Gemination of stops in Tamil: implications for the phonology-syntax interface^{*}

HEMALATHA NAGARAJAN

1 Introduction

This paper investigates the syntactic determination of the phonological rule of gemination of stops in Tamil and attempts to account for the phenomenon using three crucial notions: (1) degree/depth of embedding, (2) directionality of application and (3) a locality restriction. Contrary to popular claims in the literature on the phonology-syntax interface about the existence of intermediary prosodic structures (Selkirk (1984), Nespor and Vogel (1982, 1986)), this paper shows that direct reference to the s-structure of a sentence is necessary and sufficient for stating the rule of gemination in Tamil.

A short digression about the language and the rule would be in order here. Tamil is a major member of the Dravidian family of languages, spoken in South India, parts of Sri lanka, Malaysia and Singapore. It is diglossic (Ferguson (1959)) i.e. it has a marked 'high' variety used in writing and formal speaking and a 'low ' variety used for purposes of conversation. The rule of gemination discussed in this paper belongs to the 'high' or literary variety of Tamil. It is a phonological rule well documented in orthography by Tamil grammarians. As this rule is part of 'high' Tamil, it is not 'native' to any speaker but has to be formally learnt through a well established system of orthography.

Section 2 provides a backdrop for the gemination rule — the phonological requirement and the levels of application. Section 3 presents the syntactic triggers of the rule. The inadequacy of the two major solutions proposed in the literature- the relation based theory (Selkirk (1984), Nespor and Vogel (1982,1986)) and the direct syntax theory based on the notion of c-command (Kaisse (1985)) is pointed out. Recent theories which incorporate the feature 'focus' to establish phonological domains of rule application- though appealing, are discarded as they overgenerate in Tamil. Section 4 discusses an alternative, viable theory which refers to the s-

^{*}I would like to thank Misi Brody and John Harris for comments. I am extremely grateful to Neil for his constant encouragement and sharp comments and to Vijayakrishnan for readily providing the data and 'long distance guidance'. I am also thankful to the audiences at the Departmental seminar at UCL and at the Conference on Interfaces in Phonology at FAS in Berlin, where different versions of this paper were presented.

structure of a sentence and looks at the degree of embeddedness of a constituent (Cinque (1993)). The most deeply embedded constituent(s) with respect to the head of a phrase not only act(s) as a trigger but also set(s) the parameter for the direction of application of the rule. These notions are shown to neatly account for the basic sandhi environments. Section 5 deals with certain ticklish issues like serial verbs, finite vs. non finite clauses, dative vs.nominative subjects, and scrambling facts where the same analysis is extended. Thus, this phonological phenomenon is shown to be 'directly syntax-driven, without recourse to prosodic theoretic notions such as the phonological phrase or intonational phrase' (Cinque (1993)).

2 Phonological requirement

2.1 Introduction

Word initial stops are geminated in Tamil when preceded by either a short vowel or a glide.

- (1) pustakattay = $kkuDu^{1}$ book (acc) give Give the book.
- (2) pustakattay avanukku = kkuDu book (acc) he to(dat) give Give the book to him.

The rule is inapplicable when preceded by either a consonant or a long vowel.

- (3) pustakam # kuDu book give Give the book.
- (4) pustakamee # kuDu book (emph) give Give the BOOK.

¹The symbol = henceforth indicates a sandhi domain and the symbol #, the blocking of the rule.

However, in particular syntactic environments, this rule does not always apply even when these phonological conditions are met.

2.2 Levels of application

The rule of gemination of stops (henceforth GS) applies at both the lexical and postlexical levels.

In early Lexical Phonology (Kiparsky (1982), Mohanan (1986)), phonological rules were assigned to different levels of the lexicon or treated as post-lexical. Kiparsky (1984,1985) on the other hand suggests that phonological rules are available right at the outset, but can be 'turned off' at some point in the derivation. Three general principles constrain rule application- the Strong Domain Hypothesis, the Strict Cycle Condition and Structure Preservation. Following Vijayakrishnan (1988), if we assume the Tamil lexicon to have the three levels in (5) (followed by the post-lexical level), the rule of GS applies right through the lexical and the post lexical levels.

(5)



Vijayakrishnan (1985) argues that the rule is partly determined lexically by the feature $[\pm G]$ which is specified for all verb roots and affixes.

As the focus of this paper is the identification of domains which are larger than the phonological word- namely, the post-lexical level, I ignore the intricacies of rule application at the lexical level.

3 Earlier accounts

The rule of GS is mentioned in most ancient grammars of Tamil. What is generally cited is a random, apparently arbitrary list of constituents which trigger the rule.

If word initial k, t, p is preceded by a noun in the accusative case, a noun in the dative case, a demonstrative or interrogative adjective (inta, anta,enta), an infinitive or one of the adverbs ippaDi 'like this', appaDi 'like that', eppaDi 'how', the realisation is with a voiceless plosive, which is normally of sufficient duration to justify the use of a length or half-length mark in a phonetic transcription (Asher 1982: 236).

The following are the basic environments:

- (6) anta = ppustakattay = kkuDu that book (acc) give imp. Give (me) that book. (Dem=N,Obj.NP=V)
- (7) avanukku = ppaNam kuDutteen him (dat) money give past 1s.
 (I) gave him money.
 (I.O=D.O)
- (8) raamanukku = ppacikkiradu ram (dat) hunger pres 3p. Rama is hungry. (Dative subj.=V)
- (9) avaL aRagaaha = ppaaDinaaL she (nom) melodiusly sing past 3psf. She sang melodiously. (Adv=V)
- (10) avan miinay = ppool kiDantaan he (nom) fish post. lay past 3psm. He lay like a fish. (NP=Post.)

(11) raajaa kumaaray = kkuDikka = kkaTTaayapaTTaan king (nom) kumar (acc) drink inf. force past 3psm. The king forced Kumar to drink. (Obj NP=V,Inf V=V)

Sandhi is blocked in the following environments:

- (12) oru # payyan a boy (Indef.Det#N)
- (13) nalla # payyan good boy (Adj.#N)
- (14) naan paartta # payyan I (nom) see rel. boy The boy I saw (Rel.cl #N)
- (15) avan neeTTru # poonaan he (nom) yesterday go past 3psm. He went yesterday. (Sentential adv.#V)
- (16) raajaa kumaaray = kkuDikkanum enru # kaTTaayapaDuttinaan king(nom) kumar(acc) drink should that force past 3psm.
 lit. The king forced that Kumar should drink.
 (Comp #V)

Two general questions are involved in a study of syntax-phonology interaction. (1) What is the information in syntax that is relevant for phonology; and (2) How is this information encoded or incorporated into phonology ?

The answer to question (2) provides a clue for (1). Two possibilities have been suggested for encoding syntactic information in phonology- the first is the Direct Syntax Approach (DSA) followed by Kaisse (1985), Kenstowicz (1987), who believe that rules of phonology that require syntactic information make direct reference to syntactic structures. Thus, the notions relevant for syntax are found to be relevant for phonology too. (c-command,government etc.). Cinque (1993) in his recent paper 'A

null theory of phrase and compound stress', suggests this as a plausible approach for the treatment of external sandhi but his paper is restricted to a study of prominence at the phrasal and compound levels.

The second approach assumes a two step process for encoding syntactic information. The syntactic structures are transformed to prosodic structures which in turn are changed to phonetic representations-this has been referred to in the literature as the Indirect Syntax Approach (ISA). The chief exponents of this approach are Selkirk (1984), Nespor and Vogel (1982,1986), Hayes (1989).

I will briefly sketch the main proposals- the c-command approach (Kaisse(1985)) and the relation based theories (Selkirk (1984)), Nespor and Vogel (1982,1986) and note their drawbacks.

Kaisse (1985) assumes that the rules of the post-lexical component operate at two levels- P1 and P2, such that P1 comprises the rules of external sandhi and P2, the rules of fast speech, with pause insertion rules operating between them. While the rules at P1 use notions such as the c-command and the edge condition, those at P2 do not. Hence, not all types of syntactic information may be relevant for post-lexical rules and not all types of post-lexical phonological rules might need to refer to syntactic notions.

Kaisse claims that the two parameters for unmarked rules of external sandhi i.e.the P1 level rules are (1) the c-command condition and (2) the edge condition. Kaisse uses the stricter version of c-command, domain c-command, which states that no constituent can c-command out of its own maximal projection. The edge condition states that the words participating in the rule must lie at the edge of the constituent that contains them.

These two notions are empirically inadequate to account for the above Tamil data. They can account for only a few cases in Tamil(obj=verb, NP=post.). The unexpected behaviour of demonstratives and dative subjects, finite vs. non-finite clauses and scrambling facts (which I present in the last section) are all unaccounted for. In recent syntactic proposals (Fukui (1988), Contreras (1987)), where the subject is believed to originate in the VP internal [Spec, VP] position and moves to [Spec,AGR_s P] ,the verb would c-command everything else from AGR_s and yet this is never a sandhi domain in Tamil.²

Selkirk (1984) claims that all interaction of syntax and phonology can be taken care of by a theory of 'juncture'. The rules she proposes, look at syntactic trees and create metrical grids where juncture is translated as demibeats. The prediction is that the lesser the number of demibeats, the greater the likelihood of sandhi. Nespor and

²Even if the subject does not move, the subject in [Spec,VP] will be c-commanded by the verb in the VP.

Vogel (1982) posit prosodic structures like phonological phrases, intonational phrases and utterance groups. These prosodic structures are constructed in a pyramid like fashion. The head of a phrase and its modifiers (with a complement optionally) form a phonological phrase. The subject and the predicate form separate intonational phrases. Postposed and topicalised NPs are also considered to form separate intonational phrases. All intonational phrases are ultimately joined to form utterance groups. The prediction is that if a sandhi rule applies within a higher domain (say, intonational phrase), it should definitely apply across lower domains(phonological phrases).³ Both these theories are proved wrong in Tamil where sandhi applies in a higher domain (object NPs and the verb) and yet is blocked in a lower domain (modifiers of the noun and the noun).

Recent studies on prosodic phrasing have incorporated the feature 'focus' to account for certain puzzling facts of sandhi in languages like Hungarian (Vogel and Kenesei (1987)), Korean (Cho (1990)), Chichewa (Kanerva (1990)), Bengali (Hayes and Lahiri (1991)). In early papers (Nagarajan (1989,1991)), I identified and defined the domain of gemination using the notion of focus. Though apparently descriptively adequate, this theory 'leaks'. It makes strong predictions about the possibility of sandhi in all environments of a focussed constituent. This, unfortunately, is not true.

- (17) seethu # cenDraan, ravi illay Sethu (nom) go past 3psm, Ravi not [+F] It is Sethu who went, not Ravi.
 (Subj NP #V)
- (18) avan nalla # payyan illay, keTTa # payyan he (nom) good boy not bad boy[+F] He isn't a good boy, he is a bad boy. (Adj.#N)
- (19) kuRandaykku = kkaranDiyay = kkuratti # kuDuttaaL child to (dat) ladle (acc) gypsy (nom) give past3psf. It is the gypsy who gave the child the ladle.
 (Scrambled sentence with preverbal nominative subject [+F])

4 The theory of embeddedness

³I have just presented the gist of the relation based approach and not the exact details of Selkirk (1984).

This section shows how certain basic notions can capture the sandhi domains in Tamil in an insightful and elegant manner. In doing so, I suggest the following parameters for the determination of the Tamil sandhi rule GS:

- 1 Assign a feature [+G] to all constituents embedded deeper than the head in a syntactic domain.
- 2 Apply the sandhi rule from left to right.
- 3 Apply the rule cyclically.

I use the null theory of Cinque (1993) and show how the interaction of phonology and syntax need not always be indirect.

Cinque (1993) proposes a null theory of phrase and compound stress in which he claims that stress prominence is crosslinguistically a 'mere reflection of depth of embedding' (ibid:245). A phrase which is dominated by more projections is perceived as more prominent than its neighbours. The rightmost or leftmost location of main stress is shown to fall out of this notion of embeddedness (which is dependent on the direction of branching). Thus, the Nuclear Stress Rule (NSR) and the Compound Rule (CR) are shown to be redundant.

Before applying this theory to account for GS in Tamil, I provide some information about Tamil syntax. Tamil is an SOV language. Auxiliaries follow the verb and it is strictly postpositional. Word order is relatively free, conditioned by pragmatic factors. Nominative case ending is null, dative and accusative case endings are postpositional-the dative case marker is '-ku' and the accusative marker is '-ay'.⁴

Let us start with the simplest domain about which this theory makes predictions: the direct object and the verb. Let us take a simple, monotransitive sentence:

(20) naan aapiLLay = kkaditteen. I (nom) apple (acc) bit past 1ps. I bit the apple.

Let us assume in keeping with the VP internal subject hypothesis, that the arguments are generated VP internally- the subject in [Spec,VP] and the internal arguments in a Larsonian shell structure (Larson (1988)).

(21) AGR_sP

(1) naan oru pustakattay / pustakam paDitteen.
 I(nom) a book(acc) book(0) read past 1p.
 I read a book.

⁴The accusative marker is optional with inanimate nouns.



The verb theta marks and case marks the object structurally- so it does not have to move to get case. The subject 'naan' moves to $[Spec,AGR_s P]$ to receive case from $[AGR_s]$ by Spec-Head agreement. As is evident from the tree, the object NP is lowermost in the structure with respect to the head (the verb). In Cinque's theory, this makes it the best candidate for prominence placement and in my account too- it wins over as the most plausible potential trigger for GS. Thus, it is assigned the feature [+G] and the rule applies automatically to the following constituent (the verb in this case).

Let us now consider a ditransitive verb:

(22) naan kuRandaykku = ppustakattay = kkuDutteen.
I (nom) child to (dat) book (acc) give past 1ps.
I gave the book to the child.



As both the object NPs are lower than the verb, they are both assigned the feature [+G] and thereby induce gemination on the item to their right.

Thus, an argument of a head will always form a sandhi domain with the head as it is always lower in the tree - this prediction is borne out in the following cases:

- (24) avan miinay = ppool kiDantaan he (nom) fish like post. lay He lay like a fish. (NP =Post.)
- (25) avan puhaRvadarkku = ttahundavan he (nom) praise for (dat) worthy He is worthy of praise. (NP =Adj.)

The different behaviour of VP adverbials as opposed to sentence adverbials can also be explained straightforwardly. Larson (1988) and Stroik (1990) provide evidence that objects asymmetrically c-command VP adverbials at s- structure- a fact which indicates that the adverb would be the most deeply embedded constituent in the VP.

(26)	avaL	paaTTay	aRagaaha =	ppaDinaaL
	she (nom)	song (acc)	melodiously	sing past 3psf.
	She sang the	song melodi	ously.	

(27)SPEC avaL



A sentential adverb, on the other hand, would be under a higher projection (perhaps adjoined to $AGR_s P$) and therefore not in a relevant domain for sandhi.

(28)	neeTTru #	kumaar	uurukku =	ppoonaan.
	yesterday	Kumar (nom)	town to (dat)	go past 3psm.
	Kumar went	t to town yesterd	ay.	



As was noted in the previous section, indefinite determiners, adjectives, possessives and relative clauses are blocking domains for sandhi.⁵ Following Abney (1987), I assume that the functional category D heads the DP which takes the NP as its complement. However as Dasgupta and Bhattacharya (1993) note:

The South Asian D head of the DP is a declension and follows the NP. The South Asian determiner which carries deictic features and precedes the N must be analysed as the Spec of N, not as the head of DP. The decision to recognize a DP final element as the real D helps maintain the generalization that South Asian functional projections exhibit the head final pattern throughout the language type.

⁵Demonstratives and wh-words are the only items within an NP that induce sandhi. They are discussed in the next section.

If we take a Noun phrase in Tamil like the following:

(30)	ennooDaya #	pattu #	paccay #	puDavaygaL
	my	ten	green	sarees

we note that sandhi is not possible between the possessive 'ennooDaya' and the numeral 'pattu', the numeral and the adjective 'paccay' and the adjective and the Noun 'puDavaygaL'.

Assuming the Tamil DP to be the following (fig.31.), it is possible to provide a viable explanation for the inapplicability of the sandhi rule between the Adjective-N, the Numeral-Adj., and the Poss-Numeral.



In Chomsky (1992), two economy principles are introduced to 'regulate' movement operations- Greed and Procrastinate. The principle Greed forces any feature that is 'strong' to be checked overtly. If a feature is 'weak', Procrastinate would delay movement. As there is no overt number agreement in Tamil, I assume Tamil Number to be weak and thus does not require overt N-raising in syntax (the economy principle 'procrastinate' would delay N-raising to LF (Chomsky (1992)). This means that the noun will remain in situ as the most deeply embedded constituent in the DP. The noun being the lowermost in the tree can induce sandhi, but it cannot group with the items

preceding it as the sandhi rule is sensitive to directionality of application (it does not apply right to left). Hence, none of the items preceding the noun form a sandhi environment. When the DP itself is the object of a verb within the VP, the noun can induce sandhi on the verb on the higher cycle.

(32) ennooDaya # pattu # paccay # puDavaygalay = ppaar. my ten green sarees see Look at my ten green sarees.

To sum up, the important steps in sandhi rule application in Tamil are the following:

- 1 Locate the most deeply embedded constituents in a syntactic domain and assign them the feature [+G].
- 2 Apply the rule of GS from left to right.
- 3 Apply the rule cyclically from the lowermost to the highermost syntactic domains.

5 Problematic environments

5.0 In this section, certain problematic environments will be -discussed where my theory scores over the other two approaches- namely, scrambling, dative vs. nominative subjects, serial verbs, infinitives and definite determiners.

5.1 Scrambling

Scrambling is a phenomenon that alters word order in one way or the other. An interesting interaction of scrambling and external sandhi in Tamil was first noted in Vijayakrishnan (1985). The basic word order of Tamil can be altered for pragmatic reasons. In a basic sentence like the following

(33)	kuratti	kuRandaykku = kl	karanDiyay =	kkoDuttaaL.
	gypsy (nom)	child to (dat)	ladle(acc)	give past 3psf.
The gypsy gave the child the ladle.				

gemination applies between the two objects and between the verb and the direct object. Consider now the scrambled versions:⁶

⁶These examples are taken from Vijayakrishnan (1984).

- (34) a. kuRandaykku = kkuratti # karanDiyay = kkoDuttaaL.
 - b. kuRandaykku = kkaranDiyay = kkuratti # koDuttaaL.
 - c. kuRandaykku = kkoDuttaaL # karanDiyay = kkuratti.
 - d. karanDiyay = kkuratti # kuRandaykku = kkoDuttaaL.
 - e. koDuttaaL # karanDiyay = kkuRandaykku = kkuratti.

This is a problem for prosodic structure analyses as postposed / topicalized NPs are considered to form separate Intonational phrases. So, now we are forced to say that U (the Utterance group) is the domain of sandhi in Tamil, as the scrambled object NPs induce sandhi. Yet, this leads to a paradox- the Utterance group is both a sandhi domain and a blocking domain! Though the object NPs trigger the rule irrespective of their position in the sentence, the subject NP never forms a sandhi domain. Thus the Intonational phrase (I) is not a domain even when the Utterance group (U) is !

With regard to gemination, I have already mentioned that the internal arguments of the verb form a sandhi domain as they are most deeply embedded in the tree. Yet even when scrambled out of the VP, they trigger gemination. I attribute this to a reconstruction effect: as if the object phrases were put back in their original position to receive the feature [+G] to license gemination.⁷ This has been independently motivated in the literature to account for binding facts, the occurrence of parasitic gaps etc.

In early accounts, scrambling was treated as a PF rule. Saito (1985) proposed scrambling as an A-bar movement i.e. as s-adjunction. In a recent study, Mahajan (1990) discusses scrambling as two separate operations- argument shift and adjunction to XP. Argument shift is argued to yield a structure that enters binding theory and is not reconstructable. Adjunction to XP, on the other hand, does not provide new binders and is reconstructable. Evidence from binding facts (cf.35a and b.) and parasitic gaps (cf.36) provide evidence that scrambling in Tamil is a case of adjuction to XP and is therefore reconstructable.

⁷In Bare Phrase Structure (1994), Chomsky proposes precisely this- movements leave 'copies' and not 'traces'.

500 Hemalatha Nagarajan

(35)	a.	* avan _i he (nom) He _i loves	raama Ram's Ram's	nooDa _i (gen) i sister.	tar sis	ngayay = ter (acc)	kkaadalikkuraan. loves.
	b.	* raamane Ram's (ge	ooDa _i en)	tangayay sister (ac	c)	avan _i he (nom)	kaadalikkuraan. loves.

Ram's_i sister, he_i loves.

Sentence 35a. is clearly a Binding Condition C violation as the pronoun c-commands the name and coreferentiality is ruled out. But 35b. is unexpected, as the pronoun presumably does not c-command the name anymore and yet co-referentiality is not possible. This can be explained in terms of reconstruction.⁸ The object phrase behaves as if it were in its original position for the binding conditions.

The second piece of evidence for reconstruction is that a scrambled phrase can license parasitic gaps:

(36)	enda	ppusta	kattay	moohan	ninaittaan	[raaman	[PRO
	which	book		Mohan	thought	Ram	
	paDikka	amay]	e	tuuki	poTuvi	iTTaan]	enru.
	reading		withou	ut throw	y away		that
	Which b	ook did	Moha	n think th	at Ram thre	ew away v	without reading?

Mahajan (1990) argues that if the preposed direct object can license the parasitic gap, then it must be in a non L- position (an adjoined position). This is added evidence for scrambling as an adjunction operation in Tamil, which is reconstructable.

5.2 Non-finites vs. finites

The following contrast has been observed with regard to sandhi: non-finite clauses form transparent domains whereas finite clauses are opaque.

(37)	a.	kumaar Kumar (non Kumar wen	siitaavay n) Sita (acc) t to see Sita.	= ppaark) see.inf	ka = ppoonaan . go past 3psm.
	b.	naan ku	umaaray	vara =	cconneen.

⁸This was noted by Jayaseelan (1991) for Malayalam, a sister language of Tamil.

I (nom) Kumar (acc) come.inf. say past 1ps. I asked Kumar to come.

c.	kumaarukku =	ccaikil	ooTTa =	tteriyum
	Kumar (dat)	cycle	ride.inf.	know
	Kumar knows to	o ride the	cycle.	

- d. raajaa kumaaray = kkuDikka = kkaTTaayapaTTaan king (nom) Kumar (acc) drink.inf. force past 3psm. The king forced Kumar to drink.
- (38) a. kumaar vantaanaa enru # paar Kumar (nom) come past 3ps. that see See if Kumar has come.
 - kumaar taNNiir varavillay enra # pukaray eRutinaan.
 K. (nom) water come not that complaintwrite past 3psm.
 Kumar wrote a complaint that water wasn't coming.

The complementizers commonly used to embed finite clauses are 'enru', 'enra' and 'pool'. The presence or absence of the complementizer has no effect on the sandhi rule. For the c-command approach this is a problem- as the matrix verb would c-command the lower verb in either case and hence sandhi should be possible in both the environments. For the prosodic structure hypothesis too it is a problem as both finite and non-finite clauses would be grouped alike into separate Intonational phrases.

The question that arises is: What is unique to the structure of non-finites which makes it a possible domain for sandhi? The answer is that non-finite clauses are not 'full clauses' or CPs but just VPs. Lehmann (1989) notes the following restrictions on the use of non-finite clauses. Firstly, verbs that take interrogative complements cannot embed non-finite clauses eg. keeL 'ask', vicaari 'enquire'.

(39)	a.	kumaar kumar (They) as	eppotu when sked when	varaan come fut. Kumar was o	enru that coming	keeTTaarkaL ask past 3ppl. g.
	b.	* kumaar K. (nom)	eppootu when	varavatay come (acc)	keeT ask pa	FaarkaL ast 3ppl.

Sentences like the following are not possible in Tamil:

(40) I asked him what to do. I asked him where to go.

This is evidence to show that non-finites are not CPs.

Secondly, Lehmann (ibid;311) notes that 'a negative clause cannot be embedded before a manipulative verb on the non-finite verb form strategy. Thus a sentence like *I told him not to go there* cannot be realized with this strategy in Tamil'. This is added evidence to show that non-finites are not 'full clauses'.

Thirdly, non-finites cannot be passivized. Nor can they take aspectual markers. Sentences like the following are impossible in Tamil:

- (41) a. The king forced wine to be drunk.
 - b. The minister ordered English to be taught.

Fourthly, the subject position of the infinitive can take a reflexive or a pronoun coreferential with the matrix subject.

(42) kumaar taan/avan pooha virumbinaan.Kumar self/him go inf. prefer past 3psm.Kumar preferred (self/him) to go.

These factors point to a process of clause reduction or clause union in non-finites. The choice of a non-finite as opposed to a finite clause as a sandhi domain can now be reduced to a locality restriction.

(43) Sandhi rule GS (Gemination of Stops) is clause bound.

5.3 Serial verbs

Serial verb constructions are those where two or more verbs occur together, sharing the same arguments and having a unique selection of tense/ aspect; i.e. it is encoded only on one of the verbs. Serial verb constructions are commonly used in Dravidian languages.

(44) kuratti kuRandayay = kkiLLi # konjinaaL gypsy (nom) child (acc) pinch past caress past 3psf. The gypsy pinched and caressed the child.

In such constructions, the subject as well as the objects (depending on the argument structure of the verb) are identical. The first verb is inflected only for tense whereas the second verb shows both tense and agreement.

There has been much debate in the literature on serial verb constructions whether they are subordinated or coordinated structures (Baker (1988), Lefebvre (1991), Cormack and Smith (1994)). What interests us here is that serial verbs can never be a sandhi domain in Tamil. Assuming the dual headed structure suggested by Baker (1988), then 'konjinaaL' in 45. below, does not create a sandhi domain.



This is obviously a problem for both the c-command and the prosodic structure theory. However, in our approach, it is immediately clear why sandhi is blocked here. The second verb is lowermost in the tree and cannot trigger gemination to its left (due to left to right application of the rule).

5.4 Dative vs. nominative subjects

The following verbs take dative subjects in Tamil:

Verbs expressing physical states

(46) enakku = ppacikkiradu me (dat) hunger pres 3ps. I am hungry.

Verbs of mental experience

(47) enakku avavay = tteriyum me (dat) him (acc) know pres3ps. I know him.

Verbs of emotion

(48) enakku avanay = ppiDikkum me (dat) him (acc) like pres 3ps. I like him.

Verbs of possession

(49) kumaarukku iranDu # payyangaL irukkiraargaL
K. (dat) two boy (nom, pl) be pres 3p pl.
Kumar has two sons.

In contrast to nominative subjects, which never trigger gemination, dative subjects can form a sandhi domain.

- (50) kuratti # poonaaL gypsy (nom) go 3psf. The gypsy went.
- (51) enakku = kkumaaray = ppiDikkum. me to (dat) Kumar (acc) like I like Kumar.

This poses a problem for the c-command and prosodic analyses: if the dative subject is a 'true' subject and is in the subject position, it should behave like a nominative subject. It would neither be in a c-command domain nor would it form a prosodic unit for sandhi. If so, it is a puzzle in my theory too. But this is not so! Jayaseelan (1990) argues persuasively that dative NPs are basically oblique arguments of the verb with a' pro' in the subject position. These oblique arguments may be topicalized to a presentential position from their VP internal base generated position. Thus, on the surface, they seem to occupy the 'subject' position.



Using this analysis, there is a straightforward account for GS in this environment: The oblique argument, being lower in the tree than the verb- is predictably, a sandhi domain for GS and does license gemination. Using this structure, it is possible to account for gemination in the other two approaches too. Thus as Cinque (1993) points out '.... it shows how careful one must be in putting a hypothesis to test. Insufficiently analysed structures may easily lead to incorrect conclusions concerning the hypothesis to test'. Therefore, reference to correct syntactic structures is crucial for GS in Tamil.

5.5 Demonstratives and wh-words in DP

It was noted in section 3 that noun modifiers are blocking domains for GS. However, demonstratives and wh-words are exceptions to this.

(53) inta = a. ppayyan this boy b. anta =ppayyan that boy c. enta = ppayyan which boy

Tamil has a trinary set of morphs i-,a- and e- which represent proximity, remoteness and interrogativity respectively. These three morphs occur either as stems or bound morphemes. They can affix to third person personal pronouns:

a-v-an	'that one, he'
i-v-an	'this one, he'
e-v-an	'which male person'

They can be prefixed to nouns as bound determiners:

ak-kaDitam	'that letter'
ip-pootu	'this time'
ev-v-aLavu	'which extent'

Finally, they occur before nouns as demonstrative determiners and wh-word 'inta',' anta' and 'enta'.⁹

It has been noted in the literature (Valois (1991), Carstens (1991), Srivastav (1991)) that demonstratives and wh-words behave alike with regard to extraction in French, Kiswahili and Hindi. Srivastav notes: 'Thus DEM seems to be on a par with variables created by movement'. Hence, they are assumed to occupy the same position within the DP.¹⁰ Let us call this [Spec,*P] and assume they have a strong feature [+def.] which requires overt checking. The noun thus raises, taking along with it,the prenominal adjective and numeral (if present). Valois (1991:161) suggests a similar analysis for prenominal adjectives in French.'The adjective first adjoins to the left of the noun which itself moves to D.'

The phrase

(54) ennooDaya inta = pattu # paccay # puDavaygaL my this ten green sarees

would have the following structure:

⁹Caldwell (1974:170) notes: 'These appear to have been formed from the neuter demonstrative pronouns a-du, i-du and the interrogative e-du by the insertion of the euphonic nasal'.

¹⁰Valois (1991) and Carstens (1991) assume they are in [Spec, DP].



The demonstrative/interrogative in [Spec,*P] being lowermost in the tree, can create a sandhi domain. Selkirk and Shen (1990) note a similar behaviour of definite determiners in Shanghai Chinese.

Thus, to summarize this section, certain unexplained facts with regard to GS are now straightforwardly accounted for by making crucial reference to syntactic structures and assigning the feature [+G] to the most deeply embedded constituent(s).

6 Conclusion

In this paper, I have shown that the external sandhi rule GS needs to scrutinize the s-structure of a sentence and that the translation of syntactic structures to prosodic units is unnecessary and insufficient. The rule of GS is in a true sense a P1 level rule (Kaisse 1985).

References

- Abney, S. 1987. The English NP in its sentential aspect. Unpublished dissertation, MIT.
- Asher, R.E. 1982. *Lingua Descriptive studies-Tamil.* Amsterdam: North Holland Publishing Company.
- Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: Chicago University Press.
- Caldwell, R. 1974. *A Comparative Grammar of the Dravidian or South Indian family of languages.* New Delhi: Oriental Books Reprint Corporation.
- Carstens, V.M. 1991. The Morphology and Syntax of Determiner Phrases in Kiswahili. Unpublished dissertation, University of California.
- Cho, Y.Y. 1990. Syntax and Phrasing in Korean. In Inkelas, S. & Zec, D. (eds), *The Phonology-Syntax Connection*. Chicago: University of Chicago Press.
- Chomsky, N. 1992. A minimalist program for linguistic theory. MIT Working Papers.
- ——. 1994. Bare Phrase Structure. MIT Working Papers.
- Cinque, G. 1993. A Null Theory of Phrase and Compound Stress. Linguistic Inquiry 24: 239-297.
- Contreras, H. 1987. Small clauses in English and Spanish. *Natural Language and Linguistic Theory* 5: 225-243.
- Cormack, A & Smith, N. 1994. Serial Verbs. UCL Working Papers in Linguistics.
- Dasgupta, P. & Bhattacharya, T. 1993. Classifiers and the Bangla DP. In Davison, A. & Smith, F.M. (eds), *Papers from the Fifteenth South Asian Language Analysis Roundtable Conference*. Iowa.
- Davison, A. 1994. Binding and Control properties of dative experiencers. Talk given at ISIS, CIEFL, Hyderabad.
- Ferguson, C. 1959. Diglossia. Word 15: 325-340.
- Fukui, N. 1988. Deriving the differences between English and Japanese:case study in parametric syntax. *English Linguistics* 5: 249-270.
- Hayes, B. 1989. The prosodic hierarchy in meter. Orlando: Academic Press.
- ——. & Lahiri, A. 1991. Bengali Intonational Phonology. *Natural Language and Linguistic Theory* 9.1: 47-96.
- Jayaseelan, K.A. 1990. The Dative Subject construction and the Pro-drop parameter. In Verma, M.K. & Mohanan, K.P. (eds), *Experiencer Subjects in South Asian Languages*. Stanford: CSLI
- ——. 1991. The Pronominal System of Malayalam. CIEFL Occasional Papers in Linguistics 3: 68-107.
- Kaisse, E.M. 1985. Connected Speech. The Interaction of Syntax and Phonology. Orlando: Academic Press.
- Kanerva, J.M. 1990. Focusing on Phonological Phrases in Chichewa. In Inkelas, S & Zec, D. (eds), *The Phonology-Syntax Connection*. Chicago: University of Chicago Press.
- Kenstowicz, M. 1987. Tone and accent in Kizigua-a Bantu language. Paper presented at the Cortona Workshop in Phonology.
- Kiparsky, P. 1982. Lexical morphology and phonology. In Yang, I.S. (ed.) *Linguistics in the Morning Calm.* Seoul: Hanskin.

——. 1984. On the lexical phonology of Icelandic. In Elert, C.C. et al. (eds), Nordic Prosody III. University of Umea.

- ——. 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2: 85-138.
- Larson, R. 1988. Light Predicate Raising. Cambridge, MA: MIT Press.

- Lefebvre, C. 1991. Serial Verbs: Grammatical, Comparative and Cognitive Approaches. Amsterdam: John Benjamins.
- Lehmann, T. 1989. A Grammar of Modern Tamil. Pondicherry: Pondicherry Institute of Linguistics and Culture.
- Mahajan, A. 1990. The A/A-Bar Distinction and Movement Theory. Unpublished dissertation, MIT. Mohanan, K.P. 1986. *The Theory of Lexical Phonology*. Dordrecht: Reidel.
- Nagarajan, H. 1989. *A Theory of Post-syntactic Phonology*. Doctoral dissertation published 1994, Madras: T.R. Publications.
- ——. 1991. A Grid Based account of External Sandhi. In *Sound Patterns for the Phonetician*. Madras: T.R. Publications.
- Nespor, M. & Vogel, I. 1982. Prosodic domains of external sandhi rules. In Hulst, H. van der & Smith, N. (eds), *The Structure of Phonological Representations (Part I)*. Dordrecht: Foris.
- —. 1986. Prosodic Phonology. Dordrecht:Foris.
- Saito, M. 1985. Some asymmetries in Japanese and their implications. PhD dissertation, MIT.
- Selkirk, E. 1984. *Phonology and Syntax: the Relation between Sound and Structure*. Cambridge: MIT Press.
- —— & Shen, T. 1990. Prosodic Domains in Shanghai Chinese. In Inkelas, S & Zec, D. (eds), *The Phonology-Syntax Connection*. Chicago: University of Chicago Press.
- Srivastav, V. 1991. The Syntax and Semantics of Correlatives. *Natural Language and Linguistic Theory* 9: 637-686.
- Stroik, T. 1990. Adverbs as V-sisters. Linguistic Inquiry 21: 654-661.
- Valois, D. 1991. The Internal Syntax of DP. Unpublished dissertation, University of California.
- Vijayakrishnan, K.G. 1985. The Interaction of Syntax and External sandhi: evidence from Tamil. *CIEFL Working Papers in Linguistics*. 2:1.
- ——. 1988. The parameters of phonological rule application. *CIEFL Occasional Papers in Linguistics* 5.1.
- Vogel, I. & Kenesei, I. 1987. The interface between phonology and other components of grammar: the case of Hungarian. *Phonology* 4: 243-263.