are no specific PNG features acting on the verb to control its subject's PNG features. The morphological component related to Agreement does not include any constraints into the PNG block. On the contrary, constraint equations denoting selectional restrictions, on the object NP and the verb's Subclass should be included. This could be done as follows:

iru: *V, SUBCLASS Exist, (↑ OBJ) = c[CASE NOM].*
veeNTum: *V, PRED Want(SUBJ, OBJ) 'SUBCLASS Modal,*
(↑OBJ) = c[CASE NOM].

veeNTum can also take a nominative subject as in

(21) *naan* *poo.k-a* *veeNTum*
I-nom go-inf need
'I need to go'

This is true of other modals too. How to differentiate this from the dative subject? How does the parser recognize this subject? How are the consistency conditions written? As (21) contains only one NP the Tamil GF that is also recognized by the MORPHA as Nominative, the GF module will assign (SUBJ ↑) to this nominal.

With psych verbs

(22) *ivan-ukku* *aTippaTai.y-ee* *puri.y-aa-tu*
he-dat fundamental-emp understand-neg-3sn
'He will not understand the fundamentals'

puri: V, Subclass Psych, PRED 'Understand (SUBJ,OBJ)',
(↑OBJ) = c[CASE NOM].

(23) *avaL tanakku$_i$ aatarvaaka peecu-kiR-aaL-a a allatu tan$_i$ kuDumbattai*
she self-dat support talk-pres-3fs- int or self family-acc
keeli cey-kiR-aaL-aa enRu caarumat-ikkup$_i$ puriya.v-illai
tease do-pres-3fs-int that Carumati-dat understand-neg
'Carumati couldn't understand if she was talking on self's side or
(was she) criticizing self's family?

(23) shows that the dative nominal antecedes the reflexive in an embedded clause. So, one could say that the psych verb constructions of Dative-Nominative pattern trigger a dative subject.

With verbs of knowing

The verbs of knowing trigger a dative subject as illustrated in (24-26): it can antecede a reflexive as in (25). The infinitival construction involving *teri* 'know' requires a dative nominal as in (26) which serves as a subject.

(24) *nidi-kku atanpin enna naTa-nt-at-enRu cari.y-aak teri.y-a.v-illai*
Nidi-dat that-after what happen-pst-inf-that correct-adv know-inf-neg
'Nidi does not know exactly what happened afterwards'

(25) *nidi-kku$_i$ tan$_i$ viiT-ee teri.y-a.v-illai*
Nidi-dat self house-emp know-inf-neg
'Nidi doesn't know self's house'

(26) *unakku paaT-a.t teri.y-a-lai-daa*
you-dat sing- inf know-inf--neg-cli
'You don't know to sing'

The lexical entry for *teri* is:

teri:V, Subclass Know, PRED'Know(SUBJ,VCOMP)',
(↑VCOMP SUBJ) (↑SUBJ).

'*vaa*' constructions

These types of constructions involve a noun, adjective or a nonfinite verb and a verb with an impoverished semantic structure (Mohanan and Mohanan, 1990:47). Grimshaw and Mester (1988), and Mohanan and Mohanan (1990) call these as Light verbs. These always trigger a dative subject and have most of the subject properties:

They are experiencers in (27)-(29).

(27) *appaTi enna veRi? manitan-ukku appaTi.k kuuTa veRi varum-aa?*
kind what violence man-dat kind also violence come-int
'What kind of violence? To man is it possible to get such wildness?'

(28) *enakku.k kuuTa koobam varum aanaal avan-ukku*
I-dat too anger come but he-dat
koobam-ee var-aa-tu
anger-emp come-neg-3sn
'Even I lose temper but he doesn't lose his temper at all' (Lit: To me also anger comes but to him anger doesn't come at all'

(29) *anta marapont-ulee jinsi kizi-kku tuukkam-ee var-a-lai*
That t tree-hole-inside jinsiparrot-dat sleep-emp come-inf/neg
'Jinsi parrot did not get sleep at all inside that tree-hole' '(Lit: To Jinsi Parrot sleep didn't come at all inside that tree-hole)'

It controls PRO as in (30 and can antecede the reflexive as in (31) and (32).

(30) *mutal taTavai.y-aa uma.v-ukku [PRO] eetoo tappu*
first time-adv uma-dat some wrong
cey-tu-TTa-uNarvu va-nt-atu
do-vbp-leave-feeling come-pst-3sn
'For the first time, Uma had a feeling that she had done something wrong'

(31) *raam.v-ukkui tan$_i$ miit-ee koobam va-nt-atu*
Ram-dat self upon-emp anger come-pst-3sn
'Ram got angry at himself

(32) *anu.v-ukku$_i$ tan$_i$ kaNavan meel kaatal va-nt-atu*
anu-dat self husband-nom upon love come-pst-3sn
'Anu felt lov*e towards her husband'*
(Lit: To Anu love came upon self's husband)

Sometimes the light verb 'vaa' is not expressed overtly as in (33).

(33) *anu.v-ukku aacciriyam*
anu-dat surprise
'To Anu surprise came'

The '*vaa'* verbs show agreement with the nominative NP present in the sentence (3sn or 3pln). However, since the dative animate NPs are experiencers, they behave like surface subjects.

Distributed binding and null anaphora

Cuerro (1999) has pointed out that if dative nominals are subjects we would expect them to have the binding properties of other subjects: she presents evidence in Spanish for the case of a "hidden" distributive quantifier that is restricted to subjects. Here, we have evidence for "overt" distributive bound variable reading of a possessive pronoun in Tamil, as in (34).

(34) *unkaL-ukk-ellaam suuparvisor gaardu-nu aaLaaL-ukku*
You-dat-all supervisor guard-that person-dat
eTTaayiram sabalam
eight thousand salary
'To you, as supervisor, guard each one (gets) eight thousand as salary'

(35) *eLuttaLan-ukku (tan) ovvoru vaartai.y-um mukkiyam*
writer-dat self every word-cli important
'To a writer, (his) every word is important'

(35) is a case where there is a null anaphora; *tan i*s bound by *eLuttaLanukku*. So, it is clear that the dative nominals in (34) and (35) are subjects.

Conjunct Dative NPs

The conjunct Dative NP has a distributive reading too and hence is the subject in (36)

(36) *pinu.v-ukk-um pirutu.v-ikk-um avang-avanga*
Pinu-dat-and Piru-dat-and their-their
niiLattu.k-ee azak-aana vaal
length-dat-emp beauty-adj tail
'Pinu and Piru have beautiful tails according to their own length'

(37a) *amma appa.v-iRkk-um oNNum teriyalai* unakk-um
mother father-dat-and nothing know-neg you-dat-and
oNNum teriyalai
nothing know-neg
'Mother and father know nothing, you too know nothing'.

(37b) *amma appa.v-iRkk-um unakk-um oNNum teriyala*
mother father-dat you-dat-and nothing *know-neg*
'Mother, father and you don't know anything'

(37a) and (37b) are situations where conjunct-reduction is possible. This assures that the dative conjunct nominal in (37b) is the subject.

The dative nominal allows constituent co-ordination too as in (38)

(38) *amma.v-ukku ciripp-um azukai.y-um oNNaa va-nt-atu*
mother-dat laughter-and cry-and together come-pst-3sn
'To mother, laughter and cry came at the same time'

Here *amma* is the SUBJ of both laughter and cry.

Conjunct-datives need not always have a distributed reading as in (39). It could be I-nominal.

(39) *oru eLuttaaLan-ak-um vasanakartaa.v-ukk-um niRaya*
a writer-dat-and dialogue-writer-dat lots
veeRupaaTu uNTu
difference be
'There are a lot of differences between a writer and a dialogue-writer'

The subject of an infinitival clause is a dative as in (40) and (41).

(40) *enakku ameericaa poo-ka.p piTiikkum*
I-dat America go-inf like
'I like to go to America'

(41) *enakku pooka veeNTum*
I-dat go-inf want
'I want to go'

With temporal aspects

The dative in a temporal construction is usually an experiencer.

(42) *enakku innam-um reNdu maNi neeram -irukku*
I-dat more-and two hour time-have
'I have two more hours left'

Possessives

The possessor is the subject in dative-possessor constructions.[7] Like Telugu, the verb in these sentences is always *iru* ' be'.

(43) *enakku kaNNu-meelee oru kaTTi irukku*
I-dat eye-on a cyst be
'I have a cyst on my eye'

(44) *enakku reNTu kaal irukku*
I-dat two leg be
'I have two legs'

(45) *enakku reNTu tangacci-nga irukk-anga*
I-dat two youngersister-pl be-pl
'I have two sisters'

It is very common to use *namakku/namma* to denote I person singular noun in colloquial Tamil as shown in (46) and (47). Look at the word order in these two constructions. As Tamil is a relatively free word order language, this kind of word order is possible.

In (46), the existential verb '*iru*' is not overtly present.

(46) *viTinju* *entiric-aa* *naay* *pollappu-teen* *nam-akku*
at dawn get up-cond dog life-emp we-dat
'I have a dog's life everyday' 'Lit : To me it is dog's life everyday'.

(47) *namma* *samsaarta.t-ukku* *uTambu* *sari.y-il-enga*
our wife-dat health good-neg-2phon
'My wife is not keeping good health' (Lit: To my wife good health is not there')

Lehmann (1990:192) argues that in the Dative-Nominative constructions, the nominative NP shows agreement with the verb and the Dative NP has other subject-like properties (being the antecedent of the reflexive *taan*, controlling PRO etc). He is of the view that the DAT-NOM pattern does not have a 'subject' or 'no-subject' but two 'subject-like' NPs can be identified where the dative nominal can be more subject-like than the nominative one. But the data presented above show that not all nominative NPs present in the DAT-NOM constructions show agreement with the verb. I claim that only certain abstract nouns in the nominative case that are inanimate show agreement with 3 person neuter singular or plural verbs. Agreement cannot be used a criteria in these cases to assign (SUBJ) to the dative nominals. However, I have used other syntactic and lexical semantic devices to show that most of the dative nominals (except the one in 39) in the dative nominative patterns do function like subjects.

In the next section, we will explore the computational implications related to the dative constructions.

4. Unification, Morphological Blocking and Dative Subjects

This section describes the role of subsumption and unification in analyzing the dative subject constructions. Feature structures are used to capture the grammatical relationships in any natural language construction. These normally consist of a set of attribute-value pairs, where both attributes and value may be atomic or complex. The notions of subsumption and unification are used for the computer implementations of feature matching. Subsumption states that a category 'A' subsumes 'B' if the set of features 'A' is a subset of those in 'B'. In other words, a feature structure is subsumed by another fstructure, if it has more information than the other one. For example, if SET A has PERS III, it cannot subsume SET B if 'B' had PERS I but if 'A' has PERS unspecified then 'B' can be subsumed by it.

> "Unification is a process of combining grammatical information present in two feature structures into one. If the two feature structures contain incompatible information, unification fails. When parsing a sentence, a grammar uses unification to assemble pieces of information gathered from the lexicon, morphological component, etc into a single structure (f-structure in LFG terms). The grammaticality conditions of an f-structure are ensured by unification. Andrews (1990) used both unification and subsumption to formulate his principle of morphological blocking which states that: "Suppose the structure S has a preterminal node P occupied by a lexical item l1, and there is another item l2 such that the f-structure determined by the lexical entry of l1 properly subsumes that determined by the lexical entry of l2, and that of l2 subsumes the f-structure associated with P in S (the complete structure, after all unifications have been carried out). Then S is blocked*"* (Andrews 1990: 319).

This is a unification-based version of elsewhere Principle found in generative phonology: the constraints or conditions under which a particular suffix will unify with those of the syntactic elements in a construction are specified in a hierarchy (the more specific, less specific, specific, elsewhere). This is useful in blocking certain syntactic elements. The purpose of introducing feature structures and unification is to provide a way to express the morpho-syntactic constraints in an elegant way that would otherwise be very difficult. Our next step is to specify ways of using feature structures to represent the constraints imposed on certain syntactic properties. Let us see how this works for dative constructions.

4.1 Agreement

Agreement is a complex feature. I use PNG as a block related to subject-verb agreement. It consists of

PERS *with 3 possible values (I, II and III)*
NUM *with 2 possible values (SG and PL)*
GEND *with 3 possible values (MASC, FEM and NEU)*

These are given in the form of constraints into the morphological rules. For instance, a rule such as

(aan → (↑SUBJ PERS) = III, (↑SUBJ NUM) = SG, (↑SUBJ GEND) = MASC)

blocks the concatenation or unification of this feature to all subjects. It requires its subject to be of III person, Singular number and Masculine gender.

(48) * *raamu va-nt-aaL*
Ram-nom come-pst-3sf
'Ram came-she'

(48) is ungrammatical because the feature GEND MASC of Ram is not compatible with the Gender specification of *aaL*. (*aaL* includes (↑SUBJ GEND) = FEM) . Unification fails thereby yielding no parse.

As dative subjects do not show any overt agreement with the verbs, the morphological specifications of the person-number-gender features of Tamil cannot be applied directly. But, of course, it does block the use of a full-verb (the one with a PNG marker specified) in a dative subject construction. But there is a default specification (↑SUBJ PERS) = III, (↑SUBJ NUM) = SG, (↑SUBJ GEND) = NEU) unified with the nominative NP's f-structure iff the other NP has a Dative case marked on it.

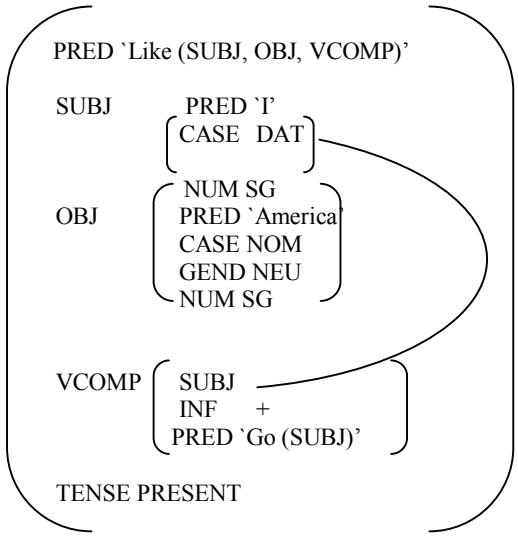
4.2 Morpho-syntactic Processing and GF recognition

The MORPHA developed for Tamil assigns the CASE labels to all the NPs present in the input sentence and the GF module assigns the respective GFs to these NPs. If an NP gets multiple GFs, it is tested for certain semantic properties. In the case of Dative subject constructions, we would expect the parser to identify its pattern (DAT-ACC or DAT-NOM) irrespective of the word order. The argument specifications of the verb present in the sentence are also tested. In case of DAT-ACC pattern, the default GF specification, *ukku*: (↑CASE) = DAT, (SUBJ ↑), is unified into the f-structure of the Dative nominal. Of course, this is only a possibility where there are no other details such as reflexives, PRO, etc. to provide additional evidence for the subject-like behavior of the dative nominal. It should be noted that the above specification consists of a defining equation. This equation introduces the information about the CASE. This does not require that the information about CASE be present in some other form (as in the case of constraint equations).

4.3 Handling PRO

This involves the unification of features present in the lexical entry of the control verbs and the PS rule. The PS rule expands to allow a to-complement. The lexical entry contains the information whether the subject or the object controls the VCOMP. In (42) the entry for piTi 'like' would contain: (↑VCOMP SUBJ) = (↑SUBJ), PRED '*piTi* (SUBJ, VCOMP)'.This information will be copied on to the f-structure of the infinitival verb. The SUBJ of *pooka* in (40) is *enakku* 'I-dat' which is the result of the unification of the features present in the lexical entry on to the VCOMP's SUBJ slot in the f-structure. The equation above indicates that only SUBJ can be the controller of the verb,

pooka ‘see’. The arguments of teri ‘know’ in (26) are also of the same type (SUBJ, VCOMP) and there is a defining equation that indicates that the SUBJ is the controller of the VCOMP. The f-structure of (40) is:



4.4 Anaphoric control

Just in case, the reflexive *taan a*nd its case-variants are present in a dative construction, a check for the antecedent could be made. Referential tables could be automatically generated for each nominal/pronominal expression to be used by the anaphoric binging algorithm. If the dative nominal serves as the antecedent of the reflexive *taan*, then it is realized as the subject. The semantic classes of the verbs also contain constraints that specify that the (SUBJ CASE = DAT). This information gets unified into the f-structure.

4.5 Conjunct-Dative NPs

I add a new operation to the LFG formalism called ‘conjoin’. Conjoin takes two feature structures and merges them to give a derived tree. It compares the feature specifications of both the structures and looks for the different ‘PRED’ values’ in the subject column. Given that the ‘case’ is the same for both, it simply adds the special schema to the derived tree. The tree obtained will be a lexicalized tree, with the lexical anchor as the conjunction: *-um, -oo* etc. The parser locates *–um* in one of the NPs present in the string and tries to find another NP showing the same. It creates a conjunct-NP that could serve as a single argument. For example, when ‘conjoin’ is applied to the feature

structures in (39a), it will yield the following:

FS 1:

SUBJ	PRED `Father'
	CASE DAT
	CONJ `AND'
PRED	Know (SUBJ, OBL)
	NEG +
OBL	PRED `Nothing'
TENSE	PRESENT

FS 2:

SUBJ	PRED `Mother'
	CASE DAT
	CONJ `AND'
PRED	Know (SUBJ, OBL)
	NEG +
OBL	PRED `Nothing'
TENSE	PRESENT

Unification of FS1 with FS2 yields:

SUBJ	PRED `Father_and_Mother'
	CASE DAT
	CONJ `AND'
	CONJUNCT +
PRED	Know (SUBJ, OBL)
	NEG +
OBL	PRED `Nothing'
TENSE	PRESENT

The attribute-value pair CONJUNCT + that appears in bold is added after unification by default. This is the f-structure for (37b). The dative nominal can serve as the SUBJ of two constituents as in (38). Applying 'conjoin' we get the following structure:

SUBJ	PRED `Mother'
	CASE DAT
PRED	come (SUBJ,OBL)
	SUBCLASS LIGHT +
OBL	PRED `Laughter_Cry'
CONJUNCT +	
TENSE	PAST
ADJUNCT	PRED `time'
	Mod: same

4.6 Miscellaneous

In DAT-NOM constructions, we also look for the following:

$$\begin{pmatrix} \text{Animate} + \\ \text{Definite} + \end{pmatrix}$$

When interpreting these constructions, we first look for the semantic class of the verbs and then move on to the syntactic properties of the elements under study. The nominative NPs of this pattern are always [ANIMATE –] and hence are eliminated in getting the value SUBJ. So (SUBJ ↑) is unified into the f-structure of the Dative NP.

5. Conclusions

This paper focused on the dative subject constructions in Tamil. It presented a descriptive method for the computation and recognition of Dative nominals within the LFG framework. It showed that case-marking and GF assignment in Tamil is morphosyntactic. The syntactic, lexical semantic and computational issues of the dative subjects were also discussed at length. It was shown that not all dative nominals could be interpreted as subjects. The datives that do not behave like subjects could be I-nominals and this needs to be studied in depth.

Endnotes

[1]Tamil is the pre-eminent member of the Dravidian Language family and has one of the longest unbroken literary traditions of any living language in the world. It is one of the two classical languages of India. It is the only language in India which has continued to exist for over two thousand years. Ninety-two percent of its speakers live in Tamilnadu (in South India), where it is spoken by 48 million first language speakers. In northern Sri Lanka between three and four million people, about 20 percent of the population of that island state, speak Tamil. Elsewhere there are several hundreds of thousands of speakers in each of South Africa, Malaysia, and Singapore; some 6,000 live in Fiji. There are significant minorities in Mauritius, Great Britain, the US, and Canada. Total speakers, including second-language speakers, number about 66 million (Grimes 1992). Tamil distinguishes three stages: Old Tamil, Middle Tamil and Modern Tamil, each with distinct grammatical characteristics. This paper is based on Modern Tamil – The language that exists since the last 400 years and which has undergone considerable changes. All Tamil speakers, including the uneducated, use two varieties of the language: literary and spoken Tamil. A high status variety is used in most writing, the media--including radio and television broadcasts--political speeches and other similar occasions. In contrast, a low status variety is used in every day discourse and conversations, films and fiction to familiarize with the audience. Examples for this study

are drawn from various sources including contemporary Tamil magazines and newspapers which cover a subset of the varieties of Modern spoken and written Tamil.

[2]Lehmann (1989:14) mentions that some nouns do not have oblique stems. In such cases, the case suffixes are added directly to the noun stem. 5 He classifies Tamil nouns into four: 1.Oblique stems with –ttu, 2. Oblique stems with –aRRu, 3. Oblique stems formed by doubling of the consonant and 4. all other nouns that do not belong to (1), (2) and (3) mentioned above.

[3]The LFG Package Andrews (1990) is used to construct the Tamil LFG grammar. It was used to analyse simple sentences in Tamil from a LFG perspective (Umarani 1996). The notations and the formalism used in this study are adapted from the same. Added value to the package is the enrichment of the morphological parser and the syntactic component to accommodate Dative subjects in complex constructions too.

[4] In Relational Grammar, some dative nominals have been analyzed as initial subject and final indirect object in the Inversion Constructins. Moore and Perlmutter (2000) call such dative nominals as the I nominals.

[5]The lexical entries are of the following format: 'WORD: Wordclass, Subclass, argument specification, featural annotations'. The format is the same as Andrews (1990) except that additional information about the Subclass of a word is also included.

[6] siitaai, tann.-ai$_i$.y-ee /avaL-ai..y-ee nontu-koNT-aaL
sita-nom self-acc-emp/she-acc-emp blame-refl-3fs
'Sita blamed herself/her

[7] The dative nominals in these constructions show a striking similarity with Malayalam (Jayaseelan 2001) and Telugu (Subbarao 2001).

Acknowledgements

An earlier version of this paper appeared in the Pre-proceedings of the "International Symposium on Non-nominative subjects" held at Tokyo, December 18-21, 2001. I thank Peri Bhasakarao, K.V.Subbarao, Alice Davison, Nick Evans, Howard Lasnik, Halldor Sigurdsson, and Lakshmi Bai for their feedback which helped me to improve this paper. All errors that remain are solely mine.

References

Andrews, Avery. 1990. Unification and Morphological Blocking. *Natural Language and Linguistic Theory*. Volume 8(4): 507-557.

Arden, A. H. 1942. *A Progressive Grammar of the Tamil Language*. Madras: Christian Literature Society.

Asher, R. E. 1982. *Tamil. Lingua Descriptive Series,* Vol. 7. Amsterdam: North Holland Publishing Co.

Bresnan, Joan. 1982. Control and Complementation. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*, 282-390. Cambridge, Mass., MIT Press.

Bresnan,Joan. 1995. *Morphology competes with syntax: explaining typological variation in Weak Crossover effects.* MS, Stanford University. Written version of paper given at the MIT Workshop on Optimality in Syntax, in May 1995.

Cuerro, Maria Cristina. 1999. Quirky but Not Eccentric: Dative subjects in Spanish. *MIT Working papers in Linguistics*, 34, 213-227.

Caldwell, R. 1856. (rep. 196*1). A Comparative Grammar of the Dravidian or South-Indian Family of Languages.* Madras: University of Madras.

Fillmore, Charles. 1968. 'The case for case.' in Bach, E. and Harms, R., *Universals in Linguistic Theory*, pp. 1-90 New York: Holt, Rinehart & Winston.

Hale, Kenneth. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory 1, 5-47.*

Grimes, Barbara F., ed. 1992. *Ethnologue Index [Twelfth edition]*. Dallas: Summer Institute of Linguistics.

Grimshaw, J. and A.Mester. 1988. Light verbs and Theta Marking. *Linguistic Inquiry*. 19 (2). 205-232.

Jayaseelan, K. 2001. Malayalam Non-nominative Subjects. In *The International Symposium on Nominative Subjects. Working Papers*, Peri Bhasakararao (ed.), 181-197. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Kachru, Jamuna. 1990. Experiencer and Other Oblique Subjects in Hindi. In. Verma, Manindra and K.P.Mohanan. Ed. *Experiencer Subjects in South Asian Languages*, 59-76. CSLI publications, Stanford, CA.

Kaplan, Ronald M. 1989. 'The formal architecture of Lexical-Functional Grammar'. In Chu-Ren Huang and Keh-Jiann Chen, editors, *Proceedings of ROCLING II,* pages 1--18. Also in Journal of Information Science and Engineering, vol. 5 no. 4, pp. 305-322.Available as a PostScript file Reprinted in Dalrymple, Kaplan, Maxwell, and Zaenen, eds., Formal Issues in Lexical-Functional Grammar, 7--27. Stanford: Center for the Study of Language and Information. 1995.

Kaplan and Bresnan. 1982. Lexical Functional Grammar: A formal system for grammatical representation. In Joan Bresnan (ed.). *The Mental Representation of Grammatical Relations*, 173-281.Cambridge, Mass., MIT Press.

Kothandaraman, R. 1986. 'With reference to sociative case.' *ISDL Working Papers in Linguistics* Vol. 2, No. 1, pp. 27-46.

Lehmann, Thomas. 1989. *A Grammar of Modern Tamil*. Pondicherry Institute of Linguistics and Culture. Pondicherry. India.

Lyons, J. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.

Mohanan, K.P. and Tara Mohanan. 1990. Dative Subjects in Malayalam: Semantic Information in Syntax.. In. Verma, Manindra and K.P.Mohanan. Ed. *Experiencer Subjects in South Asian Languages,* 43-58. CSLI publications, Stanford, CA.

Moore, John and David. M. Perlmutter. 2000. What Does It Take To Be A Dative Subject. *Natural Language and Linguistic Theory* 18 (2): 373-416.

Neidle, Carol. 1994. 'Lexical-Functional Grammar'. *Encyclopedia of Language and Linguistics*. New York: Pergamon Press.

Rachel Nordlinger. 1997. Morphology building syntax: Constructive Case in Australian languages. *Proceedings of the LFG97 Conference*. Stanford : CSLI Publications.

Rachel Nordlinger. 1998. Constructive Case. Evidence from Australian Languages. Stanford: CSLI Publications.

Sadler, Louisa. 1996. 'New developments in LFG'. In Keith Brown and Jim Miller, editors, *Concise Encyclopedia of Syntactic Theories.* Elsevier Science, Oxford.

Schiffman, Harold. 1979. *A Grammar of Spoken Tamil*. Madras: Christian Literature Society. Revised, 1995.

Schiffman, Harold. 1985. 'Deferential speech acts and the pragmatics of politeness in Tamil: From case to aspect.' In J. Fishman et al. (eds.), *The Fergusonian Impact. Vol. II, Sociolinguistics and the Sociology of Language.* pp. 23-33. Berlin: Mouton de Gruyter.

Shanmugam, S. V. 1971. Dravidian Nouns (a comparative study). (Annamalai University Department of Linguistics, Publication No.25.) Annamalainagar: Center of Advanced Study in Linguistics.

Subbarao, K.V. 2001. EPP, Agreement and Non-nominative Subjects in Telugu. In Bhaskararao (ed). In *The International Symposium on Nominative Subjects. Working Papers*, Peri Bhasakararao (ed.), 118-155. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Tolkappiyam. The most ancient grammar of Tamil.

Umarani. P. 1996. A Model for Computer Analysis of Tamil Sentences. Unpublished Ph.D. Dissertation submitted to the University of Mysore. India.

Zaenan, Annie, Joan Maling and Hoskuldur Thrinsson. 1985. Case and grammatical functions: the Icelandic Passive. *Natural Language and Linguistic Theory* 3, 441-483.

Zvelebil, K. 1982. 'Tolkaappiyam Collatikaaram, Part II. Morphology-Syntax. Chapter II: On Case.' *Journal of Tamil Studies* 21, pp. 9-19.