

A finite-state morphological analyzer using foma for Klingon, a constructed language

Matthew David Dreselly Thomas

University of Colorado at Boulder

matt@bocosoft.net

Abstract

This paper presents a morphological analyzer for Klingon, constructed language, developed following the official description of the morphology in the Klingon Dictionary (Okrand, 1992).

The open-source tool *foma* for constructing finite-state automata and transducers is used for building the morphological analyzer.

1 Introduction

Klingon was created by Marc Okrand, an American Linguist, for the 1984 movie *Star Trek III*. Klingon is spoken by the fictional Klingons in the *Star Trek* universe, so Okrand designed the Klingon language to sound alien by using grammatical constructs and sound combinations infrequently used in natural language (KLI, Development; KLI, Sounds).

Klingon has a complete grammar that can be spoken and used as a real language. Marc Okrand formalized the Klingon grammar and vocabulary in *The Klingon Dictionary* (Okrand, 1992).

The number of fluent Klingon speakers is not known, but has been estimated between a few dozen (Okrent, 2009) to around 100 (Kelly, 2013). *The Klingon Dictionary* (Okrand, 1992) has sold more than 300,000 copies and has been translated into 4 languages. Klingon has entered into popular culture, and literary works such Shakespeare's *Hamlet*, *Gilgamesh*, and "A Christmas Carol" have been translated into Klingon. The Bible is currently being translated (KLI, Bible) but is a difficult task due to limited Klingon vocabulary. The current lexicon contains almost 3,000 words (Zrajm, 2012) and only grammar and words coined by Marc Okrand are considered canonical by the Klingon speaking community (KLI, Canonical).

2 Klingon Morphology

Klingon is an agglutinative language that uses Object-Verb-Subject word order. Okrand categorizes down the Klingon word types into Nouns, Verbs, and Other.

2.1 Nouns

Nouns can take five types of suffixes. Each suffix type, if present, must be appended to the noun stem in a specific order and there can be no more than one suffix per type. The suffixes that a noun can take also depend on the type of noun. Nouns are split into three types; body-parts, beings capable of language, and general nouns.

Type 1 Augmentative/diminutive

Type 2 Number

Noun Type	Suffix
Body-Part	-pu'
Language-Capable	-Du'
Generic	-mey

The *-mey* suffix may also be used with body-part and beings-capable-of-language nouns, but the resulting word acquires an additional meaning "*all over the place*" on top of the plural meaning.

Klingon	English
jup	friend
jupDu'	friends
jupmey	friends all over the place

Table 1: Generic Plural Suffix Usage with Inherently Singular Nouns

There are some nouns that are inherently singular or inherently plural. For example the Klingon word *peng* means torpedo (singular)

but *cha* means torpedos (plural). The singular forms of these types of nouns may also take the *-mey* generic plural suffix, but the resulting word also acquires the "all over the place" meaning.

Klingon	English
peng	torpedo
pengmey	torpedos all over the place
cha	torpedos
*chamey	<i>Not allowed</i>

Table 2: Generic Plural Suffix Usage with Language-Capable Nouns

Lastly, nouns in Klingon are not required to take any plural suffix in order to be plural. The singular form of a noun may be used where the plural form would normally be used, and this is perfectly grammatically correct. The meaning is disambiguated using the context of the word in the sentence. Therefore *jup* means both "friend" and "friends" depending on the context.

Type 3 Qualification

Type 4 Possession/specification

Possession suffixes are selected depending on whether the *object* being referred to is language-capable.

Possessive Suffixes	Generic	Language Capable
my	-wIj	-wI'
your	-IIj	-II'
his, her, its	-Daj	-Daj
our	-maj	-ma'
your (plural)	-raj	-ra'
their	-chaj	-chaj

Type 5 Syntactic markers

2.2 Verbs

Verb can take both prefixes and suffixes. Prefixes specify both the subject and object of the verb, including whether or not a subject and/or object exists. Verbs can use either the **Indicative mood** or the **Imperative mood**, which can also affect whether some suffixes can be used.

Verbs are categorized into **stative** and **dynamic**. Klingon does not use adjectives so stative verbs

are used to describe nouns. For example to say "red ship", one would use the stative verb *Doq* meaning to be red/orange, "*Duj Doq*".

Verbs conform to nine ordered suffix types. Just like nouns, each suffix type, if present, must be appended to the verb stem in a specific order. There can be no more than one suffix per type.

Type 1 Reflexive, Reciprocal

Type 2 Volition, Necessity

Type 3 Inceptive, Inchoative

Type 4 Causative

Type 5 Undefined Subject, Capability

Type 6 Perfection, Uncertainty

Type 7 Aspect

Type 8 Honorific

Type 9 Syntactic markers

In addition, a "rover" suffix type can be appended immediately after the verb stem or between the nine suffix types, however rover suffixes cannot be appended after a Type-9 suffix. There are no restrictions on the number of rover suffixes that may be used. A rover suffix's meaning applies to the suffix immediately proceeding it. For example the **negating** rover suffix *be'* can be appended immediately after the verb stem and/or after any other suffix type.

Klingon	English
Dub	to improve
Dubbe'	to not improve
DubmoH	to cause to improve
DubmoHbe'	to not cause to improve
Dubbe'moH	to cause to not improve
Dubbe'moHbe'	to not cause to not improve

Table 3: Verb Rover Suffix Usage

Rover suffixes can even be appended after other rover suffixes, e.g. *Dubbe'be'* meaning *to not not improve*.

2.3 Other Words

Pronouns

Numbers

Conjunctions

Adverbials
Exclamations
Names
Questions

3 Klingon Morphophonology

Klingon completely lacks morphophonology with one exception. The endearment noun suffix *-oy* is prepended with a glottal stop ' if the noun stem ends with a vowel.

Noun Stem	Noun with Endearment Suffix -oy
jup (friend)	jupoy (dear friend)
cha (torpedo)	cha'oy (dear torpedo)

Table 4: Endearment Suffix

4 FST Implementation

Used lexc file. Used foma file for rewrite rule(s).

References

Marc Okrand. 1992. *The Klingon Dictionary*, 2nd edition. New York: Pocket Books, Simon and Schuster Inc.

Development and Use of the Klingon Language. The Klingon Language Institute <http://www.kli.org/about-klingon/klingon-history/>

The Sounds of Klingon. The Klingon Language Institute <http://www.speakklingon.info/about-klingon/sounds-of-klingon/>

Canonical Sources. The Klingon Language Institute http://www.kli.org/wiki/Canonical_sources

Klingon Bible Translation Project. The Klingon Language Institute http://www.kli.org/wiki/Klingon_Bible_Translation_Project

Arika Okrent. 2009. *In the Land of Invented Languages*, 1st edition. New York: Spiegel & Grau

Zrajm. 2012. *Klingon Word Database*, 1st edition. Klingonska Akademien

Samantha Murphy Kelly. 2013. *How Klingon Became a Universal Language* Mashable.com <http://mashable.com/2013/05/18/klingon-star-trek/>

5 Examples of text

The proceedings are designed for printing on A4 paper. yusuke@nii.ac.jp *pdflatex dvips*

Hello there