

# Interactions of South African languages: Case study of Tsotsitaal

**Karolina Brook**

Princeton University

Program in Linguistics

kbrook@princeton.edu

## Abstract

South Africa, viewed to be one of the most multilingual countries in the world, recognizes eleven official languages. This paper represents a sociolinguistic analysis of Tsotsitaal, an unofficial creole language that formed as an amalgamation of Afrikaans and isiZulu. Today, Tsotsitaal has evolved into a collection of distinct yet mutually understandable languages projected from the speaker's native language. It is marked by a high degree of codeswitching and overlexicalisation. I discuss the challenges that these properties present in developing a WordNet for a creole such as Tsotsitaal.

## 1 Introduction: The birth of Tsotsitaal

South Africa is viewed to be one of the most multilingual countries in the world, recognizing eleven official languages (van der Merwe and van der Merwe 2006:2). In addition to the officially-recognized languages, Tsotsitaal is a creole not recognized in official censuses (Mesthrie 2002:12; Makhudu 2002:399). Creole languages are defined as being young dynamic languages, only one to three centuries old (see McWhorter 2005 for a review of creoles).

Tsotsitaal is integrally tied with its birthplace and subsequent area of development. The language developed largely among racially segregated residential areas, located in Johannesburg around the gold reef (consult Molamu 2003:xviii for map of area). Despite the attitude to segregate races, in particular blacks from whites, these three communities shared extensive social interactions and came to develop some sense of overarching community (xviii-xix). The subculture that developed was vibrant, dynamic and distinctive, characterized by dance parties, music, alcohol and escalating violence (xx). During the 1950s black people were relocated during forced removals to provide housing for working white people. The exact origin of Tsotsitaal is

largely unclear, although it is possible that Tsotsitaal developed (and continues to be associated with) various criminal gangs that were active on the Witwatersrand in the early to mid 1900s (Slabbert and Myers-Scotton 1996:321), although it is also believed that Tsotsitaal came into fruition around the 1950s and 60s (Motshegoa 2005:1). The language developed for several intertwined reasons: it was born as a language of creativity and passion as well as an expression of the sadness, anger and resentment felt by these people dislocated from their sense of identity; it acted as a bridge among young segregated communities that spoke several distinct languages (Molamu 2003:xxi, xiii); it functioned as the *lingua franca* of male social interactions (Slabbert and Myers-Scotton 1996:322); and finally, it served as a means to organize illicit activities yet remain unintelligible to police who could not speak the language (Motshegoa 2005:1).

The 'culture' of Tsotsitaal was characterized by wearing clothes of particular labels, often-times flamboyantly expensive (Motshegoa 2005:1-2). In this essentially macho culture, knowing Tsotsitaal separated boys from men, and was a means to show streetwise character (1). This is exemplified by the alternate name for Tsotsitaal, namely Flaai- or Fly Taal, which would denote someone who is slick and city-wise (Makhudu 2002:402). Additionally, a breed of music was born that came to be known as Kwaito, or 'bubble gum music' (Motshegoa 2005:3). An integral part of the beliefs of Tsotsitaal speakers was characterized by an interest in crime and violence. Tsotsitaal was largely used as a means of communication among those interested in crime-related activities (Molamu 2003:xiii, xxii). The derivation 'tsotsi' as part of the name describing this language is descriptive of this function of the language; today, the word has largely come to refer to a hooligan, thug or gangster (xxii-xxiii). In response to the oppression enforced on the black communities by the Afrikaners, Tsotsitaal was largely a secret lan-

guage that was used as a means of rejection of Afrikaans, but also to ensure Afrikaans-speakers, particularly the police and other law-enforcers, did not understand the activities the youths were organizing (Molamu 2003:xiii, Motshegoa 2005:2). Afrikaans indeed formed the basis of Tsotsitaal, retaining some grammatical rules of Afrikaans while completely violating other grammatical constructions (Molamu 2003:xxv). The derivation *taal*, comprises the second part of Tsotsitaal, and is the Afrikaans word meaning 'language' (xxii). Thus, Tsotsitaal lay in the middle ground between these two separate worlds: the Afrikaans people who forced their language on this oppressed group, and the ensuing world of crime and violence.

## 2 The nature of Tsotsitaal today

The slang or street language that emerged from this political and social melting pot came to transcend the physical segregation of the original three communities in which the language developed, and today spans areas and townships around Johannesburg (Molamu 2003:xxii, Makhudu 2002:399). In these areas, Tsotsitaal may also go by different names, but confusion has arisen in the literature whether these are different languages or in fact the same language (refer to Ntshangase 1995, Makhudu 2002, and Molamu 2003 for opposing arguments). Mesthrie (2008: 95-109) recently proposed a unified account of Tsotsitaal, describing it as a 'slang lexis' which incorporates all of its varieties including Tsotsitaal associated with crime (as discussed) to simply street speech spoken by urban youth today, thereby including the Tsotsitaal based on Afrikaans, Zulu, Sotho and Xhosa, and even an English Tsotsitaal that Mesthrie proposes (see 103-108) as existing among Indians and Coloureds living in KwaZulu Natal.

In clarifying this issue, I will first discuss the aspects of the dependence of Tsotsitaal on the speaker's native language (L1), as well as the property of code switching inherent in this language, before moving to discuss my own fieldwork and conclusions.

### 2.1 L1-dependence of Tsotsitaal

The individuals who responded to my search for Tsotsitaal speakers had mother tongues including Zulu, Xhosa, Sotho, Tswana and English. What I discovered was that these native speakers all spoke a different version of Tsotsitaal that appeared to depend on the speaker's native lan-

guage, while incorporating lexicon from Afrikaans as well as isiZulu and isiXhosa, and, to lesser degrees, Sesotho, Setswana, Sepedi, Tshivenda and Xitsonga. In addition to this dependence on the speaker's native language, the Tsotsitaal spoken also seemed to depend on the speaker's main residential area (which in turn is related to the major language spoken in that area) as well as the speaker's age, with older-generation Tsotsitaal speakers being more familiar with the 'original' Afrikaans-based Tsotsitaal.

A consequence of this is that the same sentence spoken in Tsotsitaal can differ depending on the speaker's L1. For instance, 'Let's talk' can be said in Tsotsitaal as the Afrikaans-based *Let's wiet*, or as the Sotho-based *Let's qamtha*. As one speaker told me, 'There is an own Tsotsitaal [sic] in every place – as many as there are beds'. But what is even more astounding is that, as confirmed by my speakers, all these versions of Tsotsitaal are mutually understandable.

### 2.2 Code switching in Tsotsitaal

Code switching and code mixing are two separate phenomena, the former more specifically referring to 'the alternation of elements longer than one word, from two languages or dialects', while the latter referring to 'the alternation ... of shorter elements, often just single words' (McCormick 1995:194-195; Gough 1996:69; see Slabbert and Finlayson 2002 for a review of code switching in South African townships). In this paper, I will refer to both as 'code switching'.

In a country where most speakers are multilingual, code switching has become prevalent in slang use among the youth of South Africa. Slang is an idiosyncratic method by which the youth in South Africa express their identity, in a city 'characterized by a high degree of language diversity' (Bembe and Beukes 2007:464). Code switching is often used in the township environments, where there is no one single *lingua franca*. Practically, people living together in the townships in South Africa speak a variety of languages, and in order to communicate effectively, it is imperative to be able to speak several languages and be able to swop from one language to another, as this quote (taken from Slabbert and Finlayson 2000:124) illustrates (emphasis changed from the original):

'... [L]et's take for instance, ngikhuluma isiZulu lesi, lesi, lesi; and wena ungumPedi, fanele, we must communicate, sizocommunicatea kanjani? Fanele ngi compromise, ngimixe with your language ukuthi uzongizwisisa, uyabona.'

‘... [L]et’s take it that I am speaking Zulu and you are a Pedi, it is imperative that we communicate – how will we do that? I will have to compromise and mix my language with yours so that you can understand me, you see.’

### 2.3 Selected sentences in Tsotsitaal

Follows are sentences taken from my fieldwork provided by native speakers of Tsotsitaal. They serve to illustrate the L1-dependence of Tsotsitaal, as well as the high degree of code switching present in the spoken language.

#### 2.3.1 Zulu-based Tsotsitaal

<b>Bengi vaye</b>	<b>ne-mpinshi</b>
1sg go.PAST	with-friends
<b>yami si-yozama</b>	<b>i-leve ejozi.</b>
3sg-looking	for.life in.town

‘I went with my friends looking for life in town.’

#### 2.3.2 Afrikaans-based Tsotsitaal

<b>Sharota is ‘n</b>	<b>lekker mshoza.</b>
PROP.N is INDEF	nice woman.

‘Sharota is a pretty woman.’

#### 2.3.3 Tswana-based Tsotsitaal

<b>Bafwetho</b>	<b>arevayeng</b>	<b>re</b>
Guys	let’s go (to many people)	to
<b>tuba</b>	<b>die gamula</b>	<b>nyoko.</b>
take/rob	DEF rich.man	money.

‘Guys, let’s go rob the rich man of his money.’

#### 2.3.4 Sotho-based Tsotsitaal

<b>If ore</b>	<b>haona</b>	<b>yona</b>
(If) 2sg.said	2sg.NEG	have it
<b>ngaye fife</b>	<b>yanga ke</b>	
give.to.1sg five (5)	50-cent so	
<b>elo baya</b>	<b>toss.</b>	
1sg.can buy	matches.	

‘If you said you don’t have it give me 5 50-cent coins so I can buy matches.’

### 2.4 Tsotsitaal: one language or several?

These characteristics of Tsotsitaal present two problems relevant to WordNet. Firstly, the same concept can be expressed in Tsotsitaal using different words solely based on a different L1. Should such lexemes be considered synonymous? Or should we categorize the lex-

icon as entirely separate languages, i.e. Afrikaans Tsotsitaal, Zulu Tsotsitaal, etc? Should the fact that the speakers claim the different versions to be mutually understandable be taken into account?

Secondly, the issue of code switching represents an interesting dilemma for the definition of this language. Is Tsotsitaal a language that possesses a high degree of code switching as a fundamental property? Or rather, should we view it as a code switching phenomenon present in a large body of languages – rather than a language in its own right? How will a decision on this matter affect Tsotsitaal’s representation within WordNet?

## 3 Tsotsitaal lexicon

In this section, I present several examples of lexical items that have been documented in Tsotsitaal. The speakers also confirmed their use in Tsotsitaal.

### 3.1 Selected lexical items in Tsotsitaal

**tsotsi** [tsɔtsi<sup>[4]</sup>]: *n* In general, the word refers to a ‘violent, usually young criminal who usually lives by his wits ... [and has] broadened to incorporate an entire youth gang subculture’<sup>[5]</sup>. The connotations vary between a criminal who steals and kills to just a hooligan (as opposed to a ‘thug’). The exact origin is unknown, and several possibilities exist: 1) it originates from the **zoot suits** that were worn by Americans in the 1940s, and their stereotype is that tsotsis play the part of the zoot-suitor: ‘[t]hey are street-corner dandies, lounging in the doorways of vacant stores, idling in the train stations and bus terminals, giving passersby the hard eye’<sup>[5]</sup>; 2) it has been argued that the word was used to describe the stove-pipe trousers that were in fashion in the 1940s and as such may be derived from the Afrikaans word **stof-pipe** that described these tight-fitting trousers, or from **tso(lo)**, which meant the trousers they wore were sharp (fine or nice)<sup>[5]</sup>; 3) it could possibly originate from South Sotho **ho tsotsa**, a verb to engage in crime and thuggery, or **tsotsana** which means ‘to fight’ in Sesotho<sup>[5]</sup>, although it has been contended that while the word has ‘a Sotho phonemic structure, it is not a Sotho lexical item’<sup>[2]</sup>; 4) the !Kung San, an ethnic group in southern Africa, refer to non-San people as **zo si**, which refers to dangerous animals, and consequently they refer to such people as angry or dangerous<sup>[5]</sup>; 5) the reputation of **Tutsi** warriors in Rwanda and Burundi, who were involved in vicious and brutal crimes in the 1930s/40s, came to be known in the South African townships and came to refer to gangsters who were extremely physically violent<sup>[5]</sup>. **Tsitso** is a nickname that is modified from the word **tsotsi**; **tsotsigeid** (or similarly, **tsotsi-ism** or **tsot-**

**sism**<sup>[2]</sup>) uses the Afrikaans **-geid** suffix that is equivalent to **-ness** in English in order to mean ‘thuggery’, the criminal acts that are carried out by tsotsis<sup>[5]</sup>, or just the lifestyle and behaviours of the tsotsis<sup>[2]</sup>.

**kwaito** [kwAltəu<sup>[4]</sup>] *n* Popular township dance music ‘featuring rhythmically recited vocals over an instrumental backing with strong bass lines’<sup>[4]</sup>, or alternatively slowed-down US/UK house music with added African vibes<sup>[6]</sup>. The style of music emerged during the time of apartheid in the townships, and was a creative outlet during an economically-depressed time. Kwaito continues to be popular today with artists such as Mandoza or Zola. The noun can also possibly be used to refer to boys who sing in the streets. The word originates from **Amakwaito**, purportedly a group of gangsters in Sophiatown<sup>[1]</sup>, or **amakwaitosi**, which supposedly means ‘gangster’; this in turn derives from the word **kwaai** [kwaI<sup>[1,2]</sup>] (alternatives: **kwaad**, **kwaaje**, **kwaat**, **kw(a)i**, **quaai**, **quei**). Kwaai is a modern-day Afrikaans word that refers to something bad-tempered or angry, as in ‘... the lions on the opposite side were more **kwaad** ... than those where we now were’<sup>[2]</sup>. Kwaai also has slang usage to refer to something that is great or fantastic, as in, ‘The music is **kwaai** but nobody wants to [dance]’<sup>[2]</sup>. This connection seems apt, particularly since the music was a form of expressing the township dwellers’ anger and frustration during their oppression.

**malalapipe** [mala:lapaIp/mələləpaIp<sup>[2]</sup>]: *n* Literally refers to someone who sleeps in drainage pipes, or more generally, someone who sleeps rough or on the streets<sup>[5]</sup>, though possibly specifically referring to a homeless child<sup>[2]</sup> (which is supported by the related word **malunde**, which is a term used to refer to sleeping out without parental permission<sup>[5]</sup>), as a result of being chased away from home and having nowhere else to stay. Possibly, the person’s intent may be to ambush someone or perform a criminal act. It is possible the word originates from the isiZulu **umalalepayipini**, meaning ‘homeless child’, which can be broken down into **uma-** = one who + **lala** = sleeps + **epayipini** = in a drain, which itself may be derived from the English word **pipe**<sup>[2]</sup>. **Pipe**, in addition to referring to a tubular piece of metal or plastic, also may have connotations of referring to a vagabond<sup>[5]</sup>, which may be through the use of the word to refer to the instrument for smoking drugs and the sexual connotations associated with the word<sup>[3]</sup>. Another possibility is that the word is a combination of the isiZulu word **ukulala**, which means to sleep, plus **pipe**.

**chicken murder**: *adj* a person having a relationship with a younger person<sup>[5]</sup>. This ‘relationship’ varied in nuance, from an older man sleeping with a younger woman, or even rape. ‘chicken’ is possibly used in reference to a younger, more innocent person; perhaps the relationship is viewed as taking away the

innocence of a younger person. This expression would be equivalent to the American expression ‘cradle robber’ or ‘cradle snatcher’<sup>[6]</sup>. There might be a relation between this expression and that of **chicken run**, an expression that was used to describe the exodus of white people and businesses from Zimbabwe and South Africa for fear of their future in the country<sup>[1,2,4]</sup>.

### 3.2 Overlexicalisation in Tsotsitaal

Overlexicalisation, refers to the enormous number of words that can be used to refer to the same thing, and can be subdivided into two types: words that differ based on their language of origin, and words that differ based on their region of use, although these are not necessarily mutually exclusive (Makhudu 2002:402-403).

The examples presented were obtained from cross-referencing dictionary [5]. The prevalence, from highest to lowest, was determined from the speakers I interviewed.

**girlfriend/woman**: cherie/cheriekie/cherietjire, mathara, mshoza, tlakadula, tletletle, baby, rwatla, show, giedieng, tlakawibit, wherrie, charowibits, wibit.

**young (streetwise) male**: jieta, ma-Gents, ma-G-men autie, clever/klevaa, mfo, mri/mrio, deaden/laaities basie, bap,b[r]oer, (vaal)japie, ducktail, juba.

**money**: bucks, nyoko, miering, zak, chien, smackers, magegeba, kuneining, tsang, dummy, koezat.

### 3.3 Lexical issues for WordNet

This sampling of lexical items suggests further potential issues of Tsotsitaal for WordNet. Firstly, as a hybrid language, spellings in Tsotsitaal have not been officially finalized, and thus numerous variations of a word, even more than the ones provided here, exist.

Secondly, in keeping with the original purpose of the language, the majority of the words pertain to crime, sex and violence. The lexicon also includes concepts and words not present in other languages. Thirdly, the enormity of overlexicalisation presents a unique problem for WordNet. It is difficult to assess the synonymy of the overlexicalised items: while they seemed broadly similar, I did detect slight differences in the connotations of the words. It would be necessary to confirm that these connotations are widespread among Tsotsitaal speakers.

On the other hand, while Tsotsitaal is lexically rich in these concepts, it may be lacking in others. For instance, ‘sun’, ‘moon’, ‘grass’,

'dog', 'to dream', and 'to give birth', are just a few words that simply do not exist (Mesthrie 2008:102).

#### 4 Conclusion

While WordNets are being developed for so-called "official" languages, creoles such as Tsotsitaal should also be studied and integrated into the Global WordNet Grid. The variable nature of Tsotsitaal does however present interesting challenges to the development of a Tsotsitaal WordNet. These issues can be summarized as follows: 1) how do we view synonymy that depends on the speaker's L1?; 2) how do we define the language based on code switching?; 3) how do we handle the high variability in spelling?; 4) we must be aware that the language will be lexically rich in some concepts, while poor in others; and 5) we must be careful in deciding whether overlexicalised items are truly synonymous.

#### Acknowledgements

The author would like to thank both J.T.Katz and C.Fellbaum of Princeton University for their guidance, advice and support in the writing of this manuscript. Field research was funded through the Fred Fox 1939 Grant. This paper would not have been possible without the language input from the native speakers, to whom the author extends her gratitude. Finally, thanks go to my parents for helping me locate native Tsotsitaal speakers while so far from home

#### References

Please note: The works preceded by bracketed numerals ([1], [2], etc) were used in the various dictionary sections for clarity.

Bembe, Magdeline Princess, and Beukes, Anne-Marie. 2007. The use of slang by black youth in Gauteng. *Southern African Linguistics and Applied Language Studies*. 25:4, 463-472.

[1] Dictionary of South African English (DSAE). 1987. Third Edition. Oxford: Oxford University Press.

[2] Dictionary of South African English on Historical Principles (DSAE-HP). 1996. Oxford: Oxford University Press.

Gough, David. 1996. Black English in South Africa in *Varieties of English Around the World: Focus on South Africa*. Ed. de Klerk, Vivian. Amsterdam: John Benjamins Publishing Company. 15, 53-78

Makhudu, K.D.P. 2002. An Introduction to Flaaitaal (or Tsotsitaal) in *Language in South Africa*. Ed. Mesthrie, Rajend. University of Cape Town, Cambridge University Press. 398-406.

McCormick, K. 1995. Code-switching, Code-mixing and Convergence in Cape Town in *Language and Social History*. Ed. Mesthrie, Rajend. Cape Town & Johannesburg, David Philip. 193-208.

McWhorter, John H. 2005. *Defining Creole*. Oxford, University Press.

Mesthrie, Rajend. 2002. South Africa: A Sociolinguistic Overview in *Language in South Africa*. Ed. Mesthrie, Rajend. University of Cape Town, Cambridge University Press. 11-26.

---. 2008. 'I've been speaking Tsotsitaal all my life without knowing it' – towards a unified account of tsotsitaals in South Africa in *Social lives in language--sociolinguistics and multilingual speech communities : celebrating the work of Gillian Sankoff*. Meyerhoff, M., Nagy, N (ed). Philadelphia: John Benjamins.

[5] Molamu, Louis. 2003. *Tsotsi-Taal: A Dictionary of the Language of Sophiatown*. Pretoria, University of South Africa.

Motshogo, Lebo. 2005. *Township Talk: The language, the culture, the people*. The A-Z Dictionary of South Africa's Township Lingo. Cape Town, Double Storey Books.

Ntshangase, Dumisani Krushchev. 1995. Indaba yami i-straight: Language and Language Practices in Soweto in *Language and Social History*. Ed. Mesthrie, Rajend. Cape Town & Johannesburg, David Philip. 291-297.

[3] Oxford English Dictionary (OED). 2008. Oxford: Oxford University Press. Accessed from <http://www.oed.com/>

Slabbert, Sarah and Finlayson, Rosalie. 2000. 'I'm a cleva!': The Linguistic Makeup of Identity in a South African Urban Environment in *Language and Ethnicity in the New South Africa*. Ed. Kamwangamalu, Nkonko M. *International Journal of the Sociology of Language*. 144, 119-135.

---. 2002. Code-switching in South African Townships in *Language in South Africa*. Ed. Mesthrie, Rajend. University of Cape Town, Cambridge University Press. 235-257.

Slabbert, Sarah and Myers-Scotton, Carol. 1996. The Structure of Tsotsitaal and Iscamtho: Code Switching and In-Group Identity in South African Townships. *Linguistics*. 34:2, 317-342.

[4] South African Concise Oxford Dictionary. 2002. Ed. The Dictionary Unit for South African English. Cape Town, South Africa, Oxford University Press Southern Africa.

[6] Urban dictionary. Accessed from [www.urbandictionary.com](http://www.urbandictionary.com)

van der Merwe, I.J. and van der Merwe, JH (ed). 2006. *Linguistic Atlas of South Africa: Language in Space and Time*. Stellenbosch: SUN PRESS.