The Toki Pona Language: an overview and some hacks

Renato Fabbri

`renato.fabbri@gmail.com`

University of São Paulo,
Institute of Mathematical and Computer Sciences
São Carlos, SP, Brazil

November 11, 2017

Abstract

1 Introduction

Toki Pona is a minimalist conlang (constructed language) with only 126 words.

1.1 Resources on the web

http://tokipona.net

1.2 Historical note

1.3 Natural and constructed languages

2 Overview of the language

2.1 Phonology

Words in Toki Pona are written using only 14 letters:

- Vowels a (open), e (mid front), o (mid back), i (close front), u (close back).

- Consonants j, k, l, m, n, p, s, t, w:

  - Nasal: m (labial), n (coronal).
  - Plosive: k (dorsal), p (labial), t (coronal).
  - Fricative: s (coronal).

– Approximant: j (dorsal), l (coronal), w (labial).

There are standard guidelines for pronunciation, but the language allows for considerable allophonic variation. For example, /p t k s l/ might be pronounced [p t k s l] or [b d g z ɾ].

Syllables are of the form (C)V(N): an optional consonant, a vowel and an optional nasal consonant. Non word-initial syllables must follow the pattern CV(N). The following sequences are forbidden: ji, wu, wo, ti, mn, nm, mm, nn.

2.2 Syntax

As in other natural languages, colloquial Toki Pona might have incomplete sentences and deviate from the norm. The basic structure of sentences are in the form: ¡subject¿ li ¡predicate¿ e ¡object¿. The li might be repeated to associate more than one predicate to the subject. The particle li is omitted if the subject is a simple mi (I or us) or sina (you). A discussion about problems with this rule and how I deal with them is in Appendix A.

The e might be repeated to associate more than one object to a predicate. Sentences might be related though la, 'sentence' la 'sentence', where the second sentence is the main sentence, and the first sentence is a condition to the first. Multiple la-s are not described in literature, but I assume that one might assume the last sentence being a conditional to the next, except in cases where the context strongly suggests otherwise.

Noun and verb phrases are built with the non-particle words. The first word is the noun and phrase and subsequent words qualify the noun or phrase. The pi particle might be used to separate sequences of words to be evaluated before the relations yield by pi: As pi is often ill understood and used, the following structures might be handy for newbies and as a reference:

- No pi, 'word word word': word ← (qualifies 1) word ← (qualifies 2) word.

- One pi, 'word pi word word': word ← (qualifies 2) [word ← (qualifies 1) word].

- Two pi-s: 'word pi word word word pi word word': word ←5 [word 2 word] 3 word ←4 word 1 word; or: word ←5 [word 1 word] 2 word ←4 word 3 word.

Notes on the usage of pi:

- In a sequence of words, without pi, the second word qualifies the first, the third word qualifies the phrase yield by the first two words, the fourth word qualifies the noun yield by the first three words and so on.

- It is redundant to use pi before the last word in a noun or verb phrase if there is no other pi, reason why it is most often omitted. Its use in this case is regarded as wrong [?, ?], but, as one might notice, it does not introduce any ambiguity.

- The book by jan Pije [10] describes another use for pi: after li to mean possession, e.g. 'soweli li pi sina' (your pet). This employment of pi might be regarded as correct, but are promptly written as a noun phrase (e.g. 'soweli sina') and is not mentioned by the official book [?].

All the words except the structural particles (li, e, la, pi) are usable in noun and verb phrases. Notice that the phrase expresses a noun in a noun phrase (subject or object) or a verb (in the predicate).

At this point, the only missing syntax rule is related to the prepositions: kepeken, lon, sama, tan, tawa. They might appear at the end of noun phrases, should be followed by another noun phrase, and require no particle. E.g. 'toki tan jan Pije li pana e sona tawa mi'.

Other particles are: a or kin, o, taso, anu, en, nanpa, seme. mu?,

Vocatives should have an o after the noun phase.

2.3 Further notes

The only synonyms on Toki Pona are: a and kin; lukin and oko; sin or kamako; ale or ali.

toki e ni, wile e ni, etc. 'e ni' can be omitted and : used by itself.

Names are by default transliterated,

2.4 Main references for the language

- The official book is "Toki Pona: The Language of Good" and is authored by Sonja Lang, the creator of the language.

- The book "o kama sona e toki pona!", from jan Pije, is the other main reference for the language [10].

- [11].

3 Analysis and hacks for the language

3.1 Hacks from other people

The `tokipona.net` has a number of tools, just as to transliterate names into Toki Pona phonemes, and search in corpus.

4 Conclusions and further work

- Relate Toki Pona to Wordnet: should one Toki Pona word be related to more then one synset of the English language?

- Understand how the corpus is gathered in `tokipona.net`.

- Know about previously existing words that were used for Toki Pona (e.g. suno and suwi might come from sun and sweet), and about the reasons that lead Sonja (and maybe other people) to choose the 14 letters and the syllable structure. This might require a dedicated communication with the speaker community and the documentation authors.

- Corpus-based analysis.

- Publication of original texts and translations.

Acknowledgments

A My usage of Toki Pona

I use the standard sounds, but often use [z] for s. I often translate texts to Toki Pona (e.g. biblical excepts) and create new texts as poems and short stories. Most of them are in [?]. I omit the li particle after subjects sina and mi, in accordance with the norm, but sometimes I use them when there are many predicates. E.g. sina li wawa li pimeja li lukin pona li moku e kasi mute. In such cases, the first li is sometimes omitted. Also, sometimes I use li before mi and sina where I find that there is unwanted ambiguity, e.g. sina moku pona e jan (might be sina li moku pona e jan or sina moku li pona e jan).

Names are by default transliterated, but I advocate that, as in other languages, names might be used as they are in the correspondent mother tongue. E.g. the name Erdös is used in English and Portuguese although the standard alphabet does not contain ö in such languages. I also tend to legitimate the use of English (or German) words in Toki Pona texts if it is the case, as happens often in scientific writing (kernel is a German word used in English, webpage is an English word used in Portuguese).

References

[1] Anthropological physics and social psychology in the critical research of networks. Complex Networks Digital Campus (CS-DC'15). Available at https://youtu.be/oeOKYc3-nbM

[2] Fabbri, R. What are you and I? [anthropological physics fundamentals], 2015. Available at https://www.academia.edu/10356773/What_are_you_and_I_anthropological_physics_fundamentals_

[3] Ouroboros. (2017, November 2). In Wikipedia, The Free Encyclopedia. Retrieved 22:19, November 9, 2017, from https://en.wikipedia.org/w/index.php?title=Ouroboros&oldid=808392809

[4] Fabbri, R. (2017). A reasonable vimrc file. Available at https://raw.githubusercontent.com/ttm/vim/master/vimrc

[5] Ex (text editor). (2017, March 22). In Wikipedia, The Free Encyclopedia. Retrieved 22:22, November 9, 2017, from https://en.wikipedia.org/w/index.php?title=Ex_(text_editor)&oldid=771621020

[6] Fabbri, R. (2017). A Toki Pona Python Package and Vim Syntax Highlighting. Available at https://github.com/ttm/tokipona

[7] Lang, S. (2014). Toki Pona: the language of good. Tawhid Publishing. ISBN-10: 0978292308, ISBN-13: 978-0978292300.

[8] Fabbri, R., Fabbri, R., Vieira, V., Penalva, D., Shiga, D., Mendonça, M., Negrao, A., Zambianchi, L., & Thumé, G. (2013). AA: The Algorithmic Autoregulation (Distributed Software Development) Methodology. RESI. From https://arxiv.org/abs/1604.08255

[9] Fabbri, R. (2017). The Algorithmic-Autoregulation (AA) Methodology and Software: a collective focus on self-transparency. ENMC2017. From https://github.com/ttm/ensaaio/raw/master/emc/article.pdf

[10] Knight, B. (2017) o kama sona e toki pona! Available at: http://tokipona.net/tp/janpije/okamasona.php

[11] Toki Pona community (2017). Wikipesija. Available at: http://tokipona.wikia.com