

**Dravidian Influence on Indo-Aryan: The Case
of the Dative-Subject Construction**

Undergraduate Thesis

Presented in Partial Fulfillment of the Requirements for Graduation “with Honors
Research Distinction in Linguistics” in the Undergraduate Colleges of
The Ohio State University

by
Daven Hobbs

The Ohio State University
May 2016

Project Advisors: Professor Brian Joseph and Professor Donald Winford,
Department of Linguistics

ABSTRACT

Historical language contact between the Indo-Aryan and Dravidian languages of South Asia has been written about extensively. Previous literature, however, has focused largely on possible Dravidian contact-induced features in Old Indo-Aryan (namely Vedic and Classical Sanskrit). Much less work has been done to sort out the extent of Dravidian influence on Indo-Aryan at later stages in its development. In many cases, as well, the previous literature has not adopted a sufficiently principled approach to the identification and investigation of these putative Dravidian contact-induced features. This thesis addresses these shortcomings in the previous literature by first developing a general theoretical framework for the holistic study of historical contact-induced features in language, and then by beginning to apply this framework to the case study of the origin and spread of the dative-subject construction in Indo-Aryan, an areal feature of South Asia which I argue began emerging in early New Indo-Aryan due, in part, to contact with Dravidian.

1. INTRODUCTION

This thesis investigates the origin and spread of the dative-subject construction within the Indo-Aryan languages of South Asia. That this construction is present in many, if not most, of the Indo-Aryan languages as well as those of the genealogically unrelated but geographically contiguous Dravidian language family raises questions as to its status as a contact-induced feature. This thesis develops an account of this construction, arguing that its presence in Indo-Aryan is, in part, a product of contact with Dravidian.

In this section, I present the relevant background for this study. I introduce South Asia and the languages that characterize this region, and provide a brief sketch of the region's prehistory. I then review some of the previous literature pertaining to historical Indo-Aryan/Dravidian language contact and the dative-subject construction, identify the shortcomings in this literature, and conclude with the point of departure for this thesis.

1.1 South Asia and its (linguistic) prehistory

South Asia is a vast geographic region comprising several present day nations—most notably India, Pakistan, Bangladesh, Nepal, and Sri Lanka—which are often considered collectively on the basis of the shared history, and genetic, cultural, and linguistic heritage of their inhabitants. South Asia is one of the most linguistically diverse regions of the world, though the language families to which its many languages belong are relatively few in number. The four major language families of this region are Indo-Aryan, Dravidian, Munda, and Tibeto-Burman.

Indo-Aryan and Dravidian are by far the largest of these four, both in terms of number of speakers as well as in terms of their spatial distribution across South Asia. The Indo-Aryan languages occupy roughly the northern two thirds of the subcontinent. They are members of the larger Indo-European language family, and include languages like Hindi-Urdu, Bengali, and Marathi. The Dravidian languages are concentrated in the roughly one-third southern tip of the subcontinent, and include languages like Tamil, Malayalam, Kannada, and Telugu. Despite many attempts to relate the Dravidian languages to other known languages or language families, no single proposal has amassed a sufficiently strong body of evidence in support of its claims.

The Munda languages are members of the Austroasiatic family. They are therefore related to languages like Vietnamese and Cambodian. The Tibeto-Burman languages are members of the Sino-Tibetan family, and are thus related to the Chinese varieties, such as Mandarin and Cantonese. The Munda and Tibeto-Burman languages of South Asia are mainly spoken in the northeastern regions of the subcontinent.

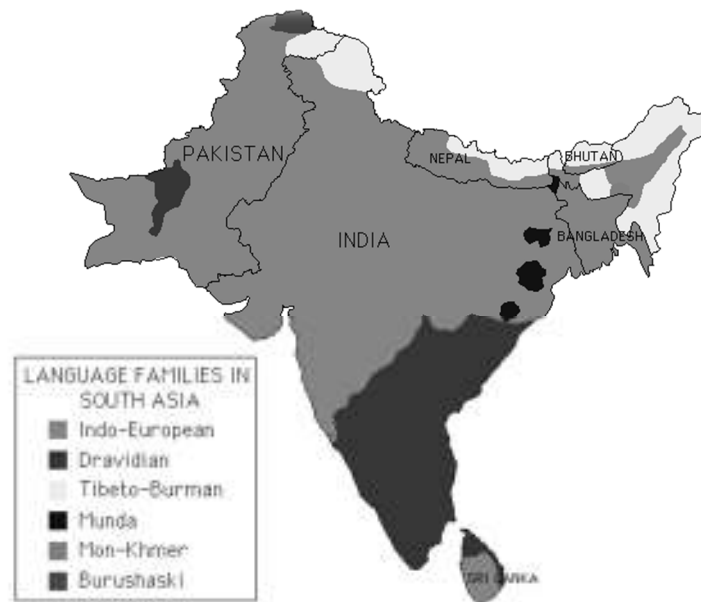


Figure 1: Distribution of major language families across South Asia

(from <http://linguistics.buffalo.edu/people/faculty/dryer/dryer/map.sasia.gif>)

The Dravidian, Munda, and Tibeto-Burman language families are assumed by some scholars to have been present in South Asia prior to the arrival of Indo-Aryan speakers (the so called “Aryans”) into the subcontinent around 2000 BCE (Cardona and Jain 2003: 33)¹. The Aryans, having left the region that now comprises Iran and its neighboring countries, likely migrated into the subcontinent gradually, in a series of waves. These waves of migration are thought to have ended by 1500 BCE (Sjoberg 1992). Theories of a militant Aryan invasion of South Asia, once very popular, are now no longer generally accepted. As the Aryans spread further across the northern regions of the subcontinent, they would have presumably encountered

¹ The view of South Asian prehistory set forth here, including the dates, is more or less speculative. It reflects the generally accepted view held by scholars of this area. As it is not the primary subject matter of this paper and is merely meant to serve as background, it is given without proper justification.

speakers of the unfamiliar Dravidian, Munda, and perhaps even Tibeto-Burman languages. The social situation that developed thereafter between the Aryans and South Asia's indigenous populations is the subject of much debate among archaeologists, literarians, and historical linguists.

The textual history of South Asia begins with the Vedas, a collection of religious hymns composed in an early form of the Indo-Aryan language, dubbed Vedic Sanskrit. Although the earliest extant Vedic manuscripts date back only to the middle ages, it is clear from linguistic analyses of Vedic Sanskrit that the hymns must have been composed at a much earlier period and transmitted orally from teacher to student as part of religious education for centuries before first being committed to writing. As the Vedas are the earliest compositions we have from South Asia, they provide valuable information about the people who composed them, the Aryans, and about the time period during which they lived. Such information includes their social organization, their cultural practices, and, of course, their language. The Rigveda is the oldest hymn in the Vedic canon; a plausible window for its composition is from 1500 to 1250 BCE (Cardona and Jain 2003: 32, Sjoberg 1992: 513). The language of the Rigveda is thus the oldest directly attested form of Indo-Aryan available to linguists for study.

1.2 Indo-Aryan/Dravidian language contact

The discovery of seemingly non-Indo-European elements in Rigvedic Sanskrit that incidentally closely resembled elements found within the Dravidian languages led some scholars (Emeneau 1954, Kuiper 1967, among others) to posit a contact scenario in which Dravidian languages were the source of these convergent features. Frequently cited examples of such features include the presence of a phonemic contrast between dental and retroflex stops and nasals, similarities in the usage of certain syntactic structures like quotatives and enclitic particles, the presence of "conjunctive" gerunds, etc.

Under the view espoused by Emeneau and others, these contact-induced features must have become conventionalized in the Old Indo-Aryan language (i.e. pre-Rigvedic Sanskrit; hereafter OIA) prior to the composition of the Rigveda. In other words, these scholars believed that the stark parallels in the phonology and syntax of Rigvedic Sanskrit and the Dravidian languages came about by means of a Dravidian *substratum influence* on OIA. This would

imply that the social situation in pre-Rigvedic South Asia must have been such that Dravidian speakers were shifting in large numbers to speaking the OIA language, and in doing so, imposing some of their native language (Dravidian) structures onto OIA as they learned it. These impositions then must have eventually become conventionalized as stable features within OIA as a whole.

This sort of contact scenario is common in multilingual settings where speakers of stigmatized languages feel pressure to adopt the language of the more prestigious social group. The substratum theory accorded with what scholars thought they knew about the social dynamic that existed between the Aryans and South Asia's indigenous populations from analyses of the Vedas, namely that the Aryans were the more prestigious, powerful, and dominant social group, and that they subjugated indigenous populations like the Dravidians. Backed by extralinguistic textual evidence of this social dynamic, the substratum theory has gained considerable support, though it is not universally accepted. Hock (1996: 22-23), for example, argues that textual evidence better supports a view of "approximate social equality" between the Aryans and Dravidians in pre-Rigvedic South Asia, and attempts to demonstrate how many of the proposed Dravidian contact-induced features in OIA can be explained through purely internal developments.

This debate over the extent of Dravidian influence on OIA has occupied the attention of works on Indo-Aryan/Dravidian (hereafter IA/Dra.) language contact, and the vast majority of work done in this area focuses solely or primarily on data from the OIA stage. While there has been some recognition that evidence for Dravidian contact-induced features is stronger at later stages in Indo-Aryan's development, there have been no attempts as yet to pursue this avenue of research in any depth.

Sjoberg (1992), for example, in her survey of the literature on IA/Dra. language contact, writes, "Although there seems to have been no systematic effort to investigate the impact of Dravidian on the Prakrits [i.e. Middle Indo-Aryan (MIA); ~200 BCE to 1100 CE], that effect was substantial." She then dedicates a few paragraphs to addressing what additional evidence there might be for Dravidian contact-induced features in MIA, focusing on features like the simplification of OIA phonology in ways which apparently accorded more with

Dravidian phonological patterns, the simplification of OIA's inflectional morphology, and the increased usage in MIA of "participial constructions," which she claims are also used extensively in Dravidian. On the whole, her arguments are not convincing, not because they are implausible claims, but because of her very surface-level treatment of the subject matter. The actual data she cites is sparse, and her analyses are too superficial to make a solid case for Dravidian contact being involved in the development of these features in MIA.

The influence of Dravidian on the modern (or new) Indo-Aryan languages (NIA; ~1100 CE to present) has also been noted by some, but has not yet been the subject of any detailed study. Most notably, Southworth (1971, 2005) and Klaiman (1997) have proposed several features within NIA languages that are of a probable Dravidian origin. The features they cite include postpositions derived from verbs meaning 'leave,' and 'stay/be,' complementizers derived from a form of the verb 'say,' and certain kinds of serial verb constructions, among others. Though these works certainly have made progress toward further understanding the extent of Dravidian influence on the NIA languages, they are nonetheless inadequate in several respects. These studies aim at breadth rather than depth when it comes to identifying and analyzing possible convergent features. This approach compromises the detail necessary for presenting a compelling argument. The actual data they provide to exemplify the features in question is minimal, and often consists of just one or two examples of each feature in one Indo-Aryan language and one Dravidian language so as to demonstrate their supposed parallels. For this reason, none of the features they mention can be considered to have been firmly established as having a Dravidian origin.

Moreover, these studies have done very little to uncover when these proposed contact-induced features first started appearing in the Indo-Aryan languages, who the agents of the transfer were, and under what kinds of social circumstances the transfer took place. This kind of information is necessary for developing a more complete picture of IA/Dra. language contact. Such a picture is best achieved through the examination of each putative contact feature individually in order to trace its unique history. This thesis is a first step towards achieving this goal—I focus on just one feature in Indo-Aryan, and undertake a detailed and principled investigation of its emergence as a product of contact with Dravidian.

1.3 The dative-subject construction

The particular construction under investigation in this thesis is characterized by expressions involving a dative-marked verbal argument which exhibits at least some of the semantic properties and syntactic behaviors typically ascribed to the grammatical relation of ‘subject.’ At its highest level of schematicity (in the terminology of Croft 2003, Barðdal, Kristoffersen, & Sveen 2011, Barðdal & Gildea 2015), there are two basic forms that this construction can take within the Indo-Aryan and Dravidian languages. These two possibilities are given below in (1a-b).

- (1) a. [NP-DAT ... V_{FULL}(+AUX)]
b. [NP-DAT ... NP-NOM V_{LIGHT}(+AUX)]

This construction is made up of a dative-marked noun phrase (that is, the logical subject) and its predicate. The predicate can be either simple (1a) or complex (1b). Simple predicates are made up minimally of a (content-)full verb, that is, a verb which has an easily definable meaning of its own, as in (2). Optional additional arguments and adjunct phrases may also be expressed in between the dative-marked noun phrase and the verb, as indicated by the ellipsis in (1a), and by the argument in parentheses in (2).

- (2) **unakku (avane) teriyumā²** [Tamil]
you.DAT he.ACC know.3SG.Q
‘Do you know (him)?’

Complex predicates, on the other hand, involve semantically light verbs, like ‘be,’ ‘happen,’ ‘come,’ etc. which must minimally occur with a nominative-marked noun phrase which supplies the descriptive content to the predicate, as in (3). Again, additional arguments and adjunct phrases may optionally occur in complex predicates as well, as indicated by the ellipsis in (1b), and by the postpositional phrase in parentheses in (3).

- (3) **mujhe (sāmpoñ se) ñar hai** [Hindi-Urdu]
I.DAT snake.PL.INS fear.SG.NOM be.3SG
‘I am afraid (of snakes).’

² All languages cited are transliterated into a slightly modified version of the International Alphabet of Sanskrit Transliteration (IAST). Some modifications to this system have been made in order to accommodate sounds of other languages not found in Sanskrit.

Although I have provided an example of a simple predicate in Tamil (a Dravidian language) and an example of a complex predicate in Hindi-Urdu (an Indo-Aryan language), it is important to note that dative-subject constructions in Indo-Aryan and Dravidian occur in *both* of these schematic forms. The Indo-Aryan languages contain dative-subject constructions with simple and complex predicates, as do the Dravidian languages.

The types of meanings that are expressed by way of this construction can vary somewhat across the languages that employ it, but crosslinguistically its prototypical use seems to be to express unintentionality and lack of agentivity or control. Masica (1976: 160) writes that this construction is most often used within Indo-Aryan and Dravidian to express “liking and disliking, states of health or sickness, happiness and unhappiness, dreaming, feeling, remembering, thinking, embarrassment, pity, doubt, pain, thirst, hunger, sleepiness, anger, urgency, and ‘knowing’ itself.” Though not exhaustive, this list serves to highlight some of the common predicate types found in the languages of these two families.

The dative-subject construction has been recognized as an areal feature of South Asia (Masica 1976), but little work has been done examining its historical development within the languages of the area. There is some recent work which has looked into its emergence in Indo-Aryan (e.g. Butt & Deo 2013, Montaut 2013, Verbeke et al. 2015), but these studies have not dealt with the contact dimension to its emergence at all, beyond suggesting that contact may have been a factor in facilitating and/or reinforcing an internal development that was already in progress. No study has taken as its primary subject matter the emergence of this construction in Indo-Aryan as product of language contact. Although the construction is mentioned in Southworth (1974, 2005) and Klaiman (1997) as possibly being a Dravidian contact-induced feature in Indo-Aryan, the arguments they put forth for all of the features they mention, as noted previously, are not developed fully enough to be convincing.

1.4 Motivations for this thesis

This thesis is a modest attempt to begin addressing the gaps in the literature identified above. To summarize, previous literature on Indo-Aryan/Dravidian language contact has focused almost exclusively on the Dravidian influences in Old Indo-Aryan. There have been very few studies which have examined the contact relationship between these language

families at later stages. Moreover, none of the work on IA/Dra. language contact thus far has adopted a sufficiently detailed approach to the identification and examination of Dravidian contact-induced features in Indo-Aryan. Such an approach would involve, among other things, (1) focusing on each putative contact-induced feature individually, (2) adhering to a principled methodology for demonstrating that the feature in question was in fact contact-induced, (3) being explicit about the time frame during which the feature was transferred from the source language(s) to the recipient language(s), (4) identifying who the agents of the transfer were, and (5) providing a synthesis of the major factors, both 'internal' (i.e. psycholinguistic and structural factors) as well as 'external' (i.e. sociolinguistic factors), which contributed to the actuation, propagation, and conventionalization of the transferred feature in question.

The primary goal of this thesis is to propose a theoretical framework for studies of historical language contact which fulfills the 5 considerations listed above, and to begin implementing this proposal, focusing on the origin and spread of the dative-subject construction within the early New Indo-Aryan languages as a case study. Further discussion and elaboration of the proposed theory and methodology is given in Section 2.

2. THEORY AND METHODOLOGY

In this section, I develop a general theoretical framework within which I argue all studies of historical language contact ought to be couched. This framework is a novel synthesis of some previous work done by scholars working within the subdisciplines of historical-comparative and contact linguistics. I will employ this framework to guide my investigation of the data that I have collected on the dative-subject construction. In doing so, I place this particular case study within the context of theorizing about language change and language contact in general.

2.1 A general framework for understanding language change

That all living languages change over time is an undeniable fact. For this reason, if we ever hope to understand how language works and why it works the way that it does, we must necessarily address questions related to how and why languages change. In this way, a theory of language change is integral to our understanding of the nature of language as a

whole. Scientific theories are, of course, constructed based on interpretations of actual data. The empirical basis upon which theories of language change are constructed is comprised of documented language changes that we know to have taken place at some point within the world's languages.

It is useful, when studying individual language changes, to delineate boundaries within which the change in question has taken place. The “life” of a language change can be conceived of as involving three roughly chronological stages: actuation, propagation, and conventionalization. These stages correspond to the emergence of a change in (an) idiolect(s), the spread of that change across idiolects and its implementation in the linguistic structure, and the general acceptance and maintenance of that change by speakers of the language, respectively.

The actuation stage involves the inception of a language change, that is to say, an innovation. A given language change can target any aspect of the linguistic system, be it lexical, phonological, morphosyntactic, semantic-pragmatic, etc. Any such change, however, must come into existence somehow. Croft (2006: 116) describes the actuation of a language change as often being functionally motivated—that is, the innovation often arises due to some functional “need,” such as that of greater ease in the production or processing of language. This idea will be returned to in Section 2.2.1.

Once an innovation is introduced within a speech community, whether it arose in the speech of a single individual or through several individuals contemporaneously, it may, under certain circumstances, spread to the speech of other members of that speech community. This is referred to as its propagation. For this to happen, however, the variant must be socially-valued in some way; other members of the speech community must view this variant as valuable or useful enough to adopt it into their own speech. In this way, the propagation stage in the life of a language change is largely socially motivated (Croft, 2006: 116). The conventionalization of a language change has occurred when it has been adopted into regular language use by some subset of the speech community, or, in some cases, the speech community at large. Further explanation of these two stages within the context of language contact is provided in Section 2.2.2.

Much of the work within historical-comparative and contact linguistics centers around understanding different aspects of these three stages in the life of a language change. To fully account for a change is to provide sufficiently detailed explanations for questions related to the actuation, propagation, and conventionalization of the change in question. This thesis, too, is an attempt to investigate these stages in the emergence of the dative-subject construction, not just within one language, but within the many languages of the Indo-Aryan language family.

2.2 Contact-induced change within the general framework

Since the dative-subject construction is an areal feature of South Asia (Masica, 1976), and since any satisfactory account of its origin and spread must for this reason address its propagation across multiple languages, it is necessary to further elaborate on the general framework of language change described above to specifically address instances of contact-induced change. Contact-induced change refers to the diffusion of linguistic elements from one language to another (for a more precise definition, see Section 2.2.1). Joseph (1983: 180), in his study on infinitive-loss within the languages of the Balkan Sprachbund, spells out this additional dimension to the framework. He summarizes the kinds of considerations that ought to be of concern to those developing a full account of the origin and spread of a language change across a linguistic area (*Sprachbund*) in particular, though his points can be reinterpreted to fit all instances of language contact more generally. He writes, “a discussion of ‘origin’ and ‘spread’ of such a change actually subsumes three interrelated issues: (a) the factors which brought on in a given language the conditions necessary for this development to begin (b) the factors which allowed this change to take hold and spread within a given language to new contexts (c) the factors which led to this change occurring in language after language in the area.”

Joseph’s point (a) corresponds to the actuation stage described above. An account of the actuation of a contact-induced change will, among other things, make reference to the specifics of the contact scenario which allowed the change in question to be transferred from a source language to a recipient language. In situations involving the diffusion of a feature across many languages, each instance of transfer may require a unique explanation, as the

particular circumstances which allowed for the transfer, and therefore the actuation of the change, may be different in each case. Joseph's point (b) corresponds to the propagation and conventionalization stages together, and calls for an examination of the factors involved in the spread and maintenance of the change in (each of) the recipient language(s). This will necessarily involve specifying information regarding the *embedding* of the change in the linguistic structure, as well as the *evaluation* of the change by members of the speech community (in the terminology of Weinreich, Labov, & Herzog 1968). Joseph's point (c) is relevant only to situations of language contact in which there is diffusion across more than two languages. It calls for an investigation of the circumstances—sociocultural, geographic, etc.—which caused the feature in question to diffuse in the particular way that it did.

Joseph's formulation of these three issues (p. 180) and his discussion of them that follows (pp. 180-212) are important for the purposes of this thesis, not only because his focus is specifically on situations of long-standing language contact between multiple languages, like those of the Balkans and South Asia, but also because he draws attention to the fact that in order to arrive at the most holistic account of a feature's origin and spread, one must unite both the major 'internal' and 'external' motivations at play. To argue for either a purely externally-motivated or internally-motivated explanation is not productive, as any change in language is likely to have been brought on by a variety of different factors, internal and external alike.

The idea of multiple causation, Joseph argues, has been overlooked in previous accounts of the origin and spread of innovations across the languages of the Balkans. Such is also the case for studies of IA/Dra. language contact in South Asia. What is ultimately bound to be most correct, and thus what we ultimately must strive toward, are accounts which present a synthesis of both the major internal and external motivations involved in the emergence of new features in languages.

2.2.1 Proving language contact has occurred

When undertaking an investigation into the origin and spread of a putative contact-induced feature of a language, it is first necessary to prove that this feature was, in fact, contact-induced before one can delve into the specifics of the contact scenario which brought it

about. It is thus necessary to develop a principled method for demonstrating that contact-induced change has occurred in retrospect. To this end, the criteria developed in Thomason (2001) are useful. Thomason (2001: 91-95) proposes a set of five diagnostics all but the first of which must be satisfied in order to make a convincing case for the contact origin of a feature within a given language.

First, Thomason argues, one ought to look for multiple instances of transfer between the languages in question. She writes, “even if our interest is focused on a particular grammatical construction, an argument for a contact origin will only be convincing if it is supported by evidence of interference [i.e. transfer] elsewhere in the language’s structure as well.” Isolated cases of transfer look odd since the outcomes of language contact are rarely, if ever, just one transferred word, sound, or grammatical construction.

Second, a source language needs to be identified for the putative contact-induced feature in question. All contact-induced features enter an RL from an SL. If an SL cannot be identified, then there is not a strong basis for claiming that a feature within a language was contact-induced. The proposed SL must also be shown to have been in close contact with the RL, such that transfer from the SL to the RL was possible. The contact between these two languages must also be shown to have occurred during the right time, that is, during the time when the feature in question first starts appearing in the RL.

Third, the feature that the RL shares with the SL must be similar to such an extent that it cannot plausibly be attributed to separate independent developments, and thus coincidence. In demonstrating this, it is necessary to adopt a linguistic approach, comparing the structural properties of the feature in question within both the SL and the RL, and identifying shared similarities between the two. A case for transfer of the feature becomes more convincing if certain shared oddities or idiosyncrasies can be identified in the SL and RL.

Fourth, one must demonstrate that the feature in question *was not* present in the RL before contact with the SL, and fifth, one must demonstrate that the feature in question *was* present in the SL before contact with the RL. Thomason writes that these two diagnostics are possibly the most difficult to satisfy, and it may not even be possible to satisfy them in many cases.

According to Thomason, however, if all five of these diagnostics cannot be satisfied, a strong case for the transfer of a feature from one language into another cannot be made. It is thus important that each one of these diagnostics be addressed in turn, to the extent possible.

2.2.2 The actuation of contact-induced changes

As Joseph notes, an explanation for his point (a), that is to say, an explanation of the factors which led to the emergence of an innovation within a particular language of a linguistic area, will usually identify language contact as being a primary impetus for the innovation. This is because the feature in question may be found in other languages of the area, and a comparison of the feature across these (potentially unrelated) languages might reveal certain points of similarity which make its prevalence unlikely or impossible to explain in terms of separate independent innovations or a shared retention from a common ancestral language. Instead, in such cases an explanation in terms of diffusion from one language to another is more feasible.

Languages can change in dramatic ways via the incorporation of linguistic elements from other languages. Such changes are the result of language contact. According to Weinreich (1958: 1), two or more languages are said to be in contact “if they are used alternatively by the same persons.” Thus, the locus of language contact is the bilingual (or multilingual) individual, where “bilingual” here is used broadly to include even those individuals with only very limited familiarity with a language other than his/her native language. The reasoning behind this is clear: purely monolingual individuals cannot possibly be instigators of contact-induced language change, as they do not have access to any other language with which to initiate transfer.

Language contact is one potential catalyst for the actuation of new features in language. The framework of language contact developed by Van Coetsem (1988) posits two mechanisms of contact-induced change. These mechanisms make reference to the processes involved in the mind of a bilingual individual when transferring material from one of his/her languages to another. In this way, Van Coetsem adopts a psycholinguistic perspective on language contact, focusing on the role that the individual plays as the initiator of contact-induced change, and on the processes involved in the actuation of the change.

Van Coetsem refers to all instances of the exchange of linguistic material across languages as instances of *transfer*. He distinguishes just two mechanisms of transfer, which he terms *borrowing* and *imposition*. The two languages in contact fulfill the role of either the source language (SL), or the recipient language (RL), and the directionality of transfer is, of course, always from the SL to the RL. The notions of *dominance* and *agentivity* are also important within Van Coetsem's framework. Dominance refers to degree of linguistic proficiency. The language in which a bilingual is most proficient is considered his/her dominant language. One's dominant language is most often one's native language (L1), though this need not necessarily be the case. Agentivity refers to the individual(s) that initiate transfer, and whether they are SL-dominant or RL-dominant bilinguals. If SL-dominant speakers initiate transfer (*SL agentivity*), the mechanism of transfer is imposition. If RL-dominant speakers initiate transfer (*RL agentivity*), the mechanism of transfer is borrowing. An example of borrowing would be a dominant Sanskrit speaker using some Dravidian vocabulary items while speaking Sanskrit. An example of imposition would be a dominant Dravidian speaker imposing his/her articulatory habits of retroflexion while speaking Sanskrit.

The types of features that are transferred depend on what Van Coetsem calls the *stability gradient* of language. This idea refers to the fact that certain domains of a language are more stable than others, and are therefore less likely to change. Generally, in borrowing, the less stable domains of the RL, such as its lexicon, are more amenable to change via transfer, while the more stable domains, such as phonology or aspects of grammatical structure, are more resistant to change or replacement. Thus, the borrowing of vocabulary items from an SL into an RL is very common, while the borrowing of SL structure is not. In imposition, it is the SL's more stable domains that are more likely to be preserved even when the bilingual individual is attempting to speak the RL. For this reason, SL structures, like its phonology and syntax, tend to be imposed onto the RL more than vocabulary.

The insights and implications of Van Coetsem's framework for studies of language contact have been increasingly recognized (e.g. Winford 2005, 2007, 2013; Lucas 2012, 2014, etc.). Winford (2007: 23-24) argues that Van Coetsem's framework offers the most promising approach for unifying the disparate work on language contact that has been produced by

linguists of diverse backgrounds—including historical linguistics, creole linguistics, second language acquisition, and code-switching—by sorting out the problems of definition and classification that have resulted from its divided and disjointed history as an object of study within the discipline, and also by clarifying questions related to the processes and principles underlying contact-induced change.

Van Coetsem’s framework is important for the broader framework expounded here because it addresses the actuation of contact-induced changes, specifically from a psycholinguistic perspective. Van Coetsem’s framework, however, does not include a sociolinguistic component. Including such a component would of course be beneficial, as it would allow for a more thorough analysis of the sociocultural context in which contact-induced changes take place. This additional dimension to the general framework developed here will be addressed further in Section 2.2.3.

2.2.3 Social context and the actuation of contact-induced changes

Van Coetsem’s framework, while useful for gaining an understanding of the psycholinguistic and structural aspects of actuation, does not address any sociocultural factors relevant to actuation at all. In order to fully understand why a given feature was introduced into the speech of one or more individuals at a particular time, in a particular place, and by a particular (group of) individual(s), it is not enough to appeal solely to psycholinguistic and structural explanations. The social context of transfer is just as relevant to providing full answers to questions related to actuation.

These sociocultural factors can be quite varied and are always, to a certain extent, situation-specific. Winford (2003: 25) writes that these factors, broadly speaking, include “...the types of community settings, the demographics of the populations in contact, the codes and patterns of social interaction among them, and the ideologies and attitudes that govern their linguistic choices,” as well as “...the degree of bilingualism among the individuals and groups in contact, the history and length of contact, the power relationships between the groups, and so on.” An in-depth discussion of these sociocultural factors will not be undertaken here. It is nonetheless important to recognize their role in contributing to actuation, and to thus incorporate them into accounts of historical contact-induced language changes.

2.2.4 The diffusion of linguistic innovations

Once an innovation has been introduced into a language, by way of contact, for example, it may then propagate from the speech of a single individual (or several individuals, as the case may be) to that of other individuals within the speech community. The spread of an innovation from speaker to speaker is referred to as its propagation, as noted previously. The conventionalization of the change happens incrementally, as it is accepted into regular language use by more and more individuals within the speech community. Joseph's point (b) calls for an account of these stages when undertaking an investigation into the origin and spread of individual language changes. He writes that an account of these stages will involve specifying the factors which led to the change "taking hold" within the language. This, in turn, will involve specifying how the change came to be embedded within the larger grammatical system of the language, how the speakers of the language evaluated the change, and why they evaluated it the way that they did. As was discussed with reference to Joseph's point (a), in situations of language contact such as linguistic areas, where changes diffuse across many languages, separate accounts are necessary to describe the factors that were involved in the propagation and conventionalization of the feature in question in each of the languages that adopts it, as these factors are likely to be different in each case.

According to the framework laid out in Croft (2006), the actuation stage of language change introduces an innovation, and an innovation constitutes a new linguistic variant. When introduced into (an) idiolect(s), a variant alternates in speech with already existing functionally equivalent variants, resulting in overlapping distribution. Croft (2006: 98) refers to these kinds of variants, which are completely interchangeable with one another, as instances of first-order variation. Variants which become socially-valued by at least some subset of the speech community are then considered to be instances of second-order variation. "Socially-valued" here means that speakers adopt these variants as a means to index certain aspects of their social identity, such as indicating (sub-)cultural affiliation, and in conveying other kinds of metamessages, like accommodation, formality, etc.

Within this framework, propagation is often a result of the process by which first-order variation becomes second-order variation. For this reason, Croft describes the propagation

stage as socially motivated—a variant will most often take hold and spread among members of the speech community if it is seen as communicatively valuable in some way. The study of how and why certain variants take on social meaning and what kinds of factors contribute to the adoption of new variants occupies much of the work within the field of sociolinguistics.

Sociolinguistic studies of this type, however, usually examine case studies in real time, where ample data is available for analysis. Historical sociolinguistic studies, on the other hand, must make do with much more limited data. Many details about the individual(s) who actuated the innovation in question, and of the complex web of social networks which allowed that innovation to spread from the speech of one (or several) individual(s) to others' within and across speech communities are unfortunately unrecoverable in many cases.

It may be possible in historical cases, however, to analyze distributions of the feature in question from available texts which showcase the language at different historical stages in its development in order to get a sense for the chronology of the feature's spread in terms of its embedding within the linguistic structure. Under some circumstances, it may also be possible to locate the "epicenter" of the innovation (i.e. where it first arose geographically), through texts and perhaps even get a sense for how it was evaluated by individuals at the time. The kinds of conclusions that can be made about the details of propagation and conventionalization of changes in historical cases, if any, are highly dependent on the data sources available for analysis. The best strategies for analyzing and drawing conclusions from such data sources would be situation-specific to a large extent, and so a general theory for proper inference-making regarding sociolinguistic information on the basis of historical data will not be elaborated upon here.

2.3 Methodology and data sources

The general framework I have just developed will be used as a guide in my investigation into the emergence of the dative-subject construction within Indo-Aryan. Adopting this general framework allows for the interpretation of the new data presented in this study within already existing theories which help make sense of this data. I do not intend to address every aspect of this framework here. Such a task is far outside the scope of this study, and I leave it for future research. Presenting the framework in full, however, even if I do not plan address

every aspect of it, does serve to demonstrate how those aspects of the framework that I do address fit into the larger picture. It also makes it clear what still remains to be done.

For this thesis, I have collected examples of the dative-subject construction primarily from four Indo-Aryan languages (Hindi-Urdu, Marathi, Gujarati, and Bengali) and four Dravidian languages (Tamil, Malayalam, Kannada, and Telugu). These data have been collected from a variety of sources, including reference grammars and pedagogical materials, as well as from native speaker informants. I include very little actual historical data in this study, due to my limited access to it at this time. For this reason, I rely almost exclusively on examples of the construction from the modern languages in order to develop my arguments.

3. THE SCOPE OF TRANSFER FROM DRAVIDIAN

Having outlined a general framework for the study of historical contact-induced language changes, I now turn to the application of this framework to the case study being investigated in this thesis, that of the origin and spread of the dative-subject construction within Indo-Aryan. In this section, I address the first two of Thomason's (2001) diagnostics for proving that language contact has occurred between two languages (or language families).

3.1 Identifying other instances of transferred features

To prove that a particular feature of a language was contact-induced, it is first useful to begin looking for other features within that language which also might have emerged as a product of contact at roughly the same point in time. If multiple convincing cases of contact-induced features originating from the same SL can be identified in the RL, the case for a contact origin of the particular feature under investigation becomes more plausible, as isolated instances of transfer of just one sound, word, or construction, etc. is fairly unlikely. In the case of IA/Dra. language contact, there is a large body of literature that focuses on possible Dravidian influenced on Indo-Aryan in its earliest documented form, but only those Dravidian contact-induced features which arose contemporaneously with the dative-subject construction are relevant to this thesis, since evidence of contact between these language families that occurred before (or after) the emergence of this construction does not in itself provide support for the contact-induced origin of this construction in particular.

Anticipating a conclusion I will justify further in Section 5.1, I claim that the dative-subject construction emerged in early New Indo-Aryan (~1100 CE). In order to satisfy Thomason's first diagnostic, then, additional instances of Dravidian contact-induced features which emerged in Indo-Aryan during this same general time period must be identified. As noted previously, there has already been some work dedicated to this enterprise (Southworth 1974, 2005; Klaiman 1997). Several of the features these scholars have examined will be mentioned here. To be sure, each of these features deserves its own detailed account, but for the purposes of this paper, they will serve as additional possible instances of Dravidian influence on early New Indo-Aryan, just insofar as to support the idea that transfer between the languages of these families at this time period was very possible. And, of course, the more similar features that are identified, the less likely it is that these similarities are due to coincidence.

First, with regard to sound similarities, Southworth (1974: 261-262) notes that there are several aspects of the New Indo-Aryan sound system which suggest contact with Dravidian in its early stages. Beyond the well-known case of Dravidian contact-induced retroflexion in OIA, Marathi and Oriya have developed sets of dental and palatal affricates (e.g. [ts] and [dz]) which are found in the neighboring Dravidian languages Kannada and Telugu. OIA had only palato-alveolar affricates (e.g. [tʃ] and [dʒ]). Additionally, Southworth (1974), citing Bloch (1919: 33), notes the tendency in less prestigious dialects of Marathi to diphthongize word initial [e] and [o] to [je] and [wo] respectively, which is common in many dialects of the Dravidian languages (c.f. Tamil **enna** 'what' is commonly pronounced [jennə]). Another relevant sound modification is the loss of aspiration in word-medial and -final aspirates in dialects of both Old (Master 1964: 15) and Modern (Apte 1962: 9) Marathi. Dravidian languages do not traditionally have aspirates in their phoneme inventories, and so it is quite possible that the loss of aspiration in Marathi is a result of Dravidian influence.

Within the realm of morphosyntax, there are even more glaring similarities which suggest that transfer between these two families occurred around the time the New Indo-Aryan languages were beginning to emerge. Several of these features were mentioned previously (Section 1.2), but will be repeated here with examples to illustrate.

One of the morphosyntactic similarities mentioned above was that there are postpositions in several Indo-Aryan and Dravidian languages that are derived from verbs meaning ‘leave,’ and ‘stay/be.’ For example, Hindi-Urdu has **choḍke**, the conjunctive gerund form of the verb root **choḍ-** meaning ‘leave.’ This form of the verb can serve as a postposition after a nominal marked in the dative case, where it has the meaning of ‘excluding,’ e.g. **isko choḍke** ‘excluding this.’ Marathi similarly has **soḍun** meaning ‘apart from/except for’ (lit. ‘having left’), and Bengali has **chāṛā**, meaning the same. As a means of comparison, Tamil has the postposition **viṭa** ‘excluding,’ which is also a gerund form of a verb meaning ‘leave’ (Southworth 2005: 173, 189). Another such example includes postpositions meaning ‘from’ which are derived from verbs meaning ‘stay/be.’ Marathi has **houn** (lit. ‘having been’) from the verb root **ho-** ‘be.’ Bengali similarly has **theke** from the verb root **thāk-** ‘stay/be,’ and Hindi-Urdu has **hoke** meaning ‘via,’ from the root **ho-** ‘be.’ Tamil has **-il iruntu** from the root **iru-** ‘stay/be’ (Southworth 2005: 172). These are clearly early New Indo-Aryan innovations, as the development of postpositions only happened after the case systems and “preverbs” of OIA and MIA started to become obsolete.

Another parallel feature found in several Indo-Aryan and Dravidian languages is that of a complementizer derived from a verb meaning ‘say.’ Note the examples in (4a-c), taken from Southworth (2005: 172).

- (4) a. **āpni dhākāy jācchen bole āmi suneche** [Bengali]
 you Dhaka.LOC go COMP I have-heard
‘I have heard that you are going to Dhaka.’
- b. **saraḷ jāun ujkade valā mhaṇun tyāni sāngitle** [Marathi]
 straight having-gone right-side turn COMP him-by said
‘He said, ‘having gone straight, turn right.’
- c. **pāl illai enru avan conṇān** [Tamil]
 milk is-not COMP he said
‘He said, ‘there’s no milk.’

Another well-known possible example of morphosyntactic transfer between early New Indo-Aryan and Dravidian, which is mentioned in Southworth (1971), Masica (1976), etc., is that of certain kinds of serial verb constructions. These constructions are characterized by verbal

sequences which consist of a main verb, that is a verb which supplies the main semantic content, followed by another verb which adds an additional aspectual or adverbial meaning to the meaning of the main verb. The verbs that can be used in this way are a small and closed class, and also generally have concrete meanings of their own outside of these constructions. The verbs' concrete meanings include 'take,' 'give,' 'go,' 'put/place,' etc. Compare (6a-b) and (7a-b), from Southworth (1971: 26) and (1974: 208) respectively, and (8a-b):

- (5) a. **karun ghe** [Marathi]
do take
'do (something to [or for] oneself)'
- b. **vāngi ko** [Tamil]
buy take
'buy (for yourself/according to your wish)'
- (6) a. **kah rakho** [Hindi-Urdu]
tell put
'tell (once and for all)'
- b. **colli veccu** [Tamil]
tell put
'tell (once and for all)'
- (7) a. **ī kāryam eḷuti tannāl mati** [Malayalam; from M. & M. (1967: 306)]
this matter write give.COND is-enough
'Please write this for me.'
- b. **joi de** [Gujarati; from Cardona & Suthar (2003: 688)]
look give
'take a look (for someone else)'

There are many other parallels between these constructions in Indo-Aryan and Dravidian to note other than what is given above, but because they are not the main focus of this thesis, I restrict my attention on them to just these three examples. These constructions are not found in either OIA or MIA, and are therefore NIA innovations.

Other morphosyntactic similarities include the lack of a verb for 'have,' with an alternative strategy of expressing possession by literally saying that the possessed is 'near' the possessor; the lack of grammatical gender and presence of numeral classifiers in some

southeastern Indo-Aryan and Dravidian languages; the presence in some Indo-Aryan languages, like Marathi and Konkani, and the Dravidian languages, of “negative verbs” like ‘not be,’ and ‘not need/want;’ the presence of differential object marking, such that direct objects in many of these languages are only ever marked overtly with accusative morphology if they are definite; etc. These features found in the New Indo-Aryan languages are also all absent from Old and Middle Indo-Aryan.

There are identifiable semantic similarities between these languages as well. Southworth (1971: 266-267) presents several words in Marathi and Tamil which express similar cultural concepts with regard to clothing, kinship terminology, etc. that “have no early Indo-Aryan parallels.” Some examples include distinct verbs to describe the act of putting on or wearing a traditional unstitched one-piece garment like a *sari* or *dhoti*, and the act of putting on or wearing any other type of clothing, distinct words for cooked and uncooked rice, and distinct words for older and younger male and female siblings. Examples of lexical distinctions based on cultural concepts such as these are clearly indicative of some sort of cultural contact between Indo-Aryan- and Dravidian-speaking peoples, if not also bilingualism on either or both of their parts.

I have focused only on a small number of possible convergent features in New Indo-Aryan and Dravidian in order to demonstrate that transfer between these two language families was very likely to have taken place around 1100 CE, when the New Indo-Aryan languages were first beginning to take shape. For further discussion of features shared between these families, refer to the works cited within this section. Though these works themselves are not entirely satisfying in the arguments they make for the contact-induced origin of the features they mention, they do focus on a larger number of features than will be listed here.

3.2 Evidence for Dravidian as the source language family

Thomason’s second diagnostic requires a source language to be identified for the particular feature under investigation. In the preceding section, I outlined some similarities between Indo-Aryan and Dravidian languages that started appearing in Indo-Aryan around the same time as that of the dative-subject construction. Although this concise enumeration of similar

features suggests that transfer between these two families was taking place at the proper time, it still remains to be seen what the directionality of that transfer might have been.

I propose that the source language family for the dative-subject construction, and for most, if not all, of the features mentioned in the preceding section for that matter, is Dravidian. Not only is there numerous convincing evidence of Dravidian contact-induced features entering Indo-Aryan at earlier stages, but several of the features noted above, possibly including the dative-subject construction (see Section 5.2), can plausibly be reconstructed for Proto-Dravidian. It is also relevant to note that these innovations in early New Indo-Aryan were concurrent with an overall major restructuring of the Indo-Aryan grammatical system from more synthetic and inflectional to more analytic. The Dravidian languages, on the other hand, have changed relatively little structurally over the course of their written history (Sjoberg 1992). They preserve the case system and the robust agglutinating structure of their shared ancestral language, Proto-Dravidian. This provides further support for the idea that Indo-Aryan was the recipient language family for many Dravidian-origin features.

4. COMPARISON OF THE CONSTRUCTION

It is necessary, when making an argument for the contact-induced origins of some feature within a language, to conduct a detailed comparison of the RL feature and the analogous feature within the SL, so as to demonstrate their close similarities. This is Thomason's third diagnostic—the RL feature and the SL feature must be shown to parallel each other in such a way as to make it clear that the similarities between them could not possibly be due to coincidence. To this end, I present the similarities between the dative-subject construction as it appears within Indo-Aryan and Dravidian in this section.

The first similarity to note is simply the forms that this construction takes within the Indo-Aryan and Dravidian languages. The languages of both families contain dative-subject constructions with both simple and complex predicates, as indicated in (1a-b). Moreover, the kinds of light verbs that can occur in complex predicates in these languages is also identical to a large extent. They both include the verbs 'be,' 'become,' 'come,' 'feel,' 'strike,' sometimes 'fall,' etc.' Another similarity of form is in the word order—the dative subjects' unmarked position is the same as if it were a canonical subject in these SOV languages.

The types of meanings expressed by the dative-subject construction across the Indo-Aryan and Dravidian languages, as mentioned previously, all tend to fall within a restricted set of semantic classes. I will adopt the classification of dative-subject-taking predicates provided in Barðdal (2004). According to this classification, predicates can be grouped broadly into either experience-based or happenstance predicates, within which subclasses of predicates can also be identified.

4.1 Experience-based predicates

Perhaps the most common subclass of experience-based predicates in Indo-Aryan and Dravidian are those that express emotions or idiosyncratic attitudes, such as happiness, sadness, anger, surprise, amazement, fear, embarrassment, regret, worry, uneasiness, urgency, doubt, pity, liking, jealousy, irritation, etc. Two examples of this subclass of predicates is given in (8) and (9) below:

(8) **sudhālā kālji āhe** [Marathi; from Pandharipande (1990: 162)]
 Sudha.DAT worry is
'Sudha is worried.'

(9) **enikku kopam vantatu** [Tamil]
 I.DAT anger came
'I became angry.'

Another subclass includes predicates which express some meaning related to cognition, such as intelligence, memory, forgetfulness, knowledge, opinion, belief, understanding, etc. See examples (10) through (12).

(10) **malā tyāci aṭhvaṇ yete** [Marathi; from Pandharipande (1990: 167)]
 I.DAT his memory comes
'I remember him.'

(11) **unakku puriyumā** [Tamil]
 you.DAT understand.Q
'Do you understand?'

(12) **āmār ei kathā bissās hay nā** [Bengali; from Klaiman (1980: 278)]
 I.GEN this matter belief becomes not
'I don't believe this.'

Another relatively small subclass involves predicates which describe aspects of perception, like seeing, hearing, or smelling, etc. Some predicates of this subclass are illustrated in examples (13) through (15).

(13) **kyā āpko dikhāi nahi detā** [Hindi-Urdu]
 Q.WORD you.DAT sight not give.PRES
'Do you not see (it)?'

(14) **malā vimān dislā** [Marathi; from Pandharipande (1990: 163)]
 I.DAT airplane saw
'I saw an airplane.'

(15) **vādiki kaḷḷu sariggā kanipincavu** [Telugu; from S. & Bh. (2004: 168)]
 he.DAT eyes properly not-visible
'He cannot see properly.'

The final subclass to be described here involves those predicates which describe bodily states, such as heat, cold, health, illness, pain, hunger, thirst, tiredness/sleepiness, itching, physical irritation, etc. See examples (16) through (19).

(16) **tār asukh holo** [Bengali; from Klaiman (1980: 281)]
 he.GEN illness became
'He became unwell.'

(17) **pratimaki jalubu cesindi** [Telugu; from S. & Bh. (2004: 167)]
 Pratima.DAT cold did
'Pratima got a cold.'

(18) **mujhe nīnd ā rahi hai** [Hindi-Urdu]
 I.DAT sleep is-coming
'I am feeling sleepy.'

(19) **nāku nidra vastundi** [Telugu]
 I.DAT sleep comes
'I am feeling sleepy.'

4.2 Happenstance predicates

The second broad semantic class of predicate are those that can be crudely subsumed under to label of “happenstance.” As was the case with experience-based predicates, subclasses of happenstance predicates can also be identified. For example, there are predicate which

express gain—for example, of physical objects, relief, etc.—which are given in examples (20) and (21).

(20) **mane paisā male che** [Gujarati; from Dwyer (1995: 88)]
I.DAT money gets
'I earn money.'

(21) **rājakumārikki māle keṭeccatu** [Tamil; from Schiffman (1999 :105)]
princess.DAT necklace got
'The princess got a necklace.'

Another subclass of predicates is that which describes personal properties, like physical characteristics, personality, abilities, etc. Some examples are given in (22) and (23).

(22) **usko hindī bolnī ātī hai** [Hindi-Urdu]
(s)he.DAT Hindi speak comes
'(S)he knows how to speak Hindi.'

(23) **mādhuriki roṭṭelu paḍavu** [Telugu; S. & Bh. (2004: 167)]
Madhuri.DAT Indian-bread does-not-suit
'Indian bread does not suit Madhuri.'

The final subclass of happenstance predicates to be mentioned here is that which expresses obligation (both internal and external), necessity, wanting, 'should,' etc. Some examples are given in (24) through (27).

(24) **mujhe cay cahiye** [Hindi-Urdu]
I.DAT tea want/need
'I want/need tea.'

(25) **avrige hana beku** [Kannada; from Schiffman (1985: 79)]
he.DAT money wants/needs
'He wants/needs money.'

(26) **enikku viṭṭil pogaṇam** [Malayalam]
I.DAT home.LOC go-want
'I want to go home.'

(27) **malā gharī zāylā pahije** [Marathi; from Pandharipande (1990: 163)]
I.DAT home go should
'I should go home.'

4.3 Additional similarities

In addition to the similarities in form and meaning of the dative-subject construction in Indo-Aryan and Dravidian noted above, there are other facts to take into consideration when assessing the extent of their parallelisms. For example, across the languages surveyed, there are certain patterns in the light verbs that are used to express particular meanings. So, for instance, to express things like remembering (10), sleepiness (18 and 19), anger (9), ability (22), smells, among other experiences/states, all Indo-Aryan and Dravidian languages use the light verb ‘come.’ However, experiences like hunger, happiness, sadness, etc. generally do not “come to” the experiencer in these languages, nor do physical objects “come to” their possessors. To express this, there is inevitably a verb ‘get/obtain/find’ which is used (which also takes a dative subject—see examples 20 and 21).

An additional consideration, which bears further looking into, is that the similarities in this construction, and in other putative contact-induced features, that Indo-Aryan shares with Dravidian seemingly become more pronounced in the Indo-Aryan languages at the IA/Dra. border. Marathi and Konkani especially exhibit many features in common with Dravidian not shared with other, more geographically distant Indo-Aryan languages. For example, both of these languages contain many more content-full verbs which take dative-subjects in comparison to other Indo-Aryan languages like Hindi-Urdu, Gujarati, and Nepali, in which the complex predicate type of dative-subject construction is far more common, and the simple predicate type is restricted to a very small set of verbs. Dravidian languages, too, have many more of the simple predicate type, comparatively.

Also, in Konkani, unlike in other Indo-Aryan languages, there is just one main word **zāi**, meaning ‘is wanted/needed,’ ‘should,’ etc. which expresses obligation, necessity, and desire, as is also the case in the Dravidian languages. It is used alone as a (defective) verb along with a noun to express the wanting or needing of that noun (see 28a). Alternatively, it can be suffixed to the root of a verb to express the wanting or needing to perform or experience the action of state described by the verb (see 28b). Parallels with the analogous Dravidian constructions are apparent (see 29a-b).

- (28) a. **makā devāci kurpa zāi** [Konkani; from Maffei (1882: 91)]
 I.DAT God.GEN grace need
'I am in need of the grace of God.'
- b. **āuṅvem kiteṃ karizāi** [Konkani; from Maffei (1882: 431)]
 I.INS what do-must
'What must I do?'
- (29) a. **enikku kāppi veṅam** [Malayalam; from M. & M. (1967: 45)]
 I.DAT coffee want
'I want coffee.'
- b. **uṭane pog-aṅam** [Malayalam; from M. & M. (1967: 126)]
 right-away go-must
'I must go right away.'

Also, the negative form of this verb in Konkani is **nākā**, a form not obviously related to **zāi**. This is also the case in the Dravidian languages, where the negative forms of these defective verbs of obligation, volition, etc. are unpredictable (though clearly related) to their “positive” counterparts, e.g. Malayalam has **veṅam** ‘wanted/needed’ and **veṅṭā** ‘not wanted/needed’, Kannada has **beku** and **beḍā** respectively. Negative verbs which are unpredictable in form given the corresponding positive verb are characteristic of Dravidian.

Stronger similarities like this and others are to be expected in the Indo-Aryan languages on the border with Dravidian if one posits a diffusion of features from Dravidian in the south up northward. These features would have first entered the languages on the border, and then from there would have diffused further northward into the Indo-Aryan languages that aren’t directly contiguous with Dravidian.

5. FURTHER EVIDENCE FOR TRANSFER

It is clear from the preceding sections that not only do New Indo-Aryan and Dravidian languages share common features in many different domains, including phonology, morpho-syntax, and semantics, but also that Dravidian is the most likely candidate for the source language family from which these features originate. An detailed examination of the dative-subject construction in particular also reveals stark parallels in both the forms that this construction takes and in the range meanings that it expresses within the languages of these

two families. In this section, I address the last two of Thomason’s diagnostics in relation to the dative-subject construction.

5.1 Absence of feature in Indo-Aryan prior to contact with Dravidian

Thomason’s third diagnostic requires that the feature under investigation must be shown to have been absent in the RL prior to contact with the SL. This is not entirely possible in the case of Indo-Aryan and Dravidian, as contact between these families has been hypothesized by most to have begun even before the composition of the Vedas, the oldest attested form of Indo-Aryan. However, it can be shown that the dative-subject construction as it appears in the modern Indo-Aryan languages is not simply a development from an earlier construction in OIA, but that it is a completely novel construction which only began appearing in the early New Indo-Aryan languages, and thus was not feature of Indo-Aryan before the period of contact with Dravidian which took place as the NIA languages were emerging.

Hock (1990) examines several instances of verbs in Sanskrit which in some cases take oblique subject-like arguments, but ultimately concludes that Sanskrit did not have a dative-subject construction. Barðdal & Eythórsson (2009), on the other hand, believe that it did, and that the construction can moreover be reconstructed for Proto-Indo-European. Regardless of the position one takes, however, it is clear that these verbs which took oblique subject-like objects were few and far between in Sanskrit. Moreover, these verbs are not the same one’s which develop dative subject-marking in later stages. For this reason, it is the general consensus (Butt & Deo 2013, Montaut 2013, Verbeke et al. 2015) that the dative-subject construction in the modern languages is not a continuation of a preexisting construction in earlier Indo-Aryan, but that it developed independently. Verbeke et al. (2015) write that some of the earliest clear traces of this construction in Indo-Aryan come from texts of Old Hindi from around the 16th century (30).

(30) **tākūm bahuri na lāgī piyāsa** [Old Hindi; from Verbeke et al. (2015: 30)]
he.DAT again not be-attached thirst
‘He never got thirst again.’

Klaiman (1980: 282) also writes that at the earliest recorded stage of Bengali (1200-1400 C.E.), dative-subject constructions were prevalent, and the subjects of these and other similar

constructions were not confined to being marked in the dative case, but also other non-direct cases like the locative (cf. genitive-marking in these constructions that is most common in modern Bengali—see examples 12, 16).

(31) **mota baṛa dayā lāge baṛāyi dekhiām** [Middle Bengali; from K. (1980: 282)]
I.LOC big pity affects Granny see
'I feel great pity looking at granny.'

Butt & Deo (2013) further support the idea that the dative-subject construction was an early NIA innovation by showing the development in argument structure and case frame of several verbs from Sanskrit (OIA) into Old Marathi, and subsequently into Modern Marathi. They demonstrate how certain verbs in Sanskrit, like **kal-** 'perceive,' which had an <experiencer, theme> argument structure and a nominative-accusative case frame, preserved the same argument structure while gradually shifting to a dative-nominative case frame into Old and Modern Marathi.

5.2 Presence of feature in Dravidian prior to contact with Indo-Aryan

Thomason's fifth criterion is a bit more difficult to satisfy. It requires that evidence of the feature in question be found in the proposed SL prior to contact with the RL. This cannot be done through textual data in the case of the dative-subject construction, considering that the earliest textual evidence for Dravidian dates back only to 254 BCE (Steever, 1998: 6), which is, on most accounts, quite a while after Dravidian speakers would have first come in contact with Indo-Aryan speakers.

However, if one accepts the view that the dative-subject construction was an NIA innovation, it just has to be shown that the Dravidian languages had this feature before its development in NIA. Sridhar (1979) writes that there is mention of this construction by Kannada grammarians from as far back as 1260 C.E. This would have been around the time that the NIA languages were first starting to emerge. One might infer that the construction was well established in Kannada at this time, as traditional Indian grammarians dealt mainly with prescriptive language use. Earlier evidence of this construction might be found in Old Tamil literature (~300 C.E.), as in (30).

- (30) **nin-akk=o ari.y-un-al** [Old Tamil; from Lehman (1988: 89)]
 you.DAT.INT know.NPST.3SF
'Do you know her?'

This same verb is found in modern Malayalam, where it invariably takes a dative subject.

- (31) **avaḷkku giṭṭār vāykkān ariyām** [from M. & M. (1967: 136)]
 she.DAT guitar play.INF knows
'She knows how to play the guitar.'

It might also be possible to reconstruct the dative-subject construction for Proto-Dravidian based on the fact that several verbs in the modern languages are clear cognates both in their form and in their meaning. Such examples include a verb for 'know' (32), 'feel/seem' (33), and others. Though the examples are suggestive, this line of research deserves further attention before a compelling case can be put forth.

- (32) a. **unakku teriyumā** [Tamil]
 you.DAT know.Q
'Do you know?'

- b. **adu nanage sariyāgi tiḷiyal illa** [Kannada; from Steever (1998: 147)]
 that I.DAT really know be-not
'I didn't really know that.'

- c. **vāru evarō nāku telusu** [Telugu; from Krishnamurti (1998: 235)]
 he who I.DAT know
'I know who he is.'

- (33) a. **enakku paṭi tōnrum** [Tamil; from Fabricius (1972: 226)]
 I.DAT so seems
'It seems so to me.'

- b. **enikku vaḷare hrudeya vēdana tōnni** [Malayalam; from M. & M. (1967: 334)]
 I.DAT much heart pain felt
'I felt a great sense of grief.'

- c. **nāku tōcindi idi** [Telugu; from Gwynn (1991: 250)]
 I.DAT thinks this
'This is what I think.'

These are just two possible verbs which in Proto-Dravidian had at least the option of marking their subject, or subject-like argument, in the dative case. A more thorough investigation of

these as well as other potential verbs like them (such as **kiṭṭu (kiṭṭi-)** (DED etym. 1538), for example) needs to be undertaken, as this would greatly strengthen the case for the Dravidian origin of this construction in Indo-Aryan.

6. SPECIFICS OF THE CONTACT SCENARIO

Having demonstrated that the dative-subject construction within Indo-Aryan was most likely induced through contact with Dravidian, it still remains to be examined what kind of contact scenario brought this change about. With regard to actuation, if Dravidian was indeed the source language family for the dative-subject construction (as well as the other features mentioned in Section 3.1), then Dravidian speakers almost certainly would have been the agents of transfer, judging from the domains of language that imposition generally targets (Winford 2013). This contact scenario, then, reflects imposition, or SL agentivity on the part of Dravidian dominant individuals, in Van Coetsem's terminology. The implications of this for our understanding of the social dynamic between the Indo-Aryan and Dravidian speakers during the time of transfer cannot be discussed concretely in the absence of additional data, and so I leave this to future work. The specifics of the propagation and conventionalization of the dative-subject construction, which, as was mentioned previously, would require detailed textual data from different stages in the development of the NIA languages, cannot be developed here. I leave this to future work as well.

REFERENCES

- Barðdal, Jóhanna (2004). The Semantics of the Impersonal Construction in Icelandic, German, and Faroese: Beyond Thematic Roles. In *Focus on Germanic Typology* [Studia Typologica 6], 105-137. Ed. Werner Abraham. Berlin: Akademie Verlag.
- Barðdal, Jóhanna, and Thórhallur Eythórsson (2009). The Origin of the Oblique Subject Construction: An Indo-European Comparison. In *Grammatical Change in Indo-European Languages*, 179-193. Eds. Vit Bubenik, John Hewson, and Sarah Rose. Amsterdam: John Benjamins.
- Barðdal, Jóhanna, and Spike Gildea (2015). Diachronic construction grammar: Epistemological context, basic assumptions, and historical implications. In Barðdal, Smirnova, Sommerer, and Gildea (eds.) *Diachronic Construction Grammar*. John Benjamins
- Barðdal, Jóhanna, and Kristian Emil Kristoffersen and Andreas Sveen (2011). West Scandinavian Ditransitives as a family of constructions: With a special attention to the Norwegian 'V-REFL-NP' Construction. *Linguistics* 49(1): 53-104.
- Butt, Miriam, and Ashwini Deo (2013). A historical perspective on dative subjects in Indo-Aryan. *Proceedings of LFG13*. CSLI Publications, Online.
- Cardona, George, and Dhanesh Jain (2003). *The Indo-Aryan languages*. Routledge.
- Cardona, George, and Babu Suthar (2003). Gujarati. In Cardona and Jain (eds.) *The Indo-Aryan languages*. Routledge.
- Croft, William (2003). Lexical rules vs. constructions: A false dichotomy. In Hubert Cuyckens, Thomas Berg, Rene Dirven, and Klaus-Uwe Panther (eds.) *Motivation in language: Studies in honour of Gunter Radden*, 49-68. Amsterdam: John Benjamins.
- Croft, William (2006). The relevance of an evolutionary model to historical linguistics. *Different models of linguistic change*, ed. Ole Nedergard Thomsen. Amsterdam: John Benjamins.
- Dwyer (1995). *Teach Yourself Gujarati: A Complete Course for Beginners*. Hodder Headline Plc.
- Fabricius, Johann Philipp. J.P. Fabricius's Tamil and English dictionary. 4th ed., rev. and enl. Tranquebar: Evangelical Lutheran Mission Pub. House.
- Gwynn, John Peter Lucius. A Telugu-English dictionary. Delhi; New York: Oxford University Press.
- Hock, Hans Henrich (1984). (Pre-)Rig-Vedic convergence of Indo-Aryan with Dravidian? Another look at the evidence. *Studies in the Linguistic Sciences* 14: 1. 89-108
- Hock, Hans Henrich (1990). Oblique subjects in Sanskrit. *Experiencer Subjects in South Asian Languages*, eds. Manindra K. Verma and K.P. Mohanan. The Center for the Study of Language and Information, Stanford University.

- Hock, Hans Henrich (1996). Pre-Rig-Vedic convergence between Indo-Aryan (Sanskrit) and Dravidian? A survey of the issues and controversies. *Ideology and status of Sanskrit: Contributions to the history of the Sanskrit language*, ed. J. E. M. Houben. Leiden: Brill.
- Joseph, Brian D. (1983). *The synchrony and diachrony of the Balkan infinitive*. Cambridge University Press.
- Klaiman, M.H. (1977). Bengali syntax: possible Dravidian influences. *IJDL* 6: 303-17.
- Klaiman, M.H. (1980). Bengali dative subjects. *Lingua* 51: 275-95.
- Krishnamurti, Bhadriraju (1998). Telugu. In Steever, ed. *The Dravidian languages*. Routledge.
- Krishnamurti, Bhadriraju (2003). *The Dravidian Languages*. Cambridge University Press.
- Lehmann, Thomas (1998). Old Tamil. In Steever (ed.) *The Dravidian languages*. Routledge.
- Maffei, Angelus Francis Xavier (1882). *A Konkani Grammar*. Mangalore: Basel Mission Book & Tract Depository.
- Masica, Colin P. (1976). *Defining a Linguistic Area*. The University of Chicago Press.
- Moag, Rachel, and Rodney Moag (1967). *A course in colloquial Malayalam*. U.S. Peace Corps.
- Montaut, Annie (2013) The rise of non-canonical subjects and semantic alignments in Hindi. *Studies in Language Companion Series*, 2013, 140 (1), pp.91-117. <halshs-00962420>
- Pandharipande, Rajeshwari (1990). Experiencer (dative) NPs in Marathi. *Experiencer Subjects in South Asian Languages*, eds. Manindra K. Verma and K.P. Mohanan. The Center for the Study of Language and Information, Stanford University.
- Schiffman, Harold F. (1985). *A reference grammar of spoken Kannada*. Seattle: University of Washington Press.
- Schiffman, Harold F. (1999). *A reference grammar of spoken Tamil*. Cambridge: Cambridge University Press.
- Sjoberg, Andree (1992). The Impact of Dravidian on Indo-Aryan: An Overview. *Reconstructing Languages and Cultures*. Trends in Linguistics Studies and Monographs 58. Mouton de Gruyter.
- Southworth, Franklin C. (1971). Detecting prior creolization: An analysis of the historical origins of Marathi. *Pidginization and Creolization of Languages*, ed. Dell Hymes. Cambridge: Cambridge University Press.
- Southworth, Franklin C. (1974). Linguistic stratigraphy of North India. In Southworth and Apte (eds.) *Contact and Convergence in South Asian Languages*. IJDL.
- Southworth, Franklin C. (2005) *Linguistic Archaeology of South Asia*. Routledge Curzon.
- Sridhar, S.N. (1979). Dative subjects and the notion of subject. In *Lingua*, 49, 99-125.
- Steever, Stanford B. (1998). *The Dravidian Languages*. Routledge.

- Subbarao, Karumuri Venkata, and Peri Bhaskararao (2004). Non-nominative subjects in Telugu. In Bhaskararao and Subbarao (eds.) *Non-nominative Subjects*. Vol 2. Amsterdam/ Philadelphia: John Benjamins
- Thomason, Sarah G. (2001). *Language Contact*. Edinburgh: Edinburgh University Press Ltd.
- Van Coetsem, Frans (1988). *Loan Phonology and the Two Transfer Types in Language Contact*. Foris Publications.
- Verbeke, Saartje, Leonid Kulikov, and Klaas Willem (2015). Oblique case-marking in Indo-Aryan experiencer constructions: Historical roots and synchronic variation. In *Lingua*, 163, 23-39.
- Verma, Manindra K. and K.P. Mohanan (eds.) (1990). *Experiencer Subjects in South Asian Languages*. The Center for the Study of Language and Information, Stanford University.
- Winford, Donald (2003). *An Introduction to Contact Linguistics*. Blackwell Publishing Ltd.
- Winford, Donald (2005). Contact-induced Changes: Classification and Processes. In *Diachronica*, Vol. XXII, No. 2. John Benjamins Publishing Company.
- Winford, Donald (2007). Some issues in the study of language contact. In *Journal of Language Contact*, THEMA 1.
- Winford, Donald (2013). On the unity of contact phenomena: The case for imposition. *In and Out of Africa Languages in Question Vol 1: Language Contact and Epistemological Issues*, ed. Carol de Feral.