

1 Introduction

This thesis aims to describe substrate influence in Kriol, a creole language of Northern Australia. A method that has been trialed only twice before (see Siegel 1999; Siegel, Sandeman and Corne 2000), the Transfer Constraints approach (see §1.4), is used to investigate the processes involved. This thesis also, therefore, aims to determine its suitability for creolistic studies. Following a discussion of the sociohistorical context of creole emergence, the possibility of feature transfer is considered for verb and nominal features, as well as semantic category and nominal modification features.

Kriol is spoken by approximately 20,000 Aboriginal people across Northern Australia and as such has more Aboriginal speakers than any other language in Australia, bar English. It is a contact language that developed and emerged from contact between English speaking colonisers, or the superstrate language speakers, and the Indigenous traditional owners of the land, or the substrate language speakers. The Kriol speaking region (see map 1) spans a large geographic area that has spread over the language territories of at least 35 Indigenous languages. In order to maintain the viability of the comparative analyses of substrate languages it has been necessary to limit the scope of this thesis, which therefore concentrates on the variety of Kriol known as Roper Kriol and the substrate languages spoken in the Roper River region of the Northern Territory.

This chapter begins by introducing the Roper River region and its languages in §1.1. The background information to the thesis, such as the fieldwork, consultants and data, as well as orthographic conventions used in this thesis for both Kriol and the substrate languages, is presented in §1.2. A discussion of relevant previous research follows in §1.3, which necessarily includes a review of relevant research in creolistics, including research on Kriol in Australia. The theoretical construct employed for this thesis, the Transfer Constraints approach, is presented in §1.4, followed by the chapter outline in §1.5.

1.1 The Roper River region and its languages.

The Roper River is a major river system in the gulf ‘country’¹ of the Northern Territory (see map 2). It is fed by freshwater springs in Mangarrayi country, winding its way east through Ngalakgan country it grows in size until reaching Roper Bar. This is where the saltwater and tides enter the system and is a sacred area for the Ngalakgan people. Further east it is fed by the Wilton River, flowing down from Arnhem Land through Rembarrnga and Ngalakgan country and further on feeds the Hodgson River, which flows south through Alawa country. As it continues to wind its way east, now salty and tidal it forms a natural boundary for the Ngandi on the northern side and the Marra on the southern side. Towards the coast it nears the Cox River and Limmen River delta to the south, in Marra country, and the Rose River to the north, in Nunggubuyu country. The most easterly sections flow through Warndarrang country, including the mouth of the river at Port Roper, where it flows into the Gulf of Carpentaria.

The Roper River region is similar in some ways to other large river systems in the Northern Territory but is also unique to others. As with all such river systems, the Roper River is a rich source of food and resources, particularly the barramundi fish, which continues to be well utilised by the Indigenous inhabitants. Such a rich food source allows for greater population numbers and a complex ceremonial tradition. Where this region differs, however, is in its contact history. Around the turn of the 1900s it was the central staging post for the colonisation of the Top End of the Northern Territory. The sheer size of the river resulted in it acting as a cargo and supply route. The land directly to the south was not only the major stock route for the pastoral industry invasion, but also the major route for all traffic into and out of the Northern Territory at that time (e.g. Harris 1986) (see §1.1.1 and chapter 2 for further discussions of this topic).

¹ In Australian English ‘country’ can be used to refer to an area of land dominated by a particular geographic feature, in this case the Gulf of Carpentaria. It can also be used when the land area is dominated by a particular language or ethnic group.

Map 4 shows the contemporary nature of the Roper River region. Ngukurr is the largest Aboriginal community in the region and is recognised as its centre. This is in part due to its historical origins, as the Roper River Mission opened in 1908 by the Church Missionary Society first at Mirlingbarrwarr, but after the 1940 flood moved to its present site of Ngukurr (see map 3). The mission is credited with the spread of Christianity, which remains an important element of the spiritual landscape of the Roper River region (see e.g. Harris 1990). Ngukurr is also important because of the large number of language groups now represented, which number at least nine. Thus, Ngukurr residents have close relationships with numerous other communities where members of the same language groups reside. It is also the administrative centre for other smaller communities. Minyerri community is at present growing rapidly, although its origins lie in the station camp associated with Hodgson Downs Station, and it is particularly associated with Alawa people (see §2.5 for a discussion of station camps and §2.6 for their development into communities). Similarly, Urapunga was originally the station camp for Urapunga Station with association to Ngalakgan people and Jilkminggan was originally on Elsey Station with association to Mangarrayi people. All these communities have had land title handed back to the traditional owners.

1.1.1 Roper Kriol

Kriol is an English lexified creole language, which is a product of the invasion of the pastoral industry into the Northern Territory and Western Australia around the turn of the 1900s (e.g. Harris 1986; Munro 2000). The large area in which it is spoken (see map 1) is also therefore predominantly the same as the territory covered by the pastoral industry in the northern reaches of these two states. There are therefore numerous regional varieties of Kriol, whose differences depend on the substrate languages of each region and the various states of language shift within that region.

Roper Kriol is one such regional variety of Kriol, which is spoken in the Roper River region including the communities of Ngukurr, Minyerri, Urapunga and Jilkminggan (see map 4). It should be noted that the name, Roper Kriol, differs in this thesis from its previous use representing Kriol in general (see e.g. Dixon 1980:72). Roper Kriol here is identified as the oldest variety of Kriol in Northern Australia (see e.g. Sandefur 1986), which has also received the greatest amount of research compared to other varieties (Sandefur 1979; Harris 1986; Sandefur and Harris 1986). It has also been noted that Indigenous people throughout the Northern Territory identify Kriol with the Roper River region (Jernudd 1971) and that those from the region identify themselves via their Kriol language (Munro 1999a), which is not common in other Kriol speaking regions.

The substrate languages of Roper Kriol are those of the Indigenous language groups that maintain custodial relationships to their land, otherwise known as country, in the Roper River region: Alawa, Marra, Ngalakgan, Wandarrang, Mangarrayi, Ngandi and Nunggubuyu. The respective countries of these groups were among the first in the Northern Territory to be stocked and settled by the encroaching pastoral industry.

The superstrate language of Kriol is English, which is the official language of Australia as well as the colonising pastoralists. Autobiographical accounts, such as Costello (1930), however, suggest that the English in use in the Roper River region around the turn of the 1900s was variable in terms of regional accents and proficiency.

It has been shown that from the 1820s the pastoral industry spread progressively from New South Wales, moving into and through Queensland in the 1840s before entering the Northern Territory in the 1870s (Dutton 1983; Troy 1994; Munro 2000). A contact language had been developing throughout this spread of the pastoral industry to facilitate communication between Aboriginal employees and pastoralists, generally referred to here as New South Wales/Queensland Pidgin (NSW/QLD Pidgin) (Dutton 1983; Troy 1994; Munro 2000). The NSW/QLD Pidgin also therefore entered the Roper River region during the pastoral invasion.

A complete discussion of the sociohistorical context of creole emergence is presented in chapter 2, although some points are worth raising here. Firstly, all the languages previously mentioned (the substrate languages, superstrate language and previous contact language) provided input to the contact language that developed in the Roper River region as a result of the establishment of the pastoral industry from the turn of the twentieth century. Harris (1986) describes this contact language as Northern Territory Pidgin (NT Pidgin) and the same label is also used in this thesis. Harris (1986) further claims that while NT Pidgin developed in the pastoral stations, a creole emerged abruptly in the Roper River Mission (see §1.3.2). It will be shown in chapter 2, however, that if the sociohistorical nature of the surrounding pastoral stations is taken into consideration alongside that of the Roper River Mission, then a more appropriate explanation is that the creole emerged gradually from the entire region, through the development of the NT Pidgin. The fact that most, if not all, elderly consultants report a stage of bilingualism in the contact language and their traditional languages, during their childhood, including those from cattle stations as well as from the Roper River Mission, supports this claim (see chapter 2).

Roper Kriol is currently spoken by at least 950 people based on a language survey carried out in Minyerri and Ngukurr (Lee and Dickson 2003:68). It is spoken by all four to five generations in these two communities, as well as in Urupunga and Jilkminggan and by people from the Roper River region living in towns such as Mataranka, Katherine and Darwin. While it is generally recognised that there is some level of generational variation within Roper Kriol, research has yet to be carried out on the topic. Roper Kriol has been used extensively in the bible translation project undertaken by the Summer Institute of Linguistics. While Ngukurr rejected the offer of the implementation of a bilingual program at the Ngukurr Community Education Centre in the 1970s, opinions towards the language generally remain positive within the region (Munro 1999a). Although it had been expected that the use of Roper Kriol would eventually give way to English, or in other words that decreolisation would occur (Sandefur 1986); this has not been the case. In fact Roper Kriol maintains a vital status in the region.

1.1.2 The substrate languages

Every classification model of Australian Indigenous languages suggests that all such languages descend from Proto–Australian (see e.g. Evans 2004). Using the ‘Pama–Nyungan offshoot model’ of classification, based on the work of O’Grady (1979), Evans (2004) suggests that one relatively recent descendent is Proto–Pama–Nyungan from which all Pama–Nyungan (PN) languages are themselves descended. PN languages cover the majority of the Australian continent (see map 1) and comprise a relatively homogenous group (Blake 1988:3). PN languages, therefore, have been identified through lexicostatistical methods as forming one language family (see e.g. Blake 1988; Evans 2004). Within this classification model the following is assumed:

...prior to Pama-Nyungan expansion, the level of genetic diversity across the presently Pama-Nyungan part of the continent was comparable to that now found in the non-Pama-Nyungan area, and ... widespread language shift to Pama-Nyungan then took place, leaving at most some substrate influence from the earlier languages (Evans 2004:9).

As this suggests there is a small pocket of other older ‘daughter-level’ descendents of Proto–Australian classified as non–Pama–Nyungan (nonPN) languages (see map 1), which are identified as ‘...containing 90% of Australia’s linguo-genetic diversity in an eighth of its land area’ (Evans 2004:3). Some nonPN languages can be subgrouped into language families, such as the Gunwinyguan language family, while others are family level isolates, which are single members of separate language families. They are all referred to as nonPN languages, which tend to be agglutinating, and non-configurational, encoding structural relationships morphologically rather than syntactically. As a consequence, word order often has little or no significance in terms of grammatical relations (see e.g. Hale 1983).

As described in §1.1.1 the substrate languages for Roper Kriol are: Alawa, Marra, Ngalakgan, Wandarrang, Mangarrayi, Ngandi and Nunggubuyu. They are all classified

as nonPN languages. The Roper River region, like other nonPN language areas, therefore, is comprised of a number of distinct and diverse languages. Having said that, there are family level relationships and typological similarities between these languages (see e.g. Heath 1978b; 1987; Baker and Harvey 2003; Baker 2004; Evans 2004). The current understanding of the nonPN Roper River substrate languages is that they generally belong to two language families:

Marran	Alawa, Marra and Warndarrang
Gunwinyguan	Ngalakgan, Nunggubuyu and Ngandi.

The classification of Mangarrayi remains uncertain. Heath (1978b:8) notes, for example, that more data is required to categorise its relationship to the Marran languages. Merlan (2004) attempts to do this by providing a comparative analysis of noun class marking (see chapter 4 for an analysis of this feature) between the Marran languages and Mangarrayi, suggesting that Mangarrayi may be descended from the same Proto-Marran but has been more innovative. The current status sees it swinging in classification between either Gunwinyguan (Evans 2004) or Marran (Merlan 2004).

The Marran languages are generally presented as a family level category and Merlan (2004) attempts to provide evidence of this. However, Heath (1978b:7) says: ‘Although ... these languages constitute a genetic subgroup, and hence descend from a proto-language ... it should be made clear that the three attested languages [Alawa, Marra and Warndarrang] are rather divergent from each other, more so than Ngandi and Nunggubuyu’. The greatest similarities lie in the shared lexicon and morphology (Heath 1978b:8). Heath (1978b:9) suggests that the only way to really determine the genetic relationship between the three languages would be to ‘...determine the common origin of their verbal inflectional systems...[which] is difficult, since ... each preserves only a few directly inflectable verb stems, and those paradigms which have survived [are altered]...’. At this stage, therefore, while current research, such as Evans (2004), refers to the Marran language family, it is generally recognised that the three languages may in fact be genetically divergent (see e.g. Baker 2004).

The Marran languages are currently in different states of vitality. A recent survey was carried out by Diwurrurru-jaru Aboriginal Corporation² (DAC) as to the vitality of languages in the Katherine region of the Northern Territory, including the Roper River region (Lee and Dickson 2003). Of the 264 people surveyed in Ngukurr and surrounds who identified as Marra, 23 claim to speak the language fluently (Lee and Dickson 2003:42-43)³. Marra language speakers reside in Ngukurr, Minyerri, Borroloola and Numbulwar (see map 4). The Alawa population largely resides in Minyerri (see map 4) although there are also important Alawa families in Ngukurr. With 435 people recently surveyed claiming Alawa identity, it can be seen as one of the strongest languages in the region; of those, 37 claim to speak the language fluently (Lee and Dickson 2003:28).

Two of the Gunwinyguan languages are in different states of vitality. Nunggubuyu⁴ is the only language that continues to be learnt by children as their first language (L1), particularly in Numbulwar (see map 4), although a large population live in Ngukurr. The survey cited throughout this section centred on Ngukurr and did not include Numbulwar, therefore, the numbers of those who identify as Nunggubuyu far exceed the 363 listed and similarly those who speak the language far exceed 44 (Lee and Dickson 2003:54). Ngalakgan speakers predominantly live in Urupunga (see map 4) and Ngukurr, and of the 233 people recently surveyed who identify as Ngalakgan, 15 claim to speak the language with fluency (Lee and Dickson 2003:48). As previously mentioned, both Nunggubuyu and Ngalakgan belong to the growing Gunwinyguan language family, along with Ngandi which is most closely related to Nunggubuyu. The primary means for identifying this relationship is the comparative analysis of the verb inflectional morphology provided by

²Diwurrurru-jaru Aboriginal Corporation was formerly the Katherine Regional Aboriginal Language Centre. It is managed by an Indigenous committee who maintain control of language management issues, such as survey reports, orthographic issues and generally overseeing linguistic research in the region.

³ It should be noted that all the figures from this source may be slightly overestimated as Aboriginal people often confuse such questions as, 'what language do you speak?' with 'what is your ancestral language?' or 'which language(s) do you know a little of?'. There is therefore confusion between language identity and passive knowledge with speaking ability.

⁴ It should be noted that the language name is in fact *Wubuy* and that *Nunggubuyu* refers to the people who speak *Wubuy*. *Nunggubuyu* is used here to follow the convention that has become established in the literature.

Heath (1978b) (see also Baker 2004). In a similar way Ngalakgan is most closely related to Rembarrnga, situated north of the Roper River region (see map 2).

Warndarrang is now classified as extinct (Lee and Dickson 2003:66) based on the work of Heath (1980a). Although there are one or two living people who claim to speak Warndarrang⁵, 135 identify with the language group, which is small in number compared to other language groups (Lee and Dickson 2003:66). It is assumed that the more rapid language loss reflects originally low population numbers of the Warndarrang speakers, particularly in the early stages of language contact. While 309 people identified as Ngandi, there are only an estimated 9 speakers (Lee and Dickson 2003:50). Ngandi is, therefore, the next most threatened language in the Roper River region. Neither Warndarrang nor Ngandi are included in the comparative analyses in chapters 3-6, primarily because Warndarrang shares many typological features with Marra, and Ngandi is closely related to Nunggubuyu (see e.g. Heath 1978b). The low numbers of Warndarrang and Ngandi speakers also implies that these groups had minimal impact on language contact (see chapter 2).

Yugul is another extinct language that was spoken around the Roper River region and whose status is somewhat uncertain. Heath (1980a:1), for example, says that Yugul, ‘...is said to have been similar to Marra’. Baker (2002b:106-107) then suggests that Yugul was another distinct language of the Roper River region with country associated with it (see chapter 2 for a discussion of land/language relationships). Alternatively a consultant for this thesis, DG, describes Yugul as a lingua franca for the region and that the current name of the Ngukurr council, *Yugul Mangi*, means the Roper River groups combined. DG (D01:1, 26/3/01) says that her mother spoke to her in Yugul and describes it in the following way: ‘Like you know, in between, in between language now, that one, because other people couldn’t talk the other languages, but everyone understood Yugul’. As there are no records of this language and consultant information can differ markedly, there are no definitive answers available.

1.2 Background to the research

1.2.1 Fieldwork

During the years 1996-2000 I worked in various capacities as a linguist in the Roper River region, particularly in and around the communities of Ngukurr and Minyerri. The majority of the work was carried out on either language revitalisation programs in these communities and their schools under the direction of DAC, or on the St. Vidgeon Pastoral Lease Native Title Claim under the direction of the Northern Land Council. All these projects involved the substrate languages described in §1.1.2. In the latter years I was increasingly turning my attention to projects involving either cultural knowledge of country to promote language revitalisation, or oral history to promote Kriol literacy. All these experiences led me to believe that there is an intricate relationship between the substrate languages and Kriol, which therefore in turn contributed to this thesis.

The particular fieldwork for this thesis first included an Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) funded project in 1998. This was process driven research that aimed to determine the most effective means to encourage young members of Minyerri community to actively participate in language revitalisation of Alawa. It was found, for example, that young people first required confidence in handling their own L1 and to this end Kriol literacy was vital. Oral history was only relevant in Kriol, not Alawa, as they had no, or little, proficiency with the latter. A series of tapes was collected in Kriol from senior members of the community during this project, with particular emphasis on training young members of the community, who either collected, or assisted collection, of the tapes. In this way the Kriol spoken is as close to everyday spoken Roper Kriol in Minyerri as could be achieved. The bulk of the

⁵ These figures were first found by the author in meetings and interviews with consultants during the St. Vidgeon Pastoral Lease Native Title Claim in 1999. Lee and Dickson (2003:66) confirm them.

texts are life stories of the speaker's involvement in the pastoral industry, which necessarily includes travelling stories. These tapes, numbered 0198-0798, are lodged in AIATSIS.

The second stage of fieldwork was carried out in conjunction with the second part of this project in 2001. After transcription of the tapes, workshops were conducted at Minyerri on Kriol literacy and editing procedures. Young members of the community edited, and sometimes rewrote, the oral history stories collected during the initial part of the project. A series of six booklets of life stories of senior Minyerri people, all written in Kriol, were produced and distributed in this year and were lodged in AIATSIS. The fieldwork component of this work involved checking transcriptions and discussing the topic of this thesis with consultants and senior members of Ngukurr and Minyerri. Approval from both the Ngukurr Language Centre sub-committee of DAC and members of Minyerri community was given for this research during this time.

My final fieldtrip was carried out in June 2003, with two weeks in Minyerri and one in Ngukurr. Direct elicitation was carried out with speakers in order to clarify analyses already undertaken. A series of nine mini-disks, numbered 2003:01–2003:09 were collected. Some songs and life stories were also collected in Kriol, although the bulk of material is elicitation.

1.2.2 Consultants

The details of the primary consultants for this thesis are summarised in Table 1, below. All of these consultants provided initial texts in Kriol as well as participating in all levels of fieldwork, from checking transcripts to elicitation. They are residents of Minyerri community and most are at least over sixty. This age bracket was chosen for two reasons. Firstly, in order to minimise the impact of generational variation, only two age groups were focussed on. Secondly, there are representatives of the bilingual phase of

contact language development, which is when feature transfer from the substrate languages to the contact language is most likely to occur (see §1.4). Thirdly, there are also representatives of the first generation of creole speakers.

Table 1: Consultant speaker profiles

Initials	YOB ⁶	Place of residence(s)	Childhood languages	Adult languages
AS	1920s	Nutwood Downs Station Hodgson Downs Station	Marra English Alawa NT Pidgin	Kriol Alawa Marra English
DG	1923	Roper River Mission St. Vidgeon Station Elsey Station	Yugul/Nunggubuyu NT Pidgin	Kriol English Marra (passive) Nunggubuyu (passive)
HW ⁷	1920s	Hodgson Downs Station	Alawa NT Pidgin	Kriol Alawa
SR	1927	Roper River Mission Hodgson Downs Station McArthur River Station	Alawa NT Pidgin	Kriol English Yanyuwa Alawa
CW	1930s	Roper River Mission Nutwood Downs Station Hodgson Downs Station	Marra NT Pidgin	Kriol Alawa Marra (passive) English
DW	1940s	Hodgson Downs Station Ngukurr	Kriol Alawa	Kriol Alawa
SL ⁸	1940s	Limmen Bight Station Nutwood Downs Station Hodgson Downs Station	Kriol Marra Alawa	Kriol Alawa Marra (passive) English

As Table 1 shows, most of the consultants learnt both the contact language, NT Pidgin, and their traditional language during childhood and can also, therefore, speak Alawa and/or Marra, although now with varying proficiency. Most of these consultants were, therefore, raised with bilingualism, which will be more fully described in chapter 2. Some of the consultants spent some time in the Roper River Mission in their childhood and some were raised solely on cattle stations (see maps 3 and 4). All of the consultants, however, spent their adult lives working in the cattle industry in the Roper River region.

⁶ Approximate Year of Birth

⁷ Unfortunately this *olmen* (respected/senior man) has now passed away.

⁸ Unfortunately this *olmen* (respected/senior man) has now passed away.

The language of this industry is undoubtedly the contact language as either NT Pidgin or Kriol. There are also two consultants who acquired Kriol as their L1 and learnt their traditional language as a second language (L2). It would have been necessary, therefore, for their parents, who are peers of the other consultants, to have used the pidgin/creole for their day-to-day lives, in order to communicate with their children and to operate in the pastoral industry. The same is assumed of the first four listed consultants throughout the majority of their adult lives. In this way the contact language became their first, or primary, language and their use of traditional languages waned, either in proficiency or opportunity for use.

There are also a number of other consultants from Minyerri. Some provided Kriol texts in early fieldtrips. Others participated in elicitation in the final fieldtrip, including two assistant fieldworkers. There are also a number of other consultants from Ngukurr, who all provided texts, discussions and approval for the project, at different times.

1.2.3 Data

The Kriol data comes largely from the tapes collected in 1998. Each tape has been fully transcribed with interlinear glosses using the application ‘Shoebox’ developed by the Summer Institute of Linguistics. Where examples from these texts are provided in the thesis, they take the first letter of the first name of the consultant who provided it and the line number from my transcripts, such that (D054), for example, is line no. 54 from the text of DW. Consultant DG did not provide texts of this sort for analysis, so there can be no confusion between DG and DW. The minidisks collected in 2003 also provided some data for the thesis, particularly on points that required clarification. Some of these have been transcribed using the software ‘Sound Edit 16’ for the sound file and ‘ClanX’. Examples taken from these sources take the first letter of the first name of the consultant, followed by the minidisk number and the track number, such that (C03:8T11) indicates that consultant CW provided the data on minidisk no. 03:08 on track 11.

A full discussion of the literature available on Kriol will be provided in §1.3.2, although there are two key references that have provided examples in the thesis. They are:

Sandefur 1979

Hudson 1983a

There are two reasons why Sandefur (1979) was not used as primary data. Firstly, as Kriol is a relatively new language, significant language change may have taken place in the 25 years since it was published. While the consultants are elderly this does not exclude them from effects of language change. Secondly, there are a number of features not covered by Sandefur (1979) that required analysis in this thesis. While much of the analysis in Sandefur (1979) has been proven correct, it has become obvious that expanded analyses are required on Kriol. Hudson (1983a) was used primarily as comparative data because it is a description of an alternative regional variety, Fitzroy Kriol as spoken in the Fitzroy Crossing region of the Kimberleys in Western Australia, and is consequently inappropriate as data for Roper Kriol.

The data for the substrate languages comes almost exclusively from the grammatical descriptions available:

Alawa, Sharpe 1972

Marra, Heath 1981a

Ngalakgan, Merlan 1983

Ngandi, Heath 1978a

Nunggubuyu, Heath 1980b and 1984

Warndarrang, Heath 1980a

There were also two alternative sources used, namely Baker (2002a) for Ngalakgan and Sharpe (2001) for Alawa. It will become apparent throughout the body of the thesis that each linguist favours different terminology and grammatical descriptions of features. As far as was possible each feature discussed in this thesis has been compared and re-analysed with a consistent description. The same has been applied to the orthography, which is presented in §1.2.5. The examples of substrate data are usually taken from the body of the respective grammatical descriptions, although in some circumstances

alternative or more appropriate examples were sought from the texts provided for each language.

1.2.4 Kriol orthography

The original orthography for Kriol was initially developed by linguists in the early 1970s, (Sandefur 1979:60) predominantly to begin the task of the Kriol bible translation project, undertaken by the Summer Institute of Linguistics, that is now in its final stages. There had also been various workshops and consultations with speakers and Aboriginal teachers in the late 1970s before settling on a standard orthography (Sandefur 1979:61). That orthography has obviously served its purpose but there have been continued calls to reassess orthographical issues in order to increase access to Kriol literacy among Kriol speakers, which remains low. It is generally understood that for this to happen Kriol speakers must feel a sense of ownership over the language and that the chances of this would be increased if they were involved in orthographic decisions from the outset. Currently, there are initial steps being taken to develop a unique orthography for the regional variety of Kriol spoken in the Kimberleys in Western Australia, through the Kimberley Language Resource Centre. This initiative focuses on such inclusion and results so far are proving to be positive (p.c., S. Casson, 12/02/04). The only sustained literacy work involving Roper Kriol has been undertaken by the Summer Institute of Linguistics, who continue to use the standard orthography. No steps to change the orthography have been taken for Roper Kriol.

The orthography devised for Roper Kriol in this thesis differs only marginally from that devised by Sandefur (1979). While it is only intended to accurately present the data collected for this thesis, it may suggest minor alternatives for orthographic revisions by Roper Kriol speakers at some stage in the future. The Kriol orthography is presented in Tables 2 and 3 that follow.

Table 2: Kriol orthography – consonants

	Bilabial	Labio-dental	Dental	Alveolar	Retroflex	Palatal	Velar	Glottal
Stop	b...p			d...t	rd	j	g...k	
Fricatives		f	th	s		sh		h
Nasals	m			n	rn	ny	ng	
Laterals				l	rl			
Tap/Flap				rr				
Approximants	w			r		y		

Table 3: Kriol orthography – vowels

	Front		Back
High	i		u
Mid	e		o
Low		a	
	Diphthongs ei ai ou oi au		

While this orthography makes provision for the retroflex series of sounds, it should be noted that this series no longer appears to be productive in Roper Kriol. During the delivery of Kriol literacy courses in Katherine in 1999 and 2000 for Batchelor Institute of Indigenous Education, carried out with young adult Kriol L1 speakers, for example, it was found that these students simply could not provide example words with these phonemes. Furthermore, if the retroflex phonemes appear in the data for this thesis they are predominantly used only in words directly borrowed from the substrate languages into Kriol. This lexicon is largely made up of relationship terms, place names and personal names, although some verbs are also borrowed. The retroflex phonemes do not provide a meaningful contrast with any other series of sounds, nor are they involved in the production of new words; they are included in this orthography in order to represent these borrowed words.

The inclusion of two symbols for the stop series also requires some explanation. There is no voicing contrast in the Kriol sound system. There is, however, fairly regular allophonic variation within the stop series of sounds, which sees voiced stops used word or syllable initially and voiceless stops word or syllable finally. This is an innovation of the orthography implemented in this thesis and reflects similar changes in orthographies in the surrounding substrate languages.

1.2.5 Substrate language orthography

The substrate language phoneme inventories are generally very similar to each other, although there are some differences. The orthographic conventions that have been used for each language can, on the other hand, differ considerably. There are some orthographic changes recommended, although examples follow as closely as possible their form in their text of origin. The development of this orthography is a compromise between the orthographies of disparate grammatical descriptions by linguists and orthographic conventions in use by DAC. The orthography, for use in this thesis, is presented in Tables 4 and 5 below.

Table 4: Substrate orthography – consonants

	Bilabial	Inter-dental	Alveolar	Retroflex	Palatal	Velar	Glottal
Stops	b/p	dh	d/t	rd	j/tj	g/k	h
Nasals	m	nh	n	rn	ny	ng	
Laterals		lh	l	rl	ly		
Flap/tap			rr				
Approximants	w			r	y		

Table 5: Substrate orthography – vowels

	Front		Back
High	i		u
Mid	e		o
Low		a	
	Diphthongs ey ay uy		

The substrate languages differ in their consonant inventories. In Nunggubuyu, for example, there is a series of interdental, which includes [ɖ] and [ɗ] and rarely [ɳ] (Heath 1984). Heath (1981a:9) also presents an interdental series for Marra but says they are ‘...marginal to the system...’ because they only occur in borrowed words from languages that readily use an interdental series of phonemes, such as Nunggubuyu. Other examples of phonemes that occur in only one language are: the palatal lateral from Alawa, which is

represented as /ly/ in the orthography and the glottal stop from Ngalakgan, which is represented as /h/. This is the same symbol used for the glottal fricative in Kriol, which suggests that an alternative symbol be sought for the glottal stop here. The glottal stop and glottal fricative are, however, in complementary distribution: the glottal stop can only be syllable final, and the glottal fricative can only be syllable initial. Furthermore, using the ‘h’ to represent both sounds is the received practice at DAC where both orthographies operate at a practical level side by side. The same practice is therefore followed here. Finally, while the Alawa phoneme inventory includes a series of pre-nasalised stops that occur word initially (see Sharpe 1972:14), further research has shown that they are now almost non-existent in present day spoken Alawa, although traces do still exist (Sharpe 2001:xxi), and are therefore not included in this orthography.

The other point to make about the orthography is in regard to the two stop series. Some grammatical descriptions, such as for Ngalakgan (Merlan 1983) and Nunggubuyu (Heath 1980b), describe this as a fortis/lenis contrast in the stop series of these languages. Baker (1999) reanalysed this feature by arguing that in the case of Ngalakgan there is a one stop series with single and geminate forms. This orthographic convention is included in this orthography in order to represent the examples provided by Baker in Ngalakgan only.

Within the vowel inventory it should be noted that the phonemes /e/ and /o/ only rarely occur in the Marran languages. Secondly, Nunggubuyu distinguishes between a set of short and long vowels, the latter represented by [:] following the vowel. This convention will be used where they occur in the examples taken from the grammatical description of Nunggubuyu.

1.3 Substrate influence in pidgins and creoles

This section examines the previous research in this area. This includes theoretical constructs developed in the field of creolistics in order to study substrate influence; this is

the topic of §1.3.1. Previous research into substrate influence in Kriol in Australia is the topic of §1.3.2. There are two central processes involved in the research of substrate influence, which therefore require definition from the outset: transfer and levelling.

Siegel (2000:1-2) characterises transfer as follows: ‘In language learning in language contact situations, structures from the target language (usually the superstrate language) may be reinterpreted according to the syntactic or semantic properties of the learner’s native language (the substrate language)’. This view is maintained in this thesis, although while some researchers (Mufwene 1990; Siegel 1999) refer to this process as ‘transfer’, others use ‘relexification’ (Lefebvre 1998) (see §1.3.1). Mufwene (1990:4) also suggests that feature transfer is carried out by individuals and only when transfer of the same feature is replicated and repeated by other speakers, does it become a feature of substrate influence in the contact language. It is assumed in this thesis, therefore, that the substrate language feature transfer affecting Roper Kriol occurred during an earlier stage of development of the NT Pidgin, and that those features transferred by the greatest number of speakers at that time had a greater chance of being retained.

Siegel (1997) extends the process of levelling, originally applied to dialect contact and koineization, to the emergence of new contact languages, such as Kriol. Siegel (2000:3) describes the process thus: ‘The elimination of some variants and the retention of others is a sociolinguistic or community-level process known as levelling, which occurs during the stabilisation of a new contact variety...’. In the case of Kriol, Siegel (1997:133-134) shows that while the preceding NT Pidgin displays much feature variation, Kriol does not, or more concisely, the variation in Kriol is more systematic, which suggests that levelling has previously occurred in its development. It is in this process that ‘frequency’ plays the most important role. As described above, individual speakers will transfer substrate features to the developing contact language, which results in any number of transferred features from the substrate. When levelling occurs, this variation is minimised by the retention of certain features and the loss of others. The most important factor influencing the retention of such features is ‘... frequency of occurrence in the contact environment...’ (Siegel 2000:3; see also Siegel 1999). High frequency features

may be the result of demographics of language groups or typological similarity among the substrate languages (Mufwene 1990).

1.3.1 Substrate studies in creolistics

It is generally accepted that pidgins and creoles are derived from multilingual settings where there is no shared language among the substrate language speakers, and where access to the superstrate, whose speakers are dominant but in the minority, is restricted for substrate language speakers (Foley 1988; Le Page and Tabouret-Keller 1985; Thomason and Kaufman 1988). As these characteristics suggest, there are numerous linguistic resources available in any contact situation. Research differs as to the weight afforded each form of input. The primary sources of input may be one or a combination of the following: the substrate languages (Keesing 1988; Faraclas 1988; Lefebvre 1998; Siegel 1999; Migge 2003), the superstrate language (Bruyn and Veenstra 1993), possibly as a regional variety or colonial koine (Mufwene 1997; 2000), a pre-existing stable contact variety (Hancock 1993; McWhorter 1995), and/or universal linguistic principles (Bickerton 1981; 1984; DeGraff 1999). The methodology applied in the research into substrate influence, therefore, depends in part on the input afforded most significance.

Substrate influence in contact languages has long been acknowledged both informally, as far back as 1883 in the case of Guyana and Trinidad (see Lefebvre 1998:3), and formally within creolistics (see e.g. Hall 1966). While the Universalist position (Bickerton 1981; 1984) disregarded substrate input, particularly in attempting to explain creole genesis, substratists argued in favour of the following positions: that substrate languages did act as input to pidgins and creoles (e.g. Holm 1988), that creoles could emerge gradually (Singler 1990; 1992; Arends 1995) and that both universal principles and substrate languages could have an influence (Mufwene 1986). Increasingly the focus of research has been on the processes involved in language contact (e.g. Thomason and Kaufman 1988; Siegel 2000), as well as developing appropriate theoretical constructs to describe

them (e.g. Lefebvre 1998; Siegel 1999), rather than the nature of creole genesis specifically. Winford (2002) places pidgin and creole studies clearly within the field of contact linguistics.

Early examples of research into substrate influence in contact languages applied a methodology that took as the starting point the contact language, which was examined in order to identify any features not present in the superstrate language. A comparison between the contact language and the substrate languages followed in order to find any evidence that the features in question may have been influenced by the substrate languages. Examples include Camden (1979) and Smith (1979).

Bickerton (1981; 1984) did not apply any substrate methodology because the focus of research was on the universal linguistic principles captured in the Language Bioprogram Hypothesis (LBH). From this perspective, which dominated research in the 1980s, only 'radical creoles' or those that arose within one generation without any previous stabilised contact language, were considered 'true creoles'. Hawai'i Creole English is considered one of these rare cases. Under these extreme circumstances it was suggested that children created a creole through L1 acquisition without adequate input. They would therefore have had to rely on their innate, universal language bioprogram to create the new language. The substrate languages were therefore considered irrelevant.

One of the most important topics of research into substrate influence has been in determining whether contact languages are primarily derived from L1 acquisition by children or adult second language acquisition. This is closely connected to the speed in which contact languages, particularly creoles, can emerge. The universalist position posits that L1 acquisition by children without adequate input is a rapid process that creates a creole (e.g. Bickerton 1981). As Winford (2002:290) points out, however, a growing body of research has shown that contact languages are developed primarily through adult Second Language Acquisition (SLA), highlighted in the fact that creoles can emerge gradually from a pre-existing stable pidgin (e.g. Singler 1990; 1992; Arends 1995). This position, presented by Thomason and Kaufman (1988), places contact

languages under the same umbrella as other instances of contact induced language change, thereby reinforcing the role of adults in the process (also noted by Winford 2002). Thomason and Kaufman (1988:50) describe how creoles could emerge either through nativisation of a pidgin or through language shift in which there is 'extreme unavailability of the Target Language', which necessarily involves adults. This view has gained currency to the point that Migge (2003:4) notes: 'Most current approaches maintain that creole formation is a case of extreme contact-induced change. The most widely accepted approach maintains that it is the result of L2 acquisition with restricted access to the target language'. This view is shared in this thesis.

There is now overwhelming evidence that substrate influence does contribute to contact languages (e.g. Keesing 1988; Lefebvre 1998; Siegel 1999; Migge 2003). A significant component in research presenting such evidence is the claim that the greater the typological similarity among the substrate languages, the greater the expected substrate influence (Mufwene 1986; Singler 1988; Thomason and Kaufman 1988). Bearing in mind the definitions of transfer and levelling provided above, consider the following description by Siegel (1997:137):

When substrate languages are typologically similar, their second language versions of the superstrate language (or of the emerging vernacular target) will often be similar (if not identical) in terms of transfer. Thus when leveling occurs, such features will be common (and also reinforced), and therefore be retained in the creole. When the substrate languages are typologically heterogeneous, then second language versions with transfer will be varied. Since none of them will be very common, and since none will have strong majority support, most of them will be leveled out.

Defining the typological distance among the substrate languages will depend on the classificatory information available. As previously presented in §1.1.2, in the case of the Roper Kriol substrate languages, for example, family level relationships may belie a greater genetic distance between some languages. This thesis aims to overcome such classificatory difficulties by comparing the substrate languages, feature by feature. Only in this way can high frequency features be determined (see §1.4 for further information on this point).

Most contact language research allows for variable input. Mufwene (1986), for example, takes into account both universal principles and substrate languages, which leads to a finding (noted by Lefebvre 1998:2) that is significant to this thesis. Mufwene (1986; 1990) describes how creoles such as a dialect of Kituba of West Africa, are isolating languages, even though their substrate languages, in this case Bantu, are agglutinative. It is suggested that the universalist hypothesis may provide a framework with which to explain this, while substrate theories cannot (Mufwene 1986:135). This is important here, for as described in §1.1.2, the substrate languages of the Roper River region are also agglutinative, while Kriol is isolating, like most creoles. It is assumed in this thesis that this may be the result of a universal principle, as suggested by Mufwene (1996) in regards to Kituba. It seems likely, therefore, that affixes in general would not be transferable. However, the functions of an affix, and its position with regard to the root, may well be transferred in the form of a free grammatical morpheme.

The majority of researchers today consider input from the superstrate and the substrate languages, when considering contact languages. As Winford (2002) discusses, contemporary approaches investigate the interaction of the two language types in the contact situation. If contact language development is considered an extreme case of SLA, as previously suggested, then the superstrate language can be described as the L2, and the substrate languages as the L1s of speakers. In the language contact situation L1 speakers attempt to learn an L2, without adequate input. To compensate they use features from both their L1 and the L2. In order to help identify which features may be derived from which language at least three notions can be examined: ‘simplification’, ‘semantic transparency’ and ‘somewhere to transfer to’.

Simplification is a process that reduces the semantic, grammatical and functional roles of each feature from the input languages to a minimum, thereby increasing regularity within the contact language (Mühlhäusler 1974; Todd 1984:23; Romaine 1988:32). Features from the L1 or the L2 that display complexity on these levels may either be simplified for transfer or simply not transferred. ‘Semantic transparency’ is a property of language

features whereby the relationship between the surface structure and the semantic interpretations of it are clear and straightforward (Sueren and Wekker 1986). If semantic transparency does not exist in a feature, then it is possible that simplification may take place in order that there is a more direct relationship between the form and the meaning, or function, of a feature. Seuren and Wekker (1996) also claim that semantic transparency can determine feature transfer to creoles, whereby features high in semantic transparency are more likely to be found in creoles, whereas those low in semantic transparency are not. Finally, the notion of ‘somewhere to transfer to’, developed by Andersen (1983) with regard to SLA, is increasingly being considered in contact language studies. This notion suggests that L1 feature transfer is most likely to take place when speakers perceive there to be a similar feature in the L2. As Winford (2002:316) notes, there may be two consequences of this process: overgeneralisation of L1 feature characteristics, and reinterpretation of the L2 feature based on characteristics from the L1 feature. Each of these three notions is used in this thesis to examine whether or not feature transfer has occurred (see also §1.4).

As DeGraff (1999:508) points out, demographic and sociohistorical information also plays a vital part in substrate studies as they are often the only data that can accurately determine when the creole emerged in any contact situation. Furthermore, such information is essential in order to proportion significance to the respective languages of input, understand their relationships and hence the processes at work. Numerous researchers, among them Singler (1990; 1993), Arends (1995; 2001) and Parkvall (2000) provide models of how to best present such data. The range of demographic and sociohistorical information found to be of most importance follow:

- Nature of the contact setting
- Population figures of each language group in the contact setting
- Type of interaction between language groups, particularly between superstrate and substrate speakers
- Changing nature of language contact in time frames

This is also an area of research within creolistics that has the ability to overturn previously held assumptions about the contact language in question. Roberts (2000), for example, shows through sociohistorical and demographic data that Hawai‘i Creole

English emerged gradually, including a generation that could speak both their ancestral languages and the contact language, rather than emerging abruptly in one generation, as proposed by Bickerton. Such data therefore provides a far more accurate account of contact language development. Chapter 2 of this thesis examines such data in relation to Roper Kriol.

The most well known of the contemporary substrate positions is the Relexification approach (Lefebvre 1997; 1998; Lumsden 1999). Many of the aspects of substrate studies so far described find a place in this theoretical construct. SLA, for example, is seen as an important part of contact language development (Lefebvre 1998:35). Levelling is also taken to be a process that occurs to reduce 'idiosyncratic features' in a creole, whereby features common to all the substrate languages are the most likely to be retained (Lefebvre 1998:46). It can generally be understood then that the more homogeneous the substrate the greater the expected substrate influence. Lefebvre (1998) tests the Relexification Hypothesis on Haitian Creole by comparing it to one substrate language Fongbe (a Kwa language) and its superstrate, French.

Lefebvre (1998) identifies three processes in creole genesis: relexification, reanalysis and dialect levelling. 'Relexification is ... a mental process that builds new lexical entries by copying the lexical entries of an already established lexicon and replacing their phonological representations with representations derived from another language' (Lefebvre 1998:16). It can be seen then that this process provides an explanation for the fact that most contact languages are comprised of words, or phonetic strings, derived from the superstrate language. As such, the Relexification approach allows for superstrate input, usually at the phonetic level, although Lefebvre (1997:182) does say that there must also be some corresponding semantic component. This process is said to occur at an early stage of development and a large number of variables are created in the process.

Next, reanalysis is a process described in other cases of language change and Lefebvre (1998:41) describes it in relation to relexification as '... a mental process by which the

phonological form of one lexical entry becomes the phonological form of another lexical entry ...'. Such reanalysis often applies to 'functional category lexical items' in either nominal or clause structures, which are important because, '...there is a finite inventory of functional categories; [and] each language presents a subset of this inventory. If this subset is the same in the creole and the substratum language and contrasts with that shown in the superstratum language, this is evidence supporting the [relexification] hypothesis' (Lefebvre 1998:71). Some lexical items, however, are said to be assigned a null form during relabelling (see following paragraph), which means they are not pronounced, yet 'a relexified lexical entry that has been assigned a null form at relabelling may acquire a phonological form through reanalysis' (Lefebvre 1998:44-45). It is in this way that substrate languages are said to continue to influence the developing creole even though they may no longer have contact with it.

Finally, dialect levelling is similar to the definition already given for the process of levelling, whereby the numerous variant features are either retained or lost depending largely on their frequency within the substrate languages. Lefebvre (1998:46) says: 'It is hypothesised that, when the speakers of the creole community stop targeting the lexifier language [the superstrate] and start targeting the relexified lexicons, that is, the early creole, they begin levelling out the differences between the relexified lexicons'. The process differs to those above as it is a social process.

While the Relexification approach has influenced this thesis, particularly in its overlap with other theoretical constructs and in terms of semantic category transfer (chapter 5), it is not directly applied in this work. There are two main reasons for this. Firstly, Lefebvre (1998) uses only one substrate language in analytic comparison to Haitian Creole. In this thesis, however, as chapter 2 will show, no one substrate language can be singled out at this stage as being the most influential on Kriol. Secondly, and most importantly, the Relexification approach suggests that all the properties of L1 structures would be replicated in the L2 by copying and relabelling of the L1 lexical entries (Lefebvre 1998:16–17). It has been found, however, in creoles that grammatical morphemes with forms from the lexifier often have only some of the properties of the

corresponding structure in the substrate (see e.g. Siegel, Sandeman and Corne 2000). As chapters 3 – 6 show, this appears to be the case with Kriol, as there are no examples of grammatical features in that language with the exact properties of a corresponding substrate feature, although there are features with some properties having been transferred. The process of transfer seems to allow for this better than relexification.

1.3.2 Kriol in Australia

Previous descriptive research on Kriol was done almost solely by Sandefur (1979), and this continues to provide the only full length grammatical description of Kriol as spoken in Barunga⁹ (formerly Bamyili community, see maps 3 and 4) and Ngukurr. Steffenson (1979) also discussed reduplication in Kriol spoken in Barunga. Following on from preliminary studies such as Fraser (1977), Hudson (1983a) provides the only detailed grammatical description for Kriol as spoken in the Fitzroy River region of the Kimberleys of Western Australia. Other descriptive analyses of Kriol have been scant on the ground: Graber (1997) and Hudson (1983b) are the exceptions, both examining the verb phrase in terms of mood in the former and transitivity and aspect in the latter. Other resources are also available in Kriol, most notably the dictionary (SIL 1986) and the lesson package on conversational Kriol (Sandefur and Sandefur 1981).

A large part of the research on Kriol has been sociohistorical in nature, however, and initially it was necessary to determine the extent of the Kriol speaking region. To this end Graber (1988) explored the Barkly Tablelands, Sandefur and Sandefur (1980) the Kimberley, and Sandefur (1982) Queensland. It was found that Kriol has spread to all these regions (see map 1).

⁹ Barunga Community Education Centre ran a bilingual program in Kriol and English from the 1970s although due to policy changes within the Northern Territory Department of Education it was abandoned in 2002.

The most useful source of sociohistorical information directly related to Kriol includes Harris (1986), who provides an indepth discussion on the development of NT Pidgin in Northern Australia, its possible stabilisation and aspects of the emergence of the creole. This research is directly influenced by the work of Bickerton (1981; 1984) and as such attempts to determine that the creole emerged abruptly, in one generation, in order to define the status of Kriol as a 'true' creole. Harris (1986) suggests that NT Pidgin developed and stabilised by the turn of the 1900s. It is further suggested that the break in transmission of multilingual skills to children and the dormitory system in the Roper River Mission contributed to the abrupt development of the creole. A group of children were to some degree isolated from adults and other linguistic input, thereby having to rely on their innate linguistic abilities to create a new language. These two aspects of contact language development are described as having occurred largely independently of one another.

Sandefur (1986) also provides extensive information on such issues as identity, variation within Kriol, the use of Kriol at the time, and a historical overview of Ngukurr as a case study of a Kriol speaking community. Sandefur and Harris (1986) pay particular attention to the variation within Kriol. Variation is described as being due to generational differences, regional differences and the possibility of 'decreolisation' of the creole due to an increase in English proficiency by Kriol speakers.

More recently, Munro (1995; 2000) discusses the spread of Kriol, comparing the emergence of creole in the Daly River region in the Northern Territory and the Kimberley region in Western Australian, to the Roper River region. It is suggested that Kriol, along with NT Pidgin, may have spread throughout the Kriol speaking region, particularly through army camps and the pastoral industry. This view differs from that put forward by Sandefur (1986), Harris (1986) and Sandefur and Harris (1986) that Kriol emerged independently in various locations and later converged into one language. Munro (1995), however, found that the only place that creole emergence could have taken place independently appears to be Moola Bulla in the Kimberleys in Western Australia.

Sharpe (1972:9) first suggested that there was substrate influence in Kriol, referred to at the time as Pidgin English, by claiming the following: ‘In surface structure the languages [Alawa and Pidgin English/Kriol] are very different; in deep structure and semantically they are almost identical...’. Yet it remained just that, a suggestion, because no analysis or research on the topic was carried out until Sandefur (1979) included a comparison of the phonological systems of the substrate languages with that of Kriol. Sandefur (1979:48) found that, ‘...the sound system of Creole [Kriol] is derived from a complex combination of the sounds of AL [Aboriginal Languages] of the Creole [Kriol] area and of English’. The variations in the sound system, including features of substrate influence, were more fully described by Sandefur and Harris (1986).

The only research into morphological and semantic substrate influence is presented by Hudson (1983a) in which Fitzroy Kriol is compared to Walmajarri, a PN substrate language of that region. In many ways this research was ahead of its time and attempted to carry out an analysis of substrate transfer prior to appropriate theoretical constructs being developed. The methodology used is the same, therefore, as others of that era, which first provide a description of grammatical features of the creole, in this case the regional variety of Fitzroy Kriol. Where a Fitzroy Kriol feature differs from English and shows a similarity to the substrate language, a detailed comparison between the two is provided. The study was similar to the approach taken by Lefebvre (1998) in that only one substrate language was used in the analysis, although Hudson (1983a) also referred to general features of Aboriginal languages. A summary of the features and findings are:

- a) Kriol prepositions, similar to Walmajarri in terms of expressing a limited range of case relations,
- b) Verbless clause types, particularly Ascriptive, Equative, Possessive, Locative, Associative and Existential types, all showing some similar functions to those in Walmajarri,
- c) Stance verbs used in a similar manner as existential verbs in Walmajarri,
- d) Verbs *stap*, *leidan* and *bi*, all showing similarities in use to the innovative copula feature *nguna* in Walmajarri,
- e) Verb *git*, ‘become’ similar function as Inchoative suffix in Walmajarri

- f) Reflexive and Reciprocal roles being marked by one form, *jelp*, in Fitzroy Kriol, reflecting influence of Aboriginal languages
- g) Semantic analysis of borrowed words from Walmajarri and possibly other Aboriginal languages into Fitzroy Kriol.

Using such methods, however, there is no systematic means of selecting features to compare, nor a way to predict which substrate features may be included in the creole. Furthermore, a more representative group of substrate languages is required for a comprehensive comparison of substrate language features. Hudson (1983a:13-15) suggests that Fitzroy Kriol originally developed at Moola Bulla, an Aboriginal reserve near Halls Creek. This creole spread to Fitzroy Crossing when Moola Bulla closed in 1952 and its inhabitants dispersed. Substrate transfer would, therefore, have most likely occurred at Moola Bulla, where the substrate languages included Kija, Jaru and Walmajarri. Alternatively Hudson (1983a:9) describes two PN languages, one of which is Walmajarri, and two non-PN languages in the Fitzroy Crossing region, which may all have influenced the Kriol spoken there.

Koch (2000) also investigates substrate transfer in relation to early Australian Pidgin. One New South Wales substrate language, Ngiyampaa, is used as an example of Australian Aboriginal languages, particularly those that are PN. This investigation into substrate transfer in New South Wales is significant to creolistic studies in Australia. As previously mentioned in §1.1.1 NT Pidgin, and ultimately Kriol, was preceded by NSW/QLD Pidgin, which travelled via the pastoral industry expansion into the Northern Territory and on to Western Australia. It must be expected, therefore, that there will be features in Kriol derived from this NSW/QLD Pidgin, which in turn suggests that substrate transfer from anywhere along the route in New South Wales and Queensland may have contributed to Kriol. Harris (1986:288), for example, notes that some words in Kriol, such as *bogi* 'bath/swim' and *gabarra* 'head', originated with the Sydney area Aboriginal languages in New South Wales. Identifying grammatical features within Kriol that are due to substrate transfer elsewhere will, therefore, need to be considered in any such analysis. Koch (2000) identifies two such features in the early Australian

Pidgin, the transitive marker *-im* and the ‘adjective-*pela* noun’ construction as examples of possible substrate transfer from New South Wales languages. Both these features also exist in Kriol, which therefore leads to the assumption that they originated from substrate transfer at an earlier stage of pidgin development, such as in New South Wales, and have been retained during the development of the NT Pidgin and the emergence of the creole.

This present research is well placed to expand on the research on Kriol described in this section in a number of ways. Firstly, the grammatical analysis of Kriol provided in this thesis improves our understanding of the classificatory nature of the nominal referentiality system, the tense–mood–aspect system, the pronominal system and semantic roles encoded in the prepositions. Other features, such as reciprocal/reflexive marking and the behaviour of adjectives, while previously noted in the research, receive a more comprehensive analysis in this thesis. Secondly, the sociohistorical data provided in this thesis improves our understanding of the conditions surrounding the emergence of a creole in the Roper River region, particularly by taking into consideration the role of the pastoral industry. While outside of the study of Kriol, to some degree, this thesis also attempts a comparative typological analysis of the Roper River substrate languages, not previously carried out in Australian linguistics. This thesis also contributes to the field of creolistics by improving our understanding of the processes of transfer and levelling in a language contact situation.

1.4 Transfer Constraints approach

The Transfer Constraints approach aims to account for the presence of some substrate features in creole languages and the absence of others by examining the likelihood of transfer of substrate features at an earlier stage of development and the retention of these features in the creole. Siegel (1997) formulated transfer constraints based on research in the SLA literature. Siegel (1999) then applied them to Melanesian Pidgin. Siegel, Sandeman and Corne (2000) trialed the approach on another creole, Tayo. These studies

will be described in more detail later in this section. In this thesis I further test this theoretical approach by applying it to Kriol. This will not only provide information on the substrate transfer that has taken place in Kriol but also test the predictive capacity of the approach, particularly in relation to the nonPN substrate languages of the Roper River region.

The Transfer Constraints approach employs two primary availability constraints and a reinforcement principle to both predict possible transfer and explain what actually does occur. The availability constraints are related to perceptual salience and congruence. The process of transfer can take place only when speakers perceive a morpheme in the developing contact language or superstrate that appears to have a function or meaning similar to a morpheme in the L1. This morpheme also needs to be in a congruent syntactic position. In other words, transfer of an L1 feature can take place when there is 'somewhere to transfer to', as per the principle devised by Andersen (1983) (see §1.3.1). Features that provide somewhere to transfer to need to be perceptually salient – that is, easy to distinguish, such as stressed, free morphemes (Siegel 1999:23) or multisyllabic lexemes (Siegel, Sandeman and Corne 2000:83). Features that are congruent are those that occur in what appear to be similar syntactic structures (Siegel, Sandeman and Corne 2000:83). Siegel (1999:31) says: 'In other words, transfer can occur only if there is a feature in the L2 [superstrate] superficially similar enough to a feature in the L1 [substrate] that it can be misinterpreted or reanalyzed to correspond to L1 [substrate] rules'. Alternatively where there are no perceptually salient or congruent features in the superstrate, transfer will be constrained.

There are currently two positions as to when transfer, and hence, when the availability constraints may apply. The original position of the Transfer Constraints approach is that transfer occurs and the availability constraints operate during the early stages of development of the contact language as a result of second language learning (e.g. Siegel 1999). However, Siegel (2003) has revised this position to suggest that transfer may occur during expansion of a stable pidgin, when speakers are bilingual in the pidgin and their ancestral languages, as a result of second language use. It is also possible that

transfer can occur during both stages, in which case, both the superstrate language and the stable pidgin can be examined for perceptually salient and congruent features that could influence transfer. Chapter 2 will determine the most likely timeframe and conditions when transfer occurred to Roper Kriol.

Siegel (1999:4) notes that in SLA research positive and negative transfer have been identified. Positive transfer occurs when the L1 and L2 share the same feature, whereby L2 learners simply apply known features from the L1 to the L2. Negative transfer, however, occurs where there is a discrepancy between the L1 and the L2, whereby the SLA learner continues to apply their L1 feature even though it does not occur in the L2. In the case of contact languages, where there is a correspondence between the substrate and superstrate feature, positive transfer can have taken place. Yet it cannot be proven because the presence of the superstrate feature in the contact language might simply be a consequence of second language learning. The primary concern, therefore, is with negative transfer, in which case the substrate language speakers may perceive the superstrate feature as similar, when in fact it is not.

The reinforcement principle of frequency applies during levelling, hence at a later stage of development when the contact language is stabilising. This may be at the stage when a stable pidgin is first emerging or later when it is expanding or developing into a creole. The central aspect of this principle is that the more a particular feature occurs in the developing contact language as a result of transfer, the greater the chance of retention of that feature during levelling. As Siegel (1999:27-29) notes, there are two ways that transferred features become high in frequency. Firstly, it is expected that common structures within a substrate language will demand constant use, thereby being transferred the most often. These features will be referred to in this thesis as ‘core features’. Peripheral or rare features of a substrate language are not in use as often and therefore not so likely to transfer. Secondly, high frequency features may be due to the typological similarity of the substrate languages (see §1.3.1). Those features that are shared among the substrate languages, or L1s, will be transferred by the most speakers and hence will be high in frequency at levelling. These features will be referred to as ‘shared features’

in this thesis. It should be noted that the processes of expansion and stabilisation could happen more than once in any given contact situation. Levelling, as well as transfer, is then possible during more than one stage. Once again, chapter 2 will provide the timeframe and conditions surrounding levelling in the Roper Kriol situation.

Siegel (1999), attempting to counter Mufwene's (1990:6) criticism of the substrate approach for not devising regulating principles, provides an overview of the principles and constraints on transfer during SLA from the literature in that field. These constraints were applied to eleven core features of the Central Oceanic languages, the substrate languages of Melanesian Pidgin. It was found that they accounted for the fact that seven features occur in Melanesian Pidgin, those previously described by Keesing (1988), and four which do not. This was done in order to determine those constraints or principles that had the most explanatory power in relation to substrate transfer in pidgins and creoles. It was found that the availability constraints of perceptual salience and congruence and the reinforcement principle of frequency were best able to account for the results.

Siegel, Sandeman and Corne (2000) refine the use of these constraints and principles in applying them to Tayo, a French lexified creole of New Caledonia in the South Pacific. This study provides a vital shift in methodology because it is the comparative analysis of the substrate languages, not the description of the contact language, that provides the initial indicator of what features may have been transferred. After a brief sociohistorical overview of the contact situation, the Tense-Mood-Aspect (TMA) system of each of the three major substrate languages is presented. Next, after a description of the methodology, a list of shared features of the substrate TMA system is provided in order for predictions to be made according to the reinforcement principle of frequency. The availability constraints are also examined in relation to any corresponding features in the superstrate language, French, in order to determine whether transfer may have been constrained or not. The findings of both of these sections are combined to make five predictions of expected, as well as not expected, features in the TMA system in Tayo. These are generally borne out in the description of the language that follows. As Siegel,

Sandeman and Corne (2000:94) state: ‘Examining in detail one particular substrate language may account for the presence of some features in a creole but it cannot account for the absence of others’. This is the significant development of this approach, as it not only accounts for substrate features that occur in the contact language, but also those that do not.

The application of the Transfer Constraints approach in this thesis is similar to that just described. The sociohistorical data is first presented (chapter 2) in order to identify when the processes of transfer and levelling could have occurred in the Roper Kriol context; which substrate languages had most input and how the creole developed. The main body of the thesis discusses a broad range of linguistic features: verb features (chapter 3); nominal features (chapter 4); semantic category features (chapter 5); and nominal modification features (chapter 6). This study expands on that of Siegel, Sandeman and Corne (2000) by including not only morphosyntactic features but also features of semantic categories. It was discussed in §1.1.2 that the substrate languages of Roper Kriol comprise two language families. It may be significant that the two families are genetically distant to one another and that the languages within one of the families are also correspondingly distant to each other. In order to overcome this, a comparative analysis of the substrate languages is provided for each feature, so as to determine the core and/or shared features among them, rather than relying on incomplete classificatory information.

The method used in this thesis differs to that used by Siegel, Sandeman and Corne (2000) in two main ways. Firstly, the initial predictions are made based on the reinforcement principle of frequency alone. In this way the two processes at hand, transfer and levelling, are more clearly separated. The predictions at this stage can only claim that if transfer previously took place then the high frequency features would have been retained in levelling. As previously stated, this thesis identifies high frequency features based on core and/or shared features in the substrate languages. Siegel, Sandeman and Corne (2000) only investigated shared features between the substrate languages, rather than core

internal features. This thesis, therefore, also differs from the previous research in this regard.

The predictions, based on the identification of core and/or shared features in the substrate languages, suggest which features are expected to have been retained in Kriol during levelling, which are termed the E Features. In some circumstances, low frequency features warrant a prediction to the effect that they are not expected to have been retained in Kriol during levelling. These are the NE Features, which are usually features that are not shared between the languages but are a core feature of a particular substrate language. The findings from such predictions could suggest which, if any, substrate languages may have had a stronger influence in the contact environment.

Following the predictions, the Roper Kriol data is presented, which describes whether the predictions are borne out. Then the availability constraints are examined in relation to English, in order to determine whether transfer was constrained or not. The next section in each case therefore involves an analysis and description of any corresponding features in English in terms of perceptual salience and congruence. If perceptually salient features in congruent constructions are found in English and the predicted feature occurs in Kriol, then it can be said that transfer occurred at this earlier stage and due to being high in frequency was retained in levelling. If, however, no such features are found in English and the predicted feature does not occur in Kriol, then it can be said that the transfer was constrained, and furthermore, that the substrate feature could not have been reinforced during levelling because it had not been transferred.

Finally, simplification or semantic transparency (see §1.3.1) may also be used for the explanation of findings, particularly where only partial feature transfer has taken place. Furthermore, there may have been other influences that affect transfer possibility, such as substrate influence in other places at earlier stages of development of the preceding contact language. As previously mentioned in §1.1.1, the preceding pidgin that travelled with the pastoral industry expansion, NSW/QLD Pidgin, originated in New South Wales from the 1820s, and travelled through Queensland from the 1840s before entering the

Northern Territory or the Roper River region from the 1870s onwards. Substrate influence from any region in these two states is therefore possible. Koch (2000) discusses substrate influence in New South Wales and Baker (1993) provides the earliest attested examples of a list of contact language features in Queensland, New South Wales and various Pacific Islands. These resources may help determine where such features were first used, which may also provide evidence of previous substrate influence in NSW/QLD Pidgin.

1.5 Chapter layout

The sociohistorical nature of the contact situation around the Roper River region of the Northern Territory from the 1870s through to the 1950s is presented in chapter 2. Of particular emphasis is the nature of contact within and around the Roper River Mission, especially in relation to other pastoral stations that were being established from 1908 onwards. It is claimed that Kriol gradually emerged out of the pastoral industry that includes the Roper River Mission, from the 1910s to the 1950s. Population figures are also presented, where available, in order to determine how much access to English substrate language speakers had.

From chapter 3 onwards the grammatical analyses are presented, which follows the methodological procedure of the Transfer Constraints approach, as described in §1.4. The comparative analysis of the structural features of the verb in the substrate languages, includes the discussion of four features, that are both core and shared features, in chapter 3, as well as a core feature of one language that is not expected to be transferred. Similarly, chapter 4 presents an analysis of four core shared structural features of the nominal, although the focus shifts to examples of positive and negative transfer. One feature is also only shared by two of the four languages. As described in §1.1.2 the nonPN substrate languages are agglutinative and non-configurational, with particularly complex verb structures, which necessarily means that word order is generally not fixed

and phrase structures are limited. Providing an analysis of both the core and shared features of the verb and nominal complex is the most appropriate introduction, not only in terms of feature transfer, but to languages themselves.

The focus in chapter 5 shifts, once again, to semantic features, two from the nominal and one from the verb complex of the substrate languages. It is not an analysis of the semantics of individual lexemes, however, but rather the categorical range of three features. The substrate languages are compared in order to identify the pronoun, TMA and case marking categories that are shared between all the languages. Seven shared features are identified and discussed using the transfer constraints approach.

The complicated nature of nominal modification is presented in chapter 6, which includes modification by demonstratives and adjectives and methods of marking possession and number. The nonPN substrate languages employ a number of strategies to indicate each feature. It is important, once again, to identify the shared features, but in this case identifying the core features is, therefore, most significant. Four shared core features are discussed in relation to demonstratives. Five similar features are discussed in relation to adjectives, along with one feature that is a core feature in only one language. Five core shared possession marking strategies are also discussed, along with one core feature found in only one language. Similarly, one core feature of one language and four core shared features are discussed in terms of number marking strategies.

While a discussion section is found at the end of each chapter, an overall discussion of the findings is presented in chapter 7.