A sense of agency

Accounting for a change-in-progress in Australian Kriol pronoun distribution

Joshua Phillips
Yale University

Roper Kriol exhibits variation in the shape of the first-person singular pronoun in subject position. This paper provides an account of the numerous syntactic, semantic and pragmatic factors that appear to influence the selection of either *ai* or *mi* based predominantly on a study of a corpus of the written language. It is claimed that the synchronic distribution of *ai* and *mi* is an innovation primarily motivated by speaker reanalysis of the semantic entailments frequently associated with English subject and object arguments – effectively evidence of the partial grammaticalisation of agentivity in these varieties. This work has implications for our understanding of 'agentivity' as a cross-linguistic, cognitive category and for the dynamic relationship between semantic roles and the morphosyntactic encoding of grammatical relations.

Keywords: grammaticalisation, Kriol, Australian Kriol, Pidgins & Creoles, semantics, semantic change, corpus linguistics, contact linguistics, transitivity, agentivity, pronouns, pragmatics' grammatical relations

1. Introduction

Eastern varieties of Australian Kriol *(viz*. Roper Kriol) exhibit variation in the shape of the first-person singular pronoun when it occurs in subject position. This paper provides an account of the numerous syntactic, semantic and pragmatic factors that appear to govern the selection of either *ai* or *mi*, based on a corpus of the written language in addition to structured elicitation. Here, I claim that the synchronic distribution of *ai* and *mi* is an innovation, motivated in part by speaker reanalysis of the semantic entailments frequently associated with English subject and object arguments. This effectively constitutes evidence of the partial grammaticalisation of agentivity in these varieties and the apparent emergence of a morphosyntactic

Studies in Language 42:2 (2018), **329–368**. DOI 10.1075/sl.17025.phi
ISSN 0378-4177 / E-ISSN 1569-9978

alignment that is sensitive to the agentivity of the speaker in first-person predications: broadly, *ai* foregrounds the agentivity of the speaker whereas *mi* backgrounds this same property. An additional implication is the interplay between this apparent grammaticalisation and 'contextual allomorphy' in the pronoun paradigm (see Parrott 2009).

1.1 Background

Borne of westward frontier expansion into Northern Australia by European settlers in the late-eighteenth and early-nineteenth centuries and consequent disruption of traditional lifestyles, communities and language ecologies (see Harris 1986), Australian Kriol (hereafter 'Kriol' *simpliciter)* is an English-lexified creole language spoken by between 30,000 and 50,000 people in Northern Australia. There is a significant amount of dialectal variation between Kriol-speaking communities and between generations of speakers, in large part motivated by sustained contact with both traditional Australian languages and varieties of English (e.g. Sandefur & Harris 1986; Schultze-Berndt et al. 2013; Dickson 2016 a.o.; cf. Schultze-Berndt et al. forthcoming). The data in this paper is predominantly drawn from Kriol speakers living in Ngukurr (southeastern Arnhem Land), a place frequently considered to be the 'birthplace' of the language (Harris 1986; Sandefur 1986).[1]

Striking morphological simplification is a process that is frequently associated with creolisation or radical situations of language contact more broadly (e.g., McWhorter 2007). As a consequence, studying paths of 'complexification' in these languages provides a unique opportunity to observe in-progress grammaticalisation and meaning change, providing potential insight into the cognitive structures which underpin grammatical categories and grammatical relations in the case of the present work.

1.2 The Kriol pronoun paradigm

Existing descriptions of North Australian Kriol have described the distribution of first-person singular (1s) pronouns *ai~mi* as being largely isomorphic to what has traditionally been described as the nominative-accusative/oblique contrast that is expressed in English (Hudson 1983: 44; Schultze-Berndt et al.: 5; Meakins

1. The Ngukurr Kriol substrata are generally thought to include Marra [mec], Wandarang [wnd], Alawa [alh], Mangarayi [mpc], Ngandi [nid], Ngalakan [nig], Rembarrnga [rmb] and Ritharrngu [rit]. Ritharrngu, a Yolŋu variety, is described as the only of these languages which continues to be actively spoken as a first-language in Ngukurr households (Sandefur 1985: 208).

2009: 66, 68).[2] Indeed, for contact varieties for which an *I~me* distinction is drawn, this distinction appears to reflect a more English-like morphosyntactic alignment (e.g. Harris 1986: 327ff). That some distinction is drawn in the first place is striking insofar as this diverges with the rest of the Kriol pronominal paradigm (provided as Table 1),[3] which (outside of the first-person singular and perhaps the third-person plural) has a single invariant form for each person/number combination. Consequently, in Kriol, grammatical relations are marked entirely configurationally (i.e. the position of arguments in the clause determines their function) – a feature that diverges sharply from the 'remarkably free' word orders and fusional case morphology of most traditional Australian Languages (Dixon 1972) as well as from the (defective) vestigial case distinctions drawn in the pronominal inventory of its English lexifier. While it is striking that this component of the Kriol grammar diverges from both of its 'parent' languages, the collapse of the pronominal case distinctions is a process attested across many creole languages, and that is perhaps predictable on the basis of the claim that creoles form a typological class of languages (McWhorter 2005 a.o.)

Table 1. Ngukurr Kriol personal pronoun inventory

	s	**d**	**p**
1.EXCL	*ai~mi*	*minbala*	*melabat*
1.INCL	–	*yunmi*	*wi*
2	*yu*	*yunbala*	*yumob*
3	*im*	*dubala*	*dei/olabat*

In this paper, I provide an account for the synchronic distribution of *ai* and *mi*. I suggest below that this distribution can be described as neither (a) nominative-accusative[4] nor (b) free. Versions of this observation are implicit in the literature, including Sandefur's claim that "*mi* can be used in all subject position and object positions, while *ai* can only be used in subject positions, though not as the subject of an equational sentence" (1979: 83). Further, Nicholls (2011) claims that:

2. Hudson (1983) and Meakins (2009) are based on data from Fitzroy Valley and Kalkaringi, where the distinction described here is supposedly unattested. Schultze-Berdnt et al. (2013) identify *mi* as a subject pronoun in Ngukurr (the region on whose dialect this study is predominantly based), but its distribution is not detailed except for the claim of regional and 'intra-speaker' variation (243).

3. See also Schultze-Berndt et al. (2013) for a more elaborated discussion of variation in Kriol pronominal forms.

4. Or better, English-like, in order to remain agnostic with respect to the true morphosyntactic character of English pronominal case. See § 4.3 for further discussion of this analysis.

> *Mi* is unattested as a subject pronoun followed by the tense *bin*PST or *gada*FUT'/OBLIG […] Both *mi* and *ai* occur before verbs that lack tense or aspect marking; however *mi* is more common […] speakers claim that *mi* is possible in subject position within a tensed clause if it is clearly in focus […] this is unattested in spontaneous data. (78)

Clues for solving the distribution puzzle may be in part provided by the fact that "[n]ominals in almost all Australian languages inflect on [an ergative-absolutive] pattern" and "[e]ach Australian language makes a strict division between transitive and intransitive verbs" (Dixon 1980: 278, 286). Indeed, as is the case for the historically related Melanesian Pidgins, the vast bulk transitive verbs in Kriol are marked with a suffix *-im*,[5] distinguishing between transitive and intransitive *(i.e.* unaccusative) uses of particular verbs. I show that existing variation between *ai* and *mi* in subject-position is structured by and correlated with features that have been associated in typological work with the notion of 'transitivity.' As such, it is claimed here that the distribution of *mi* is wider than '1s clausal object' and that the distribution of *ai* is (coming to be) more narrow than '1s clausal subject.' If this analysis is right, then this phenomenon might be cited as evidence of the grammaticalisation of sensitivity to agentivity and (broadly construed) notions of transitivity; associated with (and perhaps transferred from) the substrate. This paper, then, advances a hypothesis that the form *mi* is encroaching in a systematic manner out of the accusative space and into the distributional domain heretofore occupied by *ai*.

This paper begins with a short introduction to semantic and typological approaches to defining notions of agentivity and clausal transitivity and surveys attempts to situate these categories as theoretical concepts and linguistic universals. Section 3 approaches the question of 1s pronoun distribution by providing empirical evidence, predominantly drawn from a text corpus of Kriol, and deploying both quantitative and descriptive methodologies. The observations made in this section are buttressed by judgments elicited from native speakers of Ngukurr Kriol. Finally, Section 4 contains a general discussion of the implications of this study, proposing an account of the development of the distribution observed.

5. A suffix that has itself been reanalysed from the Standard (perhaps 'Broad') Australian English form *'em*, a phonologically reduced form of 'them/him.' This suffix is unstressed, subject to various phonological processes (i.e. progressive vowel harmony) and synonymous with another, less productive suffix *-it*.

2. Agentivity & transitivity

A great deal of evidence has been presented in the semantics and psycholinguistic literature, that suggests the centrality of notions of 'transitivity' and 'agentivity' to human language. This section provides an overview of some influential attempts to capture and justify the existence of TRANSITIVITY (§ 2.1) and AGENTIVITY (§ 2.2) as universally relevant semantic categories. § 2.3 gives a synopsis of some empirical demonstrations of how these categories see overt grammatical realisation cross-linguistically. Developing a working understanding of these concepts, we will see, will support an adequate description of *ai~mi* variation in Kriol.

2.1 Transitivity: Hopper & Thompson (1980)

Based on a wide survey of cross-linguistic morphosyntax, Hopper & Thompson's (1980) paper is a fundamental early contribution to the development of a nuanced theoretical definition of TRANSITIVITY. They designate a number of 'components' which they deploy to characterise transitivity as "a relationship which obtains THROUGHOUT A CLAUSE" (266, emphasis in original), pointing to the degree to which grammars are sensitive to variation in these components.[6] A vital aspect of this paper's contribution is the suggestion that Transitivity be conceived of as a "continuum along which various points cluster and tend strongly to co-occur" (294) and thereby be defined as the *effectiveness with which an action is transferred or carried over to a patient*. These components are provided in Table 2. The essential insight of this paper, then, is the concept that transitivity be construed as a gradient, rather than binary, phenomenon.

---

6. To provide an illustrative example, Hopper and Thompson provide a Kalkatungu ([ktg] Pama-Nyungan: Queensland) example from Blake 1976 to illustrate the connection between aspect and case marking. This is adapted below from 1980: 272–3).

(i) *Kupaŋuru-ṭu caa kalpin ḻai-ṉa*
old.man-ERG here young.man hit-PST
'The old man hit the young man'

(ii) *Kupaŋuru-Ø caa kalpin-ku ḻai-ṉa*
old.man-ABS here young.man-DAT hit-IPFV
'The old man is hitting the young man'

This example is taken to be a demonstration of how an imperfective clause *(i.e.* one that describes an event which has not successfully reached its endpoint) is treated similarly as a univalent clause by Kalktungu morphology: the agent loses its ergative inflection whereas the patient is demoted to an oblique argument.

Table 2. Components of 'TRANSITIVITY' (adapted from Hopper & Thompson 1980: 252)

Component	HIGH	LOW
Participants	2 or more: '*A*','*O*'	1
Kinesis	action	non-action
Aspect	telic	atelic
Punctuality	punctual	nonpunctual
Volitionality	volitional	nonvolitional
Affirmation	affirmative	negative
Affectedness of O	O totally affected	O not affected
Individuation of O	O highly individuated	O nonindividuated

After providing empirical justifications for each of the components in Table 2, Hopper and Thompson posit the existence of a linguistic universal that originates in some general "pragmatic function" (280) that unifies the category. Specifically, they claim that their co-occurring "high Transitivity" features "predominate in the foregrounded portions of discourse" (292). It is this purportedly universal 'grounding' distinction that, for Hopper and Thompson, subtends (and motivates) grammaticalised notions of Transitivity.

2.2 Proto-Roles: Dowty (1991)

Dowty's 'Proto-Roles and Argument Selection' (1991) adopts a similar perspective to Hopper & Thompson (1980) in terms of its conception of transitivity as a property modellable as a cline. Important for Dowty is the concept that a verb entails certain semantic properties in its argument(s): *viz*. properties associated with proto'-Agentivity and -Patientivity. In doing so, Dowty suggests generalising away from the (useful if *ad hoc)* traditional set of discrete thematic roles and appealing instead to these two primitive concepts. The entailments associated with Dowty's proto-Agent and proto-Patient arguments are given in Table 3 below.

Table 3. Proto-Role entailments (adapted from Dowty (1991): 572)

Verbal entailments for: Agent Proto-Role	**Patient Proto-Role**
volitional involvement	undergoes state change
sentience/perception	incremental theme
causing event/state change in other participant	causally affected by other participant
movement relative to other participant	stationary relative to other participant

On the basis of this framework, the 'transitivity' of a given clause can be represented as the sum of the entailments from Table 3 received by its arguments: i.e. the proximity to proto-agentivity and -patientivity represented in a clause. One of the central

claims of Dowty (1991) is that, in nominative-accusative languages, "the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate" (576). Dowty provides empirical evidence of the psychological reality (and inferred universality) of these two "supergroups of entailments" (600ff): he points to data that evince the development of notions of causation in first language acquisition as well as to typological studies of case syncretism along a 'causal chain', where, cross-linguistically, syncretism occurs with an impressive frequency between case markers associated with either proto-agent or proto-patient properties (600–4).

2.3 Uniting these approaches

Building on this tradition, Åshild Næss's (2007) book *Prototypical Transitivity* marshalls additional cross-linguistic evidence for the cognitive reality of some semantically universal Transitivity phenomenon, setting out to formulate the category in terms of a prototype. Central to this notion is her 'maximally distinguished arguments' hypothesis: that is, a highly transitive sentence is one with a proto-agent (volitional instigator) and a proto-patient (affected participant) (44). For Næss, a 'maximal distinction' between agent and patient prototype roles can be understood heuristically as an interaction between three binary features, tabulated in (4) below.

Table 4. Næss's three privative features (2007: 44)

Property	**Agent**	**Patient**
Affectedness	–	+
Volitionality	+	–
Instigation	+	–

To support her proposed typology, Næss identifies and motivates all eight (2^3) possible distinct argument types, all of which she claims are distinguished through grammatical mechanisms in at least some world languages, inferring that it is in terms of these features (as opposed to specific ϑ-roles opposed by Dowty *v.s.*) that verbs (or, more accurately entire clauses) subcategorise for their arguments.[7] By extension, Næss suggests that marking of core case, for some languages (e.g. the appearance of ergative marking), is triggered by the presence of a volitional agent and an affected, fully individuated patient argument, i.e. that this morphological operation is restricted to clauses that are **fully semantically transitive** (166, emphasis in original).

7. As an example, Næss suggests that *break* subcategorises for a [+ INST] subject, but is underspecified for the volition and affectedness: such a subcategorisation has the advantage of admitting of both agent and instrument arguments (Næss 2007: 109).

2.3.1 *Optional case marking*

Accounts of variable (or 'optional') case marking (OCM), such as the 'typological-semiotic' perspective advocated by McGregor (2010), the case study of Tsova-Tush case in Holisky (1987) and that of Gurindji Kriol in Meakins (2009),[8] provide empirical evidence that bears out these theoretical predictions about the nature of the interaction between the semantics of transitivity and language-specific encoding of grammatical relations. These studies yield important insights into the centrality of agentivity and discourse grounding, and motivate the grammaticisation of these categories.

Holisky's (1987) analysis of Tsova-Tush intransitives reveals an "unusually close relationship" between its ergative-absolutive morphosyntax and the semantics of agentivity (122): she shows how appearance of ergative or nominative morphology on intransitive subjects is linked to the semantics of the verb (which assigns 'actor'-like or 'undergoer'-like roles to its arguments) but that variability occurs commensurate with the relative agentivity of the subject in a given predicate with respect to semantic roles that might be expected for such an event (118ff). In this respect, Holisky suggests that agentivity cannot be defined purely as an entailment of verbal semantics, but is rather a conventional implicature in situations where the *effector* is human – this implicature (and its inverse) can be defeated by optional undergoer (or actor) marking. Effectively, a less agentive or more agentive subject than that which a given verb might by default implicate of its subject receives explicit marking.

Meakins investigates OCM in Gurindji Kriol (GK; 2009, 2011) – a mixed language and the result of sustained contact between Gurindji (Pama-Nyungan: Northern Australia) and Australian Kriol. She describes the 'repurposing' of Gurindji's ergative case suffixes as SVO word order emerges as the dominant marker of grammatical relations in the language, displacing case morphology. Quantitatively accounting for this newly 'optionalised' case marking, Meakins shows both that ergative marking is correlated with *agentivity* (including the hallmarks of clausal transitivity)[9] as well as with *discourse prominence* (focused nominals and topic arguments).[10] In (1) below, the speaker "emphasises that the responsibility for intergenerational transmission of knowledge now lies with them, the parents" (Meakins & O'Shannessy 2010: 1705).

(1) *an* *ngantipa-**ngku*** *wi* *tok* *ngantipany* *karu* *na*
and 1p.INCL-ERG 1p.INCL talk 1p.INCL.DAT child EMPH
'And now we tell these stories to our children'
(GK, Meakins & O'Shannessy 2010: 1705)

8. Meakins (2009) also explicitly appeals to Hopper & Thomson's description of the semantics of Transitivity in predicting the appearance of ergative morphology on subject arguments in Gurindji Kriol.

9. I.e. the more 'transitive' the clause, the more likely ergative marking is to surface

10. Interestingly, it is suggested that case marking is assuming this information structural function in order to fill a functional space vacated by word order in traditional Gurindji.

In a similar vein, McGregor (2010) defends what he dubs a "usage-semiotic" account of optional ergative marking, where, again, the presence or, equally notably, the *absence* of core case marking can be equally meaningful as either asserting the prominence of or backgrounding the agency of a subject argument.[11] Underscoring a nonindependent relationship between usage and grammar, he suggests that some ergative markers "appear to have begun life as focus markers" before categoricalising.[12] Consequently, "the meaning pragmatically associated with the plain construction became coded" (1628) as a backgrounding operation that emphasises a noteworthy lack of agentive properties of the argument. This insight is particularly interesting for current purposes, given the ostensible possibility that the usage and meaning of *mi* has tracked a similar process.

2.3.2 *The semantics of case marking*

Butt and King (2002) further highlight the inadequacy of purely syntactic approaches to case assignment and grammatical relations, providing an analysis of a complex interaction of structural and semantic clausal components that give rise to case marking alternations in Georgian and Urdu. Motivating the need for such an analysis, they point to the invariant assignment of dative case to the subjects of psychological predicates in Georgian.[13] Further, the use of dative inflection on subjects in Georgian (as opposed to ergative) gives rise to evidential (heresay) and aversive readings (65).

Similarly, in Urdu, intransitive subjects optionally receive ergative case in volitional readings (69). Additionally, a series of light verb constructions occur which vary between dative subjects and ergative subjects: these are selected for by the verbs 'come' and 'do' respectively, the choice of verb entails either an experiencer or agent argument. Two minimal pairs that elucidate this variation are adapted below as (2, 3):

(2) a. *raːi-Ø* *kʰãːs-a*
Ram-**NOM** cough-PERF.3sm
'Ram coughed'

b. *raːm-**ne*** *kʰãːs-a*
Ram=**ERG** cough-PERF.3sm
'Ram coughed (purposefully)' (Butt & King 2002: 69, 75)

(3) a. *naːdya=ne* *kahaːni* *yaːd* ***k**-i*
Nadya = ERG story.NOM memory **do**-PERF
'Nadya remembered the story (actively)'

---

11. In McGregor's terms, "the contrast between use and non-use of a morpheme can be semioticised, that is, convey coded meaning" (2010: 623).

12. McGregor cites his own account (e.g. 2008) of Bunuban and Nyulnyulan.

13. As per linguistic convention, the subject is taken to be that argument which is cross-referenced on the verb.

b. *na:dya=**ko*** *kaha:ni* *yad* ***a**-yi*
Nadya=**DAT** story.NOM memory **come**-PERF.3sf
'Nadya remembered the story (spontaneously)' (Butt & King 2002: 69, 78)

The thrust of Butt and King's Lexical Functionalist analysis is that case markings themselves "provide information about the syntactic and semantic environment in which they occur, thus allowing them to influence the meaning of the clause" (80). As the sentences above show, the lexical entry of Urdu *=ne* 'ERG' as formulated by Butt and King (2002: 76) limits its appearance to external arguments and subjects. Further, it can appeal to either the syntactic structure of the clause (presence of a perfective verb and an object argument) or to the semantic structure (the specification of a volitional actor).

Incidentally, their approach promotes an understanding of case as derived from the complex interaction of both syntactic and semantic information. This information *(viz.* participant volitionality, evidentiality) can be encoded in the lexical entries for case morphemes.

We have seen in this section that *agentivity* (a feature that can be conceived of as an entailment that more 'transitive' predicates can make of their (human) arguments) is morphologically encoded cross-linguistically. Models of stable variation as a function of clausal and lexical semantics and pragmatics are instructive for our purposes in moving towards an understanding of the distributional behaviour of these Kriol pronouns and the potential grammaticalisation of a universal conceptual category. For present purposes, we will investigate a hypothesis that *mi* can occur in subject position to encode a first-person argument with *limited agentivity* modulo a series of syntactic constraints. This hypothesis is defended in § 3 and motivated in § 4 below.

3. Pronoun selection: Corpus study

3.1 Materials and method

The most significant documentation of Kriol currently available is the result of a twenty-five year translation project, which yielded a complete rendering of the Bible into Kriol in 2007. The Kriol *Baibul* is available, along with an electronic concordance, online (The Bible Society of Australia 2010 [2007]). This being the most exhaustive resource available at almost 1 million words, and having been described as reasonably representative of some Kriol varieties,[14] it represents the primary source of corpus data for this study.

14. The variety of Kriol utilised in the *Holi Baibul* predominately reflects Arnhem land speakers of Kriol, most people engaged in the project speaking the dialects of Ngukurr, Minyerri, Barunga and Beswick communities. Efforts have been made however to "optimise its readability and

By way of a series of automatic and manual processes, all verses of the *Kriol Baibul* were split into clauses, 11,753 with first-person subjects have been identified: 9,106 (77.5%) with the *ai* form and 2,647 (22.5%) with the *mi* form. These clauses were coded by the author for several syntactic and semantic variables that research on Kriol, semantically-motivated conceptions of transitivity and optional case marking have identified as potential correlates. These variables are summarised in Table 5. Book (of the Bible) was treated as a random variable to control for potential inconsistencies in pronoun-use strategies across translators and the duration of the translation project.

In addition to the text data that forms the empirical basis of this study, data collected by the author at the Ngukurr Language Centre in June-August 2016 is used to provide additional support to the findings. This data takes the form of structured elicitation, speaker judgments of (un)grammaticality and spontaneous conversational recordings from six native speakers. This data is cited in-text with consultant initials and the date of recording.

3.2 Environmental effects: Obligatory conditions

There are several categorical observations to be made with respect to some of the variables listed above. These initial facts about the distribution of *ai~mi* are presented below.

1. 'Disjunctive' ***(DISJ)*** uses of the first person (i.e. instances of left-dislocation – a topicalisation strategy which appears to fulfil a similar pragmatic function to clefting in English), usually collocated with an emphatic particle *na*[15] always surface as *mi* ($N = 804$). Examples of this usage are provided as (4) below. I assume that these uses of *mi* do not receive nominative case, rather residing in the left-periphery of the clause and being 'spelled out' with default case (which in this case is homophonous with the English object form).[16]

appropriateness as much as possible for a 'pan-regional' Kriol readership. Nevertheless, most Kriol speakers themselves would identify the variety in the Holi Baibul as aligning most closely to Roper River dialects spoken at Ngukurr and Minyerri" (Angelo & Schultze-Berndt 2016: 257 citing Margaret Mickan, p.c.).

15. Graber (1987) provides an analysis of the *na* discourse particle as indicating both topic shift and sequentiality of events.

16. This analysis following that of Schütze's 2001 treatment of English 'default case' which accounts for the surfacing of *me* in left dislocation, apposition, ellipsis, gapping, pronoun modification and subject pronoun coordination. Similarly, first-person subjects that occur with a conjunct categorically are pronounced *mi* in Kriol.

Table 5. Corpus study: coding variables

Dependent variable:	Pronoun	*ai/mi*
Independent variables:	disjunctive	[± disjunctive/left-dislocated subject]
	past	± PST auxiliary present
	modality	±NEC auxiliary present
	aspect	± IPFV-marked clause
	polarity	± NEG marker present
	interrogative	± interrogative
	valence	± overtly-marked/prototypically plurivalent verb
	stance	± stance verb
	copula	± copular (appositive ADJ/ADV/PP complement)
Random variable:	Book	(66 books of the Bible)

(4) a. *If im dum lagijat,* ***mi na*** *Yawei garra gudbinji langa im*
if 3s do thus 1s EMPH NAME NEC be.happy LOC 3s
'If he performs like that, I, God, will be pleased with him' [KB Lab1.3]

b. ***Mi na*** *det boi blanga Rabeka.*
1s EMPH the boy POSS NAME
'It is me, Rebekah's son' [KB: Jen29.12]

c. *yu nomo bradin blanga enijing Eibram, dumaji* ***mi na***
2s NEG be_afraid DAT anything NAME because 1s EMP
garra oldei lukaftumbat yu, en ***mi na*** *garra gibit*
NEC IPFV look.after.TR.CONT 2s and 1s EMP NEC give.TR
yu ebrijing
2s everything
'Don't fear for anything, Abram, because it will be me who looks after you and me who gives you everything (you want for)' [KB Jen15.1]

2. Past-tensed first-person singular subject (1s) clauses ***(PST),*** those occurring with tense auxiliary *bin*, occur almost exclusively with *ai* (N = 3653, 99.8%).[17] *Mi* is universally rejected in these contexts by native speakers.

(5) a. ***ai=bin*** *askim im blanga woda*
1s=PST ask.TR him DAT water
'I asked him for water' [KB Jen24.25]

b. ***ai=bin*** *reken ai garra jidan laibalawan en gudbinjiwei*
1s=PST think 1s NEC cop live.ATTR and be_happy.ADV
blanga longtaim
PURP long_time
'I thought "I will be alive and happy for a long time"' [KB Job29.18]

17. There are no contiguous instances of *mi bin*. The few times that past-marked clauses occur with a *mi* subject are the 78 tokens of *mi na bin* (dislocated 1s subjects) and a few (N = 8) instances where *nomo* 'NEG' intervenes between the subject pronoun and verb.

c. *Mi na King Seksis en* ***ai=bin*** *jandim main*
1s$_I$ EMPH King NAME and 1s$_I$=PST send.TR my
wekinmen olabat
worker PL
'I am King Xerxes and I sent my workers…' [KB Est1.15]

d. *Nomeda mi jidan iya langa jeilhaus, bat stil yumob oldei*
CONC 1s sit here LOC jail but still 2p IPFV
joinin garram mi langa det wek seimwei laik yumob bin
join.in with 1s LOC the work identical like 2p PST
oldei joinin garram mi wen ***ai*** *nomo* ***bin*** *jidan*
IPFV join.in with 1s when 1s NEG PST sit
langa jeilhaus.
LOC jail
'Even as I'm in jail, you join me in this work as you join me when I'm not in jail' [KB Phi11.7]

e. ***Ai=bin*** *brabli bradin blanga det greib*
1s=PST very (be.)afraid DAT the grave
'I was terrified of the grave (death)' [KB: Ps116.3]

3. 1s clauses with frequently occurring modal auxiliaries (**MOD**),[18] namely *garra* 'NEC',[19] *gin* can', *kaan* 'cannot' and *lafta~labda* 'have to', also occur almost exclusively with *ai* (*N* = 3811, 99%), *mi* categorically rejected by speakers in this context.

(6) a. *ai garra meigim gudwan daga*
1s NEC make good food
I'll make some good food [KB Gen27.9]

b. ***ai*** *gin luk olabat brom dislot hil*
1s PERM look PL from DEM:PL hill
'I can see them from the hills' [KB Nam23.9]

---

18. These modal auxiliaries all encode some shade of irrealis meaning. Nevertheless, given that these auxiliaries' semantics are not material for the argument levelled at the end of this paper and disagreement on the meaning of 'irrealis' among scholars (Maïa Ponsonnet, pers.comm.), I refer to clauses with modal auxiliaries as +MOD throughout.

19. N.b: *garra* – ostensibly dually derived from SAE *gonna* and *gotta* (Phillips 2011: 44) – among its primary functions, encodes futurate meaning. There is good reason nonetheless to analyse its semantics as modal. The question of *garra's* precise semantics, as we will see in § 4.3, is immaterial for the purposes of the current work; it is claimed that the categorical collocation of *ai* + MOD is *not* attributable to semantic information. Consequently, here I gloss *garra* as a NECESSITATIVE marker to capture a broad range of its uses (which appear to be unitable insofar as they express a *modal force of necessity*) while remaining agnostic with respect to its precise semantics (see Schultze-Berndt et al. forthcoming for a more detailed description of modal auxiliaries in Kriol – they gloss *garra~gada* as FUTURE/OBLigation).

Similarly, *kaan* appears to be the negation of *garra* (in both its future and deontic uses); it is glossed here as PROHIBITIVE.

c. ***ai kaan*** *luklaftumbat* *main braja*
1s PROH look.after.TR.CONT my brother
'I can't (be expected to) look after my brother' [KB Jen4.9]

4. Copular (attributive) clauses (**COP**), those for which some property or membership to some category is being predicated of the subject, are constructed by the apposition of a subject argument and its prepositional or nominal complement (i.e. these clauses typically contain no overt verb form (**Ø-COP**).[20] When the subject of these clauses is 1s, *mi* is almost invariantly selected, *ceteris paribus*. This variable is discussed in greater detail in the following Section (§ 3.3.3), with special reference to particular properties of a class of univalent predicates headed by ***STANCE*** verbs that also appear to function attributively (compare (7d) below).

(7) a. ***mi brabli wail*** *langa yu*
1s very angry LOC you
'I'm furious with you' [KB: Isik23.25]

b. *Wal* ***mi blanga*** *Eibrahem biginini du*
DM 1s DAT NAME children too
'I'm also a child of Abraham' [KB 2Kar11.22]

c. ***mi brom*** *najawan kantri*
1s from other.ATTR country
'I'm from another country' [KB Ruth2.10]

d. ***mi brabli strongbalawan*** *en* ***mi jidan rediwan***
1s very strong.ATTR and 1s STANCE ready.ATTR
'I'm strong and I'm ready' [KB Ais51.5]

e. ***mi/*ai*** *brabli olgamen du*
1s very old.woman too
'I'm a very old woman' [IA 27072017, KB Jen 18.12]

These observations allow for the postulation of a preliminary constraint ranking as presented in (8) below. This ranking captures the distributions of pronouns that we see in Figure 1 below.

Following the conditional inference framework in Hothorn et al. (2006), which makes use of binary recursive partitioning, each of the variables presented here is shown to be significant ($p < 0.001$), yielding the 'constraint ranking' given in (8). Making use of this constraint ranking, Figure 1 shows the relative distribution of *ai* and *mi* in clauses with the conditions described in this subsection. A diachronically-informed account of this distribution is provided in § 4.3 below.

20. While, like many other languages including many creoles, Kriol might be described as 'zero-copula', as we will see below, the Baibul corpus data provides evidence of the grammaticalisation of stance verbs, particularly *jidan* 'sit' (as well as, perhaps *jandap* 'stand' and even *wokim* 'walk') to fulfill the function of an overt copula. These are tagged separately in the corpus (STANCE).

(8) $DISJ/mi >> \left\{ \begin{array}{c} PST/ai \\ MOD/ai \end{array} \right\} >> COP/mi$

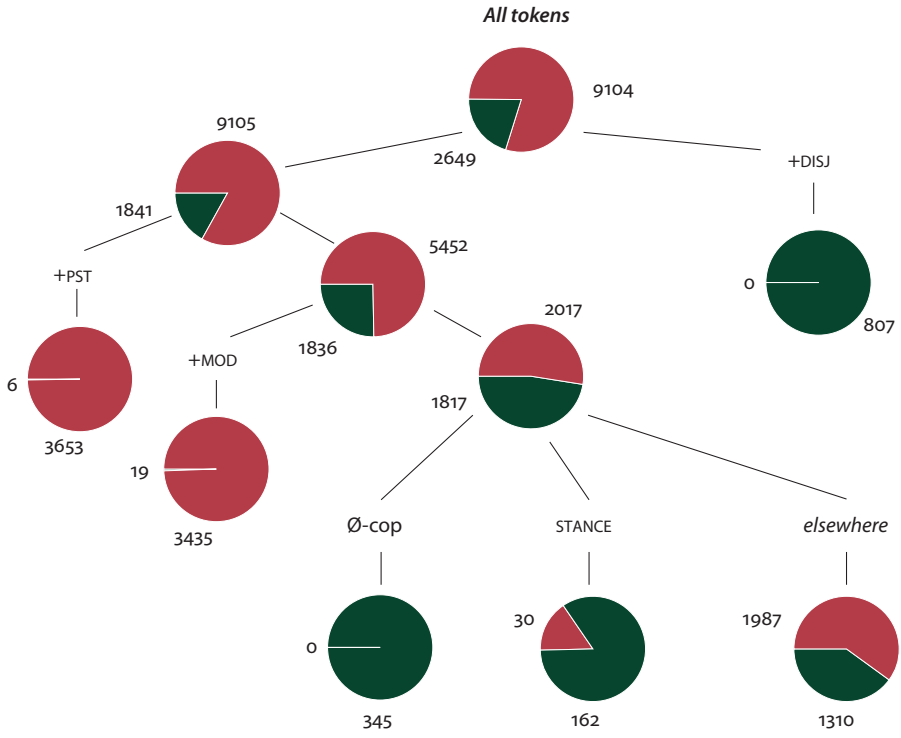

Figure 1. Categorical clausal features determining pronominal form. Red represents the proportion of *ai*, green the proportion of *mi* as exponent of 1s subject argument in the corpus. The terminal nodes of this tree represent the 'categorical variables' (PST, MOD auxiliaries and DISJ, COP constructions) discussed in § 3.2. The node labelled ***elsewhere,*** then, represents the remainder of the 1s clauses in the text corpus as discussed in § 3.3; where 1987 subjects are exponed by *ai* (~ 60%) and the remaining 1310 (~ 40%) by *mi*.

3.3 Structured variation: Semantic constraints

None of the other factors presented in Table 5 can be shown to have such a clear effect as those enumerated in the preceding section. The remaining distributional variation of 1s forms shown in the "*elsewhere*" node *(i.e.* those 1s subject clauses in which none of these 'categorical variables' occurs) in Figure 1 above remains to be accounted for. Nonetheless, this subsection will show highly suggestive correlations between some of these categories and pronoun selection. Given the high-ranked

constraints schematized in § 3.1 above, it is present-tensed clauses that will provide the best domain for such investigation: as such, datapoints in which the 'knockout' (categorical, i.e. clauses with tense and mood auxiliaries) features discussed in § 3.1 are present will be excluded from the following analysis.[21,22] This subsection will investigate each of the conditioning factors that appear to be at play in structuring pronoun variation in this domain. The full results of a linear mixed effects type regression analysis *(Rbrul,* Johnson 2009) are provided as Table 6.

Table 6. Rbrul (mixed effects variable rules analysis) of *ai~mi* distribution. Higher factor weights (FW) for a given variable (factor) indicate a higher probability for the selection of *mi* in a 1s clause (*i.e.* if *FW* = .5, pronoun selection is maximally unpredictable in those contexts.)

Present clauses[23]	***N* = 3622**			
Input probability *mi*	**0.489**			
Deviance	**3780.259**			
Factor	**FW**	**%mi**		***N* (% of data)**
Polarity			$p < 2 \times 7.25^{-11}$	
nomo absent	.594	51%		3140 (87%)
nomo present	.406	34%		482 (13%)
Valence			$p < 1.3 \times 10^{-142}$	
Zero-Copula	.86	93%		251 (7%)
Stance-Copula	.7	89%		156 (4%)
Univalent	.33	57%		1361 (37%)
Plurivalent	.12	33%		1854 (51%)
Aspect			$p < 2.56 \times 10^{-57}$	
No marking	.3	43%		3068 (85%)
Imperfective marking	.7	80%		554 (15%)
Interrogative			$p = 0.625$	
True	[.529]	57%		60
False	[.47]	49%		3562
Book *(Baibul)*			*Random*	

21. Additionally, the presence of a subject conjunct (e.g. ***mi en main san*** *garra go tharrei* 'me and my son will go that way' [KB Jen22.5]) appears to trigger *mi* as a matter of course so these tokens (*N* = 30) have also been excluded given the assumption that these pronouns are not assigned case (see also note 16).

22. The appendix to this paper [p. 368] contains an example of a *Holi Baibul* passage in which *ai* and *mi* both occur as the subjects of present-tensed clauses.

23. 'Present clauses', as elsewhere in the paper, will refer to those clauses without the categorical ('knockout') variables present – i.e. clauses without *na* (marking left-dislocation) or PST/MOD auxiliaries.

The outcome of this variable rules analysis shows highly statistically significant ($p < 0.001$ for all factor groups) effects for verbal aspect (§ 3.3.1), polarity (§ 3.3.2) and valence (§ 3.3.3) on pronoun selection in present-tensed 1s clauses. It supports a hypothesis that the selection of *mi* is more probable in 1s clauses with explicit imperfectivity marking and lower valence. There is also evidence of a small negative correlation with *mi* when negative auxiliary *nomo* occurs. Each of these factors will be discussed in turn.

3.3.1 *No endpoint: Correlation with imperfectivity*

Typological work has pointed to an empirical observation that suggests an association between imperfective clauses and lower clausal transitivity (e.g. Næss 2007; Hopper & Thompson 1980: 271ff, see note 6 *supra* on Kalkatungu). The reasoning for this is intuitive: to the extent that transitivity might be defined as the effective transfer of some controlled action between a volitional actor and an affected patient, events without an endpoint (or those which have not achieved some implied endpoint) are likely to be construed as less effective. Kriol has three morphological mechanisms by which it explicitly marks various flavours of imperfectivity:[24] (1) the presence of an auxiliary *oldei* (<SAE 'all day') or less commonly *kipgon, stil* <SAE 'keep(s) going', 'still'), (2) the presence of verbal suffix *-bat*,[25] and (3) less frequently, reduplication of the verb stem. All three of these strategies can co-occur as shown in (9) below, ostensibly emphasising the habituality/durativity of the event described by the predicate. The minimal sentence pair in (9c) shows an alternation in pronoun choice that is predicted by the transitivity-perfectivity relation. *mi oldei bidim det enemi* 'I beat the enemy' in (i) reflects a general imperfective (habitual) usage of the verb. Conversely, *raidap ai bidim ola enemi* 'until I beat all the enemies' in (ii) specifies a (prospective) accomplishment.

Coding for the presence of one of these markers across all 1s subject clauses (excluding disjunctive, tensed and modalised clauses) yielded the distribution represented in Figure 2 and Table 7 below.[26]

---

24. Semantic work remains to be done to uncover the precise semantics of each of these markers – Munro (2004: 132); Ponsonnet (to appear 2018a: § 4.3) provide some description – certainly each particle has been described as being associated with a discrete shade of aspectual meaning, but these descriptions are frequently heuristic and not entirely justified or borne out by the data available.

25. *-bat* is derived from English adverbial uses of 'about', as 'to talk/jump about', which similarly emphasises a characteristic absence of 'endpoint' for such verbs.

26. This distribution was effectively replicated (i.e. a heavy weighting towards *mi* was found) for all three of these aspect-marking strategies when coded separately.

(9) a. *Wotfo* ***mi*** ***oldei*** ***wek~wek-bat?***
why 1s IPFV work~DUR-PROG
'Why am I always working?' [KB: Ak14.8]

b. ***mi*** ***oldei*** ***bilimap*** *olabat haus garram ola gudbala ting*
1s IPFV fill.TR.up PL house with PL good.ATTR thing
'I fill up their houses with good things' [KB: Prab8.21]

c. i. ***mi*** ***oldei*** ***yusum*** *det pawa blanga God blanga fait […]* ***mi***
1s IPFV use.TR the power GEN God PURP fight 1s
oldei ***bidim*** *det enami*
IPFV beat.TR the enemy
'I use God's power to fight… and I beat the enemy' [KB 2Kar10.4]

ii. *Yu jidan iya langa main raidensaid raidap* ***ai*** ***bidim*** *ola*
2s sit here LOC my right until 1s beat.TR PL
enami blanga yu
enemy GEN 2s
'You sit at my right hand until I beat all your enemies'
[KB: Meth22.44]

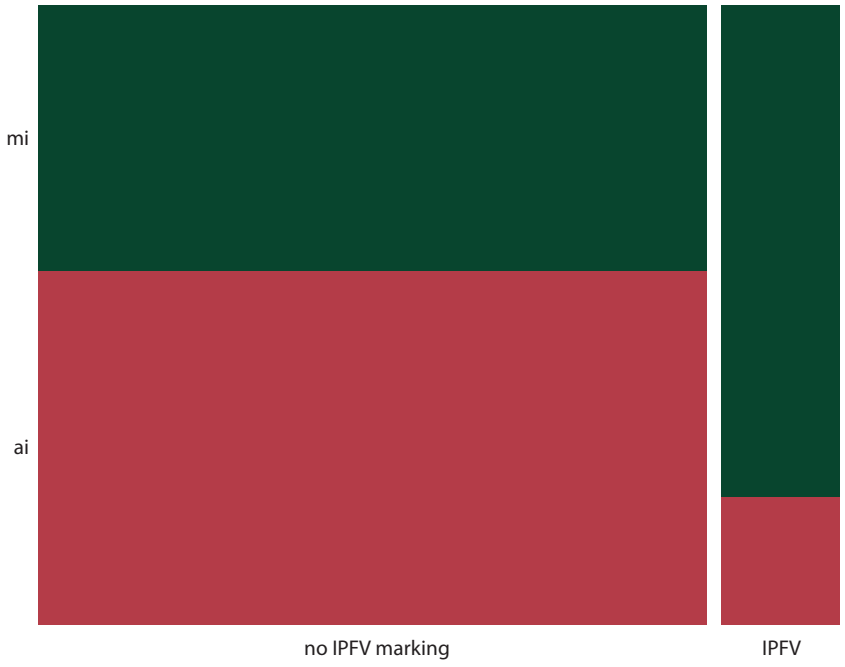

Figure 2. Mosaic chart showing the relative distribution of pronouns (*y*-axis) relative to morphological aspect marking (excluding clauses with tense and modal auxiliaries) in the corpus (*x*-axis)

Table 7. Cross-tabulation of pronominal form and imperfective marking Excluding DISJ/PST/IRR

Form	−IPFV (*N* = 3065)	+ IPFV (*N* = 577)	% + IPFV
ai (*N* = 1852)	1715 (57%)	137 (24%)	7.4%
mi (*N* = 1760)	1320 (43%)	440 (76%)	25%

On the basis of this evidence, there appears to be a significant correlation between imperfectivity and the appearance of *mi* in present-marked clauses, $\chi^2(1, 3612) = 207.01$, $p < 0.0001$. As shown above, this converts to a factor weight of 0.7 for imperfectivity in *mi-* selection.

3.3.2 *Modalities and polarity*

Hopper & Thompson (1980: 276) point to a modest amount of cross-linguistic data that suggests an association between polar negatives, modalisation and their other 'components of Transitivity.'

The above section describes constraints against *mi* in modalised clauses marked with auxiliaries *garra, kaan, gin*. Additionally, *nomo* 'NEG' appears to be weakly associated with a dispreference for *mi*, notwithstanding many counterexamples of this observation (*N* = 174, 34% of occurrences of *nomo* in 1s relevant clauses, *FW* = .4).

339 clauses (< 3%) were also tagged as interrogative: evidence of any association between interrogative and pronoun selection was not found in the corpus ($p = 0.625$).

In view of these data, the semantic properties of the modal and negative semantic domains provide no evidence in favour of a semantic motivation Kriol distribution of *ai~mi*.

3.3.3 *Valence*

In addition to those semantic features discussed in the literature, one of the most telling correlates of transitivity by any account (see § 2) is the number of participants involved in a predicated event: *transfer* necessarily requiring the presence of multiple entities (cf. Hopper & Thompson 1980: 252). As mentioned above, transitive verbs in Kriol are distinguishable by a suffix *-im* and its various allomorphs. The corpus was coded for the presence of this morpheme within 1s clauses, with automatic and manual correction for errors in frequent verbs that subvert this pattern (i.e. false positives and negatives). The results in Table 8 below, provided graphically as Figure 3, are then taken to represent the interaction between pronoun selection and valence. Note that for the purposes of the statistics below, the complement of the 'plurivalent' (i.e. univalent) category is considered to contain *all one-place predicates:* copular clauses (both with covert copulæ and STANCE verbs)

in addition to the traditional grammatical subcategory of intransitives.[27] By this definition, plurivalence is a very strong positive correlate for the selection of *ai* as 1s subject: $\chi^2(1, 3612) = 402.98$, $p < 0.001$.

Table 8. Cross-tabulation of pronominal form and valence

Form	**univalent (*N* = 1761)**	**plurivalent (*N* = 1851)**
ai (*N* = 1852)	601 (34%)	1251 (68%)
mi (*N* = 1760)	1160 (66%)	722 (32%)

As mentioned above, the *mi*+UNIVALENT category includes zero-copula equational clauses (*N* = 210) in addition to equational clauses with 'stance verbs' (grammaticalising copula, *N* = 156).

With respect to STANCE verbs *jidan* 'sit' and *jandap* 'stand', there is evidence that these are grammaticalising, adopting the function of an overt copula.[28] From 527 tokens, 237 occurred in present-tensed clauses. From these 237, 204 (86%) occurred with *mi* as against 33 (14%) with *ai*. Figure 1 above showed that present-tensed copular clauses select overwhelmingly for mi. It is notable, therefore, that the *mi*/cop condition appears to be categoricalising for this prototypical stative verb and grammaticalising copular particle. This observation can be construed as striking support for the psychological reality of a connection between *ai*/*mi* and the notion of agentivity. The examples in (10) and (11) below provide some insight into the potential contribution of pronoun selection in the context of this verb.

27. On inspection, this characterisation of ±UNIVALENT is a reasonable one, particularly taking into account the blurry line between stative verbs and copular predicates that is observed in Kriol, both syntactically and semantically. Further, Australian languages frequently treat the subjects of these predicates identically: 'The subject of an equational clause is always of nominative case' (Cunningham [Sharpe] 1969 on Alawa: 185). The following Alawa examples exemplify the difficulty of distinguishing subtypes of univalent uses.

lalama atiyin ruwu
fish trap hole tree
'The fish trap is a tree with a hole in it' (1969: 188)

ŋina ŋawaɖurruwarr
1s axe.with.1s
'I have an axe' (1969: 187)

28. Phillips (2011: 52ff) provides evidence of this change being a function of substrate transfer and a proposed reconstruction. The use of stance verbs as copula is certainly attested in Australian languages – see also Dixon 2002a: 21, b: 241ff. The grammaticalisation from lexical stance verb to copula is reported in the literature as a "path of change" e.g. Bybee & Dahl 1989: 78).

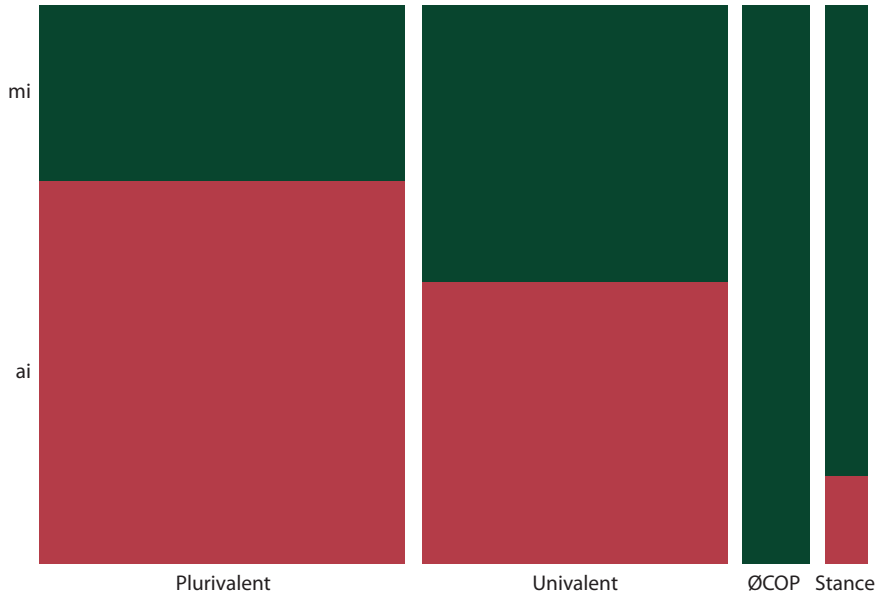

Figure 3. Relative distribution of pronouns occurring in present-tensed clauses with multiple (plurivalent) and various types of single (univalent) arguments

(10) a. ***Mi*** *oldei* ***jidan*** *sedisfaid nomeda mi hanggri en nomo garram*
1s IPFV *jidan* satisfied even.if 1s hungry and NEG have
enijing en nomeda mi bulap en garram blendibala.
anything and even.if 1s full and have plenty.
'I'm always satisfied, whether I'm hungry and have nothing or I'm full and have a lot' [KB: Fi14.12]

b. ***Mi jidan*** *ebriweya langa skai en langa graun du*
1s *jidan* everywhere LOC sky and LOC ground too
'I am everywhere; across the sky and the ground' [KB: Jer 23.24]

(11) a. *Nomeda* ***ai*** *oldei* ***jidan****-bat raitwei, yu oldei hantingbat*
CONC 1s IPFV *jidan*-PROG correctly 2s IPFV hunt-PROG
blanga mi…
DAT 1s
'Even if I behave righteously, you hunt me' [KB: Job 10.16]

b. *mi na YAWEI oldei jidanbat garram main wekinmen olabat,*
1s EMP NAME IPFV *jidan*.CONT with my worker PL
bat ***ai oldei jidan*** *wail langa main enamimob*
but 1s IPFV *jidan* angry LOC my enemies
'I, God, will be present with all my servants but I am wrathful against my enemies' [KB: Aisa66.14]

c. ***ai jidan*** *enami blanga detlot mesinja hu oldei dalimbat*
1s *jidan* enemy DAT DET:PL messenger who IPFV tell.PROG
yumob laiya blanga olabat drim
2P lie PURP PL dream
'I am the against those messengers who tell you lies about their dreams'
[KB: Jer 23.32]

While there are few tokens of *ai jidan* in the corpus, some observations can nevertheless be made. For the *mi*-sentences in (10) the speaker is predicating a characteristic trait of himself as an individual in (a), and a necessary spatial location in (b). Contrastingly, the selection of *ai* in (11) appears to foreground the agency and instigational competence of the speaker. In (a), Job draws attention to his concerted, upstanding behaviour, in (b) and (c) God is making declarations about Their policies and attitudes towards Their enemies (namely threats of punishment). While *mi* is likely available in these cases, the selection of *ai* appears to correlate with the speaker's choice to foreground their own sentience and agentivity in these cases. Additionally, this observation appears to expand to lexical uses of stance verbs as shown in the speaker judgment in (12), where the use of *ai* was considered to be infelicitous. These alternations are further discussed below.

(12) **Context:** it is a hot day
mi/#ai jandapjandap la san
1s stand.IPFV LOC sun
'I'm standing in the sun' [AD 27072016]

This section has so far presented a variable rules style analysis (*Rbrul,* Johnson 2009) of the interaction of these variables in present-tensed clauses and the relation that they bear on 1s pronoun in Kriol. The following subsection continues discussion about the apparent contribution of additional semantic variables.

3.4 Lexical semantics: Further observations

The variables discussed in the subsections above reveal significant correlations between clausal features and pronoun selection, explaining a good deal of the variation reported. To nuance our model, and explain the remaining variation, it will be informative to further investigate the effects of the lexical semantics of Kriol verbs upon pronoun selection.

Investigation of coincidence of *mi*-subject clauses and transitive verbs reveals some preliminary generalisations. Of the 619 tokens of bivalent verbs taking *mi*-subjects (some 25% of all *mi*-1s clauses), a striking proportion of them deviate sharply from the prototype of the 'highly-transitive' verb as it is laid out in the literature (e.g., Næss 2007). These include categories such as 'psychological verbs'

(those selecting for an 'experiencer' subject) such as *jinggabat* 'think',[29] *wandim* 'want' and *laigim* 'like', or verbs selecting for an 'affected subject -argument, including *dagat* 'eat', and *dringgim* 'drink', which also frequently occur with *mi* in subject position. A case can easily be made for verbs such as *bulurrum* 'obey, follow and *preisim* 'praise, worship to categorise in this way as well, given that these appear to select for [−INSTIGATOR; +AFFECTED] arguments, to borrow from Næss's feature set, see the example in (13a) below. Both of these predicates seem to implicate backgrounded agency and foregrounded humility of their subjects relative to the objects. Furthermore, *ardim* 'hurt' can select for either: an agentive subject as *aibin ardim main fren* 'I hurt (PST) my friend'/*yu ardim mi* 'you hurt me ; or an experiencer subject (e.g. (13b)) above, in cases where the object is a body part, a reflexive pronoun, or elided[30] – in all of these cases the subject is understood as the undergoer of some hurting event. Unfortunately there are no cases of present-tensed 1s agentive subjects in the corpus. The hypothesis advanced in this paper would predict that an agentive use of the verb, as in '1s hurt 2s would surface as ***ai/#mi*** *ardim yu*.[31]

(13) shows two present-tensed clauses with bivalent, but 'less transitive' predicates. Table 9 provides a list of plurivalent lexical verbs and their occurrence with *ai~mi* in the corpus.

(13) a. ***mi*** *bulurrum detlot lowa blanga yu*
1s follow DEM law POSS 2s
'I obey your laws' [KB: Saam119.173]

b. ***mi*** *ardim main hed!*
1s hurt my head
'My head hurts!' [KB: 2Kng4.19]

While it is problematic, on the basis of the available data, to make assessments about the direct effects of *ai* and *mi* on clausal semantics, some insights can be drawn upon closer examination of context. In the case of *gibit* 'give', clauses with a concrete theme (direct object) appear to be more felicitous with *ai* ($N \approx 25$) than with *mi* ($N = 1$).

29. Note the univerbation of the adverb from English phrasal verb collocation 'think about.' The presence of this preposition itself might be interpreted as formal evidence of the low transitivity of clauses with this predicate, cf. accounts for the semantics of case frames in Germanic (e.g. Næss 2007: 41, Chapter 8). Cf. Ponsonnet (2018a, b).

30. Subject elision resembles an unaccusative transformation. Interestingly, Dowty (1991) eschews the need for such an analysis, suggesting that appealing to the patientive semantics of the argument would be sufficient to account for this behaviour (as opposed to the argument's status as an syntactically underlying direct object.)

31. Speaker judgments elicited by the author in Ngukurr, June 2016 lend additional empirical support to this prediction.

(14) a. *dijan* ***mi*** ***gibit*** *yu im det nyubala olagijawan lowa*
DEM 1S give.TR 2S 3S the new.ATTR complete law
'This I give to you, it is the new perpetual law' [KB Eksa12.14]

b. ***ai*** ***gibit*** *yumob olkainaba sid daga en frut daga*
1S give.TR 2P various seed food and fruit food
'I give you all kinds of seeds and fruit' [KB Jen1.29]

meigim 'make', a verb with causative semantics, is predicted to occupy a high 'interval' on the transitivity cline, given the distinct entailments that this verb makes of its subject and object arguments (i.e. a volitional, instigative subject and an affected object). There are 16 occurrences (37% of tokens) of *mi* in subject position. From these 16, eight have explicit imperfective marking (cf. 0 tokens of *ai meigim*+IPFV) and five are reflexive (cf. 14% of *ai meigim*+REFL): both of these are operations clearly associated with transitivity reduction.

Table 9. Individual plurivalent lexical predicates by 1s pronoun selection. Lexical items in [brackets] are predicted to be more highly transitive/more resistant to *mi*.

predicate	***ai*-SUBJ**	***mi*-SUBJ**	**% *mi***
ardim 'hurt'	0	20	100%
raidimbat, raidimdan 'write (down)'	1	30*	97%
dagat, dringgim 'eat, drink'	1	10	89%
bulurrum 'follow'	5	15	75%
jinggabat 'think'	14	42	75%
trastim 'trust'	7	21	75%
askim 'ask'	25	77	75%
abum 'have' (light verb)	8	21	72%
preisim 'praise'	7	10	59%
irrim 'hear'	5	7	58%
laigim 'like'	83	60	42%
[jandim] 'send'	16	11	41%
[meigim] 'make, CAUS'	27	16	37%
[panishim] 'punish'	6	2**	25%
garram 'have'	119	36	23%
[gibit] 'give'	62	18	22%
[dalim] 'tell'	517	82	14%
wandim 'want'	279	41	13%

* *raidim* 'write' occurs with 1s subjects only with progressive marker *-bat* or with adverbial suffix *-dan* <'down'. This may confound arguments about lexical semantics given the correlation between imperfectivity and *mi*-selection detailed in the previous section. In all cases it has a direct object.
** Both tokens of *mi panishim* (*olabat*) 'I punish (them)' were marked for IPFV and describe habitual activities.

Similarly, Table 10 provides a selection of univalent verbs which show considerable internal variation in 1s pronoun selection.

The motion verb *go,* importantly, co-occurs with *ai* considerably more frequently than the bulk of these verbs. Given that the collocation of an animate argument and a motion verb gives rise to an implication of instigation and volition (i.e. agentivity) on the part of their subject, we have reason to understand that these verbs are amongst the more 'transitive' univalent predicates. An elicited example is given in (15) below: here we see a single clause where *ai* has been selected in a volitional univalent clause (the motion verb *wok* 'walk') and *mi* in a non-volitional plurivalent one. These two predicates both invite conversational implicatures: agency on the subject's part in the first clause and a lack of this property in case of the second.

(15) ***ai wok~wok-bat*** *en* ***mi burr-um but*** *sinek*
1s walk~DUR-PROG and 1s put-TR foot snake
I'm walking around and I (unwittingly) step on a snake' (AJ 14072016)

Table 10. Individual univalent lexical predicates by 1s pronoun selection. Lexical items in [brackets] are reliably predicted to implicate agentivity/be more resistant to *mi*

predicate	*ai*-SUBJ	*mi*-SUBJ	% *mi*
wek 'work' (see 17)	0	11	100%
cry 'krai'	0	9	100%
sheiksheikbat 'trembling'	0	3	100%
hatjamp 'be startled'	0	2	100%
(no) gudbinji 'be (un)happy'	2	42	95%
bradin '(be) frighten(ed)'	1	18	94%
jingat 'shout out'	5	24	83%
silip 'sleep'	1	5	83%
prei 'pray'	17	44	72%
sing 'sing'	3	5	63%
dai 'die'	10	18	64%
[pramis] 'promise'	4	2	33%
[go] 'go'	23	5	18%

The sentence pair in (16) below again permits us to compare the effects of pronoun choice. Here, the appearance of *ai* in (b) appears to correlate with an agentive and volitional subject; one who is an active participant in refraining from sleep. Such an implicature does not appear to hold for (a). Additionally, the strength of the tendency for univalent *wek* 'work' to co-occur with a *mi* subject may be connected to an apparent lack of volitionality or the presence of an oblique argument that makes explicit the beneficiary (and likely instigator) of the work which is undertaken by the speaker. This is shown in (17) below.

(16) a. *wen mi leidan, **mi nomo silip,** jis laik det bard weya im*
when 1s lie.down 1s NEG sleep just like the bird REL 3s
jidan miselp ontop langa haus
sit REFL on LOC house
'when I lie down, I don't sleep, just like a bird perching alone on a roof'
[KB: Saam 102.7]

b. *naitaim **ai nomo silip** dumaji mi wandim jinggabat wanim yu*
night 1s NEG sleep because 1s want think what 2s
oldei dalimbat mi
IPFV tell.CONT 1s
'I don't sleep at night because I want to reflect upon what you say to me'
[KB: Saam 119.148]

(17) a. *wotfo **mi oldei wekwekbat?** Ai nomo garram eni femili*
why 1s IPFV work.DUR.PROG? 1s NEG have any family
en ai nomo jidan gudbinjiwei
and 1s NEG sit happy.ADV
'What am I constantly working for? I have no family and I'm not happy.
[KB: Aklis 4.8]

b. ***mi wek** blanga yu jus laik main mami bin oldei dum*
1s work PURP 2s just like my mother PST IPFV do
'I work for you just as my mother did' [KB: Saam 116.16]

This allows us to make the a prediction that *ai* will be in felicitous in contexts like (18) below; this is confirmed by elicitation.

(18) *kwikbala na! dumaji mi/*ai gulijap dai/finish bla daga*
quick EMP because 1s almost die PURP food
'Quickly now! I'm about to die of hunger' [GT 22062016, KB Jen 25.30]

The data in Tables 9 and 10 indicate tentative support for the (partial-) reanalysis of *mi* as a 'first person-experiencer'-type pronoun and, consequently, the in-progress grammaticalisation of agentivity and semanticisation of grammatical relations in Kriol. Simultaneously, given the nature of the data (including a concomitant resistance among highly transitive verbs to occurring in the present tense, associated with their lexical aspect), it is additionally important to understand the degree to which pragmatic and discourse (e.g., frequency) effects may also be at play in determining pronoun selection. Further discussion of the implications of these findings are provided in the following section.

4. A diachronic account of *ai~mi* variation

We have seen clear evidence in the preceding sections that the distribution of Kriol *ai* and *mi* diverges from that of its English sources and cognates, *I* and *me*, a claim that has not been closely investigated in previous work. Additionally, as we have seen, it is clearly not the case that the distribution of *ai* and *mi* is completely, or even primarily, conditioned by syntactic valence. Rather, there is clear evidence of both high-ranked distributional constraints in addition to indications that clausal *semantics* (and perhaps features of the greater discourse) allow for the optional occurrence of *mi* in subject position for clauses, particularly those with a low degree of semantic transitivity. Hopper and Thompson's (1980) framework has been cited extensively as a model of parameters for semantically motivating transitivity, spawning a growing literature on this notion. It can be hypothesised, in keeping with these, that for bivalent, telic verbs that favour highly agentive subjects and affected objects, we are much less likely to see *mi* occupying subject position. Elicited speaker judgments lend support to this generalisation in the form of negative evidence. This clear judgment is shown in the alternation provided as (19) below: *mi* is judged by native speakers to be a degraded (if not altogether infelicitous) subject for the predicate *kilim ded* 'kill'.

(19) ***ai***/**#mi kili yu ded*
1s strike 2s dead
'I kill you' [AJ 14072016]

Conversely, stative predicates like *silip* 'sleep' and *fil* and less agentive intransitives like *drimdrim(bat)* 'dream' occupy the other end of the spectrum and have a demonstrably greater tendency to attract subjective-*mi*.

(20) ***mi***/*??ai fil nogudbinji*
1s feel unhappy
'I'm feel sad' [IA 27072016]

We have additionally seen the consistent selection of *mi* in subject position for copular predicates, both in zero-copula constructions and with explicit STANCE verbs. Distinguishing syntactic categories can be a nontrivial enterprise; the boundary between appositional adjectives/attributive nominals and stative verbs (which appear to associate with verbal morphology and adverbial modification) is unclear (*cf. gudbinji* '(be) happy', *bradin* '(be) afraid', *wail/einggri* '(be) angry', *sheim* '(be) ashamed').[32] Examples showing the morphosyntactic behaviour of these predicates in use are provided in (21–23).

32. See also Ponsonnet (2016), to appear (2018b) for a related take on the exponence of the 'emotion middle' in Kriol.

(21) a. ***mi jidan*** *lida langa olabat en* ***mi sheim*** ***miselp*** *na*
1s sit leader LOC PL and 1s be.ashamed REFL SEQ
'I am their leader and I am ashamed of myself' [KB Prab5.14]

b. *nomo* ***sheim~sheim!***
NEG shame~REDUP like that
'Don't be embarassed!' (Ponsonnet to appear 2018a: 9)

(22) a. ***ai garra brabli gudbinji*** *blanga tharran en ai garra larramgo*
1s NEC very be.happy DAT then and 1s NEC let.TR.go
main pipul fri
my people free
'I'll be very happy about that (a series of offerings) and I'll forgive my people's sins' [KB Isik45.17]

b. *bat nomeda mi oldei abum detkain trabul bat stil* ***mi***
but CONC 1s IPFV have.TR DEM trouble but still 1s
oldei gudbinji
IPFV be.happy
'but even if I have these problems, I still am happy' [KB 2Kar12.10]

c. *wal trubala* ***ai garra jidan gudbinji*** *langa olabat*
DM true 1s NEC sit happy? LOC PL
'And then, I will be happy with all of them' [KB Lem3.21]

(23) a. ***ai garra git*** *brabli* ***einggri***
1s NEC INCH very angry
'I will get very cross' [KB Jer4.4]

b. ***mi nomo einggri*** *kwikbala*
1s NEG anger quick
'I don't anger quickly' [KB Num14.18]

This data demonstrates that the shared semantic properties of attributive adjectives and stative intransitives, which manifests in a blurry distinction between the two syntactic categories, provides a likely conduit for the extension of the categorical occurrence of *mi* as a copular subject (in the absence of the 'higher ranked' local environment constraints) into that of an '1s-EXPERIENCER' type pronoun. The categorical selection of *mi* in copular clauses (**ai/mi gudbinji*) has facilitated its regular selection in stative predicates (**?ai/mi jidan gudbinji*). This hypothesis, one that also can be shown (e.g. Phillips 2011: 52–5) to align with morphosyntactic patterns in Kriol substrata, is elaborated in the following subsections.

4.1 Transfer

There is a growing literature that investigates the effects of radical language contact and associated large-scale, abrupt changes to the grammar of traditional Australian languages (Van Den Bos et al. 2017; Lee 1987; McConvell & Meakins 2005; Meakins 2009; Meakins & O'Shannessy 2010; Schmidt 1985 a.o.). The rich case morphology and ergative alignment that characterise Australian languages is a particularly fertile domain of inquiry in this respect. Work across these post-contact varieties has repeatedly observed the establishment of English-like, configurational strategies in marking grammatical relations as the case-marking systems erode. Schmidt (1985), in one of the first works of this kind, investigates the emergence of what she terms 'Young people's Dyirbal' (YPD) – a contact language unrelated to Kriol and a radically restructured form of a traditional Australian language – Dyirbal (†Pama-Nyungan: NE Queensland). YPD is characterised by the deterioration of traditional case marking as word order grammaticalises, while ergative marking becomes an optional, stylistic choice.[33]

Additionally, as relayed in § 2.3.2, Gurindji Kriol, in addition to Light Warlpiri, two mixed languages borne of sustained contact, deploy the traditional ergative suffix to a similar effect, marking a topicalised construction (Meakins & O'Shannessy 2010). The repurposing of this category in the presence of competition between strategies of grammatical relation marking and the ostensible connection between constituent focus and agentivity described in this work may be instructive for our purposes (see Meakins 2009: 72–7).

Given that pronominal 'case' marking in English makes distinctions in the third-person and the plural as well, one of the several pressing questions that arises is how to best account for *ai~mi* variation while the remainder of Kriol's 11-pronoun inventory is invariant for case. Evidence from the early pidgins spoken in Australia between the late-seventeenth and early-twentieth centuries – the inputs from which Kriol creolised – has demonstrated some variation in the first person but no other ostensible sensitivity to either case or number in pronouns (Harris 1986; Troy 1990). Troy describes "*me* as the case invariant first person pronoun" (102) and indeed most

---

33. Additionally, in these varieties, Schmidt suggests that "there is evidence of weakening of the pronominal-nominal distinction as the ergative affix is extended to 1st and 2nd person pronouns" (134). Australian languages are also notable for the near-ubiquitous occurrence of 'split ergativity' – most commonly manifesting as a system where first- and second-person arguments pattern accusatively whereas third-person (full NP) arguments pattern ergatively.

This empirical observation, analysed in important work by Michael Silverstein (1986 [1976]) can also likely be fruitfully explained based on Dowty's (1991) Proto-Role Hypothesis: proto-agentivity is calculated for subject assignment for 1st- and 2nd-person arguments and -patientivity in the 3rd-person. Further analytical work is needed in this area.

The levelling of this split is important for our purposes, given that it demonstrates the salience of this feature of Dyirbal grammar and concerns the pronominal domain.

individual speakers transcribed in Harris (1986) generally appear to select *me* and certainly display no variation between *I~me*. Of Kriol, Harris suggests:

> *Ai* is also found sometimes. It is probably not the case that 'me' and 'I' survived as alternatives from Northern Territory Pidgin English, but that *ai* is used by those whose Kriol usage exhibits a higher degree of recent Engish influence.
>
> (Harris 1986: 372, note 168, emphasis my own)

In an attempt to counter Derek Bickerton's vocal skepticism (e.g., 1981: 304) of substrate influence in creole languages (he claims that substratists have failed to provide a means of predicting when substrate features will or will not appear in a given creole), Jeff Siegel (2008) outlines principles – 'transfer constraints' – that have been argued to govern the likelihood of a substrate feature to be transferred into a creole language (148ff). These can be partitioned into 'availability constraints' and 'reinforcement principles.' The 'transfer constraints' theory requires that "there[…]exist a lexical morpheme in the L2 (here, the expanding pidgin or the lexifier) that can be reanalysed according to the functions of a grammatical morpheme in the L1" (166). Such a morpheme will be *perceptually salient* (easy to distinguish phonologically) and *congruent* (syntactically resembling the substrate analogue) (Munro 2004: 32). In one of the first deployments of this framework, Munro observes the presence of explicit suffixation of ergative arguments in all substrata (114) but concludes that its transfer is frustrated by the unavailability of "perceptually salient postnominal forms in English that could be interpreted as being case markers" (116).

Bearing these assertions in mind, there are multiple reasons that the transfer of a case distinction during creolisation may have been limited to *ai~mi*:

1. The English pronoun paradigm is defective for case: there is no existing lexical resource that could therefore be recruited to encode, for example, a similar distinction in the second person.
2. The phonological distinction between 'I' [ae] and 'me' [miː] relative to 'he' /hiː/ and 'him' /him/, for example, may have contributed to failure to perceive of a crucial, recruitable distinction in the third person singular.
3. The relative frequency and discourse-centrality of the first-person singular compared to other pronouns, particularly in child speech (cf. acquisition data, e.g. Harley & Ritter 2002) is likely to have augmented the perceptual salience of this distinction (particularly with respect to the nonsingular pronouns). As a result, the *ai~mi* distinction is privileged from the vantage point of recruitment, as learners attempt to infer the distributions and lexical entries of each of these frequent, discourse-central pronouns.[34]

34. Importantly, an anonymous reviewer points out that the literature on agency has observed that speakers can only reliably ascribe 'Transitive' properities such as volitionality and sensory

Given the attested simultaneous presence of vocabulary items *ai* and *mi* '1s' in the Kriol lexicon, likely a function of continued contact with English in early- and mid-twentieth century (Harris 1986, Sandefur 1982 a.o. on 'decreolization'), the following subsections propose a diachronic account of the emergence of their synchronic distribution in the contemporary language.

4.2 Pronominal 'case': A diachronic vestige as the input

Contrary to traditional synchronic descriptions of Present Day English morphosyntax, it has been suggested (Emonds 1986; Parrott 2007, 2009) that English lacks morphologically assigned case and that the 'non-transparency' of overt cues for nominativity and accusativity across the English grammar has impeded the acquisition of case morphology. That is, unlike related languages (e.g., German), case alternation does not pervade English grammar but is rather restricted to the pronominal domain. They claim that this constitutes insufficient evidence for the acquisition of grammaticalised case marking. Additionally, Emonds refers to a series of 'extra-grammatical constraints': effectively 'overcorrection tendencies' that are native to no dialect but are rather taught (i.e. prescribed). Given the non-systematicity and extremely low productivity of morphologised case variation, it is claimed that observed alternation in first- and third-person pronouns is attributable to the acquisition of *contextual variants* as opposed to being an exponent of true case. Parrott analyses the 'vocabulary items' of 1s subject and object forms as familiar rewrite rules as follows:

(24) a. [D, Pers:1, Num:s] ↔ [aj] / [$_{TP}$_[T$_{[\text{Past:}\pm]}$…]]
b. [D, Pers:1, Num:s] ↔ [mi] / *elsewhere* (Parrott 2009: 280)

Parrott's claim, then, following Emonds, is that '1s' is pronounced *ai* in clause-initial position (i.e. in the 'specifier' position of a Tense Phrase), and that otherwise it is pronounced *mi*. The primary intuition here is that [aj] and [mi] are allomorphs, exponing a singular set of features. Furthermore, additional vocabulary items are *learned*. Additional rules can then explain Emonds' "deviant prestige varieties", as in (21c) below.

(24) c. [D, Pers:1, Num:s] ↔ [aj] / [ænd]__ (*ibid.*)
PRONOUNCE A 1S PRONOUN [aj] IN THE LOCAL ENVIRONMENT OF [ænd] 'and.'

experience to themselves, further distinguishing first- from other-person marking in terms of agentivity ascription ('the empathy hierarchy see, e.g. Van Valin & Wilkins 1996 a.o.).

Additional distributional rules like (19c) above, lead to interspeaker variation and broader change in pronoun distribution, triggering the observed diversification between 'prestige' and 'standard' varieties of English (284ff).

The absence of inflectional morphology across creole languages has long been suggested to be a hallmark of typological "creoleness" (e.g., McWhorter 2005). It is proposed that inflectional morphology is levelled out during pidginisation due to its "low perceptual saliency with low import to basic communication, encouraging learners acquiring the language rapidly and informally to bypass acquiring them" (McWhorter 1998: 791). This observation is borne out across the Kriol grammar as a whole: as we saw above, Kriol's pidgin predecessors retained *me* as an invariant first-person pronoun (and additionally, this invariance is still exhibited in some dialects and across its Melanesian Pidgin cousins.) The question arises, then, of how sustained contact with English weak case cues and the pursuant competition between 1s pronouns as a consequence of the introduction of *ai* into this domain of the Kriol lexicon has yielded the structured variation described above. I propose an account of this in the following section.

4.3 Accounting for the variation

In light of these facts, the diachronic hypothesis that this paper advances is that the entry of *ai* into the Kriol lexicon as a consequence of sustained contact with English[35] at some stage in the early development of Kriol has triggered a restructuring of the case-less pronoun paradigm. Given the weak, inconsistent cues of morphological case available in English (discussed in § 4.2), these pronouns represent **congruent** and **perceptually salient** locations for transfer and/or reanalysis (cf. Siegel 2008: § 6 discussed in § 4.1, this paper).

4.3.1 *Semantic factors*

I propose in this case that the mesolectal Kriol learner, one who has limited access to Standard Australian English and partial access to a series of substrate languages, makes innovative inferences about the semantics of existing vocabulary items *ai* and *mi,* triggering reanalysis in the 1s domain. Given that *mi* occurs predominantly as the exponent of a 1s direct object, a syntactic position that English closely associates with patientive properties, this pronoun's lexical entry comes to be interpreted as being associated with any first-person argument that carries those semantic properties, independent of its grammatical relation. Notwithstanding the generalisation that object position associates with patientive arguments, verbs also frequently entail patientive properties in their subject arguments (*viz*. the 'experiencer' ϑ-role).

35. See "decreolization", described in Harris (1986) and Sandefur (1982).

Consequently, *mi* naturally comes into competition with *ai* as it begins to encroach upon the space previously occupied by the latter – a space that happens to be the 'absolutive', a category familiar in all substrata which unifies intransitive subjects with transitive objects. The claim defended in this paper is that this reanalysis is the source of the described variation.

As we have seen, the occurrence of *mi* as a copular subject is categorical: the blurry boundaries between copular complements, lexical statives and even intransitive eventive predicates, contribute to and reinforce the propagation of this variant through the grammar.

Additionally, exemplified in the sentence pair in (16), repeated below as (25), given the space that *mi* has come to occupy, the selection of *ai* in a univalent and less-transitive verb such as *silip* 'sleep' appears to have become a marked choice. If *mi* is indeed coming to be understood as marking a non-agentive subject, then a conversational implicature arises in which the speaker's decision to eschew use of this 1s form is interpreted as their intention to *foreground* their own agentivity in effecting the predicate. A consultant involved in the original translation process explains that the selection of *ai* in the first clause of (26) has a 'strengthening effect' – foregrounding the individual relevance of the speaker and their volitionality in attending to the Temple during their prayer [pers.comm., IA 27072016]. These principles have been adopted in 'dynamic' accounts of grammaticalization, e.g. Traugott (1999), Deo (2015: 24, a.o.), see also Horn 1984. This process represents Deo's 'categorialisation' – the **semanticisation** of an erstwhile conversational implicature as a primary driver of meaning change.

(25) a. *wen mi leidan, **mi nomo silip,** jis laik det bard weya im*
when 1s lie.down 1s NEG sleep just like the bird REL 3s
jidan miselp ontop langa haus
sit REFL on LOC house
'when I lie down, I don't sleep, just like a bird perching alone on a roof'
[KB: Saam 102.7]

b. *naitaim **ai nomo silip** dumaji mi wandim jinggabat wanim yu*
night 1s NEG sleep because 1s want think what 2s
oldei dalimbat mi
IPFV tell-CONT 1s
'I don't sleep at night because I want to reflect upon what you say to me'
[KB: Saam 119.148]

(26) ***ai luk** langa dijan Serramoni Pleis wen **mi** prei*
1s look LOC this ceremony place when 1s pray
'I look at the holy ground when I pray' [2Kran 6.20]

4.3.2 *Distributional factors*

That the source of this variation is borne of opaque morphological case cues in the lexifier is perhaps further evinced by the observation that the emergence of *mi* in subject position is greatly constrained by the presence of a pre-verbal inflectional auxiliary (*bin, garra etc.*). While this initially appeared an arbitrary foil to a hypothesised transitivity-centric account, it is in fact attributable to a phenomenon that is similar to the pseudo-grammatical "contextual allomorphy" rules discussed in § 4.2. That is, the invariant selection of *ai* before *bin* 'PST', *garra* 'NEC', *kaan* 'PROH' etc. is a function of a learned and highly-ranked constraint associated with the pronominal allomorphy conditions provided in (24). This is almost certainly attributable to the very high frequency of these morphs: indeed *bin* has encliticised to and is a phonological dependent of pronominal subjects: *aibin, yubin* and *imin* are all written as single words in the *Holi Baibul*. Similarly, there is evidence that *garra* has traced a similar path with the attestation of forms such as *airra* '1S=NEC' *wirra* and '1P.INCL=NEC' and lenition of the initial stop in *kaan*> [ɣan]. *oldei* – the imperfective auxiliary and one of multiple strategies of grammatical aspect marking – does not have this status: indeed a close inspection of the syntactic behaviour of this lexical item reveals that it has not been so quick to grammaticalise, having retained some of the features of its adverbial etymon. An example of this relative freedom is provided in (27) below.

(27) *ai garra deigimwei ol detlot ting blanga yumob weya meigimbat*
1S NEC remove all the.PL thing DAT you that make.CONT
yumob wori ***oldei***.
2P worry all.day
'I'll remove all those things that make you worried' [KB: Ais 54.14]

Additionally, *oldei* occurs in the corpus with roughly 32% of the frequency of *garra* and 13% of *bin*. In addition to its apparent syntactic freedom, it is likely that it provides significantly less collocational cue of grammatical subjecthood than these other auxiliaries.

These frequency of *collocation* between *ai* and the auxiliaries and other high-frequency lexical items discussed above would appear to provide a non-semantically motivated confound that deceptively appears to exclude past- and future-marked clauses from co-occurrence with mi, much in the same way as English varieties with invariant *"…and I"* independent of grammatical relation (24c). Additionally, this appeal to high frequency may explain the unexpectedly low occurrence of *mi* with some frequent psych predicates presented in Table 10 above (e.g. *wandim* 'want', *sabi* 'know'). For many speakers, the occurrence of high frequency items in the verb phrase 'protects' the selection of English-like *'ai'* in what appears to be an

'incomplete' decreolisation. That is, similarly to Emonds' "deviant prestige constructions" (e.g. the formalisation of English *'…and I'* collocation in Parrott (2006, 2009)),[36] this component of *ai*'s behaviour is a function of semantically-unmotivated, distributional-collocational (Emonds' "extra-grammatical") rules which have been learned by speakers targeting an English-like grammar. They do not point to (and indeed override) the semantic composition of $1s_{\text{AGENTIVE}}$ + PREDICATE.

4.4 Concluding remarks and future directions

We have cause, on the basis of this corpus study, to posit that the *ai~mi* distribution in Roper Kriol may be evidence of an emergent and partially-context dependent grammaticalisation of agentivity in the Kriol pronominal domain.

The emergence of such a distinction, whereby a partition of the semantic space is grammaticalised, is consistent with McWhorter's proposal of diachronic accretion of ornamentality in creole languages:

> …much grammaticalization and reanalysis simply gives overt manifestation to underlying semantic distinctions which, while compatible with Universal Grammar and human language processing abilities, are not required by it… (2001: 4)

We have seen evidence that the patientive-semantics frequently associated with the English object pronoun have provided a conduit through which *mi* has entered the exponence domain of '1s subject' in Kriol. It is presently obligatory in copular constructions and appears to be gradually expanding into the a semantic space that might be described as '1s experiencer/undergoer' – evidence that *ai~mi* does not expone a defective case system but rather encodes semantic and pragmatic information. This 'encroachement' is visualised in Figure 4.

The appearance of *ai* in the Kriol lexicon and the weakness of English morphological case cues has permitted this reanalysis/interrupted decreolisation, underpinned by universal and substrate sensitivities to agentivity.

A better understanding of the interaction of pragmatic and discourse features in pronoun selection is likely a fruitful area of future research, which promises to shed further light on the semantic import of grammatical relations and case marking and the diachronic emergence of these linguistic categories.

36. Parrott provides additional evidence of similar prescription/frequency-based effects obtaining in Danish (2009: 284ff).

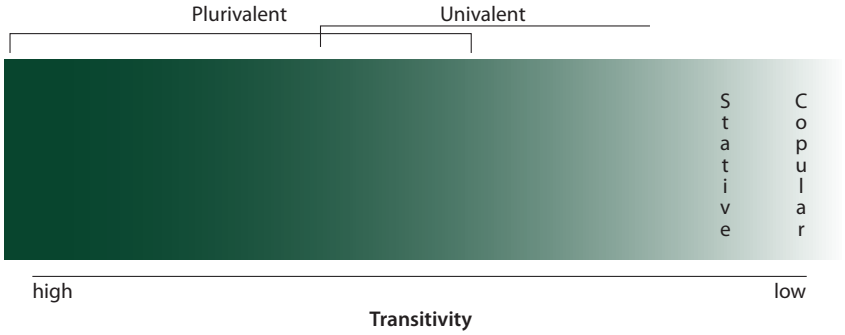

Figure 4. Clause types in Kriol. Darker areas of the diagram indicate a lower likelihood of the emergence of *mi* as a 1s subject pronoun

It is of course problematic to predict the eventual catagoricalisation of an explicit marking system of agentivity given the unpredictability of the outcomes of competition between communicative strategies[37] and given the degree of language contact between Kriol and English (continued decreolisation and other restructuring phenomena associated with sustained language contact), coupled with the wide dialectal variation and the low prestige that are associated with Kriol.

Independent, however, of the future of this alternation, a convincing demonstration of the structured variation between erstwhile subject and object pronouns stands as an interesting example of grammaticalisation with implications for our theories of language change under contact and the dynamics of encoding a gradient, but highly discourse-salient semantic primitive such as transitivity. It also provides clues into the cognitive, reanalytic processes that are likely implicated in the emergence of and transitions between 'active' and accusative morphosyntactic alignments.

Acknowledgements

I'd like to acknowledge Claire Bowern for her essential advice on project design, corpus design, analysis design and helpful comments in this process. I also thank Ashwini Deo for her feedback, useful suggestions at various points throughout the life of this project and the extended discussions about grammaticalisation and semantic change. Thanks also due to Maïa Ponsonnet, Felicity Meakins, Jim Wood, Martín Fuchs & Kevin Tang for additional support, feedback and discussion. Additionally, I owe much to Ngukurr Language Centre and my consultants for their time and patience. All errors, of course, remain my own.

37. See Deo 2015 for a treatment of this trajectory and a game-theoretic modelling of features contributing to the success or failure of innovative forms.

References

Angelo, D., & E. Schultze-Berndt. 2016. 'Beware bambai – lest it be apprehensive.' In Felicity Meakins & Carmel O'Shannessy (eds.), *Language Contact in Australia*, 255–296. Berlin: Mouton de Gruyter.

Bible Society in Australia, The. 2010. Kriol e-Baibul. Retrieved from http://aboriginalbibles.org.au/Kriol/Conc/root.htm

Bickerton, D. 1981. *Roots of Language*. Ann Arbor: Karoma.

van den Bos, J., F. Meakins, & C. Algy. 2017. Searching for "Agent Zero": The origins of a relative case system. *Language Ecology* 1(1). 4–24.

Butt, M., & T. H. King. 2002. Case systems: Beyond structural distinctions. In E. Brandner & H. Zinsmeister (eds.), *New Perspectives on Case Theory* (CSLI Lecture Notes 156), 53–87. Stanford: Center for the Study of Language and Information.

Bybee, J., & Ö Dahl. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13(1). 51–103. doi:10.1075/sl.13.1.03byb

Deo, A. 2015. The semantic and pragmatic underpinnings of grammaticalization paths: The progressive to imperfective shift. *Semantics and Pragmatics* 8(14). 1–52.

Dickson, G. 2016. Alabat tok najawei: surveying geographic variation in Kriol speaking communities east of Katherine. Poster presented at theCoEDL Huddle, WSU: Sydney, NSW.

Dixon, R. M. W. 1972. *The Dyirbal language of North Queensland*. London: London Cambridge University Press. doi:10.1017/CBO9781139084987

Dixon, R. M. W. 1980. *The languages of Australia*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511611896

Dixon, R. M. W. 2002a. Copula Clauses in Australian Languages: A Typological Perspective. *Anthropological Linguistics* 44(1). 1–36.

Dixon, R. M. W. 2002b. *Australian languages: their nature and development*. New York: Cambridge University Press. doi:10.1017/CBO9780511486869

Dowty, D. 1991. Thematic Proto-Roles and Argument Selection. *Language* 67(3). 547–619. doi:10.1353/lan.1991.0021

Emonds, J. 1986. Grammatically deviant prestige constructions. In M. Brame, H. Contreras & F. J. Newmeyer (eds.), *A Festschift for Sol Saporta*. Noit Amrofer Pub Co.

Graber, P. L. 1987. The Kriol particle 'na'. *Working Papers in Language and Linguistics* 21. 1–21.

Harley, H., & E. Ritter. 2002. Person and Number in Pronouns: a feature-geometric analysis. *Language* 78(3). 482–526. doi:10.1353/lan.2002.0158

Harris, J. 1986. *Northern Territory Pidgins and the Origin of Kriol*. Canberra: Australian National University.

Holisky, D. A. 1987. The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua*. 71 (1–4). 103–132. doi:10.1016/0024-3841(87)90069-6

Hopper, P. J., & S. A. Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56(2). 251–299. doi:10.1353/lan.1980.0017

Hothorn, T., K. Hornik, & A. Zeileis. 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics* 15(3). 651–674. doi:10.1198/106186006X133933

Hudson, J. 1983. *Grammatical and semantic aspects of Fitzroy Valley Kriol*. Darwin: Summer Institute of Linguistics, Australian Aborigines Branch.

Johnson, D. E. 2009. Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3(1). 359–383. doi:10.1111/j.1749-818X.2008.00108.x

Lee, J. 1987. *Tiwi Today: A study of language change in a contact situation*. Canberra: Pacific Linguistics.

Levin, B. 1993. *English verb classes and alternations: a preliminary investigation*. Chicago: University of Chicago Press.

McConvell, P., & F. Meakins. 2005. Gurindji Kriol: A Mixed Language Emerges from Code-switching. *Australian Journal of Linguistics* 25(1). 9–30. doi:10.1080/07268600500110456

Meakins, F. 2009. The case of the shifty ergative marker: A pragmatic shift in the ergative marker in one Australian mixed language. In J. Barddal & S. Chelliah (eds.), *The Role of Semantics and Pragmatics in the Development of Case* (Studies in Language Companion Series 108), 59–91. Amsterdam: John Benjamins.

Meakins, F. 2011. *Case-Marking in Contact: The development and function of case morphology in Gurindji Kriol*. Amsterdam; Philadelphia: John Benjamins. doi:10.1075/cll.39

Meakins, F. & C. O'Shannessy. 2010. Ordering arguments about: Word order and discourse motivations in the development and use of the ergative marker in two Australian mixed languages. *Lingua* 120(7). 1693–1713. doi:10.1016/j.lingua.2009.05.013

McGregor, W. B. 2008. Indexicals as sources of case markers in Australian languages. In F. Josephson & I. Söhrman (eds.), *Interdependence of Diachronic and Synchronic Analyses*, 299–232. Amsterdam: John Benjamins. doi:10.1075/slcs.103.15mcg

McGregor, W. B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120. 1610–1636. doi:10.1016/j.lingua.2009.05.010

McWhorter, J. H. 1998. Identifying the Creole Prototype: Vindicating a Typological Class. *Language* 74(4). 788–818. doi:10.2307/417003

McWhorter, J. H. 2001. The Rest of the Story: restoring pidgins to creole genesis theory. *Journal of Pidgin & Creole Languages* 17(1). 1–48. doi:10.1075/jpcl.17.1.02mcw

McWhorter, J. H. 2005. *Defining Creole*. New York: OUP.

McWhorter, J. H. 2007. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Oxford: OUP. doi:10.1093/acprof:oso/9780195309805.001.0001

Næss, Å. 2007. *Prototypical Transitivity*. Amsterdam: John Benjamins. doi:10.1075/tsl.72

Nicholls, S. 2011. *Referring Expressions and Referential Practice in Roper Kriol*. Armidale, NSW: University of New England PhD Thesis.

Parrott, J. K. 2007. *Distributed Morphological Mechanisms of Labovian Variation in Morphosyntax*. Washington, D.C.: Georgetown University Ph.D. Dissertation.

Parrott, J. K. 2009. Danish Vestigial Case and the Acquisition of Vocabulary in Distributed Morphology. *Biolinguistics* 3(2–3). 270–304.

Phillips, J. 2011. *Kriol and Kriolisation: exploring the creole language of Northern Australia*. Sydney: University of New South Wales Unpublished Honours thesis.

Ponsonnet, M. 2016. Emotion middle predicates in Barunga Kriol. Paper presented at the *Australian Languages Workshop*. Canberra; Kioloa, March 2016.

Ponsonnet, M. 2018a. Expressive values of jyoti reduplication in Barunga Kriol (northern Australia). *Studies in Language* 42(1). doi:10.1075/sl.00009.pon

Ponsonnet, M. 2018b. Lexical semantics in language shift: Comparing emotional lexica in Dalabon and Barunga Kriol (northern Australia). *Journal of Pidgin & Creole Linguistics* 33(1).

Sandefur, J. R. 1979. *An Australian Creole in the Northern Territory: a description of Ngukurr-Bamyili dialects, part 1*, vol. 3. Darwin: Summer Institute of Linguistics, Australian Aborigines Branch.

Sandefur, J. R. 1982. Kriol and the question of decreolization. *International Journal of the Sociology of Language* (36). 5–13.

Sandefur, J. R. 1985. Aspects of the Socio-political History of Ngukurr (Roper River) and its Effect on Language Change. *Aboriginal History* 9(2). 20.

Sandefur, J. R. & J. Harris. 1986. Variation in Australian Kriol. In J. A. Fishman, A. Tabouret-Keller, M. Clyne, B. Krishnamurti, & M. Abdulaziz (eds.), *The Fergusonian impact: in honor of Charles A. Ferguson on the occasion of his 65th birthday. Contributions to the sociology of language*, 180–90. Berlin: Mouton de Gruyter.

Schmidt, A. 1985. *Young People's Dyirbal: an example of language death from Australia*. Cambridge: Cambridge University Press.

Sharpe, M. C. [As Margaret Clare Cunningham]. 1969. *Alawa Phonology and Grammar*. Brisbane: Brisbane University of Queensland PhD Thesis.

Schultze-Berndt, E., F. Meakins, & D. Angelo. 2013. Kriol. In S. M. Michaelis, P. Maurer, M. Haspelmath, & M. Huber (eds.), *Atlas of Pidgin and Creole Structures, vol. I: English- based and Dutch-based Languages*, 241–251. Oxford: OUP.

Schultze-Berndt, E., M. Ponsonnet, & D. Angelo. (forthcoming). *The semantics of modal markers in North Australian Kriol* [manuscript in preparation]. Language Science Press.

Schütze, C. T. 2001. On the Nature of Default Case. *Syntax* 4(3). 205–238. doi: 10.1111/1467-9612.00044

Siegel, J. 2008. *The Emergence of Pidgin and Creole Languages*. New York: Oxford University Press.

Silverstein, M. (1986 [1976]). Hierarchy of features and ergativity. In P. Muysken & H. C. van Riemsdijk (eds.), *Features and Projections*, 112–170. Amsterdam: Walter de Gruyter.

Traugott, E. 1999. The role of pragmatics in semantic change. Paper presented at the Pragmatics in 1998: Selected Papers from the 6th International Pragmatics Conference.

Troy, J. 1990. *Australian Aboriginal Contact with the English Language in New South Wales: 1788–1845*. Canberra, ACT: Pacific Linguistics.

Van Valin, J., D. Robert, & D. P. Wilkins. 1996. The Case for 'Effector': Case Roles, Agents, and Agency Revisited. In M. Shibatani & S. A. Thompson (eds.), *Grammatical Constructions: Their form and Meaning*. Oxford: Clarendon Press.

Appendix A

Wal	*Aisik*	*bin*	*tok,*	*"**Mi***	*brabli*	***olmen***	*na,*	*en*	***mi***
DM	NAME	PST	talk	1s	very	old(man)	SEQ	and	1s
gulijap	***binij.***	*Wal*	*yu*	*go*	*hanting*	*garram*	*yu*	*bo*	*en*
almost	finish	DM	2s	go	hunting	with	2s	bow	and
erro	*en*	*kilim*	*bif*	*blanga*	*mi,*	*dumaji*	***mi***	*oldei*	***laigim***
arrow	and	kill	meat	DAT	1s	because	1s	IPFV	like
dagat	*detkain*	*bif.*	*En*	*yu*	*bringimbek*	*det*	*bif*	*en*	*gugum*
eat	that.kind	meat	and	2s	bring.TR.back	the	meat	and	cook.TR
blanga	*mi,*	*en*	*afta*	*wen*	***ai***	***binij***	***dagadagat***	*det*	*bif,*
DAT	1s	and	after	when	1s	CESS	eat~DUR.TR	the	meat
wal	***ai***	***garra***	***gibit***	*langa*	*yu*	*main*	*laswan*	*wed,*	*en*
DM	1s	NEC	give.TR	LOC	2s	my	last.ATTR	word	and
ai	***garra***	***hendimoba***	*ola*	*ebrijing*	*blanga*	*mi*	*langa*	*yu*	*bifo*
1s	NEC	hand.TR.over	the	everything	GEN	1s	LOC	2s	before
mi	***dai.***								
1s	die								

'And Isaac said "I'm very old now and I'm close to death. You go out hunting with your bow and arrow and kill some game for me because I like to eat that type of meat. Bring that meat back and cook for me and after when I finish eating the meat, I'll give you my blessing and I'll hand over all of my things to you before I die. [KB Gen27.2–4]

Author's address

Joshua Phillips
Yale University
Linguistics
370 Temple Street
New Haven, 06511
USA
joshua.phillips@yale.edu