2024-04-16

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **Preliminary proposal to encode sitelen pona in the UCS**
Source: **Sitelen Pona Working Group**
Authors: **Rebecca Bettencourt (jan Lepeka), Janet Black (akesi Jan), Ethan Corgatelli (jan Itan), Gregory Danielson III (jan Kekan San), Isaac Jones (jan Asiku), lipamanka, Spencer van der Meulen (jan Pensa), João Reis (jan Juwan), Ret Samys (jan Ke Tami), Shestakov Maxim (kala Asi), jan Tepo, Alison Withonel (seme Alison), Anders Yu (jan Kokan Ato)**
Status: **Individual Contribution**
Action: **For review by JTC1/SC2/WG2 and UTC**
Date: **2024-04-16**

This is a preliminary proposal to encode the *sitelen pona* script in Unicode. It provides a brief background on the script, a description of the proposed encoding model, a tentative character repertoire and related properties, orthographic details, and specimens of the script. Research is ongoing and a formal proposal is forthcoming. The information presented here is subject to change.

1. Introduction **2**
2. History **3**
3. Methodology **6**
3a. Collecting media 6
3b. Criteria for encoding logographs 6
3c. Criteria for encoding future logographs 8
3d. Collecting usage data for encoding Standardized Variation Sequences 9
3e. Criteria for encoding future Standardized Variation Sequences 10
3f. Determining collation order 11
4. Structure **12**
4a. Variant glyphs 13
4b. Combined glyphs 17
4c. Numbers 18
4d. Cartouches 19
4e. Punctuation 21
4f. Extended glyphs 22
4g. Personal glyphs and name glyphs 25
4h. Legacy characters 26
5. Copyright status **26**

6. Proposed character repertoire **26**
6a. Character names 27
6b. Code point assignment 27
6c. Unicode character properties 27
6d. Bidi class and bidi mirroring 29
6e. Line breaking 30
6f. Collation order 30
6g. Standardized Variation sequences 31
6h. Vertical orientation 31
6i. East Asian width 31
6j. ISO 15924 script code 32
7. Acknowledgements **32**
8. References **32**
Code charts **44**
List of sitelen pona logographs **46**
Figures **67**
Works by Sonja Lang 67
Works by other authors 75
Fonts and input methods for sitelen pona 174
Search results, graphs, and statistics 188

1. Introduction

This document proposes the addition to the UCS of 125 text characters and 6 format control characters (131 characters total) for the *sitelen pona* script used to write the popular constructed language *toki pona*.[1] The toki pona language was created by Canadian polyglot Sonja Lang (also known as *jan Sonja*[2]) in 2001. The sitelen pona script was first published in full in Sonja Lang's book *Toki Pona: The Language of Good* in 2014, although excerpts were published earlier. Today, there are known to be thousands of speakers of toki pona worldwide, with likely thousands more involved in the online and offline spaces where the 2022 toki pona census could not reach. According to the census, 73% of speakers know and 50% of speakers actively use sitelen pona. (See Figures 178 and 179.)

Speakers of toki pona have formed a large online community. On Discord, the largest toki pona community claims approximately 13,000 members (as of February 1, 2024), more than the largest Na'vi community at approximately 9,600 members, the largest Interslavic

[1] In toki pona, only words which function as proper names are capitalized. All toki pona words, including the name of the language and its script, are always written in lowercase. Most speakers follow this convention even when referring to toki pona or sitelen pona in English.

[2] Proper names in toki pona are preceded by a noun called a *head noun*. Speakers of toki pona often refer to each other using a head noun followed by a name that has been *tokiponized* (modified to conform to toki pona phonotactics), even when speaking another language. The most popular head noun for toki pona speakers themselves is *jan* (meaning "person," and pronounced [ˈjan] in the IPA, similarly to the English word "yawn"), although others are possible, such as *soweli* ("animal") and *kulupu* ("group"). This document refers to many people within the toki pona community using this format.

community at approximately 8,000 members, or the largest Esperanto community at approximately 7,400 members. On Reddit, the largest toki pona community claims approximately 19,000 members (as of January 31, 2024); even if only half use sitelen pona, that is still more than the largest communities of two other languages known for their unique scripts, Sindarin and Quenya, at approximately 7,500 and 7,300 members, respectively. (See Figure 169.) Unlike Tengwar and Cirth, however, sitelen pona is not potentially encumbered by restrictive copyright protections. (See Section 5 and Figures 4 and 7.)

There are notably large efforts to teach toki pona outside of Sonja Lang's book. On Discord, the second largest toki pona community, *kama sona*, claims approximately 4,600 members (as of January 2024) and is designed entirely around teaching the language. There are over 40 members identified by the community as being proficient in the language and teaching, translation exercises held multiple times weekly, and multiple highly active channels for different types of questions about the language. This is only one teaching-centric community, however; virtually every community is equipped for effective teaching, often with two or more channels specifically for learning and asking questions, and other channels for practicing. Current speakers are generally enthusiastic about teaching the language.

The community also recommends a variety of learning resources for different purposes and tailored to different types of learners. The first notable example of lessons focusing on sitelen pona is the series *o kama sona e sitelen pona* by James Flear (jan Same). This and many of Flear's related works (such as the *linja pona* font used in the lessons) were historically the most influential in the adoption of sitelen pona as a writing system. Later lessons which reference and teach sitelen pona, but do not make it central to the lessons, include *lipu sona pona* by jan Lentan in 2020 and *lipu sona mi pi toki pona* by soweli Tesa in 2021. The series *Learn Toki Pona!* by Gregory Danielson III (jan Kekan San), introduced as written and video lessons in August 2022, teaches toki pona while using sitelen pona as the primary writing system. It is organized around specific grammatical concepts like previous written lessons, but features many more examples.

Many toki pona speakers are technologically inclined, as demonstrated by the multitudes of fonts (many with advanced OpenType features), input methods, web-based converters, and other applications made for sitelen pona. However, despite best efforts, support for sitelen pona is still a challenge in environments such as Discord, Google Docs, and on mobile devices where custom fonts cannot be added easily.

2. History

In August 2001, Canadian polyglot Sonja Lang first introduced the toki pona language in a public post on her website at GeoCities. It described toki pona as "a cute, pidgin-like language designed to express basic human realities by reducing words and ideas to their most simple and universal meanings and avoiding unnecessary complexity. [...], the entire language has only 14 sounds and under 200 words." The website was later moved to tokipona.org, where it remains to this day.

Between 2002 and 2003, Bryant J. Knight (jan Pije) wrote the second set of toki pona lessons, the first being Sonja Lang's own. Knight's lessons were widely used by many speakers until their discontinuation around 2020. During this time period, the community

grew primarily on Yahoo! Groups and IRC. The Yahoo! group was abruptly closed in October 2009, but was archived onto its successor, the forums at forums.tokipona.org.

In 2004, Eliazar Parra Cárdenas wrote the Spanish-language document *toki pona en 76 lecciones ilustradas* ("toki pona in 76 illustrated lessons"), the first example of toki pona lessons written in a language other than English. These lessons were later translated to English, Esperanto, and French, and were influential in the growth and education of toki pona, being occasionally recommended as educational material until as recently as 2021. Notably, the 2013 Esperanto translation of these lessons would be the first formally published work on toki pona, being published by an Esperanto publishing house and having its own ISBN. The word list taught in these lessons has some notable differences from the list in *Toki Pona: The Language of Good*. For example, in the original version, the word *pata* ("sibling"), which is no longer used today, is included, and the words *alasa*, *esun*, and *pan* are not included as they hadn't been coined yet. Later translations would alter the word list to reflect contemporary usage.

In October 2009, Sonja Lang closed the Yahoo! group and replaced it with the Toki Pona Forums, active from then until 2020. These forums are still open today as a read-only archive. This move was done at the same time as adding a wiki to her site tokipona.org. This wiki hosted early versions of the notes Lang had written for her future book, as well as introducing a small number of new words to the language on the official dictionary page of the wiki.

In October 2010, jan Mimoku created the subreddit r/tokipona, which today is the largest toki pona community by membership, having nearly 20,000 members as of March 2024.

In May 2014, Sonja Lang published *Toki Pona: The Language of Good* (also known as *pu*[3]), the definitive guide to the toki pona language, which included the complete sitelen pona script for the first time. The section on sitelen pona lists 120 logographs for the 120 words designated *nimi pu* ("words from *Toki Pona: The Language of Good*"). (See Figures 1a and 1b.) Words outside this original set of 120 are known as *nimi sin* ("new words"; this is considered a misnomer as some of these words even predate the publication of *pu*, but the term has stuck). Three *nimi sin* mentioned in *pu* (*namako*, *kin*, and *oko*) were considered to be synonyms of three *nimi pu* (*sin*, *a*, and *lukin* respectively); today they are considered separate words to various degrees.

In 2015, Mitch Halley (jan Misali) created a series of videos, *12 Days of sona pi toki pona*, for learning toki pona. They have only recently stopped being recommended by the community due to better resources replacing them, including a newer series by the same creator started in 2021 and still being produced.

In July 2019, James Flear (jan Same) created a series of video lessons titled *o kama sona e sitelen pona*, the first example of toki pona lessons which focus on sitelen pona. Prior to this, sitelen pona was primarily taught separately from the language, such as via flash cards. In addition to proliferating sitelen pona as a writing system, it also influenced its design: Flear used his own font *linja pona* for his lessons, which was (and to some degree still is) extremely popular and influenced many later fonts. The design of *linja pona* is notably more rounded and uniform than that seen in Lang's original publication of the script.

[3] Technically, the terms *pu* and *ku* refer to anything related to or involving interaction with the books *Toki Pona: The Language of Good* and *Toki Pona Dictionary*, respectively, and not just the books themselves. However, they are often used as shorthand to refer specifically to the books as well.

In 2020, kala Asi conducted the first toki pona word usage survey, with 86 respondents. The survey has been conducted every year since then and is coordinated through the *kulupu Linku* Discord server. The 2023 survey had 868 respondents. Words are sorted into one of seven *usage categories* based on how many respondents answered they used each word. The seven usage categories are *core* (90%–100%), *widespread* (70%–90%), *common* (50%–70%), *uncommon* (20%–50%), *rare* (10%–20%), *obscure* (2%–10%), and *not notable* (0%–2%), with the final category indicating the word will be dropped from the *lipu Linku* dictionary (see Figure 80).

In 2021, Lang published the *Toki Pona Dictionary* (also known as *ku*), a collection of common translations between English and toki pona collected from community surveys on the *ma pona pi toki pona* Discord server throughout 2020. The dictionary includes a number of *nimi sin*, 17 of which were found to be in common use (recognized by at least 40% of speakers at the time) and designated *nimi ku suli* ("major words from the *Toki Pona Dictionary*"). The other 41 words, which were not in common use but were still recorded by the surveys, were designated *nimi ku pi suli ala* ("non-major words from the *Toki Pona Dictionary*"). This category would later be labeled *nimi ku lili* ("minor words from the *Toki Pona Dictionary*") by various community dictionaries, and this name is more widely recognized today.

Shortly after the publication of the *Toki Pona Dictionary*, Gabriel Tellez submitted document L2/21-137 entitled "Toki Pona for Unicode." Far from a complete proposal, it consisted only of two short paragraphs and a table of glyphs from *linja pona*. The proposal summary form offered only the phrase "it's a conlang" as context. The Script Ad Hoc rightly rejected the proposal in their recommendations in L2/21-130.

In August 2021, Spencer van der Meulen (jan Pensa) submitted change request 2021-043 to create an ISO 639-3 language code for toki pona. SIL International adopted the change request in January 2022 and assigned toki pona the language code `tok`.

In December 2021, Julius Strake (jan Juli) translated and published *Toki Pona: Die Sprache des Guten*, the German edition of *Toki Pona: The Language of Good*. With this edition, sitelen pona was released into the public domain. (See Figures 4 and 7.)

In early 2022, jan Tepo, Ethan Corgatelli (jan Itan), and Rebecca Bettencourt (jan Lepeka) formalized and published an encoding of sitelen pona in the Private Use Area of Unicode as part of the Under-ConScript Unicode Registry (UCSUR), a fork of John Cowan and Michael Everson's ConScript Unicode Registry (CSUR), neither of which are affiliated with the Unicode Consortium or the Unicode Technical Committee (UTC). Since then, the encoding has been used in many fonts by various developers, input methods for various platforms, web-based utilities to convert between the Latin and sitelen pona scripts, and tutorials for writing sitelen pona. The "sitelen pona UCSUR guide!!!" by Lily Larsen (jan Lili) lists three ways to modify the Discord chat client to enable the display of sitelen pona and seven different input methods for multiple major platforms (Windows, Mac OS, Linux, Android, and Web). The *ilo Linku* and *qbot* Discord bots, as well as many others, include functionality for rendering sitelen pona.

In October 2022, Spencer van der Meulen published *Tokipono: La lingvo de bono*, the Esperanto edition of *Toki Pona: The Language of Good*. This edition was the first to include

sitelen pona logographs designed and adopted by the community for the 17 words designated *nimi ku suli* in the *Toki Pona Dictionary*, as well as popular alternate forms of *akesi*, *namako*, and *sewi*. (See Figures 2a, 2c, and 3.)

In November 2023, a Proton Mail user named Michael attempted to submit the UCSUR page for sitelen pona as a proposal to the UTC. It was forwarded to the Script Ad Hoc, who recognized it as a proposal for the ConScript Registry and not a proper proposal for The Unicode Standard. It was then forwarded to the maintainers of the CSUR and UCSUR, who were of course already aware of its existence. This incident led to the creation of this proposal.

In February 2024, Lang published *The Wonderful Wizard of Oz (Toki Pona edition)*, the first book in the *su* series of official toki pona translations. This abridged version of L. Frank Baum's classic story is presented in both English and toki pona, with the toki pona translation written entirely in sitelen pona. (See Figure 6.) This proposal includes and takes into account the pre-release notes (with the approval of Lang herself), which include several variant glyphs. (See Figure 5.)

Given the history of attempts to encode sitelen pona, the enthusiasm of the community toward creating solutions for doing so in UCSUR, and the widespread adoption of the script among the community, we feel the only missing piece to applying for an encoding in the UCS is a formal and well structured proposal. The interest is already more than present.

3. Methodology

3a. Collecting media

The toki pona community is significantly larger and more prolific with respect to media creation than a single person or even a group can hope to explore. There are thousands of speakers across dozens of communities who use many different languages besides toki pona.

We created a short Google Form, available in both English and toki pona, for the collection of media which uses sitelen pona (see Danielson 2024a, 2024b), which allowed us to quickly filter through the media, ensuring a wide spread of content mediums and both original and translated work were demonstrated.

The survey was provided via announcement channels in all immediately accessible toki pona communities on Discord, Facebook, Reddit, and Telegram, and drew a total of 177 responses. In addition to this form, a large number of media examples were collected by those working on the proposal.

3b. Criteria for encoding logographs

The methodology and criteria described in this section are subject to a rework pending the completion of research which studies toki pona communities directly to determine usage, rather than relying exclusively on what survey respondents self-report. As such, the

information presented here is subject to change, and should not be taken as authoritative or reflective of the final result.

To determine which logographs to include in this proposal, we began by examining all pre-existing categories of words. These categories are: *nimi pu* (words discussed in *Toki Pona: The Language of Good*); *nimi pu* plus synonyms which adds the words *kin*, *namako*, and *oko*; *nimi ku suli*, which adds another 14 words; the UCSUR inventory, which adds the historical words *apeja*, *majuna*, *pake*, and *powe* to the previous categories; the *Nishiki-teki* inventory, named after a notable font which has glyphs for every word in UCSUR plus *linluwi* and *kiki*; and finally, the 259 words reported to be used by at least 2% of speakers according to the 2023 word usage survey conducted by *kulupu Linku*.

Those involved in the discussion agreed that including words such as *wuwojiti* and *yupekosi* would not be reflective of the language. The word *wuwojiti*, a mnemonic for the disallowed syllables in toki pona phonotactics, is primarily used in other languages while teaching toki pona. The word *yupekosi*, a joke word which intentionally breaks the phonotactics of the language, is reported to be used by a non-negligible number of speakers, likely due to being introduced by Sonja Lang; however, there is almost no demonstrable use of the word outside of English. Linku tracks several words like these which are descriptive of the language but generally not considered part of the language. For these reasons, we set the baseline for encoding by reported usage to at least 25% as indicated by the 2023 word usage survey, just above these words.

Based on the above, our criteria began as the following: A word must have two consecutive years of at least 25% usage among the community, as opposed to just recognition. The distinction between usage and recognition is important, as the 2020 and 2021 Linku surveys erroneously increased the recognition of words through its existence and distribution (Shestakov, 2022). If interpreted strictly, the requirement of 25% usage over two years would have excluded *apeja*, *pake*, *powe*, and *kiki*, and would have included *majuna* and *linluwi*. Notably, it would have included all of the *nimi ku suli*, which was a good indicator of the metric's usefulness.

However, it was thought that this line would not be sufficient for the initial proposal for a few reasons. Several of those involved in the proposal considered the inclusion of all words in UCSUR to be an obligation so that it would be possible to convert between the two encodings. This would require including the historical words *apeja*, *pake*, and *powe* regardless of their usage. These historical words saw usage in the community in a period further back than the Linku survey has data for, such that we cannot say with confidence what portion of the community considered these words part of their lexicon without performing a complete corpus study of what speech remains from that period. As such, the inclusion of these words came down to a judgment call built on the understanding of the historical relevance of these words, rather than strictly following the "25% usage over two years" criteria described above.

A poll was conducted among those involved in the discussion to rank all the words from Linku's *widespread*, *common*, *uncommon*, and *rare* categories from 1 (do not encode) to 5 (encode). The result showed middling to strong support for all words from the level of usage of *kiki* (29% in 2023) and above, with the expected exception of the word *nimisin* due to its thus far unproven staying power (35% in 2023, but only 13% in 2022). Notably, the words

kiki and *linluwi* actually outperformed *epiku* (one of the *nimi ku suli*) in this poll. There was also middling to strong opposition to words at or below the usage of *usawi* (22% in 2023). However, three of the four historical words (*apeja*, *pake*, and *powe*) and one of the *nimi ku suli* (*kokosila*) were between entirely ambiguous and middling opposition (see Figures 183 and 184).

After further discussion, these four were considered as a group in one last informal poll, which showed strong support for encoding all of them as a unit (see Figure 185). This result gives us exactly the Nishiki-teki inventory described at the outset.

Note that this survey will not be material to the final chosen inventory, as we are changing directions to a community study rather than a short and informal opinion study. We expect to have a result which is more reflective of the community through this, but gauging the interests and opinions of the committee involved in this proposal is still valuable.

3c. Criteria for encoding future logographs

The criteria described in this section are subject to a rework pending the completion of research which studies toki pona communities directly to determine usage, rather than relying exclusively on what survey respondents self-report. As such, the information presented here is subject to change, and should not be taken as authoritative or reflective of the final result.

For any future proposals, our criteria for proposed logographs will be the same as our starting point from the previous section: a word must have at least 25% usage among the community for a period covering at least two notable usage surveys with a year between them. Right now, the only such usage survey is the annual Linku survey. This criteria is chosen for several practical reasons. First, if a word is over 25% usage for a period of at least two years, it has demonstrated staying power. (Many words in Linku's *obscure* category as of 2023 had notably higher one-off performances in prior years, primarily due to the word's popularity around the time of the survey as opposed to its long-term effect on the community.) Second, observing more than one period of survey results helps avoid one-off surveying errors from causes such as attempts to artificially increase the popularity of a word or changes to the survey's methodology. The line of 25% is primarily chosen for being just above the usage levels of *wuwojiti* and *yupekosi*; the reasoning for this is stated in the previous section (they are not considered words in the language proper, but a teaching tool and a joke made by the creator of the language, respectively).

There is room for the two survey periods and the 25% usage line to be changed in the future. If the community were to grow significantly in size, a lower line of usage may become appropriate as relatively smaller communities form and adopt locally relevant words. It is worth acknowledging that this phenomenon is already occurring in various smaller communities, but the vast majority of words proposed in these communities are relegated to a short period of popularity, a negligible number of users, or both. This is due to the much smaller size of these communities, with membership measured in the hundreds and daily active users measured in the dozens. Alternatively, it may become necessary to consider words over periods covered by more than two surveys, such as if more organizations began

tracking usage data with different methodologies, or if the existing surveys were to change methodology or frequency.

3d. Collecting usage data for encoding Standardized Variation Sequences

> The criteria described in this section are subject to a rework pending the completion of research which studies toki pona communities directly to determine usage, rather than relying exclusively on what survey respondents self-report. As such, the information presented here is subject to change, and should not be taken as authoritative or reflective of the final result.

To gather usage data on variant glyphs (see section 4a), we sent out an ad hoc survey with two sections to many large communities. The first section was primarily for demographic analysis, asking respondents to identify how long they had been using toki pona and sitelen pona and how frequently they used sitelen pona via different mediums. The second section provided respondents with a list of notable variants for a word and asked to identify which of those variants they use (or prefer if they do not use the word) by selecting any number of the provided variants or by write-in. Respondents were also asked to indicate their relationship to the word among "I use this word in sitelen pona," "I use this word, but little or none in sitelen pona," "I do not use this word," and "I do not use this word, and I have no preference for its glyph." The first three options allowed users to select or write in descriptions of variants; the fourth option disallowed this.

This survey was designed to quickly obtain usage numbers for variants where previously we had only anecdotal evidence, which involved taking some shortcuts. For example, presenting all variants alongside the word they correspond to primed the respondent to assume the variants are already known to correspond to the given word, regardless of their prior understanding. A larger and more complete survey would ask users to identify the variant by name, instead of the survey identifying it for them, in order to avoid this bias. This would also distinguish variants which were created and persisted by font developers; for example, the star variant of *kokosila* is likely to be instead identified as a variant of *mun* by respondents who are not already familiar with that variant as a way to write *kokosila*. This more complete survey design would have required much more investment to complete on part of both the survey creator and the respondents. The Linku organization (*kulupu Linku*) is planning to conduct such a survey later this year at the same time as the annual word usage survey they already run (see section 3e).

The variants provided in the survey were collected by observing variants in many fonts and by collecting handwriting samples from images posted in various communities. While there is constant experimentation in sitelen pona, and thus a massive number of variants for most words, there are a small number of variants which have achieved enough popularity to be seen across multiple fonts and be used over multiple years. As such, identifying all of the relevant variants is a quick task; most font developers are already familiar with all of the popular variants. However, we did make one notable omission of the overlapping hearts variant of *olin*, which we initially considered too similar to the most popular variant to be included as a distinct option. This variant ended up being the most frequent write-in at 31 responses, or 33 if including a similar variant. We have chosen to accept this as a fault of our

methodology and treat these responses the same as every other response. We will address this fault in the future (see section 3e).

The survey concluded after two weeks and received 459 responses in total, roughly half that of the 2023 Linku survey. The majority of the results matched our expectations of what variants were popular. Notable exceptions include the thick-stem variant of *soko* as its most popular variant as opposed to second or third, and the *kulupu*-derived variant of *linluwi* as its second most popular variant as opposed to least popular.

We considered multiple interpretations of the data for determining our list of variants to encode. For example, we considered global usage, the same metric used to decide what logographs to encode (see Figure 186). This was considered undesirable, as it nearly prevents low-usage words such as *linluwi* and *majuna* from having variants, even though low-usage words are more likely to have multiple competing or interchangeable variants that should be included to accurately reflect the language. We also considered relative usage within words, where we compare the usage of any given variant to the most popular variant (see Figure 187). This approach is as effective on high-usage words as on low-usage words for identifying relevant variants, but does have a downside: a variant for a high-usage word such as *wile* may be unpopular compared to its most common variant, but still represent a greater number of respondents than two equally-used variants for a low-usage word. We also considered using the relative metric to derive a minimum adoption for the global metric, but this was considered too volatile due to being entirely controlled by low-usage words. Lastly, we considered relative usage proportional to global usage, in an attempt to normalize the difference between the two (see Figures 188 and 189). This approach seemed promising, but all it did was recreate the issues with the global usage metric.

After several weeks of discussion ending in a poll of those involved (see Figure 191), we chose to use the relative metric. We recommend encoding all variants which have at least one third as many users as the most popular variant for the same word. This metric represents the variants of lower-usage words and higher-usage words equally, with its only notable downside being that it can omit variants which have a higher number of total users than the variants of low-usage words. Regardless, we consider this to be highly reflective of the way sitelen pona is used.

However, Janet Black ran another survey in early March 2024 to determine whether or not the variants we had identified for the original survey were actually distinguished by those reading and writing them, or if they were instead even more subtle variances of the same glyph that would be better ascribed to handwriting distinctions than a fully distinct variant. This survey identified with confidence that there exists a distinction between the secular *sewi* variant and the Arabic-derived *sewi* variant, which matches the expectations of this committee. This same survey also identified with confidence that there exists little to no distinction between the four-legged variant and the six-legged variant of *akesi*. We would like to use a survey of this form to filter the list of variants we have determined to be in use, only encoding those which are distinguished by the community. However, as of writing, we have chosen to leave this analysis incomplete in favor of our change in direction for how we determine what logographs and variants should exist; ultimately, we recommend no variants to be encoded at this time.

3e. Criteria for encoding future Standardized Variation Sequences

The criteria described in this section are subject to a rework pending the completion of research which studies toki pona communities directly to determine usage, rather than relying exclusively on what survey respondents self-report. As such, the information presented here is subject to change, and should not be taken as authoritative or reflective of the final result.

At this time, our intention is to continue using the relative usage metric to determine what future variants to recommend for encoding, maintaining that recommended variants should be used by at least a third as many users as the most popular variant for a word. This will be maintained even if the most popular variant for a given word were to change, which is likely for words such as *akesi*. This will also be maintained for newly-encoded words. As noted previously, this methodology is more reflective of the volatile nature of variants for lower usage words, which makes it especially valuable for newly-encoded words that would presumably also be of lower usage.

One complication in continuing to use this metric is that we do not have any usage surveys studying sitelen pona variants except for the one we conducted for this proposal. The Linku organization plans to survey variant usage data starting in August 2024. Until that survey is drafted and run, the inclusion of variants will be primarily via ad hoc surveys focusing specifically on the variants we are aware of. There is only one known notable variant outside of those polled for the purpose of this proposal, so creating future surveys will be relatively easy as future notable variants would be easily identified. However, the methodology chosen by the Linku organization may not sufficiently match the structure of the survey used by the writers of this proposal. This would necessitate an alternate methodology of selecting variants, which we do not have the data to comment on at this time. In the event that we must reconsider our methodology, our intention is to choose a metric that is as close as possible to the one we chose for this proposal. Several of the writers of this proposal are involved in the Linku organization, and are intent on using the alternate survey methodology discussed in the previous section, which will not prime the respondent with the identity of the variant before they answer whether or not they use it.

There is also the possibility that we will reconsider the methodology used for selecting variants due to future changes in the community. For example, it was noted in the discussion of what logographs to encode (see section 3c) that a significant growth of the community could necessitate decreasing the minimum usage required to include a logograph. The same applies to variants, where we may want to decrease the required number of users of a variant relative to the most popular variant, or switch to a global usage or combined metric that reflects the different dynamics of the community. The most likely future change is that, as we acquire more variant usage data over multiple years, we will apply the same limit of a variant needing to appear over the course of two or more years across two or more usage surveys to be considered for inclusion. As with the discussion of logograph encoding, this would help control for changing survey methodology and recency bias in the popularity of a given variant.

3f. Determining collation order

Collation order was determined by a poll by Janet Black (akesi Jan), in which participants were asked to sort a given list of strings in sitelen pona “alphabetically” as they saw fit (see

Black 2024 and Figures 192 and 193). The poll consisted of four different lists, which differed along the following four key axes:

1. Collation of basic characters (no extended logographs or cartouches), including both freestanding logographs and punctuation. Results suggest that people sort punctuation before logographs, and logographs according to their representation in the Latin script.
2. Collation of cartouches (see section 4d) amongst logographs and punctuation. Results show that as a group, people do not have a definitive way to sort cartouches.
3. Collation of extended logographs vs. non-extended logographs (see section 4f). Results show that people do not treat extended logographs any differently from non-extended logographs.
4. Collation of extended logographs vs. cartouches vs. freestanding logographs. Results from this question were consistent with results from the previous questions.

4. Structure

The sitelen pona writing system is a logographic script written left to right. Each logograph corresponds to a single word in the toki pona language. Logographs may be either monospaced or proportional. Some speakers insist that monospaced logographs are easier to read; some argue the opposite. If possible, fonts should provide for both.

Note on terminology: Speakers of toki pona call their logographic characters "glyphs," likely as a clipping of "hieroglyphs," and this leads to this word appearing in other terms such as "combined glyph," "extended glyph," "name glyph," and so on. This is the common usage of the term and should not be confused with the term as defined by The Unicode Standard. We will attempt to avoid this use of the singular word "glyph" where possible; however, we may not be able to eliminate it completely. In most places where "glyph" appears in this document, it is only intended as a reference to the name given to a concept by the community. (One notable exception is the proper use of the terms *glyph substitution* and *glyph stream* when discussing OpenType features.)

Some logographs share common elements called *radicals*. These are unlike CJK radicals in that they are not well-defined and logographs are not composed of them; instead they were invented post hoc as an observation of common themes across existing logographs. A common radical is a set of three short lines emanating from the top of a logograph, known as *emitters*. The logograph for the word *luka* is another common radical, often called the *hand radical*. Some input methods for sitelen pona are based on radicals, such as the *Wakalito* input method. (See Figure 155.)

emitters radical	*epiku*	*kalama*	*pana*	*sona*	*toki*
hand radical	*luka*	*kepeken*	*moku*	*pali*	*pana*
triangle radical	*jelo*	*laso*	*loje*	*pimeja*	*walo*

A full list of logographs (excluding name glyphs; see section 4g) are attached to this proposal, sorted by usage category according to the Linku usage data. Only characters for words in the *core* and *widespread* usage categories are proposed. Characters for words in the *uncommon*, *rare*, *obscure*, and *not notable* categories are not proposed (See section 3b for details on how we chose which logographs to propose).

Some sitelen pona logographs resemble symbols already encoded in Unicode, such as circles, hearts, arrows, and math symbols. They should not be unified with these existing characters as they are part of a single, coherent script in which they function as letters (that is, they have general category `Lo`), not symbols. Additionally, these logographs have distinct semantics from the existing symbols they resemble (for example, *ijo*, whose logograph is a circle, means "thing," not "circle").

Since many fonts for sitelen pona have already been created using the UCSUR encoding, the following sections will often describe the various features of the sitelen pona script in terms of how existing fonts have implemented them. For the most part, we are proposing implementing these features in Unicode-compliant fonts using the same methods. This will ease the transition to a proper Unicode encoding for both developers and users of sitelen pona fonts. However, there are some glyph variants previously made available using other means but for which we are proposing Standardized Variation Sequences. These will be explicitly noted in the next section.

4a. Variant glyphs

> The criteria described in this section are subject to a rework pending the completion of research which studies toki pona communities directly to determine usage, rather than relying exclusively on what survey respondents self-report. As such, the information presented here is subject to change, and should not be taken as authoritative or reflective of the final result.

Some logographs come in multiple forms beyond those naturally expected from differences in handwriting or font design. Speakers call these *variant glyphs*. These differences vary greatly in significance.

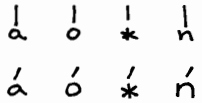
The words *jaki* and *ko* can be represented using any sufficiently chaotic scribble and any sufficiently amorphous blob, respectively. The particular shape of the scribble or blob is not significant. This leads to a nearly infinite amount of glyph variation for these two characters.

We do not propose encoding multiple variants of these characters. However, some fonts do include multiple variants and present them randomly using the OpenType `rand` feature.

jaki in *linja lipamanka* (see also Figure 2b)

ko in *linja lipamanka*

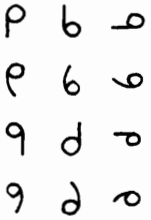
The logographs for *a*, *o*, *kin*, and *n* can have a long vertical upper line, reminiscent of an exclamation mark, or a short diagonal stroke, reminiscent of a Latin letter with an acute accent. The exclamation mark presentation is far more popular. We do not propose encoding variants for these characters. Fonts which include both variants can use the OpenType stylistic sets feature to choose between them.



a, *o*, *kin*, *n* in *sitelen seli kiwen*

a, *o*, *kin*, *n* in *sitelen seli kiwen* with OpenType feature `ss17`

The logographs for the pronouns *mi*, *sina*, and *ona* can have a straight stem or a slightly curved stem. The straight stem appears in many geometric typefaces, while the curved stem appears in most handwritten typefaces and is more popular generally. These logographs have also been seen flipped about their major axis (stem on the right for *mi* and *sina* or on the top for *ona*); although these are widely considered mistakes, they are used intentionally on occasion. These variants do not change the meaning. We do not propose encoding variants for these characters. Fonts which include variants of these characters can use the OpenType stylistic sets feature to choose between them.



mi, *sina*, *ona* in *sitelen seli kiwen*

mi, *sina*, *ona* in *sitelen seli kiwen* with OpenType feature `ss19`

flipped variants of *mi*, *sina*, *ona*

Several words have variant glyphs that range from minor differences in number of lines or dots to entirely different glyphs. These variants were created for several reasons, including: multiple ideas for glyphs for new words, attempts to avoid confusion between similar glyphs, attempts to avoid stereotypes, space limitations in pixel fonts, and personal preference. Speakers of toki pona generally do not consider these variants semantically distinct and will generally accept the use of one form in place of another, although some may not recognize less popular forms. However, for some words, variants cannot always be used interchangeably because one conveys a more specific meaning, two or more are used within the same text, or one or more are highly divisive. (See Figure 190.) Some fonts which include multiple variants currently use the OpenType character variants feature to choose between them.

akesi with six legs and *akesi* with four legs (see Figure 2a)

up arrow *epiku* and emitters *epiku*

	kala without eyes and *kala* with eyes
	three of the many forms of *kiki*; the first is by far the most popular
	toki-derived *kokosila* and star *kokosila*
	three of many forms of *ku*; the first is by far the most popular
	jo-derived *lanpan*, which is most popular; mistake *lanpan*, from when a font developer misremembered the first form; and *pana*-derived *lanpan*
	three of many forms of *linluwi*; the first is the most popular
	sin-derived *majuna*, moon *majuna*, and lotus or sunset *majuna*; moon *majuna* is rare and only appears in two older fonts
	hair or headscarf *meli*, female sign *meli*, and Latin letter *meli*
	vertical *meso* and *ante*-derived *meso*
	broad shoulders *mije*, male sign *mije*, and Latin letter *mije*
	pill *misikeke* and mortar and pestle *misikeke*
	moli; the first is most popular; the second is common in bitmap fonts due to its compactness
	animal face *mu*, used to mean "animal sound," and emitters *mu*, ostensibly used to introduce dialogue
	three lines *mute*, used to indicate "many," and four hands *mute*, used to indicate the quantity 20
	sin-derived *namako* and pepper *namako*
	disconnected *olin*, overlapping *olin*, and emitters *olin*
	Arabic *sewi*, more common but with religious undertones, and anpa-derived *sewi*, used to explicitly avoid religion (see Figure 2c)
	soko with thick stem, *soko* with thin stem, and *soko* with annulus
	three spoke *tonsi*, male and female sign *tonsi*, and Latin letter *tonsi*
	dotted *uta* and dotless *uta*
	open *wile* and closed *wile*

We have considered many possible approaches to supporting these variant glyphs:

1. Encode different forms as separate characters. This does not let the font developer choose which form will be the default, and requires the user to specify a particular form, both of which are undesirable.
2. Encode each word once and different forms using Standardized Variation Sequences. This allows font developers to choose which form will be the default and authors to use a word with or without specifying a particular form. This option provides the most freedom and flexibility for both developers and users.
3. Encode each word once and use OpenType character variants to specify different forms. Some fonts already support variant glyphs using this method. The downside to this is that there is no standard for how to assign features to characters or values to variants, and any ad hoc agreement can potentially conflict with features already

present in existing fonts. Also, the choice of variant is conveyed in formatting and lost upon conversion to plain text, which is undesirable.

We propose option 2: encoding the most popular or divisive variants (as determined by the criteria described in Section 3d) using Standardized Variation Sequences. We are not proposing separate characters (option 1) because it is often desirable to use the character without specifying a particular variant.

> In this preliminary proposal, we have chosen to not include any variants pending the completion of research which studies toki pona communities directly to determine usage. We expect there to be variants in the final result, which will be included here.

The word *ni* ("this" or "that") is represented by a logograph resembling a downwards arrow. Some speakers, however, have started using variants pointing in two, four, or even eight different directions. These variants are used to point towards the word, clause, or sentence to which a particular instance of *ni* refers. In this sense, the variance is semantically significant. Some fonts support these using ZWJ sequences.

↓ →	*ni* in *linja sike*
↓ → ↑ ←	*ni* in an old version of *sitelen seli kiwen*
↓ → ↑ ← ↙ ↘ ↗ ↖	*ni* in *linja lipamanka*

As with other variant glyphs, we have considered many possible approaches to directional *ni*:

1. Encode directional *ni* as eight separate characters. While the character properties are different, this could be seen as a duplicate encoding of the arrows U+2190—U+2193 and U+2196—U+2199. This method prevents the use of non-directional *ni* without encoding one of the directions twice and interferes with the search and matching algorithms. Although some of the proposal authors prefer this option, others worry that it would not be accepted.
2. Encode non-directional *ni* as a single character and directional *ni* using Standardized Variation Sequences. If the font does not support directional *ni*, non-directional *ni* will be shown. This is undesirable as it will result in the loss of semantic information if the font does not support it.
3. Encode non-directional *ni* as a single character and directional *ni* using ZWJ sequences, combining *ni* with one of the arrows U+2190—U+2193 or U+2196—U+2199. Some fonts already support directional *ni* using this method. If a font does not support directional *ni*, the arrows provide a fallback mechanism for indicating the direction. Fonts may support additional forms of *ni* without needing to propose a Standardized Variation Sequence or a separate code point.
4. Encode only non-directional *ni* and use OpenType character variants for directional *ni*. This is a bad idea as it uses formatting to convey information.
5. Unify directional *ni* with the arrows U+2190—U+2193 and U+2196—U+2199. This is the least desirable solution as the arrows are not letters and do not have the required character properties for sitelen pona logographs. Additionally and unfortunately, some vendors appear to be rendering these arrows as emoji by default, in flagrant violation of The Unicode Standard.

Due to its many advantages and few disadvantages, we propose option 3: encode non-directional *ni* as a single character and directional *ni* using ZWJ sequences.

Upwards directional *ni* can be confused with up arrow *epiku*. Fonts which include both are recommended to distinguish them, for example by making the arrowhead bigger for *epiku*.

4b. Combined glyphs

The sitelen pona script allows for two logographs to be combined into a single logograph called a *combined glyph*. (See bottom left of Figure 1b.) Usually, the first logograph functions as a noun and the second logograph as an adjective, introducing a semantic component to this form. This can be done with any two logographs, so it is not practical to encode these as separate characters. The use of U+200D ZERO WIDTH JOINER is recommended to represent combined glyphs. Many fonts support these. (See Figure 146.)

pilin	+ ZWJ +	*ike*	=	*pilin ike*	(“bad feeling”)
telo	+ ZWJ +	*lete*	=	*telo lete*	(“cold water”)
kala	+ ZWJ +	*lili*	=	*kala lili*	(“small fish”)
toki	+ ZWJ +	*pona*	=	*toki pona*	(“good language”)

Some combined glyphs can take one of two forms: a *stacked* form, in which the second glyph is placed *above* the first glyph; or a *scaled* form, in which the second glyph is placed *inside* the first glyph. The default form is determined by the font designer. For instances in which more control is desired over the composition of the glyph components, the two format control characters U+1DEF4 SITELEN PONA STACKING JOINER and U+1DEF5 SITELEN PONA SCALING JOINER are proposed. These are present in the UCSUR encoding where they are often supported by fonts and used by speakers. (See Figure 150.)

kala	+ stacking joiner +	*lete*	=	*kala lete*
kala	+ scaling joiner +	*lete*	=	*kala lete*
kala	+ stacking joiner +	*lili*	=	*kala lili*
kala	+ scaling joiner +	*lili*	=	*kala lili*

A very small number of combined glyphs can take one of *three* forms, with the third form being unique to that particular combination. In these cases, the unique form is represented using U+200D ZERO WIDTH JOINER while the stacked and scaled forms are represented using the stacking and scaling joiners, respectively. (See Figure 150.)

luka	+	ZWJ	+	*luka*	=	*luka luka*
luka	+	stacking joiner	+	*luka*	=	*luka luka*
luka	+	scaling joiner	+	*luka*	=	*luka luka*
nimi	+	ZWJ	+	*sin*	=	*nimi sin*
nimi	+	stacking joiner	+	*sin*	=	*nimi sin*
nimi	+	scaling joiner	+	*sin*	=	*nimi sin*

A handful of rare words created as clipped compounds of existing toki pona words can be represented by combining the constituent corresponding characters with zero-width joiners. A few fonts support these. (See Figure 152.) (We neither recommend nor discourage the use of these sequences and leave the decision to the font developer.)

kalama	+ ZWJ +	*musi*	+ ZWJ +	*tan*	+ ZWJ +	*uta*	=	*kuntu*
kama	+ ZWJ +	*pi*	+ ZWJ +	*lawa*	+ ZWJ +	*ala*	=	*kamalawala*
pana	+ ZWJ +	*e*	+ ZWJ +	*telo*	+ ZWJ +	*oko*	=	*peto*
meli	+ ZWJ +	*pi*	+ ZWJ +	*olin*	+ ZWJ +	*meli*	=	*melome*
mije	+ ZWJ +	*pi*	+ ZWJ +	*olin*	+ ZWJ +	*mije*	=	*mijomi*
o	+ ZWJ +	*moli*	+ ZWJ +	*e*	+ ZWJ +	*mi*	=	*omen*

In extreme cases, speakers may recursively combine glyphs to produce longer phrases. We are not proposing this due to the complexity of implementing it.

Additionally, a couple of fonts even allow combining sitelen pona characters with emoji or other existing symbols to represent *name glyphs* (see section 4g). A few examples are shown below. **We are not proposing name glyphs or this method of representing name glyphs** and we do not expect The Unicode Standard to go out of its way to support name glyphs or this method of representing name glyphs, especially if it would require procedures such as changing properties of characters outside of the sitelen pona script. We are including this information in this document only for the sake of completeness.

kijetesantakalu	+ ZWJ +	✡	U+2721 STAR OF DAVID	=	*lipamanka*
meli	+ ZWJ +	🐱	U+1F431 CAT FACE	=	*melinjakulekule*
waso	+ ZWJ +	🎀	U+1F380 RIBBON	=	*Keli*

4c. Numbers

The toki pona language contains only three words used specifically to describe quantity: *wan* (“one”), *tu* (“two”), and *mute* (“many”). To express specific quantities larger than two, simply repeat *tu* as many times as necessary: *wan*, *tu*, *tu wan*, *tu tu*, *tu tu wan*, *tu tu tu*, and so on.

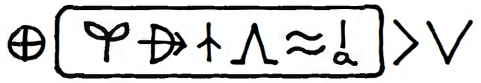
A system for larger quantities described in *Toki Pona: The Language of Good* uses *ala* for zero, *wan* for one, *tu* for two, *luka* for five, *mute* for twenty, and *ale* for one hundred. Some fonts support combining logographs for these words to represent logographs for numbers up to 20. (See Figure 151.) The character properties we are proposing assign numeric values according to this system.

The order of words for numbers is primarily big-endian (*ale mute luka tu wan* = 100 + 20 + 5 + 2 + 1 = 128). An extension of the number system called *nasin nanpa pona* uses little-endian order to indicate multiplication (*luka ale mute tu tu* = 5 × 100 + 20 + 2 + 2 = 524). This extension is used in *The Wonderful World of Oz (Toki Pona edition)* for the year of copyright (= *mute ale mute tu tu* = 20 × 100 + 20 + 2 + 2 = 2024). (See Figure 6.)

Standard Arabic numerals are also sometimes used in sitelen pona.

4d. Cartouches

Names in sitelen pona are represented by a sequence of logographs inside a rounded rectangular shape called a *cartouche*. (See bottom right of Figure 1b.) Inside a cartouche, each logograph is read as only the first phoneme (consonant or vowel) of the corresponding word in toki pona. There are multiple logographs that can represent the same phoneme; the choice of logograph is determined by the personal preference of the entity being named or, otherwise, the author of the text, perhaps taking into account the meaning of the corresponding word (somewhat similar to choosing kanji used to write names in Japanese, for example).



*ma [**k**asi **a**lasa **n**asin **a**wen **t**elo **a**] li suli*
ma Kanata li suli (“Canada is large”)

A popular extension to the cartouche feature, called *nasin sitelen kalama*, allows the use of an interpunct (middle dot) following a logograph to indicate reading the first mora[4] of the word instead of the first phoneme; *n* interpuncts to indicate reading the first *n* morae (or, if the word starts with a vowel, the first *n*+1 morae); or a colon (two middle dots stacked vertically) to indicate reading the entire word. This can sometimes allow a proper name to be represented using fewer characters.

= *jan [**m**un **o**lin **l**ukin **i**lo]* = *jan Moli*

= *jan [**mo**ku · **li**pu ·]* = *jan Moli*

= *jan [**moli** :]* = *jan Moli*

[4] A mora (plural *morae*) is a unit of speech between a phoneme and a syllable. (It is also called a *haku* in Japanese or a *mātrā* in Sanskrit.) In toki pona, a mora is a vowel, a consonant–vowel pair, or a coda -*n*. For example, the word *linja* is composed of the phonemes *l*, *i*, *n*, *j*, and *a*; the syllables *lin* and *ja*; and the morae *li*, *n*, and *ja*.

	= *[**u**npa]*	= *U*
	= *[**un**pa ·]*	= *Un*
	= *[**unpa** · ·]*	= *Unpa*
	= *[**unpa** :]*	= *Unpa*
	= *[**k**epeken]*	= *K*
	= *[**ke**peken ·]*	= *Ke*
	= *[**kepe**ken · ·]*	= *Kepe*
	= *[**kepeke**n · · ·]*	= *Kepeke*
	= *[**kepeken** · · · ·]*	= *Kepeken*
	= *[**kepeken** :]*	= *Kepeken*

The Wonderful World of Oz (Toki Pona edition) introduces a new convention for cartouches: once a name is already mentioned in full inside a chapter, it may be abbreviated only using the first logograph. (See Figure 6.) This convention was also used by at least one other author before the book's release (See Raiter 2018).

	= *jan [**t**omo **o**lin **w**ile **a**lasa **s**uwi **i**jo]*	= *jan Towasi*
	= *jan [**t**omo]*	= *jan Towasi* (implied)

To represent cartouches, four punctuation characters are proposed. These are present in the UCSUR encoding where they are already supported by fonts and widely used by speakers.

- U+1DEF0 SITELEN PONA START OF CARTOUCHE
- U+1DEF1 SITELEN PONA END OF CARTOUCHE
- U+1DEF2 SITELEN PONA MIDDLE DOT
- U+1DEF3 SITELEN PONA COLON

The middle dot and colon used in cartouches are proposed as separate characters from U+00B7 MIDDLE DOT and U+003A COLON to ensure correct behavior of cartouches (so that fallback to a sitelen pona font for rendering sitelen pona codepoints does not spring forward again, causing a break in a font run and preventing the cartouche from rendering properly) and

because they have different line breaking and East Asian width properties as well as collation order (see sections 6e, 6f, and 6i).

Around two thirds of the fonts currently using the UCSUR encoding have implemented full support for cartouches by using glyph substitution rules to give the cartouche a fully closed appearance. (See Figure 147.) An example of how this can work with the OpenType `calt` (contextual alternates) feature is shown below.

glyph stream before `calt`	jan	SOC	mun	olin	lukin	ilo	EOC
glyph stream after `calt`	jan	SOC	mun.cart	olin.cart	lukin.cart	ilo.cart	EOC
rendered output							

4e. Punctuation

Besides the interpunct and colon used in cartouches, the most common uses of punctuation in sitelen pona text are: a colon (U+003A COLON or U+1DEF3 SITELEN PONA COLON) after the word *ni*, which indicates that *ni* refers to the following sentence or phrase; an interpunct (U+00B7 MIDDLE DOT or U+1DEF2 SITELEN PONA MIDDLE DOT) used as a sentence separator (spaces or line breaks can also be used); and CJK brackets used as quotation marks. Some authors have used other ASCII or CJK punctuation marks, but this is rare. (If CJK punctuation marks are used, it is recommended to use the centered form and not the corner-justified form.)

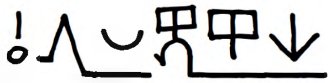
U+0021	!	EXCLAMATION MARK	proscribed; implied by *a* or *o*
U+002C	,	COMMA	occasionally used
U+002E	.	FULL STOP	occasionally used
U+003A	:	COLON	*ni* refers to the following sentence
U+003F	?	QUESTION MARK	proscribed; implied by *seme*
U+00B7	·	MIDDLE DOT	rarely used
U+3001	、	IDEOGRAPHIC COMMA	rarely used
U+3002	。	IDEOGRAPHIC FULL STOP	rarely used
U+300C	「	LEFT CORNER BRACKET	quotation; see also the rare word *te*
U+300D	」	RIGHT CORNER BRACKET	quotation; see also the rare word *to*
U+1DEF2	･	SITELEN PONA MIDDLE DOT	also sometimes used as a full stop
U+1DEF3	：	SITELEN PONA COLON	also sometimes used like U+003A

4f. Extended glyphs

Some sitelen pona logographs can be written as *extended glyphs*, in which part of the logograph is extended horizontally to appear underneath any number of subsequent logographs. This is most often done with *pi*, a grammatical particle which regroups modifiers; a *pi* written this way is called a *long pi*. In the following example, one *long pi* extends underneath *toki pona* (written as a combined glyph) and another extends underneath *pona mute*.

jan pi toki pona li jan pi pona mute

Other logographs can be extended in this way, although this is not done as often as with *pi*. In the following example, a *long awen* extends underneath *pona* and a *long kepeken* extends underneath *ilo ni*.



o awen pona kepeken ilo ni

To support extended glyphs, the two format control characters U+1DEF6 SITELEN PONA START OF EXTENSION and U+1DEF7 SITELEN PONA END OF EXTENSION are proposed. The START OF EXTENSION character is placed after the logograph being extended and before the first logograph above the extension. The END OF EXTENSION character is placed after the last logograph above the extension. The above example would thus be encoded using the following sequence of code points:

- U+1DE44 SITELEN PONA IDEOGRAPH O
- U+1DE08 SITELEN PONA IDEOGRAPH AWEN
- U+1DEF6 SITELEN PONA START OF EXTENSION
    - U+1DE54 SITELEN PONA IDEOGRAPH PONA
- U+1DEF7 SITELEN PONA END OF EXTENSION
- U+1DE19 SITELEN PONA IDEOGRAPH KEPEKEN
- U+1DEF6 SITELEN PONA START OF EXTENSION
    - U+1DE0E SITELEN PONA IDEOGRAPH ILO
    - U+1DE41 SITELEN PONA IDEOGRAPH NI
- U+1DEF7 SITELEN PONA END OF EXTENSION

The same format control characters are present in the UCSUR encoding under the names START OF LONG GLYPH and END OF LONG GLYPH; they are proposed with new names for better clarity and to avoid using the word “glyph” in a Unicode character name. Many fonts have implemented full support for extended glyphs using glyph substitution rules in a manner similar to that used for cartouches. (See Figure 148.) An example of how this can work with the OpenType `calt` (contextual alternates) feature is shown below.

glyph stream before `calt`	jan	pi	SP SOE SOE	pona	mute	SP EOE EOE
glyph stream after `calt`	jan	pi.extd		pona.extn	mute.extn	
rendered output						

Rarely, *reverse-extended glyphs* are also used. Instead of the right side of the logograph being extended to appear under subsequent logographs, the *left* side of the logograph is extended to appear under *preceding* logographs. In the following example, *kama* is extended underneath the preceding *jan* and *mi* is extended underneath the preceding *jan pona*.

jan kama li jan pona mi

It is also possible to extend both sides of a logograph. Here, *ala* is extended underneath both the preceding *pona* and the subsequent *pona*.

sina pilin pona ala pona

To support reverse-extended glyphs, two additional format control characters U+1DEF8 SITELEN PONA START OF REVERSE EXTENSION and U+1DEF9 SITELEN PONA END OF REVERSE EXTENSION are also proposed.

Separate characters are required for a reverse extension in order to resolve the following ambiguity. Given the sequence A (B) C with A, B, and C representing logographs and (and) representing the start and end of extension characters, it is unclear which of the following is the intended result:

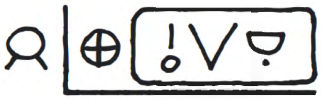
- A is rendered with a forward extension under B
- C is rendered with a reverse extension under B
- Both of the above

The same format control characters are present in the UCSUR encoding under the names START OF REVERSE LONG GLYPH and END OF REVERSE LONG GLYPH; they are proposed with new names for better clarity and to avoid using the word “glyph” in a Unicode character name. Most fonts that have implemented extended glyphs have also implemented reverse-extended glyphs. (See Figure 149.) An example of how this can work with the OpenType `calt` (contextual alternates) feature is shown below.

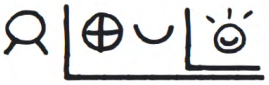
glyph stream before `calt`	SP SRE SRE	jan	pona	SP ERE ERE	mi
glyph stream after `calt`		jan.extn	pona.extn		mi.extr
rendered output					

glyph stream before `calt`	SP SRE SRE	pona	SP ERE ERE	ala	SP SOE SOE	pona	SP EOE EOE
glyph stream after `calt`		pona.extn		ala.extb		pona.extn	
rendered output							

Finally, it is also possible to nest cartouches and extended glyphs inside each other, such as by including a cartouche inside an extended glyph or an extended glyph inside another extended glyph. This is attested in handwritten sitelen pona (see Figures 123, 124, and 125).

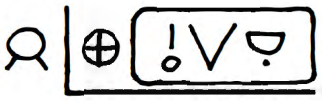


*jan pi ma [**o** suli **u**ta]*



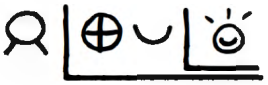
*jan pi ma pona pi **toki pona***

No fonts currently support this due to the complexity of the OpenType features that would be required. However, the UCSUR encoding and the Unicode encoding being proposed are designed to support this, and sitelen pona font developers are currently researching potential implementations. The above examples would be encoded as follows:



jan pi *ma [* ***o*** *suli* ***u****ta]*

- U+1DE11 SITELEN PONA IDEOGRAPH JAN
- U+1DE4D SITELEN PONA IDEOGRAPH PI
- U+1DEF6 SITELEN PONA START OF EXTENSION
    - U+1DE30 SITELEN PONA IDEOGRAPH MA
    - U+1DEF0 SITELEN PONA START OF CARTOUCHE
        - U+1DE44 SITELEN PONA IDEOGRAPH O
        - U+1DE63 SITELEN PONA IDEOGRAPH SULI
        - U+1DE70 SITELEN PONA IDEOGRAPH UTA
    - U+1DEF1 SITELEN PONA END OF CARTOUCHE
- U+1DEF7 SITELEN PONA END OF EXTENSION



jan pi *ma pona pi* ***toki pona***

- U+1DE11 SITELEN PONA IDEOGRAPH JAN
- U+1DE4D SITELEN PONA IDEOGRAPH PI
- U+1DEF6 SITELEN PONA START OF EXTENSION
    - U+1DE30 SITELEN PONA IDEOGRAPH MA
    - U+1DE54 SITELEN PONA IDEOGRAPH PONA
    - U+1DE4D SITELEN PONA IDEOGRAPH PI
    - U+1DEF6 SITELEN PONA START OF EXTENSION
        - U+1DE6C SITELEN PONA IDEOGRAPH TOKI
        - U+200D ZERO WIDTH JOINER
        - U+1DE54 SITELEN PONA IDEOGRAPH PONA
    - U+1DEF7 SITELEN PONA END OF EXTENSION
- U+1DEF7 SITELEN PONA END OF EXTENSION

4g. Personal glyphs and name glyphs

Some members of the toki pona community have created custom logographs which represent specific people, communities, or things. These are called *personal glyphs* when they represent their individual creators or *name glyphs* in the general case. In running text, these can be used in place of a proper name inside a cartouche. Some fonts include these. By definition, name glyphs are an open-ended category with extremely limited usage, two factors which disqualify them from encoding in The Unicode Standard. Therefore, **we are not proposing name glyphs** and are including this information only for completeness. Name glyphs should be interchanged using images or through private agreement.

Eweke

Itan

Jan

lipamanka

melinjakulekule (or *Lepeka*)

Nasulin

nata

Usawi

4h. Legacy characters

There are a handful of characters from the UCSUR encoding that are not being proposed. These were originally included as a hack to enable the representation of cartouches and extended glyphs without the use of OpenType features. However, font developers and users in the toki pona community have quickly learned how to implement and use these features and have fully embraced OpenType technology, rendering the use of these hacks unnecessary. Therefore, **the following characters from the UCSUR encoding are not proposed.**

- U+F1992 SITELEN PONA COMBINING CARTOUCHE EXTENSION
- U+F1993 SITELEN PONA START OF LONG PI
- U+F1994 SITELEN PONA COMBINING LONG PI EXTENSION
- U+F1999 SITELEN PONA COMBINING LONG GLYPH EXTENSION

5. Copyright status

Sonja Lang has released the toki pona language and sitelen pona writing system into the public domain since the publication of *Toki Pona: Die Sprache des Guten*, the German edition of *Toki Pona: The Language of Good*, in 2021. (See Figures 4 and 7.) This is reaffirmed on the copyright page for *The Wonderful Wizard of Oz (Toki Pona edition)*. (See Figure 6.) She has on multiple occasions stated that both the language and writing system belong to the community, and may be used by anyone. (See Figure 7.)

The glyphs shown on the code charts are from Lang's handwriting as seen in *Toki Pona: The Language of Good* and Spencer van der Meulen's handwriting as seen in the Esperanto edition of the same book. (See Figures 1 and 3.) Additional glyphs were provided by Lang for logographs and variants not present in these sources. (See Lang 2024c.) These have all been released into the public domain.

6. Proposed character repertoire

The tentative character repertoire presented here is subject to change pending ongoing research. For the tentative proposal, the repertoire has been reduced to a set which

unambiguously will be included in the final result. It is expected that there will be more characters recommended in the final proposal. For reference, the previous methodology found a total of 153 characters to include, 22 more than are shown here, and the final result will likely be closer to this.

The proposed character repertoire consists of the original 120 *nimi pu*, the word *tonsi*, as well as 4 characters for cartouches (see section 4d) and 6 format control characters for combined glyphs (see section 4b) and extended glyphs (see section 4f), for a total of 131 characters.

6a. Character names

Logographic sitelen pona characters are given the proposed names of SITELEN PONA IDEOGRAPH followed by the toki pona word they represent in Latin script. The use of "ideograph" instead of "logograph" follows precedent set by CJK characters. For other characters, including the 6 format control characters, we have attempted to invent names that are meaningful, unique, and conformant to WG2 and UTC guidelines.

6b. Code point assignment

Proposed characters are shown with suggested code points in a new block (1DE00..1DEFF) that is unassigned according to the "Roadmap to the SMP," revision 15.1.1. The suggested block name "Sitelen Pona" is listed in the summary form. However, it is understood that final assignment of blocks, code points, and block and character names is completely at the discretion of UTC and/or WG2.

6c. Unicode character properties

```
1DE00;SITELEN PONA IDEOGRAPH A;Lo;0;L;;;;;Y;;;;;
1DE01;SITELEN PONA IDEOGRAPH AKESI;Lo;0;L;;;;;Y;;;;;
1DE02;SITELEN PONA IDEOGRAPH ALA;Lo;0;L;;;;0;Y;;;;;          # numeric value 0
1DE03;SITELEN PONA IDEOGRAPH ALASA;Lo;0;L;;;;;Y;;;;;
1DE04;SITELEN PONA IDEOGRAPH ALE;Lo;0;L;;;;100;Y;;;;;        # numeric value 100
1DE05;SITELEN PONA IDEOGRAPH ANPA;Lo;0;L;;;;;Y;;;;;
1DE06;SITELEN PONA IDEOGRAPH ANTE;Lo;0;L;;;;;Y;;;;;
1DE07;SITELEN PONA IDEOGRAPH ANU;Lo;0;L;;;;;Y;;;;;
1DE08;SITELEN PONA IDEOGRAPH AWEN;Lo;0;L;;;;;Y;;;;;
1DE09;SITELEN PONA IDEOGRAPH E;Lo;0;L;;;;;Y;;;;;
1DE0A;SITELEN PONA IDEOGRAPH EN;Lo;0;L;;;;;Y;;;;;
1DE0B;SITELEN PONA IDEOGRAPH ESUN;Lo;0;L;;;;;Y;;;;;
1DE0C;SITELEN PONA IDEOGRAPH IJO;Lo;0;L;;;;;Y;;;;;
1DE0D;SITELEN PONA IDEOGRAPH IKE;Lo;0;L;;;;;Y;;;;;
1DE0E;SITELEN PONA IDEOGRAPH ILO;Lo;0;L;;;;;Y;;;;;
1DE0F;SITELEN PONA IDEOGRAPH INSA;Lo;0;L;;;;;Y;;;;;
1DE10;SITELEN PONA IDEOGRAPH JAKI;Lo;0;L;;;;;Y;;;;;
1DE11;SITELEN PONA IDEOGRAPH JAN;Lo;0;L;;;;;Y;;;;;
1DE12;SITELEN PONA IDEOGRAPH JELO;Lo;0;L;;;;;Y;;;;;
1DE13;SITELEN PONA IDEOGRAPH JO;Lo;0;L;;;;;Y;;;;;
1DE14;SITELEN PONA IDEOGRAPH KALA;Lo;0;L;;;;;Y;;;;;
1DE15;SITELEN PONA IDEOGRAPH KALAMA;Lo;0;L;;;;;Y;;;;;
1DE16;SITELEN PONA IDEOGRAPH KAMA;Lo;0;L;;;;;Y;;;;;
1DE17;SITELEN PONA IDEOGRAPH KASI;Lo;0;L;;;;;Y;;;;;
1DE18;SITELEN PONA IDEOGRAPH KEN;Lo;0;L;;;;;Y;;;;;
1DE19;SITELEN PONA IDEOGRAPH KEPEKEN;Lo;0;L;;;;;Y;;;;;
1DE1A;SITELEN PONA IDEOGRAPH KILI;Lo;0;L;;;;;Y;;;;;
```

```
1DE1B;SITELEN PONA IDEOGRAPH KIWEN;Lo;0;L;;;;;Y;;;;;
1DE1C;SITELEN PONA IDEOGRAPH KO;Lo;0;L;;;;;Y;;;;;
1DE1D;SITELEN PONA IDEOGRAPH KON;Lo;0;L;;;;;Y;;;;;
1DE1E;SITELEN PONA IDEOGRAPH KULE;Lo;0;L;;;;;Y;;;;;
1DE1F;SITELEN PONA IDEOGRAPH KULUPU;Lo;0;L;;;;;Y;;;;;
1DE20;SITELEN PONA IDEOGRAPH KUTE;Lo;0;L;;;;;Y;;;;;
1DE21;SITELEN PONA IDEOGRAPH LA;Lo;0;L;;;;;Y;;;;;
1DE22;SITELEN PONA IDEOGRAPH LAPE;Lo;0;L;;;;;Y;;;;;
1DE23;SITELEN PONA IDEOGRAPH LASO;Lo;0;L;;;;;Y;;;;;
1DE24;SITELEN PONA IDEOGRAPH LAWA;Lo;0;L;;;;;Y;;;;;
1DE25;SITELEN PONA IDEOGRAPH LEN;Lo;0;L;;;;;Y;;;;;
1DE26;SITELEN PONA IDEOGRAPH LETE;Lo;0;L;;;;;Y;;;;;
1DE27;SITELEN PONA IDEOGRAPH LI;Lo;0;L;;;;;Y;;;;;
1DE28;SITELEN PONA IDEOGRAPH LILI;Lo;0;L;;;;;Y;;;;;
1DE29;SITELEN PONA IDEOGRAPH LINJA;Lo;0;L;;;;;Y;;;;;
1DE2A;SITELEN PONA IDEOGRAPH LIPU;Lo;0;L;;;;;Y;;;;;
1DE2B;SITELEN PONA IDEOGRAPH LOJE;Lo;0;L;;;;;Y;;;;;
1DE2C;SITELEN PONA IDEOGRAPH LON;Lo;0;L;;;;;Y;;;;;
1DE2D;SITELEN PONA IDEOGRAPH LUKA;Lo;0;L;;;;5;Y;;;;;          # numeric value 5
1DE2E;SITELEN PONA IDEOGRAPH LUKIN;Lo;0;L;;;;;Y;;;;;
1DE2F;SITELEN PONA IDEOGRAPH LUPA;Lo;0;L;;;;;Y;;;;;
1DE30;SITELEN PONA IDEOGRAPH MA;Lo;0;L;;;;;Y;;;;;
1DE31;SITELEN PONA IDEOGRAPH MAMA;Lo;0;L;;;;;Y;;;;;
1DE32;SITELEN PONA IDEOGRAPH MANI;Lo;0;L;;;;;Y;;;;;
1DE33;SITELEN PONA IDEOGRAPH MELI;Lo;0;L;;;;;Y;;;;;
1DE34;SITELEN PONA IDEOGRAPH MI;Lo;0;L;;;;;Y;;;;;
1DE35;SITELEN PONA IDEOGRAPH MIJE;Lo;0;L;;;;;Y;;;;;
1DE36;SITELEN PONA IDEOGRAPH MOKU;Lo;0;L;;;;;Y;;;;;
1DE37;SITELEN PONA IDEOGRAPH MOLI;Lo;0;L;;;;;Y;;;;;
1DE38;SITELEN PONA IDEOGRAPH MONSI;Lo;0;L;;;;;Y;;;;;
1DE39;SITELEN PONA IDEOGRAPH MU;Lo;0;L;;;;;Y;;;;;
1DE3A;SITELEN PONA IDEOGRAPH MUN;Lo;0;L;;;;;Y;;;;;
1DE3B;SITELEN PONA IDEOGRAPH MUSI;Lo;0;L;;;;;Y;;;;;
1DE3C;SITELEN PONA IDEOGRAPH MUTE;Lo;0;L;;;;20;Y;;;;;         # numeric value 20
1DE3D;SITELEN PONA IDEOGRAPH NANPA;Lo;0;L;;;;;Y;;;;;
1DE3E;SITELEN PONA IDEOGRAPH NASA;Lo;0;L;;;;;Y;;;;;
1DE3F;SITELEN PONA IDEOGRAPH NASIN;Lo;0;L;;;;;Y;;;;;
1DE40;SITELEN PONA IDEOGRAPH NENA;Lo;0;L;;;;;Y;;;;;
1DE41;SITELEN PONA IDEOGRAPH NI;Lo;0;L;;;;;Y;;;;;
1DE42;SITELEN PONA IDEOGRAPH NIMI;Lo;0;L;;;;;Y;;;;;
1DE43;SITELEN PONA IDEOGRAPH NOKA;Lo;0;L;;;;;Y;;;;;
1DE44;SITELEN PONA IDEOGRAPH O;Lo;0;L;;;;;Y;;;;;
1DE45;SITELEN PONA IDEOGRAPH OLIN;Lo;0;L;;;;;Y;;;;;
1DE46;SITELEN PONA IDEOGRAPH ONA;Lo;0;L;;;;;Y;;;;;
1DE47;SITELEN PONA IDEOGRAPH OPEN;Lo;0;L;;;;;Y;;;;;
1DE48;SITELEN PONA IDEOGRAPH PAKALA;Lo;0;L;;;;;Y;;;;;
1DE49;SITELEN PONA IDEOGRAPH PALI;Lo;0;L;;;;;Y;;;;;
1DE4A;SITELEN PONA IDEOGRAPH PALISA;Lo;0;L;;;;;Y;;;;;
1DE4B;SITELEN PONA IDEOGRAPH PAN;Lo;0;L;;;;;Y;;;;;
1DE4C;SITELEN PONA IDEOGRAPH PANA;Lo;0;L;;;;;Y;;;;;
1DE4D;SITELEN PONA IDEOGRAPH PI;Lo;0;L;;;;;Y;;;;;
1DE4E;SITELEN PONA IDEOGRAPH PILIN;Lo;0;L;;;;;Y;;;;;
1DE4F;SITELEN PONA IDEOGRAPH PIMEJA;Lo;0;L;;;;;Y;;;;;
1DE50;SITELEN PONA IDEOGRAPH PINI;Lo;0;L;;;;;Y;;;;;
1DE51;SITELEN PONA IDEOGRAPH PIPI;Lo;0;L;;;;;Y;;;;;
1DE52;SITELEN PONA IDEOGRAPH POKA;Lo;0;L;;;;;Y;;;;;
1DE53;SITELEN PONA IDEOGRAPH POKI;Lo;0;L;;;;;Y;;;;;
1DE54;SITELEN PONA IDEOGRAPH PONA;Lo;0;L;;;;;Y;;;;;
1DE55;SITELEN PONA IDEOGRAPH PU;Lo;0;L;;;;;Y;;;;;
1DE56;SITELEN PONA IDEOGRAPH SAMA;Lo;0;L;;;;;Y;;;;;
1DE57;SITELEN PONA IDEOGRAPH SELI;Lo;0;L;;;;;Y;;;;;
1DE58;SITELEN PONA IDEOGRAPH SELO;Lo;0;L;;;;;Y;;;;;
1DE59;SITELEN PONA IDEOGRAPH SEME;Lo;0;L;;;;;Y;;;;;
1DE5A;SITELEN PONA IDEOGRAPH SEWI;Lo;0;L;;;;;Y;;;;;
1DE5B;SITELEN PONA IDEOGRAPH SIJELO;Lo;0;L;;;;;Y;;;;;
1DE5C;SITELEN PONA IDEOGRAPH SIKE;Lo;0;L;;;;;Y;;;;;
1DE5D;SITELEN PONA IDEOGRAPH SIN;Lo;0;L;;;;;Y;;;;;
1DE5E;SITELEN PONA IDEOGRAPH SINA;Lo;0;L;;;;;Y;;;;;
```

```
1DE5F;SITELEN PONA IDEOGRAPH SINPIN;Lo;0;L;;;;;Y;;;;;
1DE60;SITELEN PONA IDEOGRAPH SITELEN;Lo;0;L;;;;;Y;;;;;
1DE61;SITELEN PONA IDEOGRAPH SONA;Lo;0;L;;;;;Y;;;;;
1DE62;SITELEN PONA IDEOGRAPH SOWELI;Lo;0;L;;;;;Y;;;;;
1DE63;SITELEN PONA IDEOGRAPH SULI;Lo;0;L;;;;;Y;;;;;
1DE64;SITELEN PONA IDEOGRAPH SUNO;Lo;0;L;;;;;Y;;;;;
1DE65;SITELEN PONA IDEOGRAPH SUPA;Lo;0;L;;;;;Y;;;;;
1DE66;SITELEN PONA IDEOGRAPH SUWI;Lo;0;L;;;;;Y;;;;;
1DE67;SITELEN PONA IDEOGRAPH TAN;Lo;0;L;;;;;Y;;;;;
1DE68;SITELEN PONA IDEOGRAPH TASO;Lo;0;L;;;;;Y;;;;;
1DE69;SITELEN PONA IDEOGRAPH TAWA;Lo;0;L;;;;;Y;;;;;
1DE6A;SITELEN PONA IDEOGRAPH TELO;Lo;0;L;;;;;Y;;;;;
1DE6B;SITELEN PONA IDEOGRAPH TENPO;Lo;0;L;;;;;Y;;;;;
1DE6C;SITELEN PONA IDEOGRAPH TOKI;Lo;0;L;;;;;Y;;;;;
1DE6D;SITELEN PONA IDEOGRAPH TOMO;Lo;0;L;;;;;Y;;;;;
1DE6E;SITELEN PONA IDEOGRAPH TU;Lo;0;L;;;;2;Y;;;;;          # numeric value 2
1DE6F;SITELEN PONA IDEOGRAPH UNPA;Lo;0;L;;;;;Y;;;;;
1DE70;SITELEN PONA IDEOGRAPH UTA;Lo;0;L;;;;;Y;;;;;
1DE71;SITELEN PONA IDEOGRAPH UTALA;Lo;0;L;;;;;Y;;;;;
1DE72;SITELEN PONA IDEOGRAPH WALO;Lo;0;L;;;;;Y;;;;;
1DE73;SITELEN PONA IDEOGRAPH WAN;Lo;0;L;;;;1;Y;;;;;          # numeric value 1
1DE74;SITELEN PONA IDEOGRAPH WASO;Lo;0;L;;;;;Y;;;;;
1DE75;SITELEN PONA IDEOGRAPH WAWA;Lo;0;L;;;;;Y;;;;;
1DE76;SITELEN PONA IDEOGRAPH WEKA;Lo;0;L;;;;;Y;;;;;
1DE77;SITELEN PONA IDEOGRAPH WILE;Lo;0;L;;;;;Y;;;;;
1DE78;SITELEN PONA IDEOGRAPH TONSI;Lo;0;L;;;;;Y;;;;;
1DEF0;SITELEN PONA START OF CARTOUCHE;Ps;0;L;;;;;Y;;;;;      # punctuation start
1DEF1;SITELEN PONA END OF CARTOUCHE;Pe;0;L;;;;;Y;;;;;        # punctuation end
1DEF2;SITELEN PONA MIDDLE DOT;Po;0;L;;;;;Y;;;;;              # punctuation other
1DEF3;SITELEN PONA COLON;Po;0;L;;;;;Y;;;;;                   # punctuation other
1DEF4;SITELEN PONA STACKING JOINER;Cf;0;L;;;;;N;;;;;         # format controls
1DEF5;SITELEN PONA SCALING JOINER;Cf;0;L;;;;;N;;;;;
1DEF6;SITELEN PONA START OF EXTENSION;Cf;0;L;;;;;N;;;;;
1DEF7;SITELEN PONA END OF EXTENSION;Cf;0;L;;;;;N;;;;;
1DEF8;SITELEN PONA START OF REVERSE EXTENSION;Cf;0;L;;;;;N;;;;;
1DEF9;SITELEN PONA END OF REVERSE EXTENSION;Cf;0;L;;;;;N;;;;;
```

6d. Bidi class and bidi mirroring

Because the sitelen pona script is normally written left to right, we have given all characters a bidi class value of `L` (left-to-right). Occasionally, however, it may be written right to left or boustrophedon. Of course, Unicode already provides bidi format controls that can be used to accomplish this, and many higher-level protocols also provide mechanisms for controlling text direction.

Opinions differ on which characters should be mirrored if written right to left. However, many speakers are already comfortable reading text with all characters mirrored from interactions in *VRChat*, a 3D environment wherein sitelen pona text can be handwritten in midair and then read head-on or from behind. (See Figure 59.) For this reason, we have given all printable characters a bidi mirroring value of `Y` (mirrored).

If a font developer wishes to control the mirroring of specific characters, use of the OpenType `rtlm` (right-to-left mirrored form) feature is recommended.

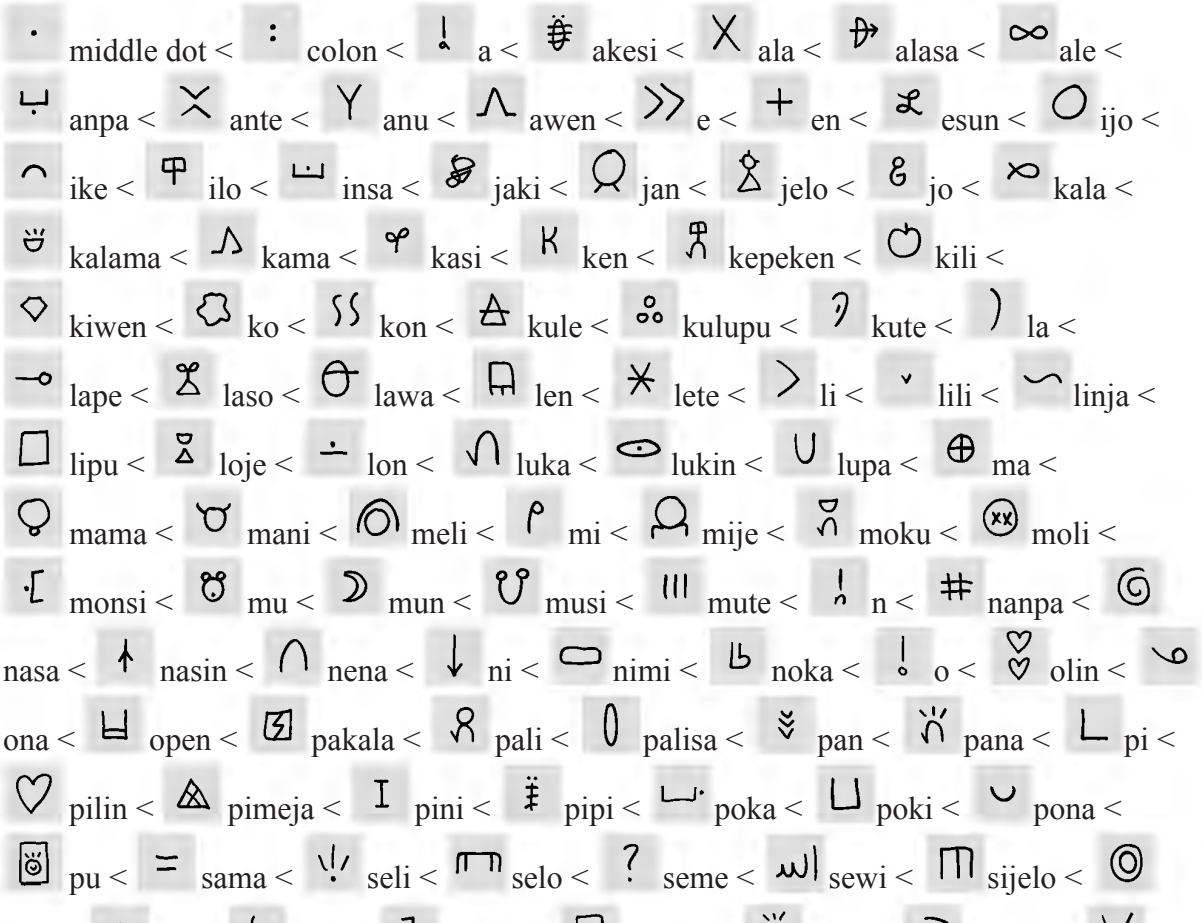
6e. Line breaking

Line breaking may occur between any two logographs. It should not occur immediately after a “start of cartouche” or immediately before an “end of cartouche,” similarly to parentheses, nor should it occur between a logograph and an interpunct or colon, immediately after an extended glyph, or immediately before a reverse-extended glyph. If line breaks are not wanted between the head word and the modifier within a phrase, the use of U+2060 WORD JOINER is recommended.

```
1DE00..1DE78   ; ID # Lo   [121] SITELEN PONA IDEOGRAPH A..SITELEN PONA IDEOGRAPH TONSI
1DEF0          ; OP # Ps         SITELEN PONA START OF CARTOUCHE
1DEF1          ; CL # Pe         SITELEN PONA END OF CARTOUCHE
1DEF2..1DEF3   ; EX # Po     [2] SITELEN PONA MIDDLE DOT..SITELEN PONA COLON
1DEF4..1DEF5   ; GL # Cf     [2] SITELEN PONA STACKING JOINER...SITELEN PONA SCALING JOINER
1DEF6          ; GL # Cf         SITELEN PONA START OF EXTENSION
1DEF7          ; CL # Cf         SITELEN PONA END OF EXTENSION
1DEF8          ; OP # Cf         SITELEN PONA START OF REVERSE EXTENSION
1DEF9          ; GL # Cf         SITELEN PONA END OF REVERSE EXTENSION
```

6f. Collation order

The collation order for sitelen pona logographs should be alphabetical by their equivalent representation in the Latin script. Interpuncts and colons, in that order, should precede logographs. Cartouches and format control characters should not affect collation order.



suno < supa < suwi < tan < taso < tawa < telo < tenpo <
toki < tomo < tonsi < tu < unpa < uta < utala < walo < wan < waso < wawa < weka < wile

6g. Standardized Variation sequences

In this preliminary proposal, we have chosen to not include any variants pending the completion of research which studies toki pona communities directly to determine usage. We expect there to be variants in the final result, which will be included here. See section 4a for further details.

6h. Vertical orientation

An example of vertical sitelen pona can be found on the book cover for *The Wonderful Wizard of Oz (Toki Pona edition)*. Logographs are kept upright while the "start of cartouche" and "end of cartouche" characters are rotated. We are proposing vertical orientation property values which reflect this.

```
1DE00..1DE78   ; U  # Lo   [121] SITELEN PONA IDEOGRAPH A..SITELEN PONA IDEOGRAPH TONSI
1DEF0          ; R  # Ps         SITELEN PONA START OF CARTOUCHE
1DEF1          ; R  # Pe         SITELEN PONA END OF CARTOUCHE
1DEF2..1DEF3   ; R  # Po     [2] SITELEN PONA MIDDLE DOT..SITELEN PONA COLON
1DEF4..1DEF5   ; R  # Cf     [2] SITELEN PONA STACKING JOINER...SITELEN PONA SCALING JOINER
1DEF6          ; R  # Cf         SITELEN PONA START OF EXTENSION
1DEF7          ; R  # Cf         SITELEN PONA END OF EXTENSION
1DEF8          ; R  # Cf         SITELEN PONA START OF REVERSE EXTENSION
1DEF9          ; R  # Cf         SITELEN PONA END OF REVERSE EXTENSION
```

6i. East Asian width

Logographs for sitelen pona are expected to be wide in a monospaced environment. A hack by Florian Cauvin (jan Polijan) called *tty_pona* patches the `wcwidth` function in the C standard library to return 2 for sitelen pona logographs in the UCSUR encoding. We propose assigning an East Asian width value of `W` (wide) to sitelen pona logographs to obviate the need for such hacks.

```
1DE00..1DE78   ; W  # Lo   [121] SITELEN PONA IDEOGRAPH A..SITELEN PONA IDEOGRAPH TONSI
1DEF0          ; W  # Ps         SITELEN PONA START OF CARTOUCHE
1DEF1          ; W  # Pe         SITELEN PONA END OF CARTOUCHE
1DEF2..1DEF3   ; W  # Po     [2] SITELEN PONA MIDDLE DOT..SITELEN PONA COLON
1DEF4..1DEF5   ; N  # Cf     [2] SITELEN PONA STACKING JOINER...SITELEN PONA SCALING JOINER
1DEF6          ; N  # Cf         SITELEN PONA START OF EXTENSION
1DEF7          ; N  # Cf         SITELEN PONA END OF EXTENSION
1DEF8          ; N  # Cf         SITELEN PONA START OF REVERSE EXTENSION
1DEF9          ; N  # Cf         SITELEN PONA END OF REVERSE EXTENSION
```

6j. ISO 15924 script code

We propose the script code `Stpn` for sitelen pona. However, it is understood that this is completely at the discretion of the Joint Advisory Committee (JAC) responsible for ISO 15924.

7. Acknowledgements

The authors wish to thank the following individuals for their work preparing materials for this proposal: Erek Butcher (soweli Eweke), kije Enki, Kelsey Higham (waso Keli), pipi Kewapi, jan Kita, jan Koko, Sonja Lang (jan Sonja; creator of the language), manœ, kulupu Menasewi, jan Toluneko, nanpa Tonpite, jan Sala, Nia W. (kulupu Nijako), one anonymous contributor, and many others in the *ma pi nasin sitelen*, *ma pona pi toki pona*, and *Ma Toki Pona VR* communities.

Additionally, we wish to thank everyone who has responded to the various surveys used to gather data for this proposal and all the creators of the media presented in this proposal.

8. References

Amraoui, Chlag. 2023. Untitled. Oil on canvas. 151 × 151 cm.

Anderson, Deborah; Whistler, Ken; Pournader, Roozbeh; Liang, Hai. 2021. "Recommendations to UTC #168 July 2021 on Script Proposals." UTC document L2/21-130. https://www.unicode.org/L2/L2021/21130-script-adhoc-rept.pdf

Baum, L. Frank; trans. Lang, Sonja. 2024. *The Wonderful Wizard of Oz (Toki Pona edition)*. https://www.amazon.com/dp/0978292375

Bettencourt, Rebecca. 2022. "ilo Nutan." Reddit. https://www.reddit.com/r/tokipona/comments/vxrda3/ilo_nutan/

———. 2023a. "Sitelen Pona Keyboard Layout." https://www.kreativekorp.com/software/keyboards/sitelenpona/

———. 2023b. "Sitelen Pona Keycaps." https://github.com/RebeccaRGB/sitelen-pona-keycaps

Billsmugs. 2022a. *jaki pona*. https://toki.pona.billsmugs.com/lipu-tenpo/2022-06-23-jaki-pona/

———. 2022b. *Late For School*. https://billsmugs.itch.io/late-for-school

———. 2023. *Kiwen Kon*. https://billsmugs.itch.io/kiwen-kon

Black, Janet. 2024. "nasin pi sitelen pona." Qualtrics. https://uky.az1.qualtrics.com/jfe/form/SV_88KWGjdP8gtHBZ4

Bradley, William S. 2024. [Photograph of a DVD case]. Google Drive. https://drive.google.com/file/d/1UIIDdBq58sOVHVjNk9NFRJ_tlUbWEkMC/view

Caramell; trans. jan Mowima [Morphimus]. 2021. "kalameletansen" [Caramelldansen]. YouTube. https://www.youtube.com/watch?v=uf4psPeNDqg

Cauvin, Florian (jan Polijan). ———. 2021. "o pana e moku tawa waso 🐦 [feat. jan Miku]." YouTube. https://www.youtube.com/watch?v=ySkF1lTIEzk

———. 2022. "Mun! ✨ Toki Pona Sci-Fi Anime (short) 🚀." YouTube. https://www.youtube.com/watch?v=aPLPLUa87uo

———. 2023. "polijan/tty_pona: HOW-TO guide and tools to use sitelen pona in terminals." GitHub. https://github.com/polijan/tty_pona

Chervil [@doginacafe]; trans. waso Atu [@wasoAtu]. 2023. “akesi Tankan en soweli Esi tan @doginacafe.” Twitter. https://twitter.com/wasoAtu/status/1594848384661217282
Corgatelli, Ethan (jan Itan). 2022. “nasin-nanpa.” GitHub. https://github.com/ETBCOR/nasin-nanpa
———. 2024. [Photo of a wrist tattoo written *ale li pona* in sitelen pona]. Discord. https://web.archive.org/web/20240227165326/https://cdn.discordapp.com/attachments/1191986138185990154/1211723349831716864/20240226_101450.jpg?ex=65ef3c4f&is=65dcc74f&hm=13d1df97d4434a7002f4b10443373b6411c0013796fedb3a50dba3b7384e885c&
Danielson III, Gregory (jan Kekan San). 2022a. “Modifiers and pi - Learn Toki Pona! Lesson 4.” YouTube. https://www.youtube.com/watch?v=J6ThX6dQyCI
———. 2022b. “Objects with e.” mun.la. https://mun.la/sona/e.html
———. 2023. “ale li nanpa.” YouTube. https://www.youtube.com/watch?v=AgkDx8dIAio
———. 2024a. “Media for Toki Pona Unicode Application.” Google Forms. https://forms.gle/tPoggrqNYpRToord9
———. 2024b. “Media for Toki Pona Unicode Application (Responses).” Google Sheets. https://docs.google.com/spreadsheets/d/197WZsi5PCWxD2K4PsQ_aJZwzAto0ugB0GKCUb5ihNfg/edit?usp=sharing
———. 2024c. “nasin Juniko | toki lili pi mun #15.” YouTube. https://www.youtube.com/watch?v=W3I_x_rpwc8
———. 2024d. “What sitelen pona glyphs do you use?” Qualtrics. https://uab.co1.qualtrics.com/jfe/form/SV_3C8Q1nZaEm30xi6
———. n.d. “Learn Toki Pona!” YouTube. https://www.youtube.com/playlist?list=PLoLQoll2U6WEeCO7C2oc1LTWgkoBC8rvT
Dawson, Scott Robert; trans. Lyman, Abigail (jan Apikela). 2022. *poki lete lili pi wan taso* [The Lonely Little Fridge]. ISBN 978-1-77-818043-9. https://www.barnesandnoble.com/w/poki-lete-lili-pi-wan-taso-scott-robert-dawson/1141864414
dec32. 2024. “Ajemi”. GitHub. https://github.com/dec32/Ajemi
Discord. “kama sona: Server Insights.” Discord Developer Portal. https://discord.com/developers/servers/969386329513295872/analytics/
———. “ma pona pi toki pona: Server Insights.” Discord Developer Portal. https://discord.com/developers/servers/301377942062366741/analytics/
Duostories. “Unofficial Toki Pona Duolingo Stories.” https://duostories.org/tok2-en
Flear, James (jan Same). 2019. “o kama sona e sitelen pona - Introduction.” YouTube. https://www.youtube.com/watch?v=Uze2LKVN-2E
———. 2020. “jan Mulan - mije wawa.” YouTube. https://www.youtube.com/watch?v=00_8_ZQVJh8
———. 2021. “Back to the future - kama sin tawa tenpo kama.” YouTube. https://www.youtube.com/watch?v=l4twpUOi0p4
———. n.d-a. “Calligraphy.” musi lili. https://musilili.net/calligraphy/
———. n.d-b. “linja pona.” musi lili. https://musilili.net/linja-pona/
———. n.d-c. “o kama sona e sitelen pona.” YouTube. https://www.youtube.com/playlist?list=PLJjeLfR6etE69KRFWQB1guBjfWIw8NN75
Fishtrouts [@fishtrouts]; trans. Butcher, Erek (soweli Eweke). 2023. “How to Be a Dragon: Extra #2.” https://web.archive.org/web/20230930025852/https://cdn.discordapp.com/attachments/340307145373253642/1076272964015759421/image.png
formicant. 2021. “soweli li mi.” https://formicant.github.io/soweli-li-mi/

Goszczycki, Maciej. 2022. “Wate.” https://wate.maciej.ie/?locale=tok-x-sp
Gotik, Rèka (jan Leka). 2020. ken pi jan ale. https://rekagotik.neocities.org/ken-pi-jan-ale
Grimm, Jacob; Grimm, Wilheim; trans. Flear, James (jan Same). 2019. “meli lili pi len loje.” musi lili. https://musilili.net/lipu/meli-lili-pi-len-loje-2/
H La Vallee. n.d. Discord. https://web.archive.org/web/20240106171300/https://cdn.discordapp.com/attachments/739692233640509504/955180667547820112/20200813_171611.jpg
Halley, Mitch (jan Misali). 2015. “12 Days of sona pi toki pona.” YouTube. https://www.youtube.com/playlist?list=PLuYLhuXt4HrQIv3xnDxZqRaLfmxB2U5rJ
Halley, Mitch (jan Misali); feat. Jules (ilo Sule). 2021. “kalama musi pi jan Peni (Penny's Song in toki pona).” YouTube. https://www.youtube.com/watch?v=e8vNliigiss
Haraguchi, Sasuke (原口沙輔); trans. jan Sikun (西困). 2023. “【人工语言·道本语】人マニア·人狂热者”. Bilibili. https://www.bilibili.com/video/BV1Hj411p7Fb/
Haywood, David; Adamsom, Peter; trans. jan Popo. 2023. *ken len pi soweli telo Alapete* [The Hidden Talent of Albert Otter]. Internet Archive. https://archive.org/details/albert-otter-reading-order (Original published in 2010)
Higham, Kelsey (waso Keli). 2022. “ante toki pi musi Seta | toki lili pi mun #2.” https://www.youtube.com/watch?v=LUzSWW40WC8
Higham, Kelsey (waso Keli); jan suli. 2024. [Screenshot of *The Legend of Zelda: Link’s Awakening*]. Discord. https://web.archive.org/web/20240109140934/https://cdn.discordapp.com/attachments/1191986138185990154/1193942731156902029/Sprite-0002.png
Ichikawa, Haruko; trans. jan Kukisulasu [u/guckyslush]. 2023. “kulupu kiwen · Houseki no Kuni · Land of the Lustrous | translation of Chapter 1 in comments.” Reddit. https://www.reddit.com/r/tokipona/comments/160jxyg/kulupu_kiwen_houseki_no_kuni_land_of_the_lustrous/
ijo Me. 2023. “toki insa mi.” https://thoughtsofmine.ca/tok/
ijo Vivi. 2024. “lipu open tan ijo Vivi” [Postcard]. lipu tawa pi sitelen pona. https://liputawa.wordpress.com/2024/02/17/lipu-open-tan-ijo-vivi/
ilo Eko. 2020. “sitelen pona etymology rough draft i guess.” Google Docs. https://docs.google.com/document/d/1bzt85HaYTPk8mWfn5tcQiSezvlOVuqnst1tVTcugju4/edit
ilo Sasi. 2021. “File:Icon for suno pi toki pona.svg.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Icon_for_suno_pi_toki_pona.svg
jan Aleko; jan Sepulon. 2023. [Photograph of a vinyl disc prop]. Discord. https://web.archive.org/web/20240211155136/https://cdn.discordapp.com/attachments/704005132958040126/1203443926359605248/20230802_191245.png?ex=65da57fe&is=65c7e2fe&hm=086aeab2f7afdfcd1638a24ea9feb69056c12be93c630eba435288fa956be3c5&
jan Ana. 2021. “pilin mi tawa ilo nanpa.” *lipu kule*. https://lipukule.org/post/2021/02/13/pilin-mi-tawa-ilo-nanpa/
jan Inkepa. 2020. “ilo sitelen.” https://web.archive.org/web/20200809091227/https://toki.increpare.com/ilo-sitelen/
jan Jami. 2023. “o esun e moku Wapa! (whopper commercial in toki pona)”. YouTube. https://www.youtube.com/watch?v=RHGQShNglBQ
jan Jontan. 2022. “sitelenponale.” https://sitelenponale.ikaluk.net/
jan Kepijona; Samys, Ret (jan Ke Tami). 2024. “Suwi suwi suwi (sitelen pona).pdf.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Suwi_suwi_suwi_(sitelen_pona).pdf

jan Kita. 2022. “akesi li wile lon nena.” https://utala.pona.la/lipu-lili/akesi-li-wile-lon-nena.html
———. 2022. “File:Ma Seka icon.svg.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Ma_Seka_icon.svg
jan Kisen [@fennaixelphox]. 2023. “mi jo e palisa sitelen sin la, mi sitelen e toki lili lon pali mi.” CoHost. https://cohost.org/fennaixelphox/post/3773701-mi-jo-e-palisa-sitel
jan Lakuse. 2023. “mama sewi pi unpa ala.” utala musi pi ma pona. https://utala.pona.la/toki-en-lipu/lipu-suli/mama-pi-unpa-ala.html
jan Lentan [rnd]. 2020a. “Donsol for PICO-8.” https://www.lexaloffle.com/bbs/?tid=37711
———. 2020b. “lipu sona pona (toki pona course).” https://lipu-sona.pona.la/
———. n.d. “ante toki tawa musi pi sitelen tawa.” lipu sona pona. https://lipu-sona.pona.la/game_patches.html
jan Lesa [zendelexa]. 2024. “wikipesija_sitelen_pona.” GitHub. https://github.com/zendelexa/wikipesija_sitelen_pona
———. 2024. [Image of lipu Wikipesija in sitelen pona]. Discord. https://web.archive.org/web/20240112202644/https://cdn.discordapp.com/attachments/994960433951801374/1195369814047346728/image.png
jan lili Enta. 2023. “ante pi toki Sonwa majuna.” https://docs.google.com/document/d/1ZRYgTBHmLYoa7Gmm3CaEy9wkx3jzkkldSz75s93AcWQ/edit?usp=sharing
jan lili Enta; jan Salo. 2024. “kijetesantakalu sewi.” Discord. https://web.archive.org/save/https://cdn.discordapp.com/attachments/301380012156911616/1203897393574187009/kijetesantakalu_sewi.pdf
jan Likipi. 2021. “nimi ale kepeken nasin Wakalito 7.3.2.” https://sona.pona.la/wiki/File:nimi_ale_kepeken_nasin_Wakalito_7.3.2.pdf
jan Lun [@lunsika]. 2023. “ilo sitelen sin.” CoHost. https://cohost.org/lunsika/post/2235659-ilo-sitelen-sin
jan Majeka [@bubblebiskit]. 2023. [Diagram of a human skeleton]. Discord. https://web.archive.org/web/20240105201041/https://cdn.discordapp.com/attachments/754933236798586882/1192898416381014146/Untitled_Artwork.png
jan Mateja. 2024. “lipu open tan jan Mateja” [Postcard]. lipu tawa pi sitelen pona. https://liputawa.wordpress.com/2024/02/10/lipu-open-tan-jan-mateja/
jan Melon. 2020. “lipu lawa pi ken jan pi ma ale - toki pona UDHR.” Imgur. https://imgur.com/gallery/QWoFSaQ
jan Nasi [Nardi]. 2021. “mi pilin e ni.” YouTube. https://www.youtube.com/watch?v=96PM_jUdS2M
jan Nijo. 2021. “len kule.” *lipu kule*. https://lipukule.org/post/2021/02/11/len-kule/
jan Niko [@soweliniko]. 2021. [Message in sitelen pona taso]. Telegram. https://t.me/sitelenponataso/2107
jan Niko [Nico8300]. 2023. “sitelen pona.” Planet Minecraft. https://www.planetminecraft.com/texture-pack/sitelen-pona/
jan Nikoleta [kqa]. 2020. “ale li ante.” Spotify. https://open.spotify.com/intl-pt/album/7A8oH2Doihj43DSX5ccKAZ
jan Opa. 2022. “open suno.” YouTube. https://www.youtube.com/watch?v=PNMN1PAkKVE
jan Pilipu. 2024. “Jonah completed!” *lipu sewi pi toki pona*. https://www.tengg-online.at/blog/2024/01/12/jonah-completed/
jan Peli. 2023. “Ce n'est pas très logique qu'elles se rencontrent, comme elles ne viennent pas du même univers, mais c'est fun haha” [Photo]. Instagram. https://www.instagram.com/p/C0Y7CAgMga1/

jan Pokoli [u/Suitable_Fishing_453]. 2024. “sina sona e mi ᵛ”. Reddit. https://www.reddit.com/r/tokipona/comments/1al2mt0/sina_sona_e_mi/
jan Polo. 2021. “File:Logo of lipu tenpo.png.” https://commons.wikimedia.org/wiki/File:Logo_of_lipu_tenpo.png
jan Poman [Bo_Tie]. 2022. “poki pi sitelen pona.” Planet Minecraft. https://www.planetminecraft.com/texture-pack/poki-pi-sitelen-pona/
jan Popo. 2023. “toki ilo pona.” Internet Archive. https://archive.org/details/albert-otter-reading-order/page/n15/mode/2up
jan Sakawi. 2024. “sitelen pona for Google Workspace.” Google Web Store. https://chromewebstore.google.com/detail/sitelen-pona-for-google-w/omjmjnmbcmilkcjadbihddaipogifnbb
jan Sepulon. 2023a. “mi moku e kiwen.” YouTube. https://www.youtube.com/watch?v=xDpgwI6Ji1w
———. 2023b. “supa lape mi li ike.” https://fwisard.github.io/supalapemi/sp.html
jan Sotan [Zotn]. 2021. “olin li tawa jan ale kin.” https://www.youtube.com/watch?v=Y2Dw4aatlEE
jan Tala. 2022. “sina ken pali e sitelen pona.” https://ilo-pi-sitelen-pona.glitch.me/
jan Tamalu. 2021. “Results of the 2021 Toki Pona census.” https://tokiponacensus.github.io/results/
———. 2022. “Results of the 2022 Toki Pona census.” https://tokiponacensus.github.io/results2022/
jan Tasija. 2023. “File:Logo for Linku.svg.” https://commons.wikimedia.org/wiki/File:Logo_for_Linku.svg
jan Tepo; Corgatelli, Ethan (jan Itan); Bettencourt, Rebecca (jan Lepeka). 2022. “Sitelen Pona ConScript Unicode Registry Proposal.” https://www.kreativekorp.com/ucsur/charts/sitelen.html
jan Tolonpon [tromboneboi9]. 2024. “ilo nanpa pi jan Tolonpon.” https://tilde.town/~tromboneboi9/ilo-nanpa/index.html
jan Toluneko. 2024. “sitelen pona lon ma Wijasa | toki lili pi mun #15.” YouTube.https://www.youtube.com/watch?v=Li5q6zU3p1o
jan Tomi Wami [u/ThatKit]. 2023. “Some sitelen i found in melbourne tonight, i think they forgot the "e" but its still cool to see this stuff in the public.” Reddit. https://www.reddit.com/r/tokipona/comments/11og49o/some_sitelen_i_found_in_melbourne_tonight_i_think/
jan Usawi. 2022. “7 (likujo).” YouTube. https://www.youtube.com/watch?v=VOSvCudF5A0
———. 2023a. “oww.” YouTube. https://www.youtube.com/watch?v=zZNhdiWBp0M
———. 2023b. “toki Gaming.” Bandcamp. https://janusawi.bandcamp.com/album/toki-gaming
jan Wesi. 2014. “sitelen pona.” Imgur. https://imgur.com/a/9B6Xx
Jenkins, Oats (jan Oto). 2023. “I Learned the World's Simplest Language.” YouTube. https://www.youtube.com/watch?v=kSl-iF9ZRLk
———. 2024. “First Look at the New Toki Pona Book.” YouTube. https://www.youtube.com/watch?v=3K2G_Hvmqns
kala pona Tonyu. 2020. “o tawa noka.” YouTube. https://www.youtube.com/watch?v=_MIMQeX_-uc
———. 2023. “ilo sona Like.” https://ilo-like.bucketfish.me/
———. 2023. “File:Ma pona pi toki pona - icon.png.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Ma_pona_pi_toki_pona_-_icon.png
kapesi Pake. 2023. “File:Ali li ale (sitelen pona).pdf.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Ali_li_ale_(sitelen_pona).pdf

kasi Lija. 2024a. “i physically cannot be normal so i translated the one omake with kobeni's sister to toki pona.” Tumblr. https://www.tumblr.com/aszles/741732035079634944/click-read-more-for-a-sitelen-pona-version
———. 2024b. [Comic panel drawn on paper]. ImgBB. https://web.archive.org/web/20240106015240/https://ibb.co/nccQrNm
kule epiku Atawan; kala pona Tonyu. 2021. “tu kuntu.” Wikisource. https://wikisource.org/wiki/Tu_kuntu
kulupu lintukamakaki [soweli Kina]. 2022. “mun.” mun monsuta. https://lipumonsuta.neocities.org/mun-monsuta/
Lamp DX. 2022. “File:Ma Toki Pona VR icon.svg.” https://commons.wikimedia.org/wiki/File:Ma_Toki_Pona_VR_icon.svg
Lang, Sonja. 2014. *Toki Pona: The Language of Good.* http://www.amazon.com/gp/product/0978292308
———. 2021. *Toki Pona Dictionary.* https://www.amazon.com/dp/0978292367/
———. 2024a. “Most Popular Conlangs Online.” https://www.tokipona.org/largest_conlangs.html
———. 2024b. “Pre-Release Notes on the *su* Style of *sitelen pona*.” https://tokipona.org/notes_on_su.pdf
———. 2024c. “Category:Sitelen Pona in Sonja Lang's handwriting in 2024.” Wikimedia Commons. https://commons.wikimedia.org/wiki/Category:Sitelen_Pona_in_Sonja_Lang%27s_handwriting_in_2024
Lang, Sonja; trans. Strake, Julius. 2021. *Toki Pona: Die Sprache des Guten (German Edition)*. https://www.amazon.com/Toki-Pona-Sprache-Guten-German/dp/B09MJ8JPV6
Lang, Sonja; trans. van der Meulen, Spencer. 2022. *Tokipono: La lingvo de bono (Esperanto Edition).* https://www.amazon.com/Tokipono-lingvo-Esperanto-Sonja-Lang/dp/9464376090
———. 2023. “File:Tokipono: La lingvo de bono - eltiraĵoj laŭ CC0 kaj CC BY.pdf.” Wikimedia Commons. https://commons.wikimedia.org/w/index.php?title=File:Tokipono_La_lingvo_de_bono_-_eltira%C4%B5oj_la%C5%AD_CC0_kaj_CC_BY.pdf
Larsen, Lily (jan Lili). 2022. “sitelen pona UCSUR guide!!!” https://github.com/Id405/sitelen-pona-ucsur-guide/blob/main/README.md
Leybikhe, Rifkele D. 2023. “Capitalism Is Bad (Toki Pona, sitelen pona) Sticker.” https://www.teepublic.com/sticker/38943537-capitalism-is-bad-toki-pona-sitelen-pona
Linku. 2023. “lipu Linku.” https://linku.la/
———. 2024. “lipu Linku” [Beta]. https://beta.linku.la/
lipamanka. 2022. “toki pona dictionary.” https://lipamanka.gay/essays/dictionary
———. 2023a. “kon pan.” https://lipamanka.gay/toki-pona-taso/kon-pan
———. 2023b. “printable dictionary.” https://lipamanka.gay/resources/dictionary
———. 2024. [Photo of a shoulder tattoo depicting a weasel with the words *soweli seli* written in sitelen pona]. Discord. https://web.archive.org/web/20240227165246/https://cdn.discordapp.com/attachments/1191986138185990154/1211838601566883890/IMG_4880.jpg?ex=65efa7a5&is=65dd32a5&hm=7de8c9551b249e9c2857ea3eccc6d4d948ca195899a148b082df0fafd63ae35b&
———. n.d-a. “linja lipamanka.” https://lipamanka.gay/linjamanka
———. n.d-b. “toki pona taso.” https://lipamanka.gay/toki-pona-taso

lipu kule. “lipu sin pi lipu kule.” https://lipukule.org/sp/
lipu tenpo. 2021a. Issue 3 (*nanpa soweli*). ISSN 2752-4639. https://liputenpo.org/lipu/nanpa-soweli/
———. 2021b. Issue 4 (*nanpa kasi*). ISSN 2752-4639. https://liputenpo.org/lipu/nanpa-kasi/
———. 2021c. Issue 5 (*nanpa pan*). ISSN 2752-4639. https://liputenpo.org/lipu/nanpa-pan/
———. 2021d. Issue 10 (*nanpa lete*). ISSN 2752-4639. https://liputenpo.org/lipu/nanpa-lete/
———. 2022. Issue 14 (*nanpa seli*). ISSN 2752-4639. https://liputenpo.org/lipu/nanpa-seli/
loje Masin [u/Red-42]. 2024. “ilo nanpa pi sitelen pona.” Reddit. https://www.reddit.com/r/tokipona/comments/1b5oza8/ilo_nanpa_pi_sitelen_pona/
ma Seka. “Homepage.” https://seka.pona.la/home
moli Momo [u/MachiToons]. 2022. “musi lili kijetesantakalu.” Reddit. https://www.reddit.com/r/tokipona/comments/y4cdc7/musi_lili_kijetesantakalu/
———. 2023a. “nasin sitelen.” Reddit. https://www.reddit.com/r/tokipona/comments/138uuil/nasin_sitelen/
———. 2023b. “"the true nature of 〃 li 〃 ".” Reddit. https://www.reddit.com/r/tokipona/comments/12k6jsi/the_true_nature_of_li/
Moniz, C. Ryan (jan inwin). n.d. “mije laso li kama” [*Sir Gawain and the Green Knight*]. https://ingwine.neocities.org/imgs/mije%20laso%20li%20kama.jpg
mun Antuwan. 2023. “Interesting stats about this subreddit (Subscribers).” https://www.reddit.com/r/tokipona/comments/16cf8xq/interesting_stats_about_this_subreddit_subscribers/
mun monsuta. 2022. https://lipumonsuta.neocities.org/mun-monsuta/
Munroe, Randall; trans. Roberts, David A. (jan Tepu). 2018. “tomo tawa sewi nanpa luka.” https://davidar.github.io/tp/tomo-tawa-sewi-nanpa-luka
Munroe, Randall; trans. jan Awisa [u/janawisa]. 2023. “xkcd 1322 tawa toki pona.” Reddit. https://www.reddit.com/r/tokipona/comments/188idln/xkcd_1322_tawa_toki_pona/
Nagabe; trans. jan Kukisulasu [u/guckyslush]. 2023. “laso pi kule ala | Monotone Blue, translation of Chapter 1 in comments.” https://www.reddit.com/r/tokipona/comments/15qx3ls/laso_pi_kule_ala_monotone_blue_translation_of/
nimi.li. “nimi.li.” https://nimi.li/
Niwatori, Shiroi (Umihotaru). 2022a. “lipu sona pi nasin kepeken pi sitelen pona.” https://umihotaru.work/nishiki_sitelen.pdf
———. 2022b. “8面非ぷトキポナダイス (leko nanpa pi pu ala) (1個)” [8-sided Toki Pona dice]. BOOTH. https://umihotaru.booth.pm/items/4344155
———. 2022c. “トキポナダイス (ilo nanpa musi pu) (1個).” [Toki Pona dice]. BOOTH. https://umihotaru.booth.pm/items/4343997
Ōta, Jun’ya (ZUN); trans. jan Tepo; ed. jan Seme. 2022. “kili ike!!” [Bad Apple!!]. YouTube. https://www.youtube.com/watch?v=AhHhM28SK80
Parra Cárdenas, Eliazar. 2004. “Toki Pona - en 76 lecciones ilustradas.” https://elzr.com/tokipona/
pipi Kewapi. 2024. “Waugh image hosting !!!” Tumblr. https://adhdo5.tumblr.com/post/738729348160487425/waugh-image-hosting-suli-la-mi-lon-poki-tawa
Phox [@fennaixelphox]. 2023 “Ouija board but in toki pona.” Cohost. https://cohost.org/fennaixelphox/post/2981431-ouija-board-but-in-t
Pochin, Timeo (jan Simejo) [@timeopochin]. 2020. “Toki Pona Keyboard.” https://github.com/timeopochin/tokiponakeyboard
Podepi. 2021. “Sitelen Pona Writer.” https://tic80.com/play?cart=1858

Poe, Edgar Allan; trans. jan Likipi. 2021. “waso” [The Raven]. Discord. https://web.archive.org/web/20240104155436/https://cdn.discordapp.com/attachments/301380012156911616/808963458838036480/waso.pdf

Potter, Beatrix; trans. Samys, Ret (jan Ke Tami); jan Pulepo. 2022a. *akesi Kalasa* [The Tale of Mr. Jeremy Fisher]. Internet Archive. https://archive.org/details/toki-soweli/beatrix-potter/sitelen%20pona/PDF/akesi%20Kalasa/

———. 2022b. *jan pali len pi ma tomo Kota* [The Tailor of Gloucester]. Internet Archive. https://archive.org/details/toki-soweli/beatrix-potter/sitelen%20pona/PDF/jan%20pi%20pali%20len%20tan%20ma%20tomo%20Kota/

———. 2022c. *soweli Kinkili* [The Tale of Squirrel Nutkin]. Internet Archive. https://archive.org/details/toki-soweli/beatrix-potter/sitelen%20pona/PDF/soweli%20Kinkili

———. 2022d. *soweli Pita* [The Tale of Peter Rabbit]. Internet Archive. https://archive.org/details/toki-soweli/beatrix-potter/sitelen%20pona/PDF/soweli%20Pita

———. 2022e. *soweli Puwan* [The Tale of Tom Kitten]. Internet Archive. https://archive.org/details/toki-soweli/beatrix-potter/sitelen%20pona/PDF/soweli%20Puwan

———. 2022f. *toki lili pi soweli Misama* [Appley Dapply’s Nursery Rhymes]. Internet Archive. https://archive.org/details/toki-soweli/beatrix-potter/sitelen%20pona/PDF/toki%20lili%20pi%20soweli%20Misama/

———. 2022g. *toki musi pi soweli Wekasi* [Cecily Parsley’s Nursery Rhymes]. Internet Archive. https://archive.org/details/toki-soweli/beatrix-potter/sitelen%20pona/PDF/toki%20musi%20pi%20soweli%20Wekasi/

Projekto Babel. 2022. *telo sitelen N-ro I*. BOOTH. ISDN 278-4-671105-01-8. https://projektobabel.booth.pm/items/4326792

———. 2023a. *telo sitelen N-ro II*. BOOTH. ISDN 278-4-671105-02-5. https://projektobabel.booth.pm/items/4569904

———. 2023b. *telo sitelen N-ro III*. BOOTH. ISDN 278-4-671105-03-2. https://projektobabel.booth.pm/items/4808124

Quine, Alexander. 2023. “lipu Ante Kulupu So.” Alexander’s Blog. https://alxndr.blog/ante-kulupu-so.html

Raiter, Brian; trans. Roberts, David A. (jan Tepu). 2018. “sona nasa pi jan Ansan.” https://davidar.github.io/tp/sona-nasa-pi-jan-ansan

Reddit. 2022. “r/place.” https://reddit.com/r/place

Reis, João (jan Juwan). 2024a. “File:Fairfax HD specimen.svg.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Fairfax_HD_specimen.svg

———. 2024b. “File:Linja Enki specimen.svg.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Linja_Enki_specimen.svg

———. 2024c. “File:Linja lipamanka specimen.svg.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Linja_lipamanka_specimen.svg

———. 2024d. “File:Linja luka specimen.svg.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Linja_luka_specimen.svg

———. 2024e. “File:Linja pona specimen.svg.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Linja_pona_specimen.svg

———. 2024f. “File:Sitelen seli kiwen specimen.svg.” Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Sitelen_seli_kiwen_specimen.svg

Revoy, David; trans. Samys, Ret (jan Ke Tami). *jan Pepa en soweli Kawa* [Pepper&Carrot]. https://www.peppercarrot.com/sp/

Roberts, David A. (jan Tepu). 2018a. "jan Kikamesi." https://davidar.github.io/tp/jan-kikamesi

———. 2018b. "o kama sona e sitelen pona kepeken sitelen." https://davidar.github.io/tp/kama-sona

Robinson-Yu, Adam; trans. jan Lilan; jan Kawen; jan Nuse; soweli Tesa; jan Talu; jan Solija; jan Nikilase; & Samys, Ret (jan Ke Tami). 2023. "tawa lili" [A Short Hike]. Internet Archive. https://archive.org/details/tawa-lili

Salvoro, Ron Raphael (jan Wan). 2024. "toki en kokosila #1." Facebook. https://www.facebook.com/photo?fbid=299794013031759

Sameyn, Emilia [u/GreyWalken]. 2023. "Toki Pona Poem: Big Sun." https://www.reddit.com/r/tokipona/comments/16s08fv/toki_pona_poem_big_sun/

Samys, Ret (jan Ke Tami). 2021a. "kijetesantakalu o! nanpa luka wan." https://janketami.wordpress.com/2021/08/03/kijetesantakalu-o-nanpa-luka-wan/

———. 2021b. *ma pi kulupu tu: You decide*. English edition. ISBN 978-3754161647. https://www.amazon.de/dp/3754161644/

———. 2021c. *ma pi kulupu tu: Du entscheidest selbst*. German edition. ISBN 978-3754161654. https://www.amazon.de/dp/3754161644/

———. 2022a. "jan pi toki pona li kama kulupu la ona li sitelen pona." Reddit. https://www.reddit.com/r/sitelenpona/comments/wpsan4/jan_pi_toki_pona_li_kama_kulupu_la_ona_li_sitelen/

———. 2022b. "kijetesantakalu li lon ma tomo Win." pali pi toki pona tan jan Ke. https://janketami.wordpress.com/2022/08/07/kijetesantakalu-li-lon-ma-tomo-win/

———. 2022c. "File:Sitelen pona pangram - telo oko li ken ante e pilin.png." Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Sitelen_pona_pangram_-_telo_oko_li_ken_ante_e_pilin.png

———. 2023a. "ante toki pona." http://antetokipona.infinityfreeapp.com/font/?i=1

———. 2023b. "ante toki pona." http://antetokipona.infinityfreeapp.com/font/autofont.html

———. 2023c. "leko nimi generator." http://antetokipona.infinityfreeapp.com/leko-nimi/?i=1

———. 2023d. "leko nimi generator." https://www.reddit.com/r/tokipona/comments/17dx88g/leko_nimi_generator/

Samys, Ret [@retsamys]; mun Antuwa [@abled47]; pipi pi walo pimeja [@pipipiwalopimeja]. 2024. "a sitelen pona li ken anu seme. musi." [Thread]. Mastodon. https://toki.social/@retsamys/111890521603967258

Sbug Games; trans. jan Popo. 2023. "mi ante e toki pi musi Webbed tawa toki pona." Reddit. https://www.reddit.com/r/tokipona/comments/11tm6k5/mi_ante_e_toki_pi_musi_webbed_tawa_toki_pona/

Shestakov, Maxim (kala Asi). 2022. "Linku li seme?" YouTube. https://www.youtube.com/watch?v=3ycg6Wl_QzU

———. 2023. "2023 Word Survey results (n = 868)." https://github.com/lipu-linku/ijo/blob/main/survey/2023/results.md

Silva, Ana Laura Beatriz Santos. 2023. "Mun ante: comunicação entre personagens sem língua em comum em jogos digitais." Repositório Institucional da UFSC. https://repositorio.ufsc.br/handle/123456789/248600

soko Piki. 2024. [Drawning written "mi ken ala toki pona e ijo la mi sona ala e ijo"]. Discord. https://web.archive.org/web/20240105231229/https://media.discordapp.net/attachments/1061485980382670911/1184265705084948490/mkatpeilmsaei.png

Somato; trans. jan Onali [u/gramaticalError]. 2023. “tomo pimeja - Chapter 1 of "Shadow's House" translated.” Reddit. https://www.reddit.com/r/tokipona/comments/16kk635/tomo_pimeja_chapter_1_of_shadows_house_translated/
sona pona. “Fonts.” https://sona.pona.la/wiki/Fonts
———. “History.” https://sona.pona.la/wiki/History
———. “Linku.” https://sona.pona.la/wiki/Linku
———. “Main Page.” https://sona.pona.la/wiki/Main_Page
———. “Name glyphs.” https://sona.pona.la/wiki/Name_glyphs
———. “*nasin nanpa pona*.” https://sona.pona.la/wiki/nasin_nanpa_pona
———. “*nasin sitelen kalama*.” https://sona.pona.la/wiki/nasin_sitelen_kalama
———. “*sitelen Juniko*.” https://sona.pona.la/wiki/sitelen_Juniko
———. “Wakalito.” https://sona.pona.la/wiki/Wakalito
———. “Writing systems.” https://sona.pona.la/wiki/Writing_systems
soweli Iseja [SirStotes]. 2022. “lipu musi: toki pona playing cards.” The Game Crafter. https://www.thegamecrafter.com/games/lipu-musi-toki-pona-playing-cards
soweli Kina [@spiders]. 2023. “california flag but toki pona.” Co-Host. https://cohost.org/spiders/post/1278590-california-flag-but
soweli nasa Alesi [glazeliights]. 2023. “hello hi Im toki ponaing again.” Tumblr. https://www.tumblr.com/glazeliights/714707216040640512/hello-hi-im-toki-ponaing-again
soweli Tesa. “lipu sona mi pi toki pona.” https://lipu-sona.kittycat.homes/en
soweli Wajole. 2023. [Photo of handwriting on a notebook]. Discord. https://web.archive.org/web/20240301183326/https://cdn.discordapp.com/attachments/733009134856699924/1166052314818293850/rn_image_picker_lib_temp_f9adea7f-6b38-4f38-885c-60e4ae0da6fc.jpg?ex=65ef32d8&is=65dcbdd8&hm=7177a600b1fad22fd28bf5cee07cc0ab7fb8e8046871ba2f170204c561b7db50&
Tellez, Gabriel. 2021. “Toki Pona for Unicode.” UTC document L2/21-137. https://www.unicode.org/L2/L2021/21137-toki-pona.pdf
telo monsuta. 2023. https://lipumonsuta.neocities.org/telo-monsuta/
Thomas, Joel (jan Telakoman); Mizrahi, Gabriel (jan Kapu). 2020. “jan li ken pana e sona ale kepeken toki pona.” https://joelthomastr.github.io/tokipona/pana-sona-ale_lp
Thorson, Madaline; trans. waso mu; jan Popo [greybeetle213]. 2023. “musi Celeste kepeken sitelen pona.” https://github.com/greybeetle213/musi-Celeste-sitelen-pona
Tomohito, Ode; trans. jan Sonki; jan San. 2024. “mi en jan pona mi li (alasa) ante toki e lipu lili nanpa wan lon lipu Komi (Komi Can't Communicate) tawa toki pona!! pakala li lon la o toki.” Reddit. https://www.reddit.com/r/tokipona/comments/1b1mguo/mi_en_jan_pona_mi_li_alasa_ante_toki_e_lipu_lili/
u/artificialstarlight. 2021. “kijetesantakalu li lon sijelo mi! ^w^”. Reddit. https://www.reddit.com/r/tokipona/comments/r8frk2/kijetesantakalu_li_lon_sijelo_mi_w/
u/enpn. 2024. “Posted this in r/language, they said it may be toki pona. Is it?” Reddit. https://www.reddit.com/r/tokipona/comments/1ahbj87/posted_this_in_rlanguage_they_said_it_may_be_toki/
u/eriksealander. 2023. “Who made this in Johnson City?” Reddit. https://www.reddit.com/r/tokipona/comments/18d4h35/who_made_this_in_johnson_city/

u/EtruscaTheSeedrian. 2024. “mi jo e selo kijetesantakalu.” Reddit. https://www.reddit.com/r/tokipona/comments/196yg2o/mi_jo_e_sitelen_selo_kijetesantakalu/

u/TromboneBoi9. 2023. “little campfire ditty I wrote, works good with "ooga chakas" in the background. feel free to use, criticize, whatever!” Reddit. https://www.reddit.com/r/tokipona/comments/17nndap/little_campfire_ditty_i_wrote_works_good_with/

u/VerySwag. 2024. “my college has a public whiteboard, i had to write something.” Reddit. https://www.reddit.com/r/tokipona/comments/19du1my/my_college_has_a_public_whiteboard_i_had_to_write/

u/y_witch. 2023. “Crêpe recipe in Sitelen Pona.” Reddit. https://www.reddit.com/r/tokipona/comments/18f5bs6/cr%C3%AApe_recipe_in_sitelen_pona/

uta monsuta. 2021. https://lipumonsuta.neocities.org/uta-monsuta/

van der Meulen, Spencer (jan Pensa). 2021. ISO 639-3 Registration Authority Request for Change to ISO 639-3 Language Code. Change Request Number: 2021-043. https://iso639-3.sil.org/sites/iso639-3/files/change_requests/2021/2021-043.pdf

———. 2021. ISO 639-3 Registration Authority Request for New Language Code Element in ISO 639-3. Change Request Number: 2021-043. https://iso639-3.sil.org/sites/iso639-3/files/change_requests/2021/2021-043_tok.pdf

W., Nia (kulupu Nijako). 2023. “ilo Petewa | toki lili pi mun #12.” YouTube. https://www.youtube.com/watch?v=CoAMJYVvCeA

———. 2024. [Screenshot of jan Tenpi writing on a wall]. Discord. https://web.archive.org/web/20240114163957/https://cdn.discordapp.com/attachments/957901157076987934/1195846450429571122/VRChat_2024-01-13_15-35-25.814_3840x2160.png

waso mu. 2023. “olin.” mun.la. https://mun.la/fanart.html

waso Talu [teepuru]. 2022. “Soweli Mu.” Itch.io. https://teepuru.itch.io/soweli-mu

———. 2024. “lipu lili mi nanpa wan.” https://waso-talu.neocities.org/lipu-lili/lipu-lili-mi-nanpa-wan

Wikimedia Commons. “Category:Application of sitelen pona.” https://commons.wikimedia.org/wiki/Category:Application_of_sitelen_pona

Wikipesija. “lipu open.” https://wikipesija.org/wiki/lipu_open

Wiktionary. “Appendix:Toki Pona.” https://en.wiktionary.org/wiki/Appendix:Toki_Pona

Wong, Cho Ching (jan Sate). 2019. “ilo nanpa.” https://ilonanpa.sadale.net/

Ye, Siyuan. 2023a. “I N T E N A S I J O N A L E.” Bilibili. https://www.bilibili.com/video/BV1Ms4y1a76M/

———. 2023b. “【kalama musi pi ma ale】独一无二、简洁而爽滑的游山恋”. Bilibili. https://www.bilibili.com/video/BV1h14y1v7Q3/

Yutai Chaguan (裕泰茶馆). 2021. “【道本语】道本语数词”. Bilibili. https://www.bilibili.com/video/BV19v41177Vt/

	1DE0	1DE1	1DE2	1DE3	1DE4	1DE5	1DE6	1DE7	1DE8	1DE9	1DEA	1DEB	1DEC	1DED	1DEE	1DEF
0	1DE00	1DE10	1DE20	1DE30	1DE40	1DE50	1DE60	1DE70								1DEF0
1	1DE01	1DE11	1DE21	1DE31	1DE41	1DE51	1DE61	1DE71								1DEF1
2	1DE02	1DE12	1DE22	1DE32	1DE42	1DE52	1DE62	1DE72								1DEF2
3	1DE03	1DE13	1DE23	1DE33	1DE43	1DE53	1DE63	1DE73								1DEF3
4	1DE04	1DE14	1DE24	1DE34	1DE44	1DE54	1DE64	1DE74								SP STJ 1DEF4
5	1DE05	1DE15	1DE25	1DE35	1DE45	1DE55	1DE65	1DE75								SP SCJ 1DEF5
6	1DE06	1DE16	1DE26	1DE36	1DE46	1DE56	1DE66	1DE76								SP SOE 1DEF6
7	1DE07	1DE17	1DE27	1DE37	1DE47	1DE57	1DE67	1DE77								SP EOE 1DEF7
8	1DE08	1DE18	1DE28	1DE38	1DE48	1DE58	1DE68									SP SRE 1DEF8
9	1DE09	1DE19	1DE29	1DE39	1DE49	1DE59	1DE69									SP ERE 1DEF9
A	1DE0A	1DE1A	1DE2A	1DE3A	1DE4A	1DE5A	1DE6A		1DE8A							
B	1DE0B	1DE1B	1DE2B	1DE3B	1DE4B	1DE5B	1DE6B									
C	1DE0C	1DE1C	1DE2C	1DE3C	1DE4C	1DE5C	1DE6C									
D	1DE0D	1DE1D	1DE2D	1DE3D	1DE4D	1DE5D	1DE6D									
E	1DE0E	1DE1E	1DE2E	1DE3E	1DE4E	1DE5E	1DE6E									
F	1DE0F	1DE1F	1DE2F	1DE3F	1DE4F	1DE5F	1DE6F									

Words from pu (nimi pu)

1DE00 SITELEN PONA IDEOGRAPH A
1DE01 SITELEN PONA IDEOGRAPH AKESI
1DE02 SITELEN PONA IDEOGRAPH ALA
1DE03 SITELEN PONA IDEOGRAPH ALASA
1DE04 SITELEN PONA IDEOGRAPH ALE
1DE05 SITELEN PONA IDEOGRAPH ANPA
1DE06 SITELEN PONA IDEOGRAPH ANTE
1DE07 SITELEN PONA IDEOGRAPH ANU
1DE08 SITELEN PONA IDEOGRAPH AWEN
1DE09 SITELEN PONA IDEOGRAPH E
1DE0A SITELEN PONA IDEOGRAPH EN
1DE0B SITELEN PONA IDEOGRAPH ESUN
1DE0C SITELEN PONA IDEOGRAPH IJO
1DE0D SITELEN PONA IDEOGRAPH IKE
1DE0E SITELEN PONA IDEOGRAPH ILO
1DE0F SITELEN PONA IDEOGRAPH INSA
1DE10 SITELEN PONA IDEOGRAPH JAKI
1DE11 SITELEN PONA IDEOGRAPH JAN
1DE12 SITELEN PONA IDEOGRAPH JELO
1DE13 SITELEN PONA IDEOGRAPH JO
1DE14 SITELEN PONA IDEOGRAPH KALA
1DE15 SITELEN PONA IDEOGRAPH KALAMA
1DE16 SITELEN PONA IDEOGRAPH KAMA
1DE17 SITELEN PONA IDEOGRAPH KASI
1DE18 SITELEN PONA IDEOGRAPH KEN
1DE19 SITELEN PONA IDEOGRAPH KEPEKEN
1DE1A SITELEN PONA IDEOGRAPH KILI
1DE1B SITELEN PONA IDEOGRAPH KIWEN
1DE1C SITELEN PONA IDEOGRAPH KO
1DE1D SITELEN PONA IDEOGRAPH KON
1DE1E SITELEN PONA IDEOGRAPH KULE
1DE1F SITELEN PONA IDEOGRAPH KULUPU
1DE20 SITELEN PONA IDEOGRAPH KUTE
1DE21 SITELEN PONA IDEOGRAPH LA
1DE22 SITELEN PONA IDEOGRAPH LAPE
1DE23 SITELEN PONA IDEOGRAPH LASO
1DE24 SITELEN PONA IDEOGRAPH LAWA
1DE25 SITELEN PONA IDEOGRAPH LEN
1DE26 SITELEN PONA IDEOGRAPH LETE
1DE27 SITELEN PONA IDEOGRAPH LI
1DE28 SITELEN PONA IDEOGRAPH LILI
1DE29 SITELEN PONA IDEOGRAPH LINJA
1DE2A SITELEN PONA IDEOGRAPH LIPU
1DE2B SITELEN PONA IDEOGRAPH LOJE
1DE2C SITELEN PONA IDEOGRAPH LON
1DE2D SITELEN PONA IDEOGRAPH LUKA
1DE2E SITELEN PONA IDEOGRAPH LUKIN
1DE2F SITELEN PONA IDEOGRAPH LUPA
1DE30 SITELEN PONA IDEOGRAPH MA
1DE31 SITELEN PONA IDEOGRAPH MAMA
1DE32 SITELEN PONA IDEOGRAPH MANI
1DE33 SITELEN PONA IDEOGRAPH MELI
1DE34 SITELEN PONA IDEOGRAPH MI
1DE35 SITELEN PONA IDEOGRAPH MIJE
1DE36 SITELEN PONA IDEOGRAPH MOKU
1DE37 SITELEN PONA IDEOGRAPH MOLI
1DE38 SITELEN PONA IDEOGRAPH MONSI
1DE39 SITELEN PONA IDEOGRAPH MU
1DE3A SITELEN PONA IDEOGRAPH MUN
1DE3B SITELEN PONA IDEOGRAPH MUSI
1DE3C SITELEN PONA IDEOGRAPH MUTE
1DE3D SITELEN PONA IDEOGRAPH NANPA
1DE3E SITELEN PONA IDEOGRAPH NASA
1DE3F SITELEN PONA IDEOGRAPH NASIN
1DE40 SITELEN PONA IDEOGRAPH NENA
1DE41 SITELEN PONA IDEOGRAPH NI
1DE42 SITELEN PONA IDEOGRAPH NIMI
1DE43 SITELEN PONA IDEOGRAPH NOKA
1DE44 SITELEN PONA IDEOGRAPH O
1DE45 SITELEN PONA IDEOGRAPH OLIN
1DE46 SITELEN PONA IDEOGRAPH ONA
1DE47 SITELEN PONA IDEOGRAPH OPEN
1DE48 SITELEN PONA IDEOGRAPH PAKALA
1DE49 SITELEN PONA IDEOGRAPH PALI
1DE4A SITELEN PONA IDEOGRAPH PALISA
1DE4B SITELEN PONA IDEOGRAPH PAN
1DE4C SITELEN PONA IDEOGRAPH PANA
1DE4D SITELEN PONA IDEOGRAPH PI
1DE4E SITELEN PONA IDEOGRAPH PILIN
1DE4F SITELEN PONA IDEOGRAPH PIMEJA
1DE50 SITELEN PONA IDEOGRAPH PINI
1DE51 SITELEN PONA IDEOGRAPH PIPI
1DE52 SITELEN PONA IDEOGRAPH POKA
1DE53 SITELEN PONA IDEOGRAPH POKI
1DE54 SITELEN PONA IDEOGRAPH PONA
1DE55 SITELEN PONA IDEOGRAPH PU
1DE56 SITELEN PONA IDEOGRAPH SAMA
1DE57 SITELEN PONA IDEOGRAPH SELI
1DE58 SITELEN PONA IDEOGRAPH SELO
1DE59 SITELEN PONA IDEOGRAPH SEME
1DE5A SITELEN PONA IDEOGRAPH SEWI
1DE5B SITELEN PONA IDEOGRAPH SIJELO
1DE5C SITELEN PONA IDEOGRAPH SIKE
1DE5D SITELEN PONA IDEOGRAPH SIN
1DE5E SITELEN PONA IDEOGRAPH SINA
1DE5F SITELEN PONA IDEOGRAPH SINPIN
1DE60 SITELEN PONA IDEOGRAPH SITELEN
1DE61 SITELEN PONA IDEOGRAPH SONA
1DE62 SITELEN PONA IDEOGRAPH SOWELI
1DE63 SITELEN PONA IDEOGRAPH SULI
1DE64 SITELEN PONA IDEOGRAPH SUNO
1DE65 SITELEN PONA IDEOGRAPH SUPA
1DE66 SITELEN PONA IDEOGRAPH SUWI
1DE67 SITELEN PONA IDEOGRAPH TAN
1DE68 SITELEN PONA IDEOGRAPH TASO
1DE69 SITELEN PONA IDEOGRAPH TAWA
1DE6A SITELEN PONA IDEOGRAPH TELO
1DE6B SITELEN PONA IDEOGRAPH TENPO
1DE6C SITELEN PONA IDEOGRAPH TOKI
1DE6D SITELEN PONA IDEOGRAPH TOMO
1DE6E SITELEN PONA IDEOGRAPH TU
1DE6F SITELEN PONA IDEOGRAPH UNPA
1DE70 SITELEN PONA IDEOGRAPH UTA
1DE71 SITELEN PONA IDEOGRAPH UTALA
1DE72 SITELEN PONA IDEOGRAPH WALO
1DE73 SITELEN PONA IDEOGRAPH WAN
1DE74 SITELEN PONA IDEOGRAPH WASO
1DE75 SITELEN PONA IDEOGRAPH WAWA
1DE76 SITELEN PONA IDEOGRAPH WEKA
1DE77 SITELEN PONA IDEOGRAPH WILE

Words not from pu (nimi pi pu ala)

1DE8A SITELEN PONA IDEOGRAPH TONSI

Punctuation

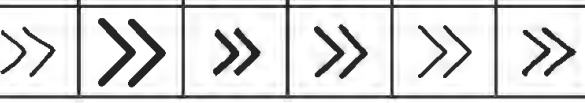
1DEF0 SITELEN PONA START OF CARTOUCHE
1DEF1 SITELEN PONA END OF CARTOUCHE
1DEF2 SITELEN PONA MIDDLE DOT
1DEF3 SITELEN PONA COLON

Format control characters

1DEF4 SITELEN PONA NESTING JOINER
1DEF5 SITELEN PONA SCALING JOINER
1DEF6 SITELEN PONA START OF EXTENSION
1DEF7 SITELEN PONA END OF EXTENSION
1DEF8 SITELEN PONA START OF REVERSE EXTENSION
1DEF9 SITELEN PONA END OF REVERSE EXTENSION

List of sitelen pona logographs

A complete list of all logographs known to the authors at the time of writing is included below, grouped by usage category according to the 2023 toki pona word usage survey. Characters are shown in six different sitelen pona fonts chosen to balance accuracy, popularity, and coverage. The first font, *nasin sitelen pu mono*, uses glyphs taken directly from Sonja Lang's first book, both the original English and its Esperanto translation which has some additional content. The second font, *nasin nanpa*, is used in *The Wonderful Wizard of Oz (Toki Pona edition)*. The penultimate font, *linja sike*, is representative of an older style that is still popular but increasingly considered outdated, even by its creator.

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
1DE00	F1900	a	core						
1DE01	F1901	akesi	core						
1DE02	F1902	ala	core						
1DE03	F1903	alasa	core						
1DE04	F1904	ale, ali	core						
1DE05	F1905	anpa	core						
1DE06	F1906	ante	core						
1DE07	F1907	anu	core						
1DE08	F1908	awen	core						
1DE09	F1909	e	core						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
1DE0A	F190A	en	core						
1DE0B	F190B	esun	core						
1DE0C	F190C	ijo	core						
1DE0D	F190D	ike	core						
1DE0E	F190E	ilo	core						
1DE0F	F190F	insa	core						
1DE10	F1910	jaki	core						
1DE11	F1911	jan	core						
1DE12	F1912	jelo	core						
1DE13	F1913	jo	core						
1DE14	F1914	kala	core						
1DE15	F1915	kalama	core						
1DE16	F1916	kama	core						
1DE17	F1917	kasi	core						
1DE18	F1918	ken	core						
1DE19	F1919	kepeken, kepen	core						
1DE1A	F191A	kili	core						
1DE1B	F191B	kiwen	core						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
1DE1C	F191C	ko	core						
1DE1D	F191D	kon	core						
1DE1E	F191E	kule	core						
1DE1F	F191F	kulupu	core						
1DE20	F1920	kute	core						
1DE21	F1921	la	core						
1DE22	F1922	lape	core						
1DE23	F1923	laso	core						
1DE24	F1924	lawa	core						
1DE25	F1925	len	core						
1DE26	F1926	lete	core						
1DE27	F1927	li	core						
1DE28	F1928	lili	core						
1DE29	F1929	linja	core						
1DE2A	F192A	lipu	core						
1DE2B	F192B	loje	core						
1DE2C	F192C	lon	core						
1DE2D	F192D	luka	core						
1DE2E	F192E	lukin	core						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
1DE2F	F192F	lupa	core						
1DE30	F1930	ma	core						
1DE31	F1931	mama	core						
1DE32	F1932	mani	core						
1DE34	F1934	mi	core						
1DE36	F1936	moku	core						
1DE37	F1937	moli	core						
1DE38	F1938	monsi	core						
1DE39	F1939	mu	core						
1DE3A	F193A	mun	core						
1DE3B	F193B	musi	core						
1DE3C	F193C	mute	core						
1DE3D	F193D	nanpa	core						
1DE3E	F193E	nasa	core						
1DE3F	F193F	nasin	core						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
1DE40	F1940	nena	core						
1DE41	F1941	ni	core						
1DE42	F1942	nimi	core						
1DE43	F1943	noka	core						
1DE44	F1944	o	core						
1DE45	F1945	olin	core						
1DE46	F1946	ona	core						
1DE47	F1947	open	core						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
1DE48	F1948	pakala	core						
1DE49	F1949	pali	core						
1DE4A	F194A	palisa	core						
1DE4B	F194B	pan	core						
1DE4C	F194C	pana	core						
1DE4D	F194D	pi	core						
1DE4E	F194E	pilin	core						
1DE4F	F194F	pimeja	core						
1DE50	F1950	pini	core						
1DE51	F1951	pipi	core						
1DE52	F1952	poka	core						
1DE53	F1953	poki	core						
1DE54	F1954	pona	core						
1DE55	F1955	pu	core						
1DE56	F1956	sama	core						
1DE57	F1957	seli	core						
1DE58	F1958	selo	core						
1DE59	F1959	seme	core						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
1DE5A	F195A	sewi	core						
1DE5B	F195B	sijelo	core						
1DE5C	F195C	sike	core						
1DE5D	F195D	sin	core						
1DE5E	F195E	sina	core						
1DE5F	F195F	sinpin	core						
1DE60	F1960	sitelen	core						
1DE61	F1961	sona	core						
1DE62	F1962	soweli	core						
1DE63	F1963	suli	core						
1DE64	F1964	suno	core						
1DE65	F1965	supa	core						
1DE66	F1966	suwi	core						
1DE67	F1967	tan	core						
1DE68	F1968	taso	core						
1DE69	F1969	tawa	core						
1DE6A	F196A	telo	core						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
1DE6B	F196B	tenpo, ten	core						
1DE6C	F196C	toki	core						
1DE6D	F196D	tomo	core						
1DE6E	F196E	tu	core						
1DE6F	F196F	unpa	core						
1DE70	F1970	uta	core						
1DE71	F1971	utala	core						
1DE72	F1972	walo	core						
1DE73	F1973	wan	core						
1DE74	F1974	waso	core						
1DE75	F1975	wawa	core						
1DE76	F1976	weka	core						
1DE77	F1977	wile	core						
—	F1980	kijetesantakalu	widespread						
—	F1979	kin	widespread						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	F1988	ku	widespread						
—	F197C	leko	widespread						
—	F1933	meli	widespread						
—	F1935	mije	widespread						
—	F197D	monsuta	widespread						
—	F1986	n	widespread						
—	F1978	namako	widespread						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
1DE78	F197E	tonsi	widespread						
—	F1983	epiku	common						
—	F197B	kipisi	common						
—	F1985	lanpan	common						
—	F1982	meso	common						
—	F1987	misikeke	common						
—	F197A	oko	common						
—	F1981	soko	common						
—	F19A1	apeja	uncommon						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	F197F	jasima	uncommon						
—	—	kiki	uncommon						
—	F1984	kokosila	uncommon						
—	—	linluwi, linuwi	uncommon						
—	F19A2	majuna	uncommon						
—	—	nimisin	uncommon						
—	—	oke	uncommon						
—	—	omekapo	uncommon						
—	F19A3	powe	uncommon						
—	—	usawi	uncommon						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	wuwojiti	uncommon						
—	—	yupekosi	uncommon						
—	—	isipin	rare						
—	—	kamalawala	rare						
—	—	kapesi	rare						
—	—	misa	rare						
—	F19A0	pake	rare						
—	—	puwa	rare						
—	—	taki	rare						
—	—	te	rare						
—	—	to	rare						
—	—	unu	rare						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	wa	rare						
—	—	jami	obscure, high end						
—	—	jonke	obscure, high end						
—	—	konwe	obscure, high end						
—	—	kulijo	obscure, high end						
—	—	melome	obscure, high end						
—	—	mijomi	obscure, high end						
—	—	mulapisu	obscure, high end						
—	—	nja	obscure, high end						
—	—	ojuta	obscure, high end						
—	—	owe	obscure, high end						
—	—	pika	obscure, high end						
—	—	po	obscure, high end						
—	—	san	obscure, high end						
—	—	soto	obscure, high end						
—	—	sutopatikuna	obscure, high end						
—	—	teje	obscure, high end						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	wasoweli	obscure, high end						
—	—	wekama	obscure, high end						
—	—	alente	obscure, low end						
—	—	alu	obscure, low end						
—	—	eliki	obscure, low end						
—	—	enko	obscure, low end						
—	—	ete	obscure, low end						
—	—	je	obscure, low end						
—	—	jule	obscure, low end						
—	—	jume	obscure, low end						
—	—	kalamARR	obscure, low end						
—	—	kalijopilale	obscure, low end						
—	—	kan	obscure, low end						
—	—	ke	obscure, low end						
—	—	kese	obscure, low end						
—	—	ki	obscure, low end						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	kisa	obscure, low end						
—	—	kosan	obscure, low end						
—	—	kulu	obscure, low end						
—	—	kuntu	obscure, low end						
—	—	likujo	obscure, low end						
—	—	molusa	obscure, low end						
—	—	nalanja	obscure, low end						
—	—	natu, aka, eki	obscure, low end						
—	—	nele	obscure, low end						
—	—	okepuma	obscure, low end						
—	—	oki	obscure, low end						
—	—	omekalike	obscure, low end						
—	—	omen	obscure, low end						
—	—	oni	obscure, low end						
—	—	pa	obscure, low end						
—	—	pakola	obscure, low end						
—	—	pasila	obscure, low end						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	pata	obscure, low end						
—	—	peta	obscure, low end						
—	—	peto	obscure, low end						
—	—	Pingo	obscure, low end						
—	—	pipo	obscure, low end						
—	—	polinpin	obscure, low end						
—	—	pomotolo	obscure, low end						
—	—	poni	obscure, low end						
—	—	sikomo	obscure, low end						
—	—	slape	obscure, low end						
—	—	tokana	obscure, low end						
—	—	tuli	obscure, low end						
—	—	waleja	obscure, low end						
—	—	wawajete	obscure, low end						
—	—	we	obscure, low end						
—	—	yutu	obscure, low end						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	ako	not notable, high end						
—	—	kikulo	not notable, high end						
—	—	kutopoma	not notable, high end						
—	—	lijokuku	not notable, high end						
—	—	mamasi, mamasina	not notable, high end						
—	—	masalo	not notable, high end						
—	—	pipolo	not notable, high end						
—	—	silapa	not notable, high end						
—	—	sipije	not notable, high end						
—	—	siwala	not notable, high end						
—	—	waken	not notable, high end						
—	—	aku	not notable, low end						
—	—	anta	not notable, low end						
—	—	apelo	not notable, low end						
—	—	awase	not notable, low end						
—	—	enepi	not notable, low end						
—	—	ewe	not notable, low end						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	i	not notable, low end						
—	—	iki	not notable, low end						
—	—	ipi	not notable, low end						
—	—	itomi	not notable, low end						
—	—	jaku	not notable, low end						
—	—	jalan	not notable, low end						
—	—	jans	not notable, low end						
—	—	kana	not notable, low end						
—	—	kapa	not notable, low end						
—	—	kikolo	not notable, low end						
—	—	kokoliko	not notable, low end						
—	—	konsi	not notable, low end						
—	—	kosikosa	not notable, low end						
—	—	loka	not notable, low end						
—	—	lokon	not notable, low end						
—	—	lo, lu	not notable, low end						
—	—	neja	not notable, low end						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	nowi	not notable, low end						
—	—	nu	not notable, low end						
—	—	nuwa	not notable, low end						
—	—	olala	not notable, low end						
—	—	panke	not notable, low end						
—	—	samu	not notable, low end						
—	—	sapelipope	not notable, low end						
—	—	sikako	not notable, low end						
—	—	sipi	not notable, low end						
—	—	ta	not notable, low end						
—	—	ten	not notable, low end						
—	—	tona	not notable, low end						
—	—	umesu	not notable, low end						
—	—	we	not notable, low end						
—	—	wi	not notable, low end						
—	—	wiwi	not notable, low end						
—	—	ana	not polled						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	api	not polled						
—	—	elen	not polled						
—	—	ene, eni	not polled						
—	—	epikule	not polled						
—	—	iseki	not polled						
—	—	kalapisituwi	not polled						
—	—	ki, koni	not polled						
—	—	kijosin	not polled						
—	—	kolo	not polled						
—	—	kulepiku	not polled						
—	—	lansan	not polled						
—	—	lato	not polled						
—	—	masu	not polled						
—	—	matula	not polled						
—	—	me	not polled						
—	—	opasan	not polled						

Proposed Code Point	UCSUR Code Point	Word	Usage Category	nasin sitelen pu mono	nasin nanpa	Nishiki-teki	linja lipamanka	linja sike	sitelen seli kiwen
—	—	saja	not polled						
—	—	si, soni	not polled						
—	—	sisi	not polled						
—	—	snoweli	not polled						
—	—	sowoli	not polled						
—	—	su	not polled						
—	—	su	not polled						
—	—	tasun	not polled						
—	—	teken	not polled						
—	—	topo	not polled						
—	—	wi, ani	not polled						
—	—	wiju	not polled						
—	—	wiki	not polled						

Figures

Works by Sonja Lang

Hieroglyphs

The letters of the Latin alphabet (presented in Lesson 1) are the most common and convenient way to write Toki Pona. There is also the ornamental *sitelen sitelen* system invented by Jonathan Gabel. (Examples of it are on page 72.)

This section presents a simple hieroglyphic system called *sitelen pona*. Each Toki Pona word is represented by a unique logogram or symbol.

Figure 1a. Chapter on sitelen pona in *Toki Pona: The Language of Good* (Lang 2014).

Figure 1b. Chapter on sitelen pona in *Toki Pona: The Language of Good* (Lang 2014).

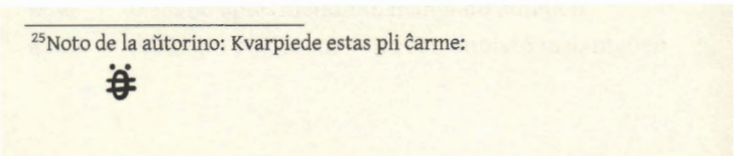


Figure 2a. Footnote from *Tokipono: La lingvo de bono* showing an alternate form of *akesi*.

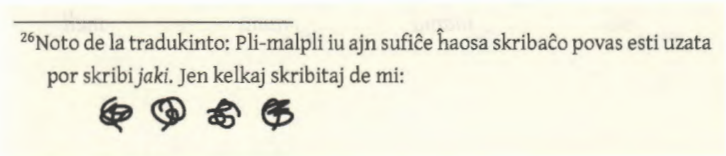
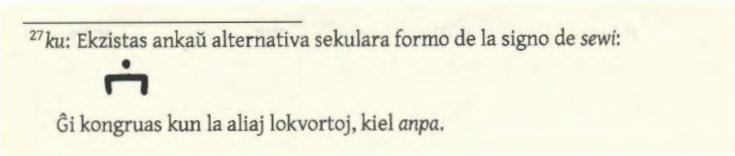
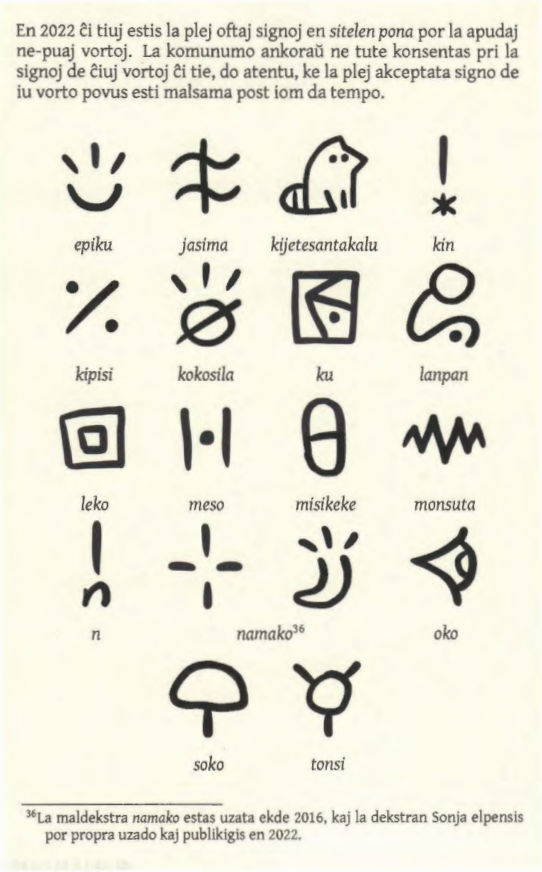

[26]Noto de la tradukinto: Pli-malpli iu ajn sufiĉe ĥaosa skribaĉo povas esti uzata por skribi *jaki*. Jen kelkaj skribitaj de mi:

Figure 2b. Footnote from *Tokipono: La lingvo de bono* showing alternate forms of *jaki*.


[27]*ku*: Ekzistas ankaŭ alternativa sekulara formo de la signo de *sewi*:

Ĝi kongruas kun la aliaj lokvortoj, kiel *anpa*.

Figure 2c. Footnote from *Tokipono: La lingvo de bono* showing an alternate form of *sewi*.


En 2022 ĉi tiuj estis la plej oftaj signoj en *sitelen pona* por la apudaj ne-puaj vortoj. La komunumo ankoraŭ ne tute konsentas pri la signoj de ĉiuj vortoj ĉi tie, do atentu, ke la plej akceptata signo de iu vorto povus esti malsama post iom da tempo.

[36]La maldekstra *namako* estas uzata ekde 2016, kaj la dekstran Sonja elpensis por propra uzado kaj publikigis en 2022.

Figure 3. Logographs for the 17 *nimi ku suli* as presented in *Tokipono: La lingvo de bono*.

Esperanto

Tiu ĉi dosiero estas kolekto de ĉiuj paĝoj el la libro *Tokipono: La lingvo de bono* de Sonja Lang (ISBN: 978-94-6437-609-8), kies permesiloj kongruas kun permesiloj de la tipo CC BY-SA. Mi publike disponebligas ilin, por ke oni povu pli facile uzi la permesilojn por krei novajn verkojn. Fine de la dosiero estas detala listo de atribuoj kaj permesiloj.

Rimarku ke paĝoj 1-8 kaj 148-149 de la libro (koresponde al paĝoj 1-8 kaj 32-33 en ĉi tiu PDF) estas nun uzeblaj per CC0 1.0, kun permeso de ĉiuj rajtohavantoj. En ĉi tiu dokumento estas ankaŭ listo de ĉiuj piednotoj en la libro, kiu estas uzebla per la permesilo CC BY 4.0.

English

This file is a collection of all pages from *Tokipono: La lingvo de bono* by Sonja Lang (ISBN: 978-94-6437-609-8) that are compatible with CC BY-SA licenses. I make them freely available to make it easier for people to use the licenses to make new works. At the end of the document is a detailed list of credits and licenses.

Note that pages 1-8 and 148-149 of the book (corresponding to pages 1-8 and 32-33 in this PDF) are now available under CC0 1.0, with permission from all rights holders. In this document there is also a list of all footnotes from the book, which are made available under the license CC BY 4.0.

Detailed credits (in English)

The page numbers in these credits refer to the page numbers of the physical book, which are located at the bottom of each page.

These are the Creative Commons licenses that were used and their abbreviations. Check the webpages for information about what each license allows you to do and what terms apply.

- CC BY 4.0 – Attribution 4.0 International
  https://creativecommons.org/licenses/by/4.0/
- CC0 1.0 – CC0 1.0 Universal Public Domain Dedication
  https://creativecommons.org/publicdomain/zero/1.0/

Texts on pages 3, 116-123, 137-147 (not including footnotes) are written by Sonja Lang (2014), and translated by Spencer van der Meulen (2022).

License: CC0 1.0 (public domain)

The tekst on pages 7-8 is written by Spencer van der Meulen (2022).

License: CC0 1.0 (public domain)

The Toki Pona text on page 95 is written by Joop Kiefte (year unknown). The introduction and Esperanto translation are written by Spencer van der Meulen (2022).

License: CC0 1.0 (public domain)

The Toki Pona text on page 96 is written by Eŭgeno; original version from 2009, this version from 2022. The introduction on page 96 and the Esperanto translation on page 97 are written by Spencer van der Meulen (2022).

License: CC BY 4.0

The poem *Ho, mia kor'* on page 98 is written by L. L. Zamenhof (1887). The Toki Pona version and accompanying explanations are written by Spencer van der Meulen (2022).

License: CC0 1.0 (public domain)

The tekst on page 148 is written by Julius Strake and Spencer van der Meulen, Esperanto translation by Spencer van der Meulen. The text on page 149 is written by Spencer van der Meulen.

License: CC0 1.0 (public domain)

The illustration on page 1 is made by Sonja Lang (2014) and (re)vectorized by Spencer van der Meulen (2022).

License: CC0 1.0 (public domain)

The *sitelen pona* illustrations on pages 116-123 (not including footnotes) are made by Sonja Lang (2014).

License: CC0 1.0 (public domain)

The footnotes marked with "*ku*" are written by Sonja Lang (2021), and translated and adapted by Spencer van der Meulen (2022), following the style of the footnotes from Julius Strake in *Toki Pona: Die Sprache des Guten* (2021).

The footnote on page 116 marked with "Noto de la aŭtorino" is written by Sonja Lang (2022).

The footnote on page 142 is written by Julius Strake (2021) and translated by Spencer van der Meulen (2022).

Other footnotes are written by Spencer van der Meulen (2022).

License: CC BY 4.0

(Footnotes on CC0 pages (7, 116-123, 137-149) are additionally available under CC0 1.0.)

The *sitelen pona* illustrations on page 149 and in the footnotes on pages 116-117 and 120 are made by Spencer van der Meulen (2022).

License: CC0 1.0 (public domain)

Figure 4. Credits from *Tokipono: La lingvo de bono*, indicating sitelen pona illustrations have been released into the public domain.

Pre-Release Notes on the *su* Style of *sitelen pona*

lipu su is an upcoming 2024 book printed almost entirely in the *sitelen pona* writing system, using a modified version of the *nasin-nanpa* font. *lipu su* will be part of the Official Toki Pona book series produced and published by Sonja Lang, creator of Toki Pona and *sitelen pona*.

Some details of the *su* style of *sitelen pona* include:

- four-legged *akesi* as default (six-legged *pu*-style akesi not used)
  sources:
  - *"akesi li suwi!"* update from *Toki Pona Dictionary (lipu ku)* p. 13
  - *Tokipono: La lingvo de bono (Esperanto Edition) (pu Epelanto)*
- *jasima*
- *lanpan*
- sunset or lotus *majuna* as default (upside-down *sin* glyph not used)
- *mute* (when is meaning is many, much, very, quantity, three)
- *mute* (when meaning is twenty)
- pepper *namako* as default (four-line *sin*-derived glyph not used)
- *sewi* (when meaning is divine, sacred, God)
- *sewi* (when meaning is upper, top, high, sky, above)
- *su*

As always, books in the Official Toki Pona series are not prescriptive. They reflect the preferred style of the language creator at the time of publication. Subsequent books also provide minor updates to previous books of the same series.

- "This is the way I use Toki Pona." (*lipu pu* p. 7)
- "Decide your style of Toki Pona for yourself." (*lipu ku* p. 23)

Sonja Lang
2024-01-06

Figure 5. Pre-release notes for *The Wonderful Wizard of Oz (Toki Pona edition)*, the latest official book by Sonja Lang, showing alternate forms of *akesi*, *majuna*, *namako*, and *sewi*. Alternate *mute* and *su* are not proposed as they are too new to have gained significant usage.

Copyrighted Material

The Wonderful Wizard of Oz

Toki Pona edition

translated by Sonja Lang
illustrated by Evan Dahm

Copyrighted Material

Copyrighted Material

Work Group

Written in English by L. Frank Baum.
Translated to Toki Pona by Sonja Lang.
Illustrated by Evan Dahm *(rice-boy.com)*.
Proofread by jan Kekan San, jan Pensa, jan Asiku, jan Sepulon and jan Polijan.
nasin nanpa font by Ethan.
With help from jan Ke Tami, Lipamanka and Emmanuelle Richard.
Published by Sonja Lang.

Toki Pona and *sitelen pona* are free.
This book is copyright Sonja Lang © 2024

ISBN: 978-0-9782923-7-9

Copyrighted Material

Copyrighted Material

Book Parts

	The Strong Wind
	Knowledge of the Little Munchkin People
	The Person of Bird Removal
	Plant Land
	The Metal Person of Plant Cutting
	The Strong Beast of Fearful Feelings
	The Land of the Deep Holes
	The Land of the Sleep Plants
	The Ruling Small Mammal
	The Protective Person of the Large Wall
	Residential Land of the Blue-Green Mineral
	The Evil Powerful Woman of the Ending Sun
	The Giving of All Possibilities
	The Winged Humanlike Beasts
	The Truth of the Mighty Person Named Oz
	Powerful Knowledge of the Entertainer
	The Air Sphere of Going Up
	The Hot Land of No Water
	The Violent Plants
	The Hard Material of Breaking Possibility
	Rule of the Animals
	The Land of the Group Called Quadlings
	The Good Powerful Woman Named Glinda
	Ah, Home!

Copyrighted Material

Copyrighted Material

Figure 6. Excerpt pages from *The Wonderful Wizard of Oz (Toki Pona edition)*.

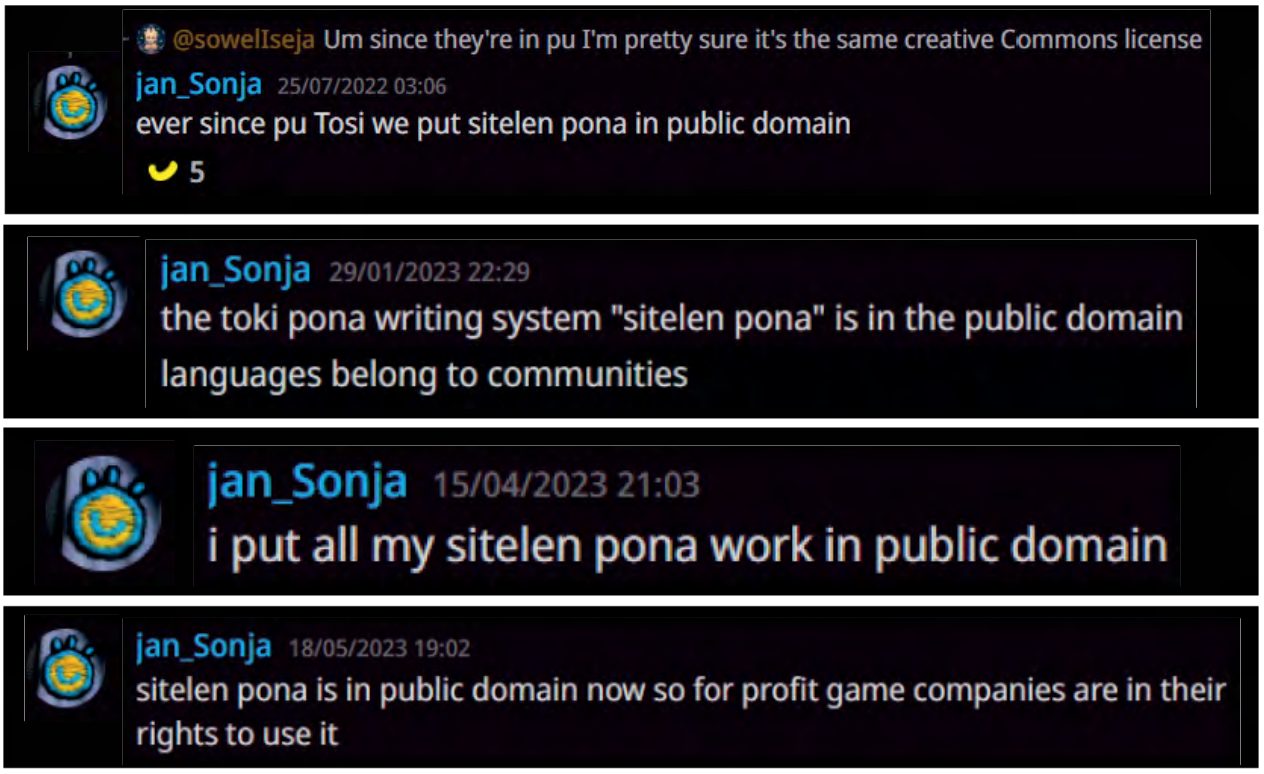


Figure 7. Statements from Sonja Lang on Discord regarding the copyright status of sitelen pona. She has released sitelen pona into the public domain. (*pu Tosi* refers to *Toki Pona: Die Sprache des Guten*, the German edition of *Toki Pona: The Language of Good*.)

Works by other authors

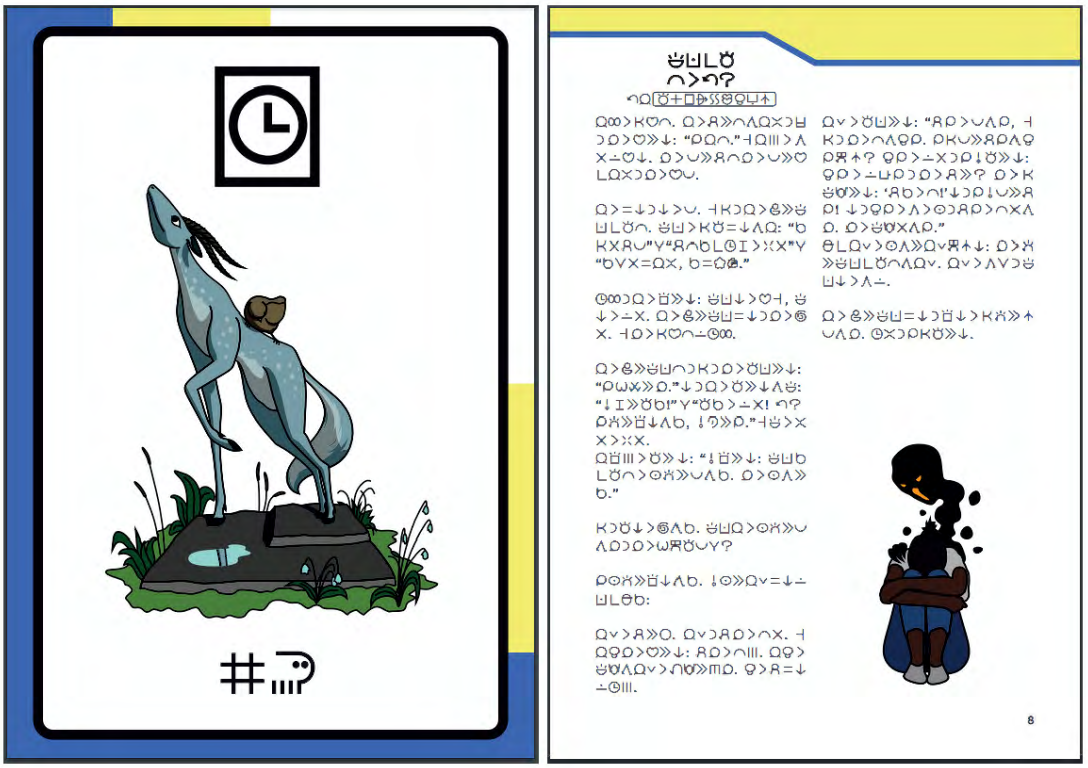


Figure 8. Cover art and a page of text from issue 3 of the toki pona magazine *lipu tenpo*. Issues 3, 4, 5, 10, and 14 had editions written completely in sitelen pona.

Figure 9. The toki pona magazine *jaki pona* written in sitelen pona (Billsmugs 2022).

Figure 10. Covers for the yearly horror zine series *lipu monsuta*.
From left to right, *uta monsuta* (2021), *mun monsuta* (2022), and *telo monsuta* (2023).

Figure 11. Covers from *telo sitelen*, a multilingual magazine written in Japanese and Esperanto, also featuring toki pona (Projekto Babel 2022, 2023a, 2023b).

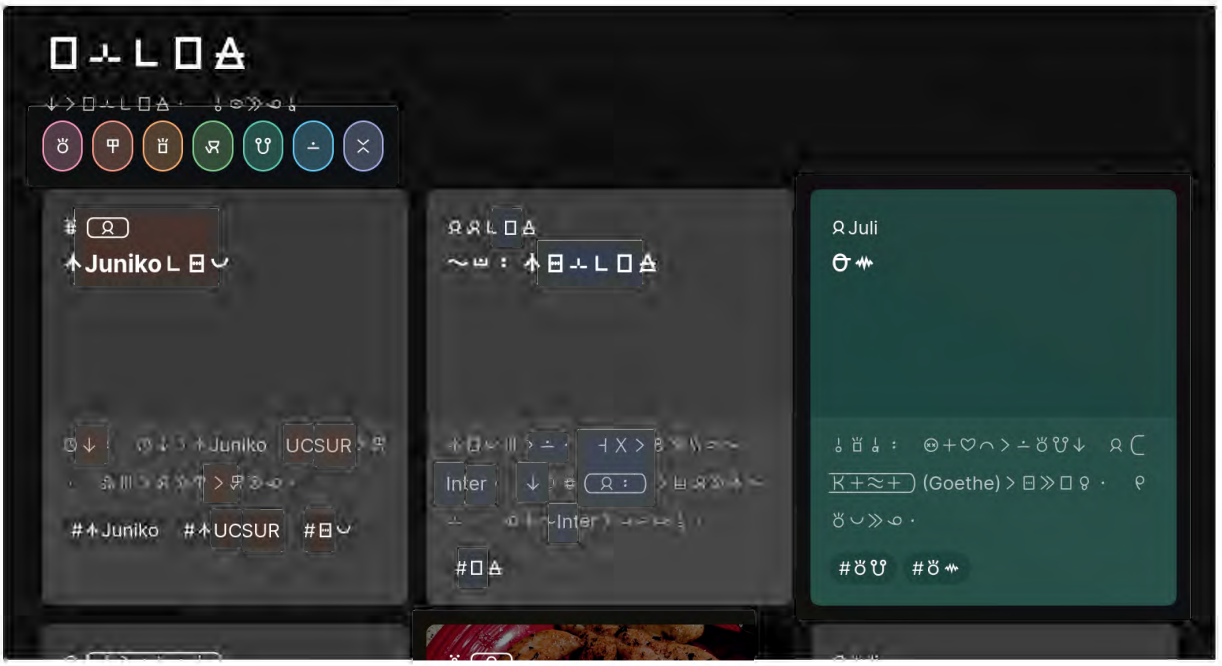


Figure 12. Screenshot from March 2024 of *lipu kule*, a collaborative blog managed by Janet Black.

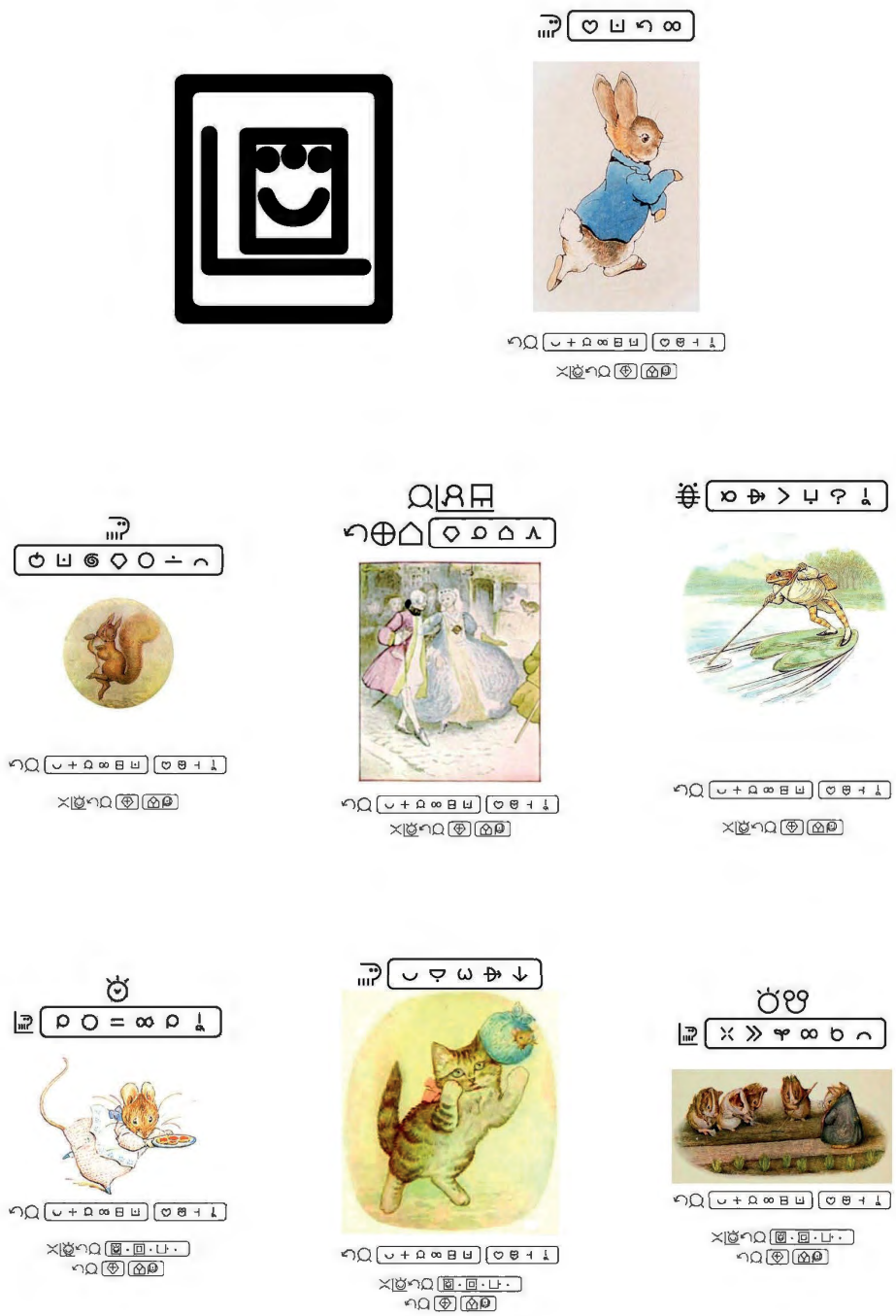


Figure 13. Covers of Beatrix Potter (2022a–g) illustrated books, translated by Ret Samys (jan Ke Tami) and jan Pulepo.

Figure 14. Pages from *meli lili pi len loje*, translation of the Grimm brothers' classic fairytale *Little Red Riding Hood* by James Flear (Grimm 2019).

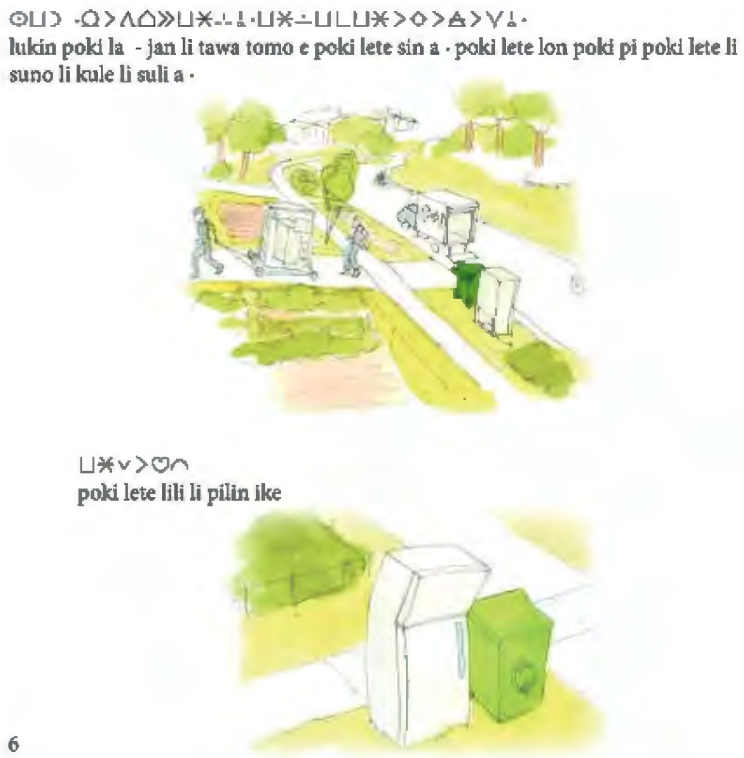
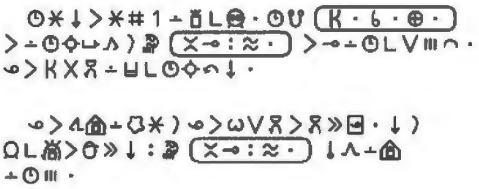


Figure 15. Page from *poki lete lili pi wan taso*, translation of *The Lonely Little Fridge* by Abigail Lyman (Dawson 2022/2022).



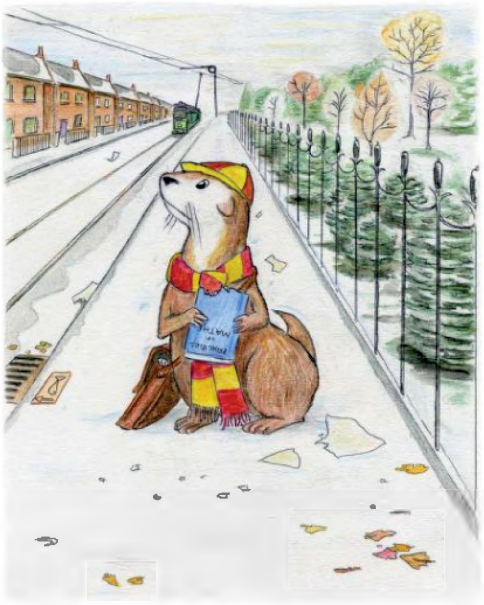


Figure 16. Page from *ken len pi soweli telo Alapete*, translation of the book *The Hidden Talent of Albert Otter* written by David Haywood and Peter Adamson, translated by jan Popo (Haywood, Adamson 2010/2023).

29

"toki," she replies. She seems ready to answer your questions.

sina pilin seme? → 31
sina pana e mani tawa mi anu seme? → 35
mi tawa → 12

30

"☉," he answers. He seems ready to answer your questions.

☋ω≫?→ 34
¡☉≫⌂L◇☒→ 36
ᑭΛ→ 11

Figure 17. Excerpt from *ma pi kulupu tu*, a "choose your own adventure" book by Ret Samys written partially in English, partially in toki pona (Samys, 2021b, 2021c).

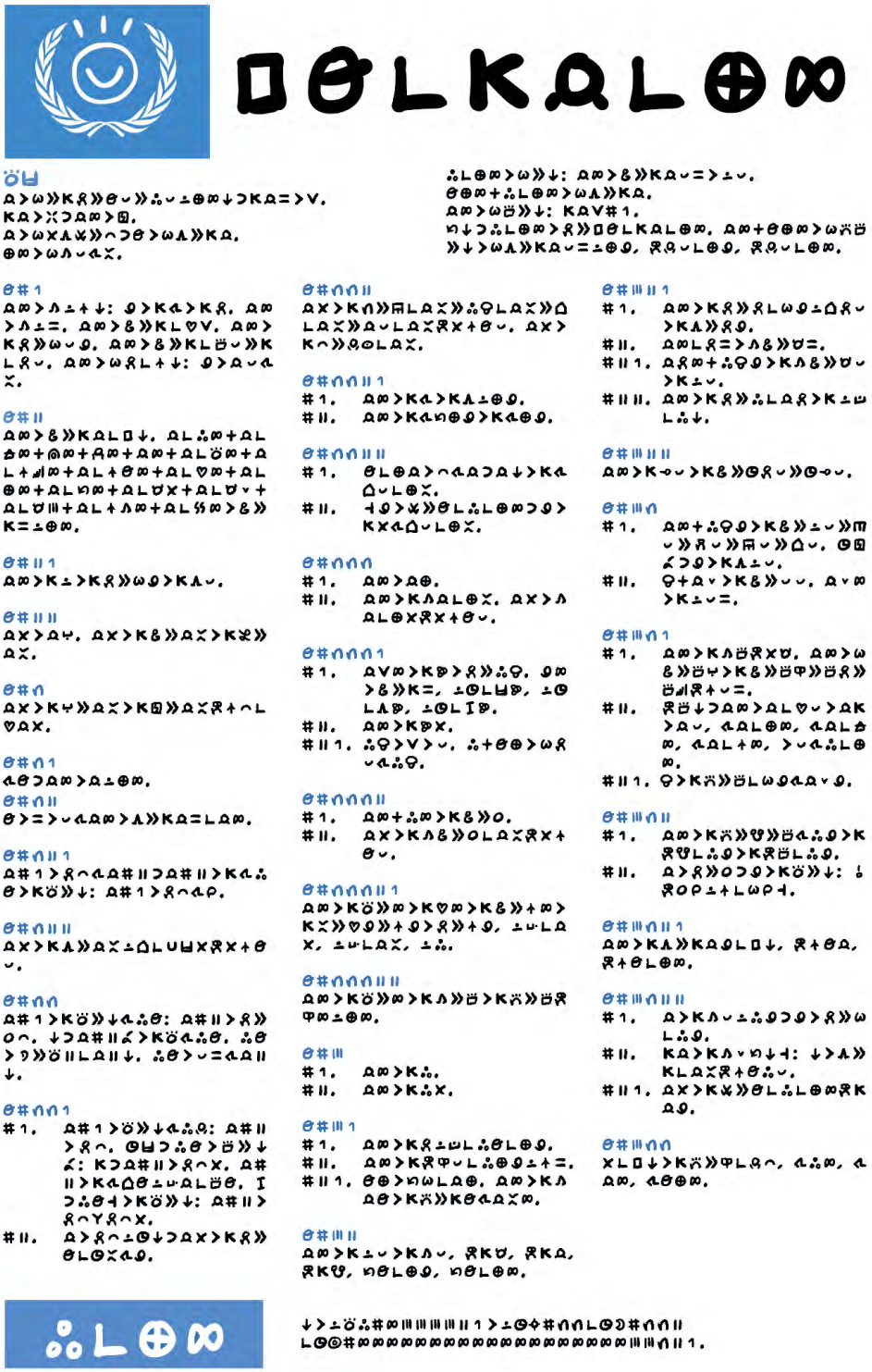


Figure 18. Unofficial translation of the Universal Declaration of Human Rights (UDHR) into toki pona, written in sitelen pona (jan Melon 2020).

Figure 19. Unofficial translation of the Universal Declaration of Human Rights (UDHR) into toki pona, compiled from previous authors (Gotik 2020).

Figure 20. Book cover and a page of text from a translation of the Bible into toki pona. The logo on the cover is stylized sitelen pona for *lipu sewi pi toki pona* (jan Pilipu 2024).

Figure 21. *supa lape mi li ike*, a web-based adventure game (jan Sepulon 2023b).

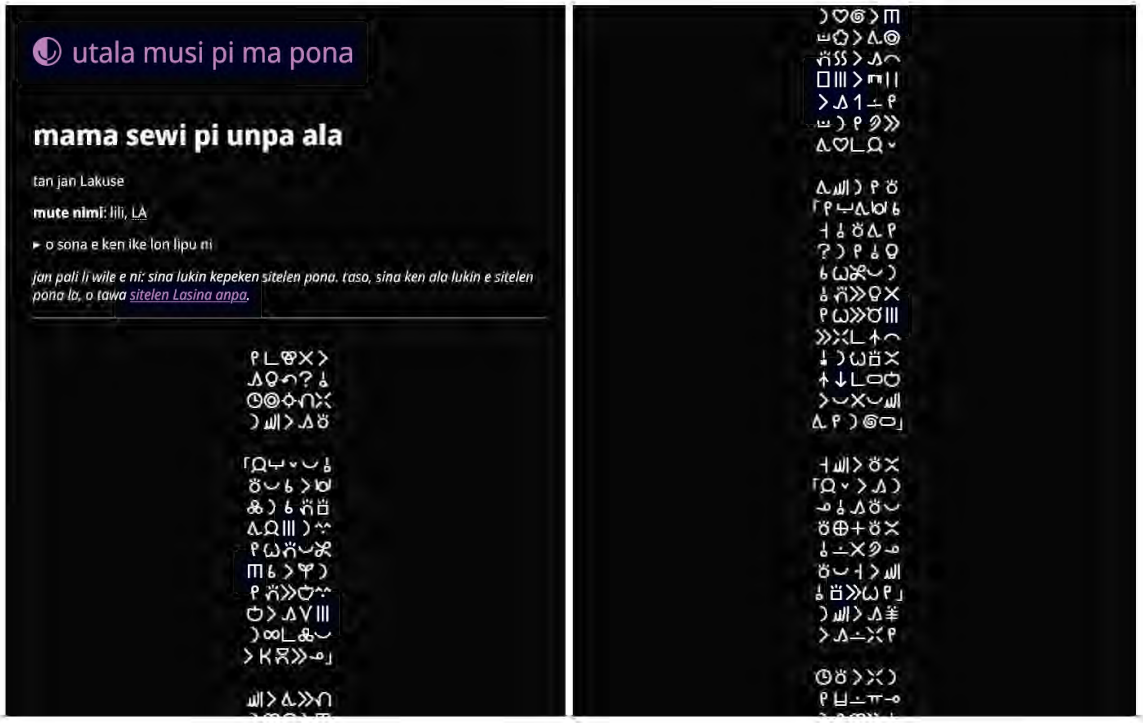


Figure 22. Excerpt from original toki pona story "*mama sewi pi unpa ala*," shown in the writing competition website *utala musi pi ma pona* (jan Lakuse 2023).

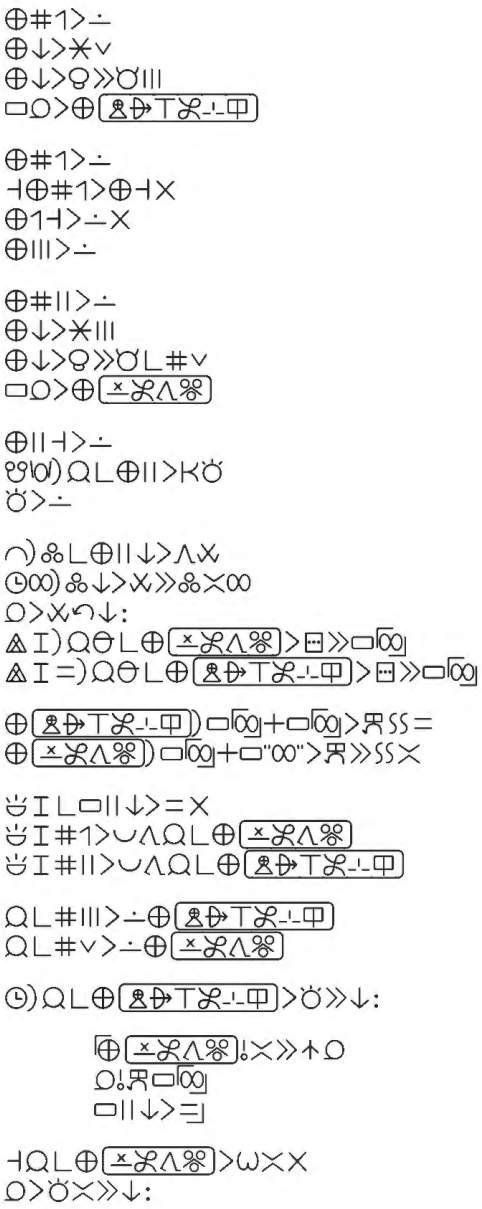


Figure 23. “ali li ale,” a fiction story written in sitelen pona (kapesi Pake 2023).

Figure 24. “telo oko li ken ante e pilin,” a story by Ret Samys which serves as a word pangram, written in sitelen pona.

Figure 25. Excerpt from “kon pan”, a traditional Jewish story, translated by lipamanka (2023a), written in sitelen pona.

Figure 26. The lyrics to “open suno,” an original song in toki pona (jan Opa 2022).

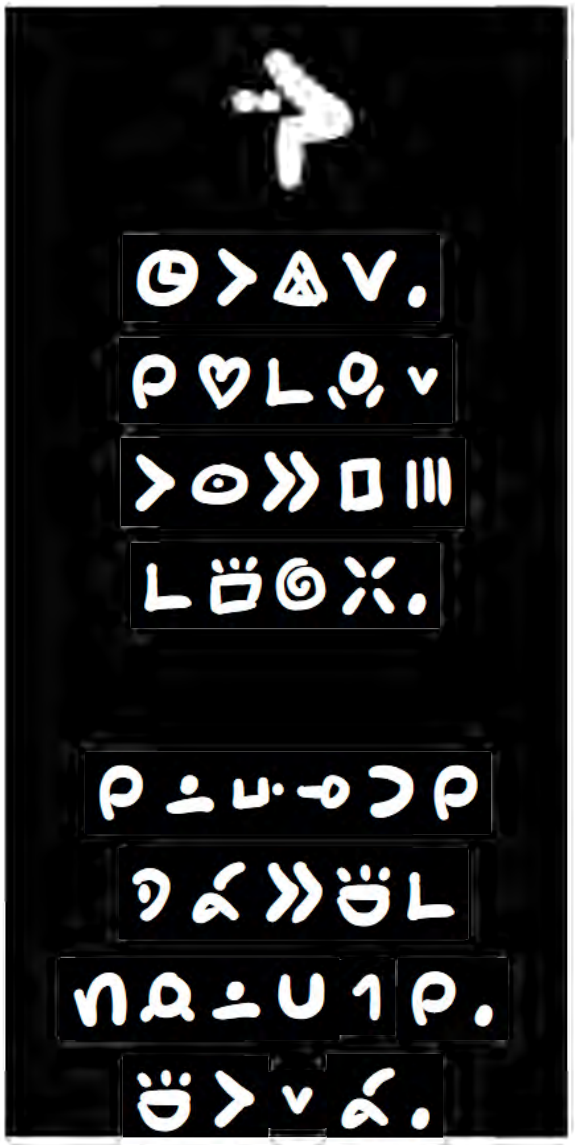


Figure 27. Excerpt from Edgar Allan Poe's *The Raven*, translated by jan Likipi (2021).

Figure 28. Excerpt from the story *akesi li wile lon nena* (jan Kita 2022).

Figure 29. Translation of the ancient Chinese text *I Ching* (*Yi Jing*),
translated by Alexander Quine (jan Lasate), in Seal Script and sitelen pona (Quine 2023).

Figure 30. Duolingo stories in sitelen pona, as a translation in the licensed unofficial language-learning resource *Duostories*.

Figure 31. Panels from *kijetesantakalu o!*, an original webcomic by Ret Samys (2021a).

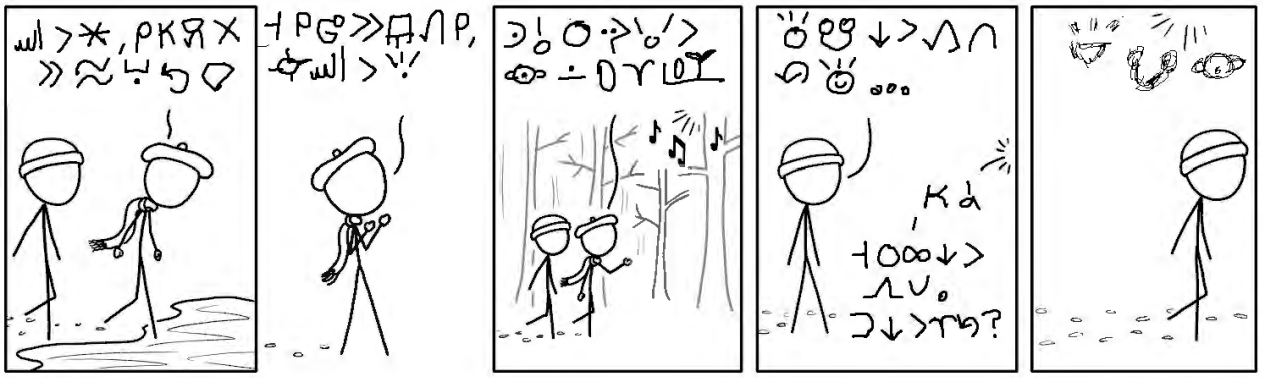


Figure 32. Translation of *xkcd* comic #1322 into sitelen pona by jan Awisa (Munroe 2023).

Figure 33. Excerpt the webcomic *Duncan & Eddie* translated into toki pona by waso Atu, displaying reversed pronoun variants (Chervil 2023).

Figure 34. Short comic by soweli nasa Alesi (2023), displaying creative and flexible cartouche usage.

Figure 35. Panels from a short comic by moli Momo (2022), displaying artistic modifications of sitelen pona glyphs.

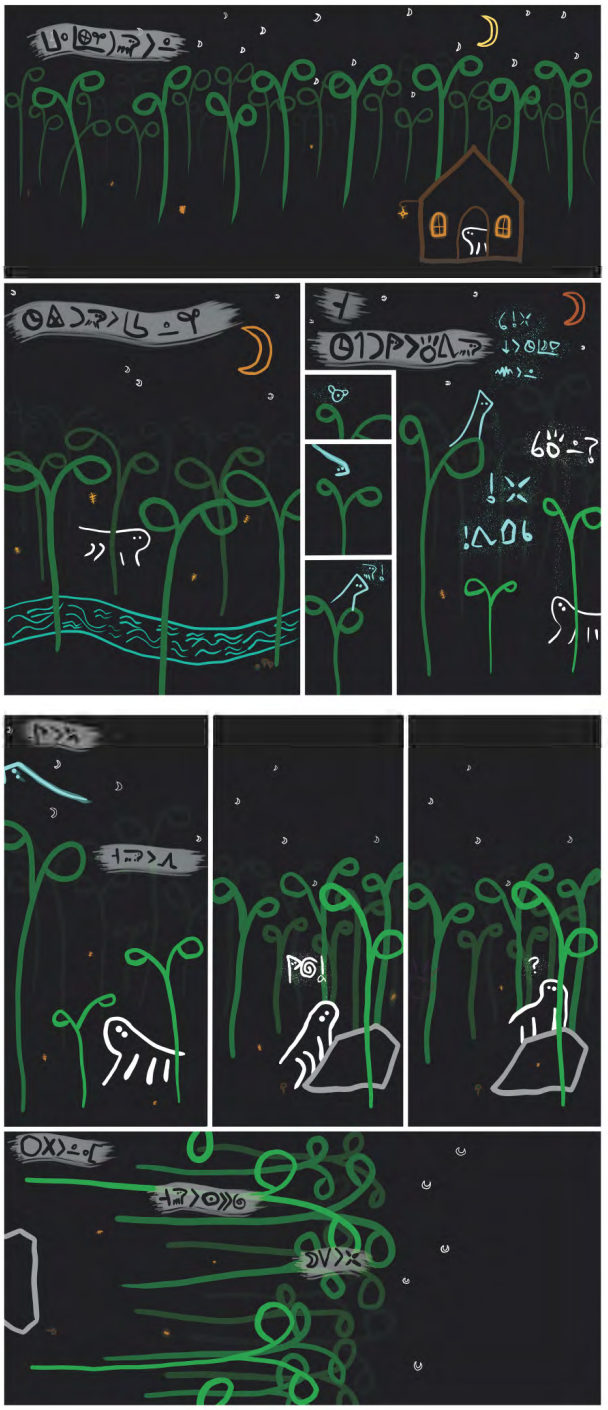


Figure 36. Short comic “mun” in the toki pona zine *mun monsuta*, written and illustrated by kulupu lintukamakaki (2022).

falktroute
troutefish
falktroute

Figure 37. Translation excerpt of the webcomic *How to Be a Dragon*, translated into toki pona by Erek Butcher (soweli Eweke; Fishtrouts 2023).

Figure 38. Page from *Chainsaw Man* translated into toki pona by kasi Lija (2024a).

Figure 39. Pages from Chapter 1 of *jan Komi li ken ala toki*, a translation of the manga *Komi Can't Communicate*, translated into toki pona by jan Sonki and jan San (Tomohito 2024).

Figure 40. Page from Chapter 1 of *laso pi kule ala*, a translation of the manga *Monotone Blue* into toki pona by jan Kukisulasu (Nagabe 2023).

Figure 41. Page from Chapter 1 of *tomo pimeja*, a translation of the manga *Shadows House* into toki pona by jan Onali (Somato 2023).

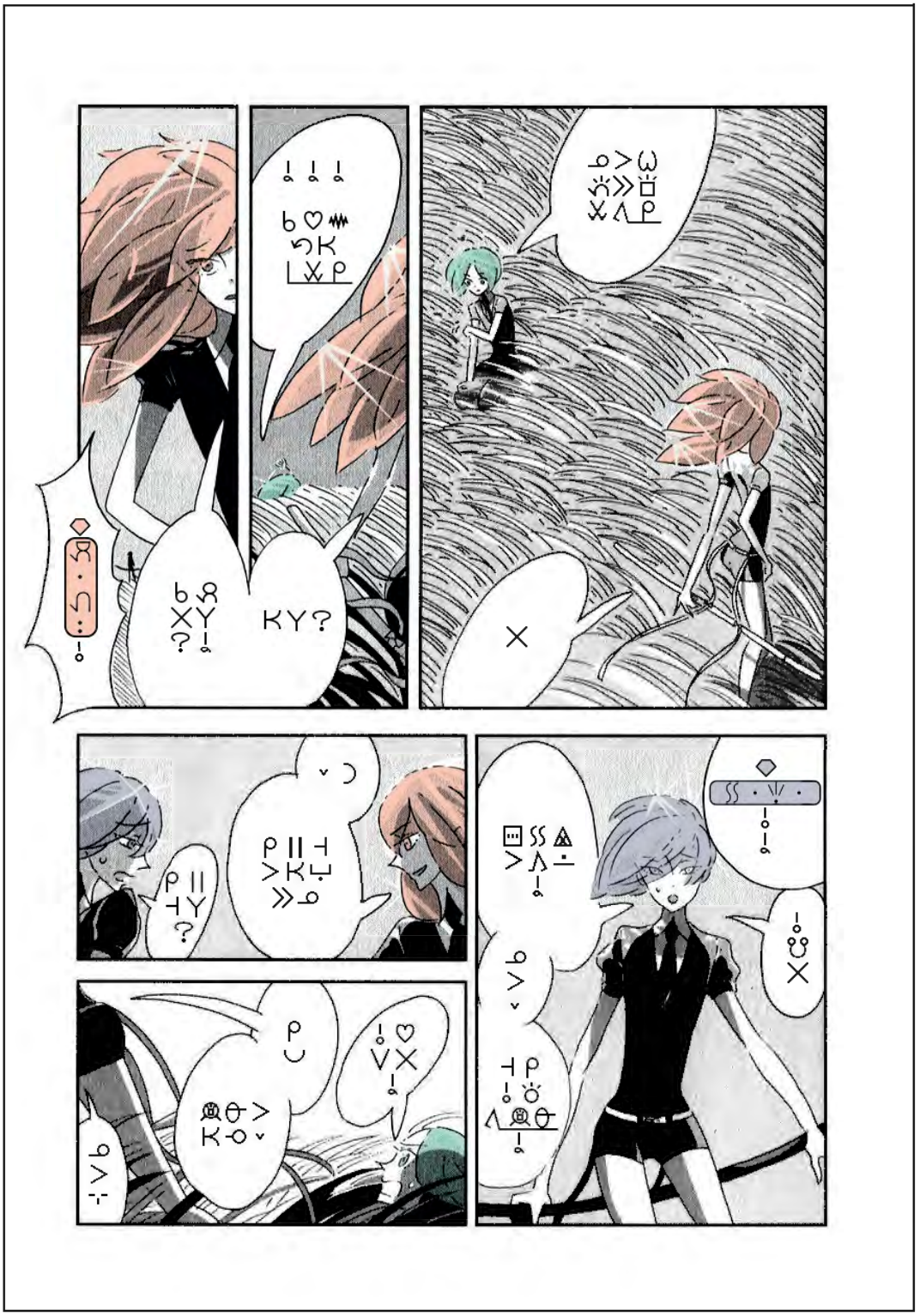


Figure 42. Page from Chapter 1 of *kulupu kiwen*, a translation of the manga *Land of the Lustrous* into toki pona by jan Kukisulasu (Ichikawa 2023).

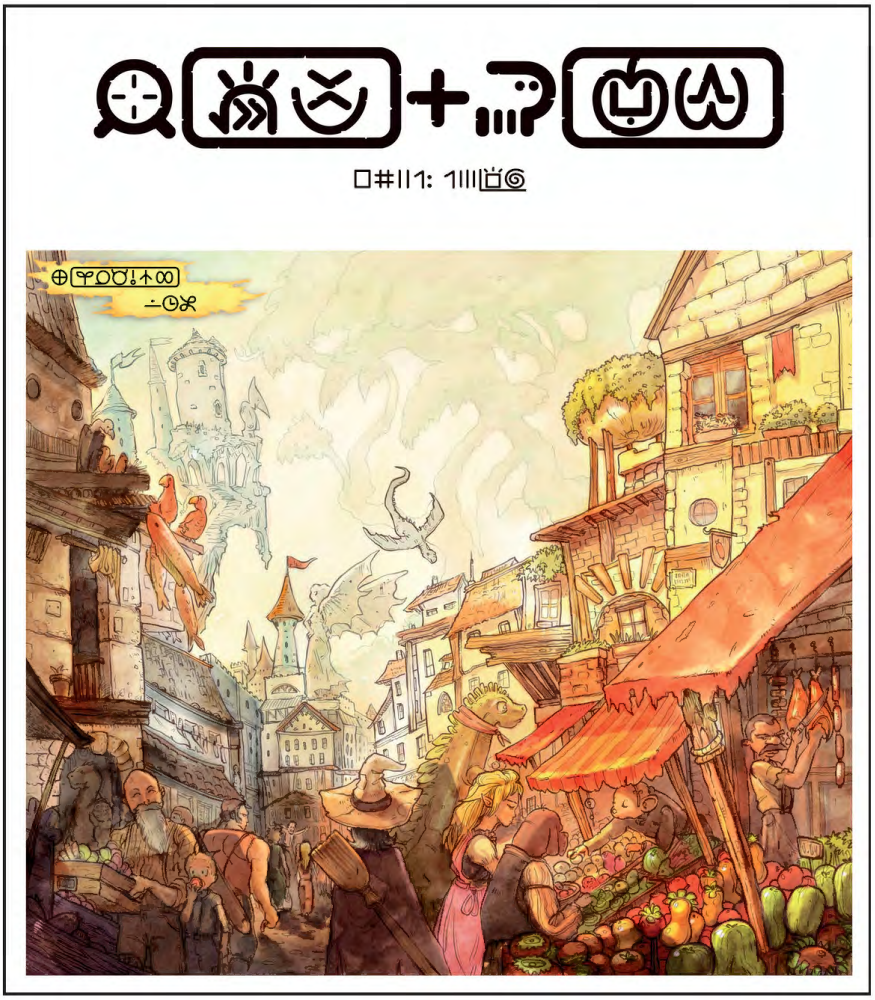


Figure 43. Excerpt from the official translation of the webcomic *Pepper&Carrot*, translated by Ret Samys (Revoy 2021).

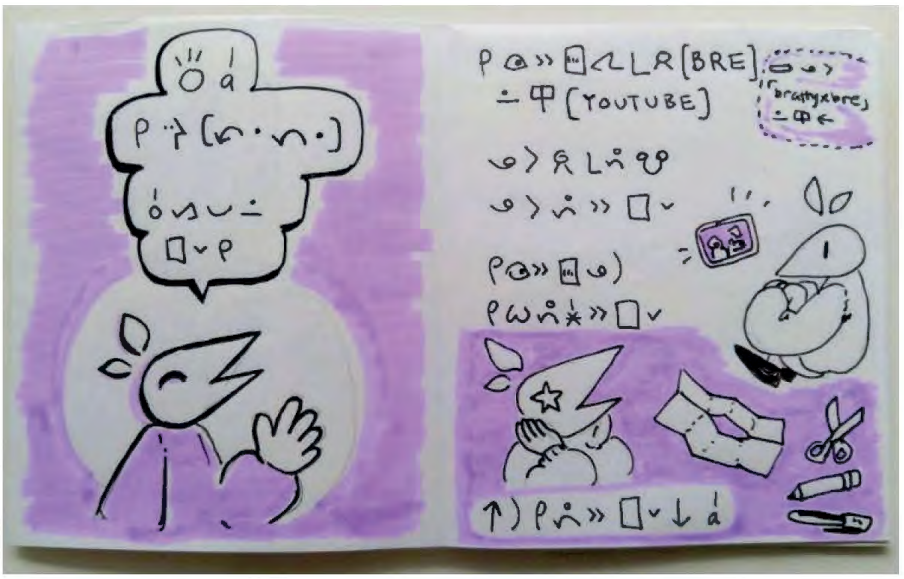


Figure 44. Pages from an original zine by waso Talu (2024).

Figure 45. Screenshot of *tawa lili*, a translation mod for the game *a short hike*, translated by several volunteers (Robinson-Yu 2023).

Figure 46. Screenshot of a translation mod for the game *Celeste*, translated by waso mu, adapted for sitelen pona by jan Popo (Thorson 2023).

Figure 47. Title screen of an unofficial translation of the game *The Legend of Zelda: Link's Awakening*, translated by Kelsey Higham (2024, waso Keli), typeset by jan suli.

Figure 48. Screenshots from the above translation by Kelsey Higham (2022), showing game dialogue using sitelen pona.

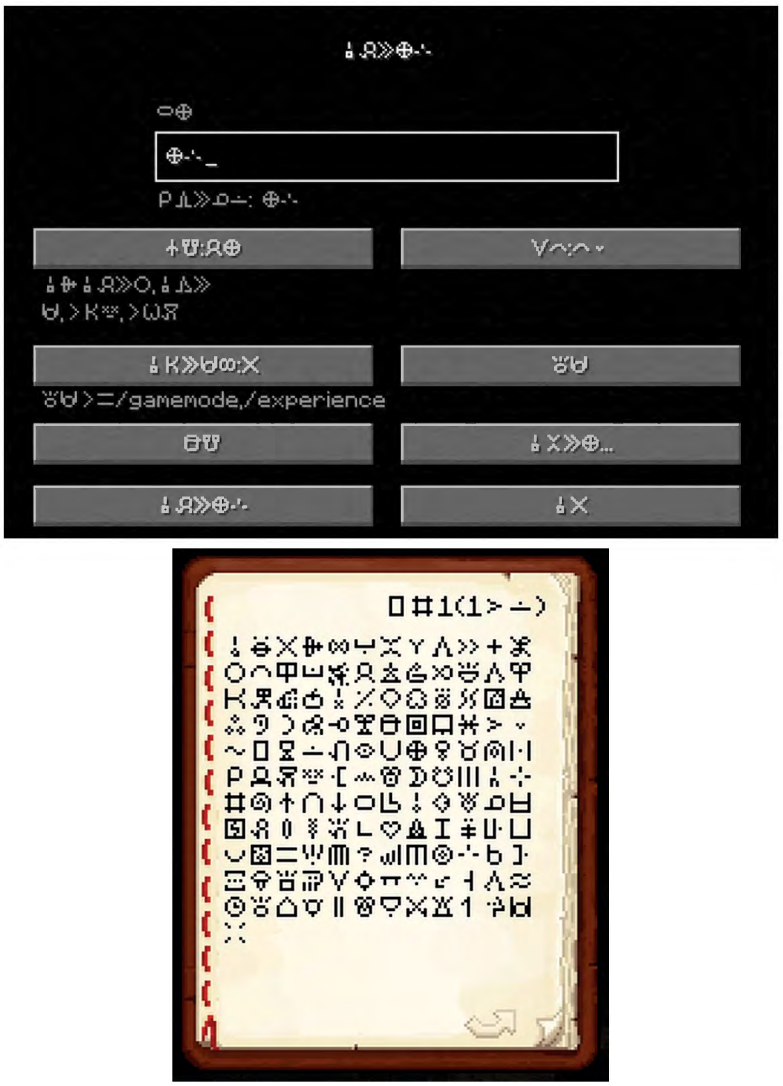


Figure 49. Screenshots of two resource packs to add sitelen pona to the official translation for the game *Minecraft*, created by (left) jan Poman (2022) and (right) jan Niko (2023).

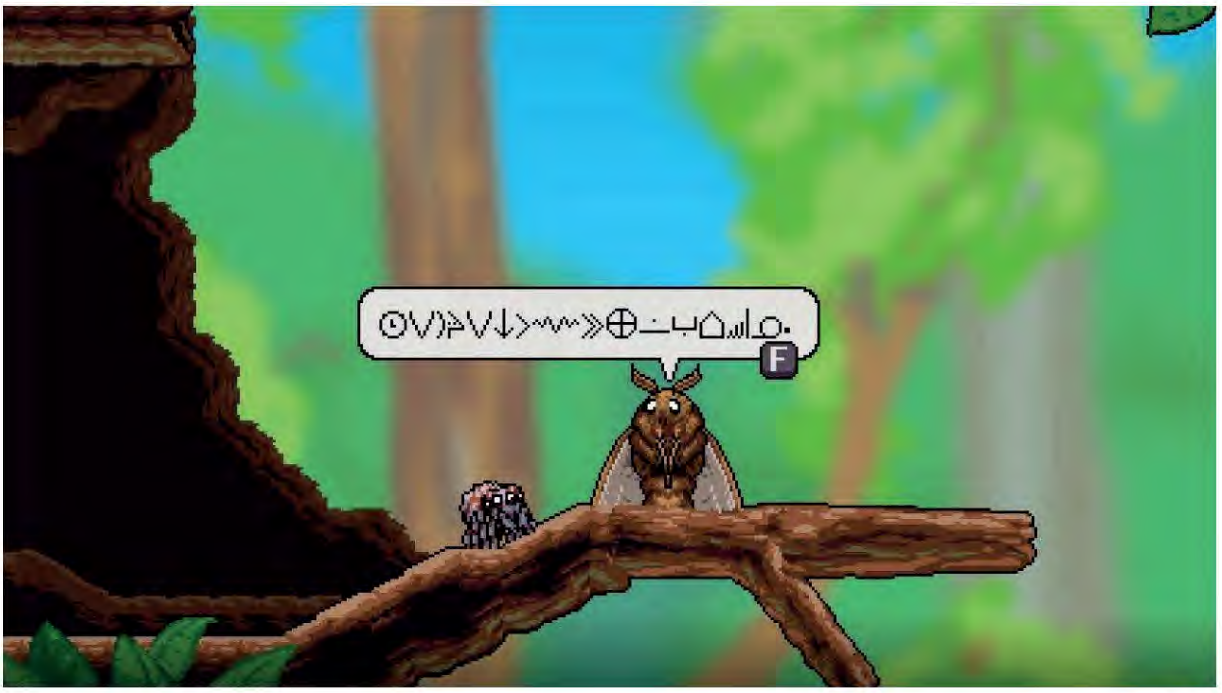


Figure 50. Screenshot of a translation mod for the game *Webbed*, translated by jan Popo (Sbug Games 2023).

Figure 51. Translations of old video games into toki pona in sitelen pona by jan Lentan (n.d.).

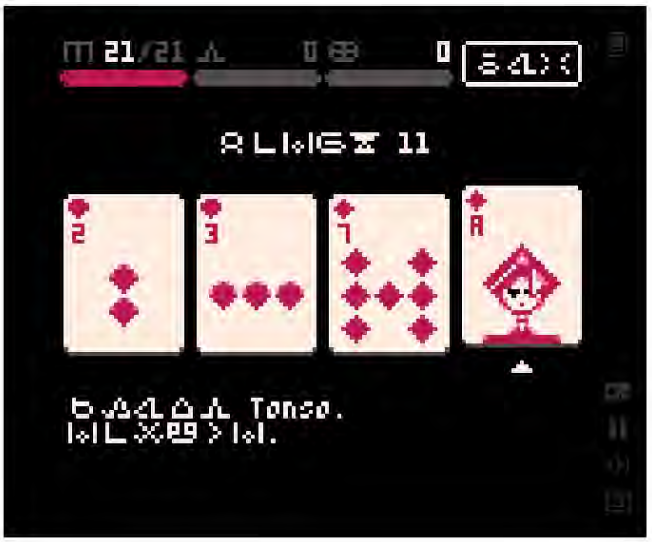


Figure 52. PICO-8 port of the game *Donsol* by John Eternal and Hundred Rabbits, translated by jan Lentan (2022).

Figure 53. Screenshot of the original game *Late for School*, created by Billsmugs (2022b).

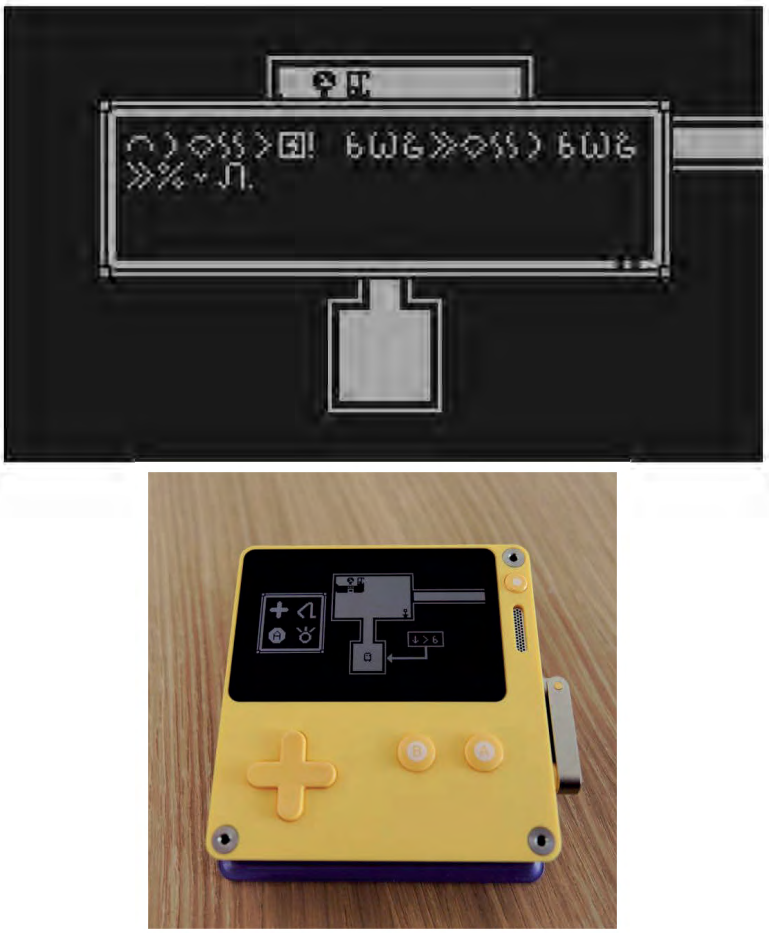


Figure 54. Screenshot of the original game *Kiwen Kon*, created by Billsmugs (2023), played on a computer and on the Playdate handheld video game console.

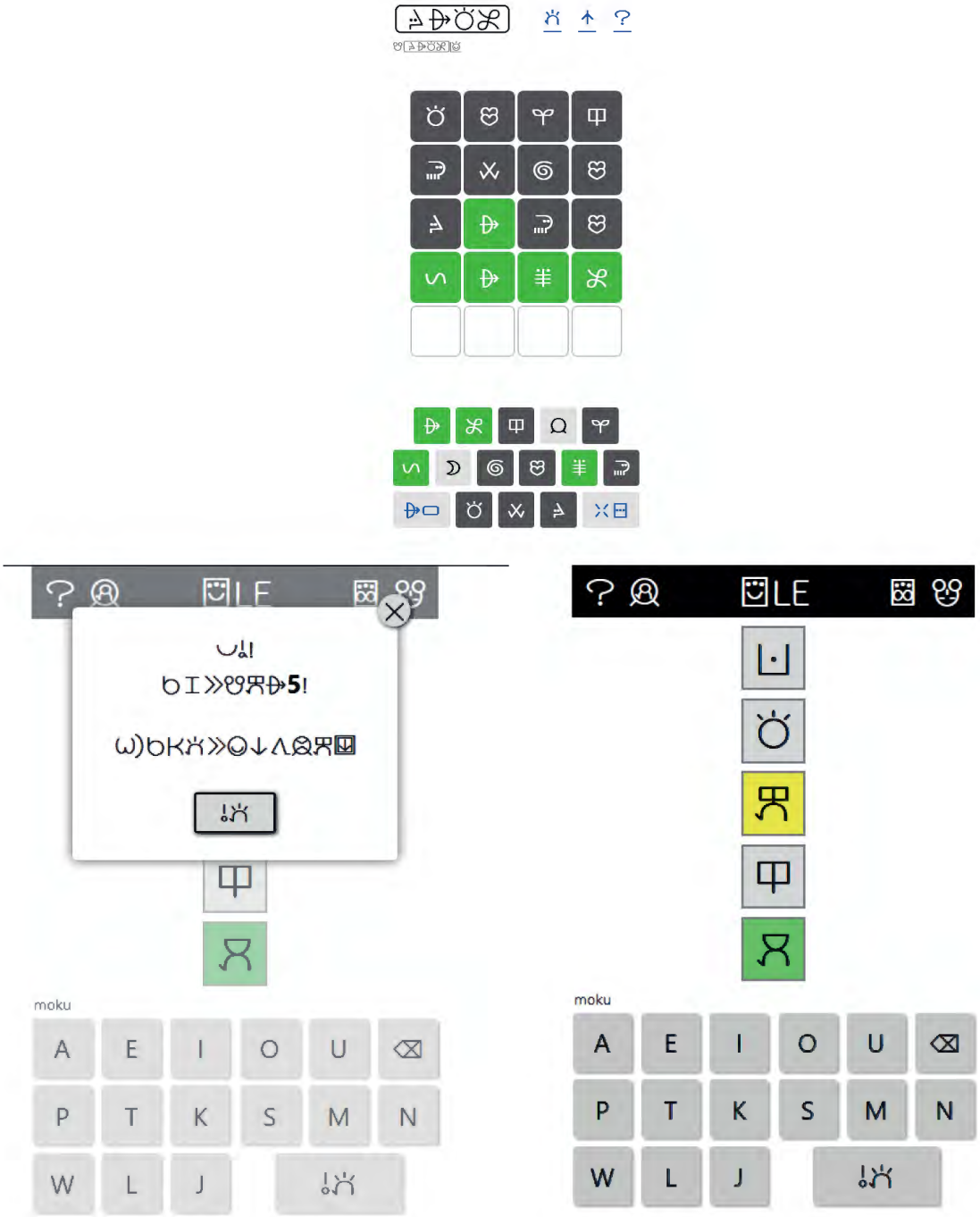


Figure 55. Clones of the web-based word game *Wordle* using sitelen pona, translated by (top) Maciej Goszczycki (2022) and (bottom) jan Jontan (2022).

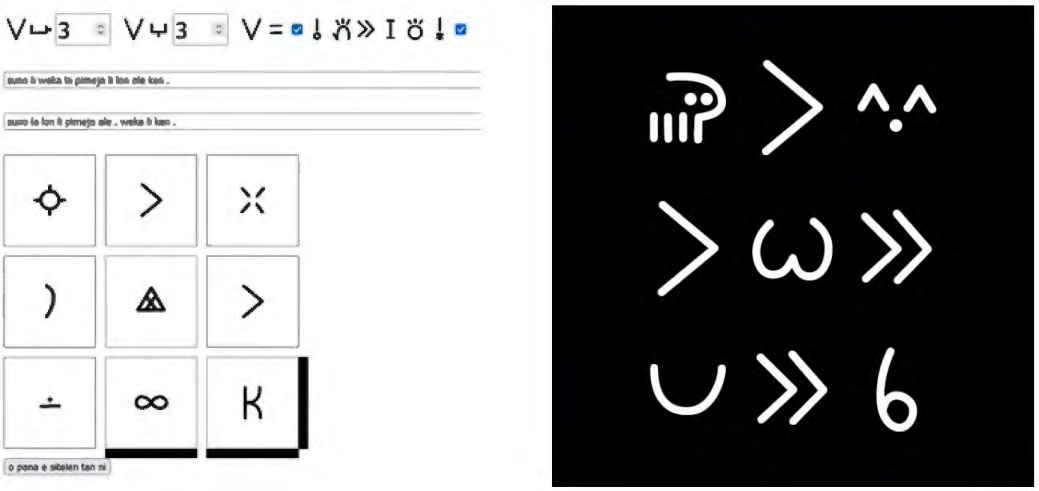


Figure 56. A generator for *leko nimi*, a form of toki pona poetry that can be read both horizontally and vertically, by Ret Samys (2023c, d).

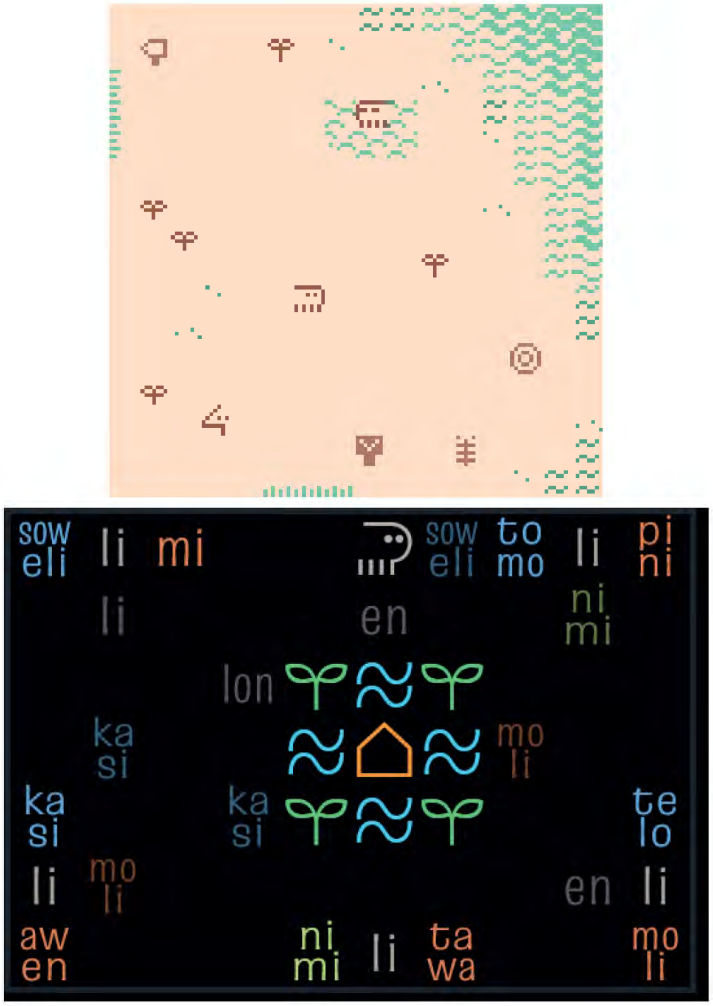


Figure 57. Original toki pona games featuring sitelen pona for artistic purposes: (left) *Soweli Mu*, by waso Talu (2022), and (right) *soweli li mi*, by formicant (2021).

Figure 58. Screenshots of the translation of the virtual reality game *Wavelength* into sitelen pona, by jan Toluneko (2024).

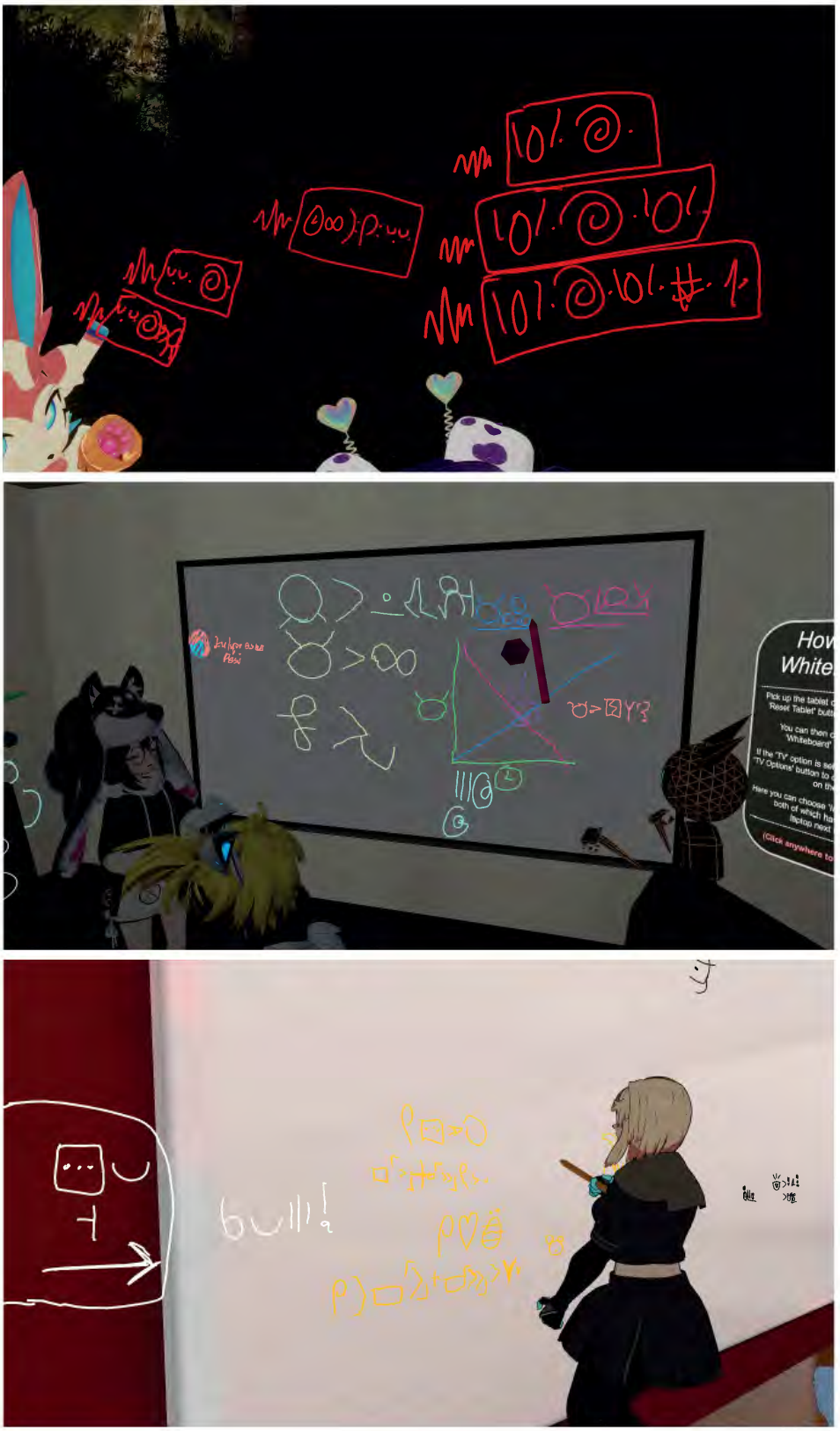


Figure 59. Screenshots from meetups between tokiponists using sitelen pona in *VRChat*.

Figure 59. Screenshots from meetups between tokiponists using sitelen pona in *VRChat*.

Figure 60. Brazilian game *Mun Ante* using sitelen pona as a narrative strategy to express dialogue between two speakers of different, unintelligible languages (Silva 2023).

Figure 61. Music video for the translation of the song “Caramelldansen” by Caramell (2021), translated and illustrated by jan Mowima, with subtitles in both sitelen pona and the Latin script.

Figure 62. Music video for the translation of the song “Bad Apple!!” by Jun’ya Ōta (2021), translated by jan Tepo, subtitled by jan Seme in sitelen pona and the Latin script.

Figure 63. Song translations in toki pona by Ye Siyuan (2023a, b) uploaded to the Chinese video streaming service, Bilibili.

Figure 64. Music video for the song *olin nasa tawa jan*, translation of "Hitomania" into toki pona by jan Sikun, with sitelen pona subtitles (Haraguchi 2023).

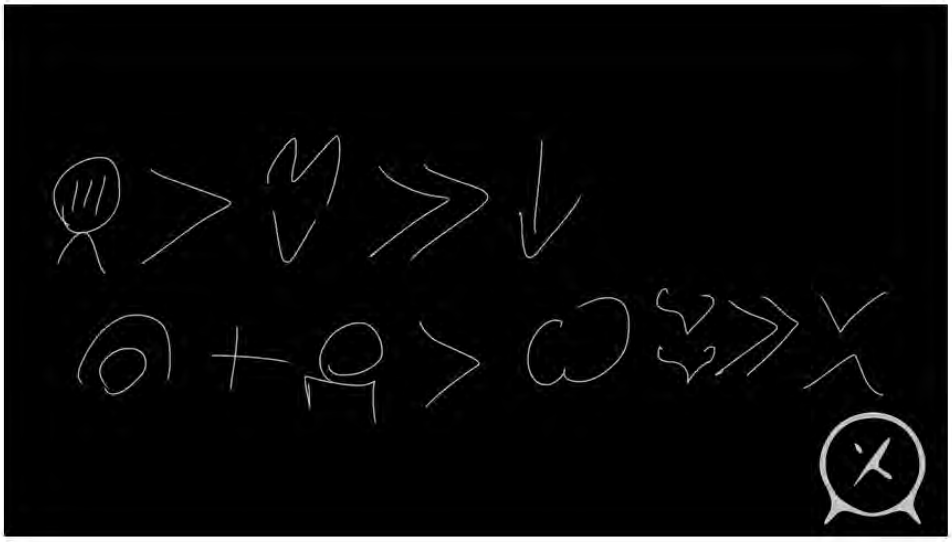


Figure 65. Music video for the original song "olin li tawa jan ale kin" by jan Sotan (Zotn).

Figure 66. Collection of cover art for original toki pona songs. From left to right: “o tawa noka” by kala pona Tonyu (2020), “ale li ante” by jan Nikoleta (2020), “mi pilin e ni” by Nardi (2021), “oww” by jan Usawi (2023a), and “mi moku e kiwen” by jan Sepulon (2023a).

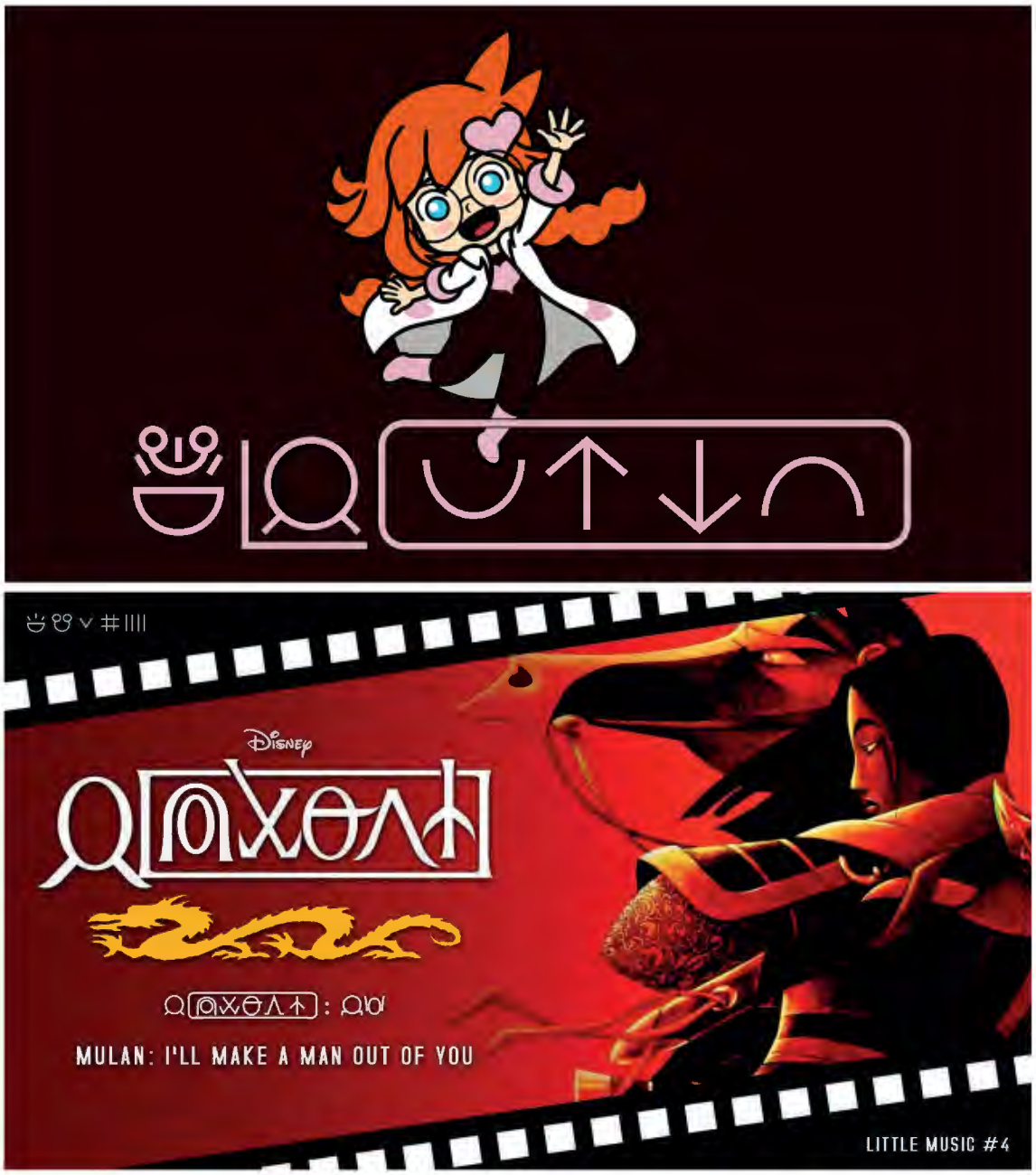


Figure 67. Collection of cover art for translated toki pona songs. From top to bottom: "kalama musi pi jan Peni" (original "Penny's Song"), translated by Mitch Halley (2021) and "mije wawa" (original "I'll Make A Man Out Of You"), translated by James Flear (2020).

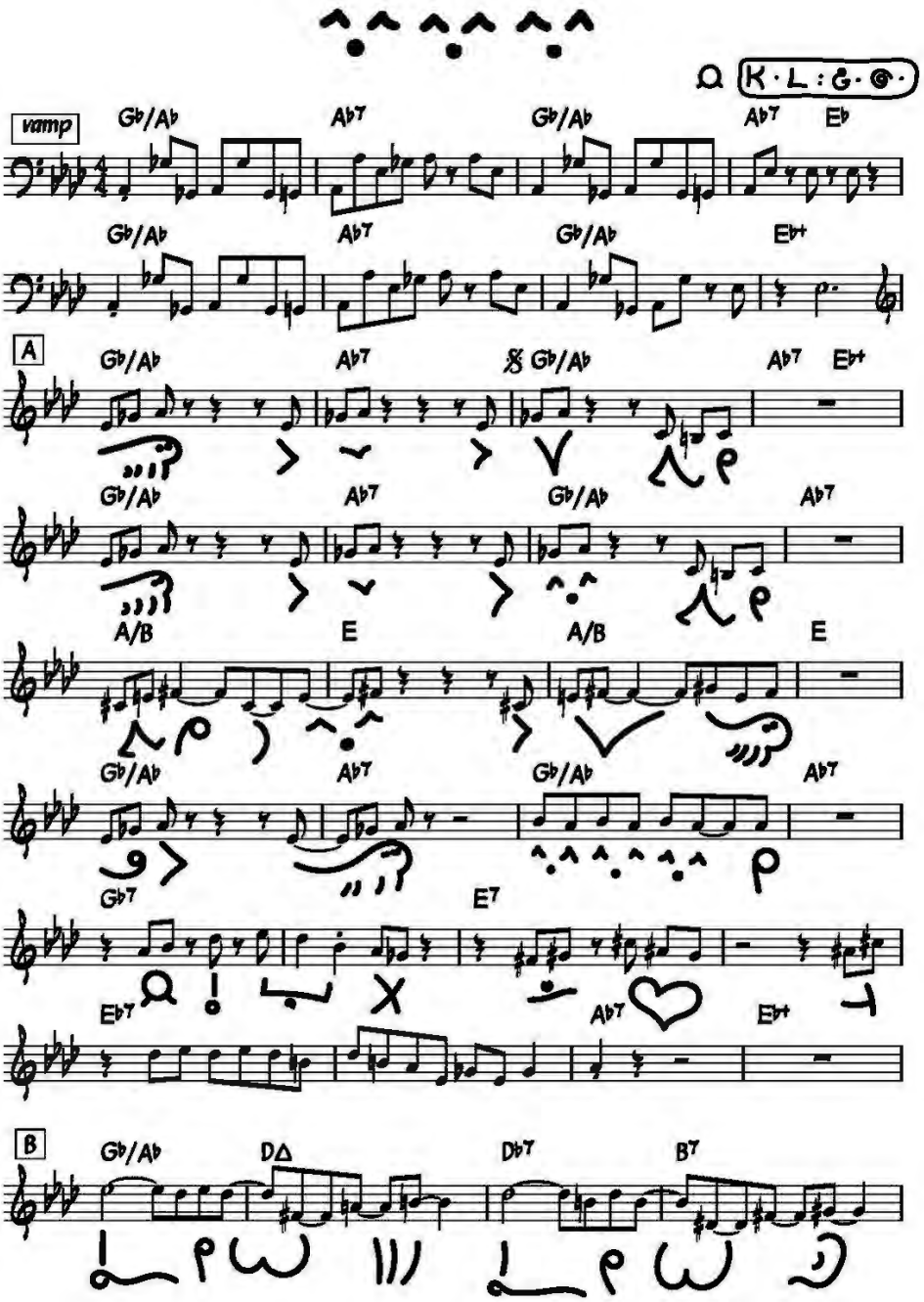


Figure 68. Sheet music for the song “suwi suwi suwi” by jan Kepijona, adapted into sitelen pona by Ret Samys (2024).

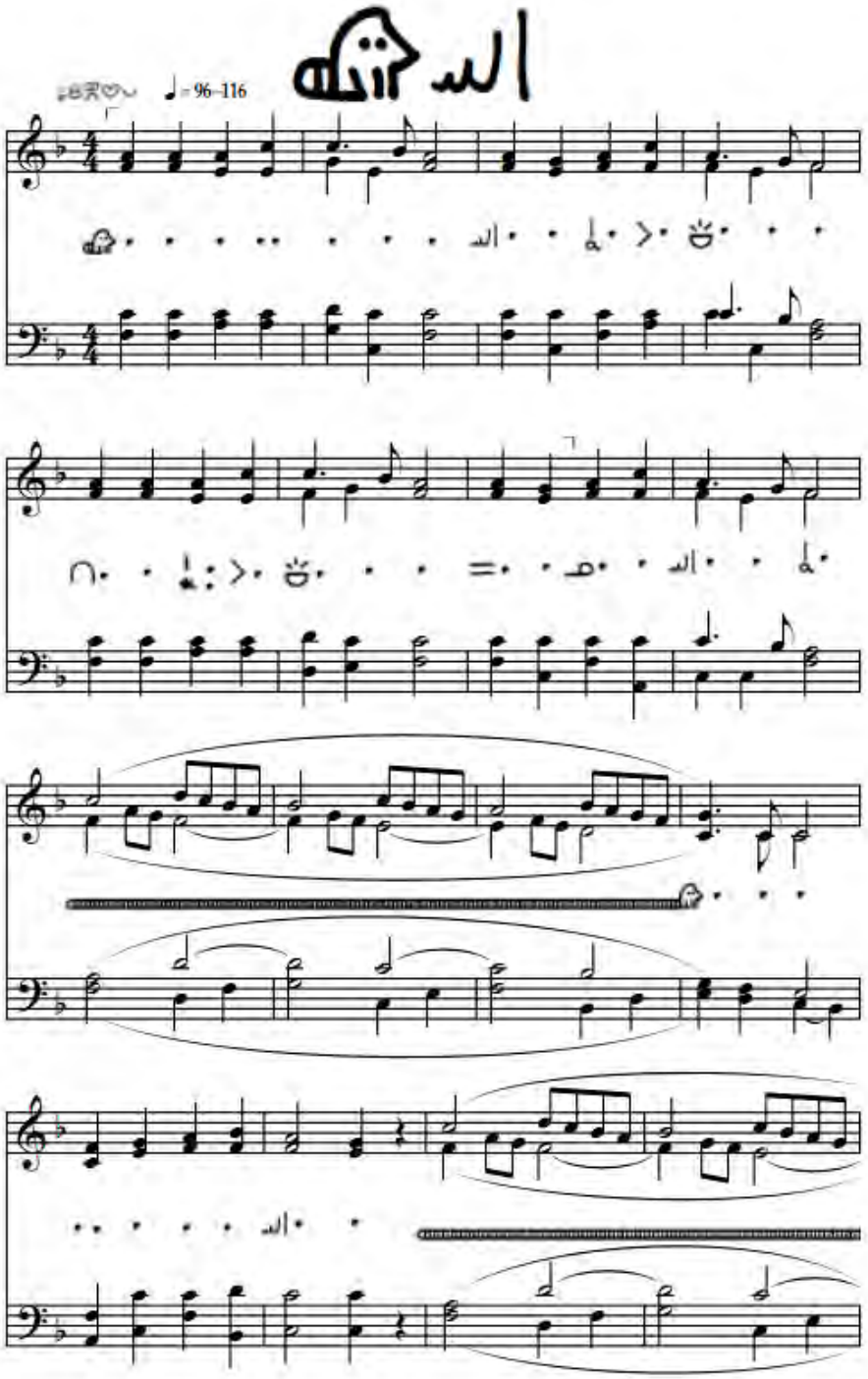


Figure 69. Sheet music for “kijetesantakalu sewi,” notated by jan lili Enta and jan Salo (2024),
using *nasin sitelen kalama* to indicate syllables, as well as extended glyphs for slurs.

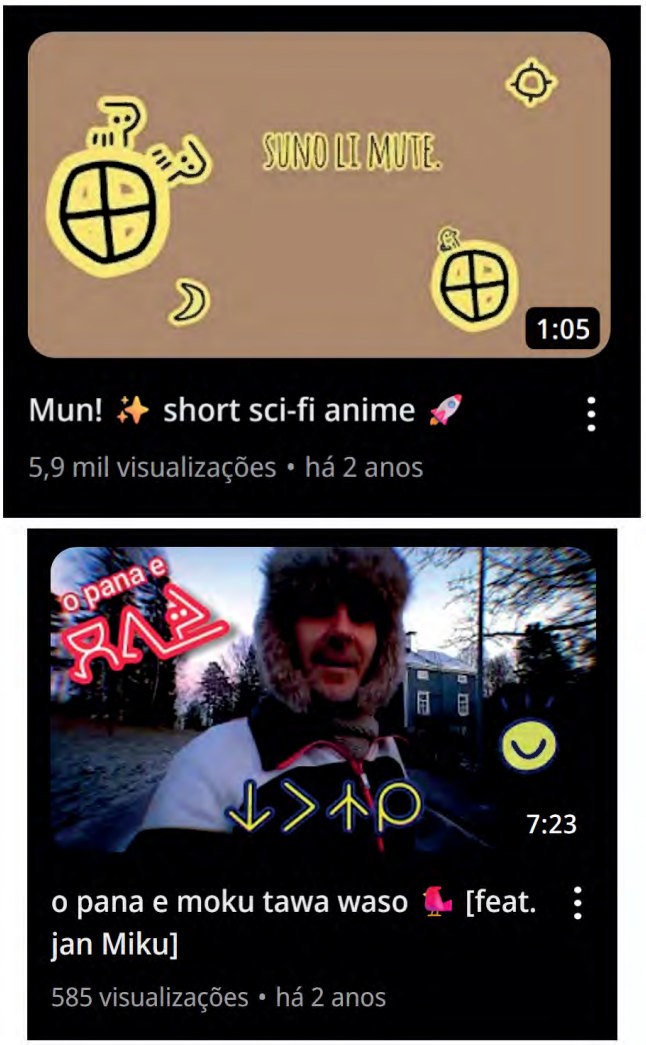


Figure 70. Thumbnails from YouTube videos by Florian Cauvin (2021. 2022) in sitelen pona.

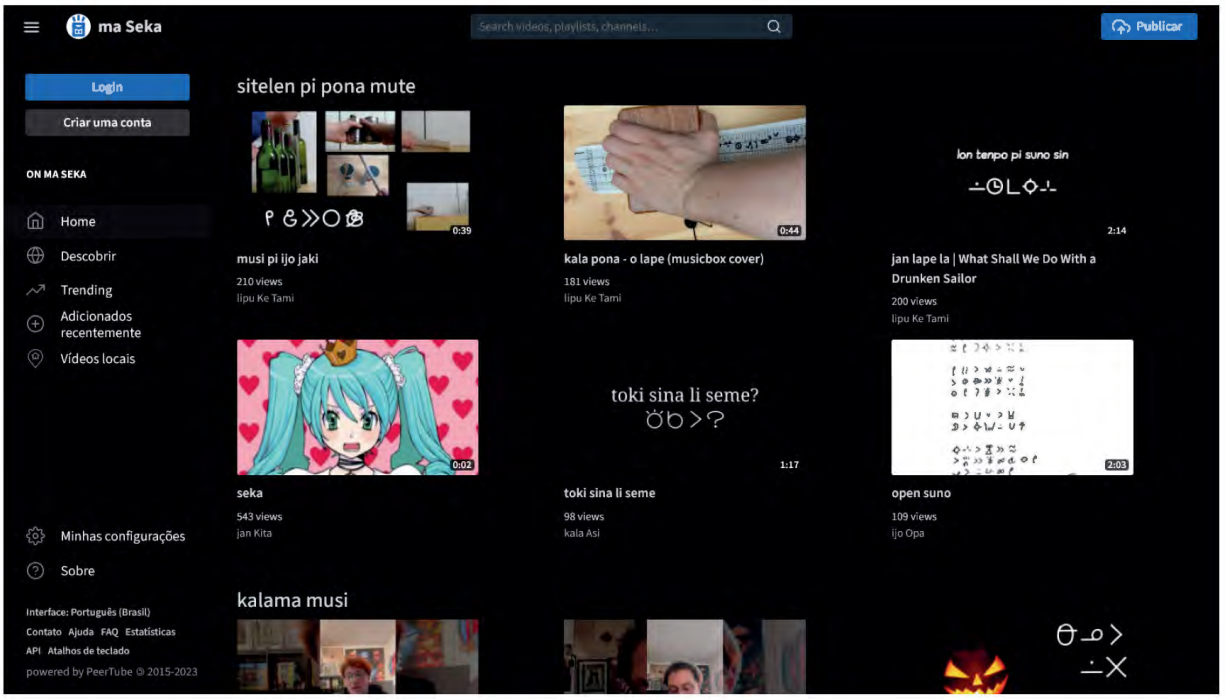


Figure 71. Thumbnails from the video sharing platform and PeerTube instance, *ma Seka*.

Figure 72. Screenshot from the movie *Back to the Future* (1985) with sitelen pona subtitles, translated and subtitled by James Flear (2021).

Figure 73. Screenshots from *The Simpsons* episode "Lisa On Ice" (1994) with sitelen pona subtitles, translated by Rebecca Bettencourt (2021).

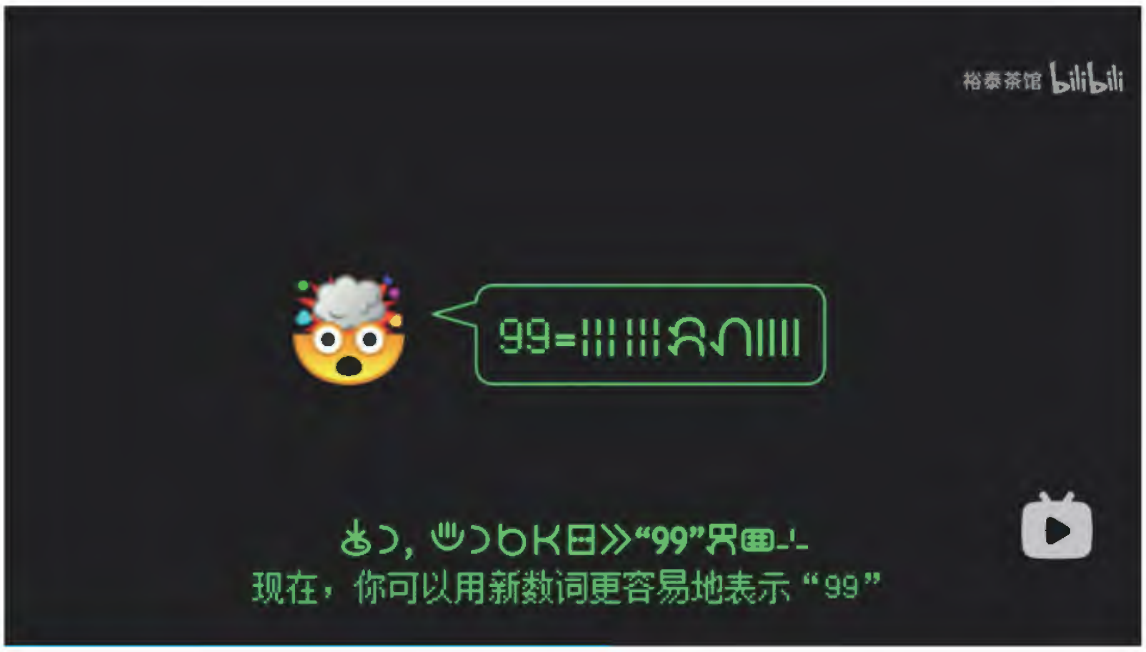


Figure 74. Introductory video on numbers in Toki Pona in both sitelen pona and Mandarin Chinese, by Bilibili user Yutai Chaguan (裕泰茶馆, 2021).

Figure 75. Logos and icons incorporating sitelen pona. From left to right: suno pi toki pona, ma Seka, lipu tenpo, Linku, ma pona pi toki pona, and Ma Toki Pona VR.
(The logos themselves are **not** being proposed.)

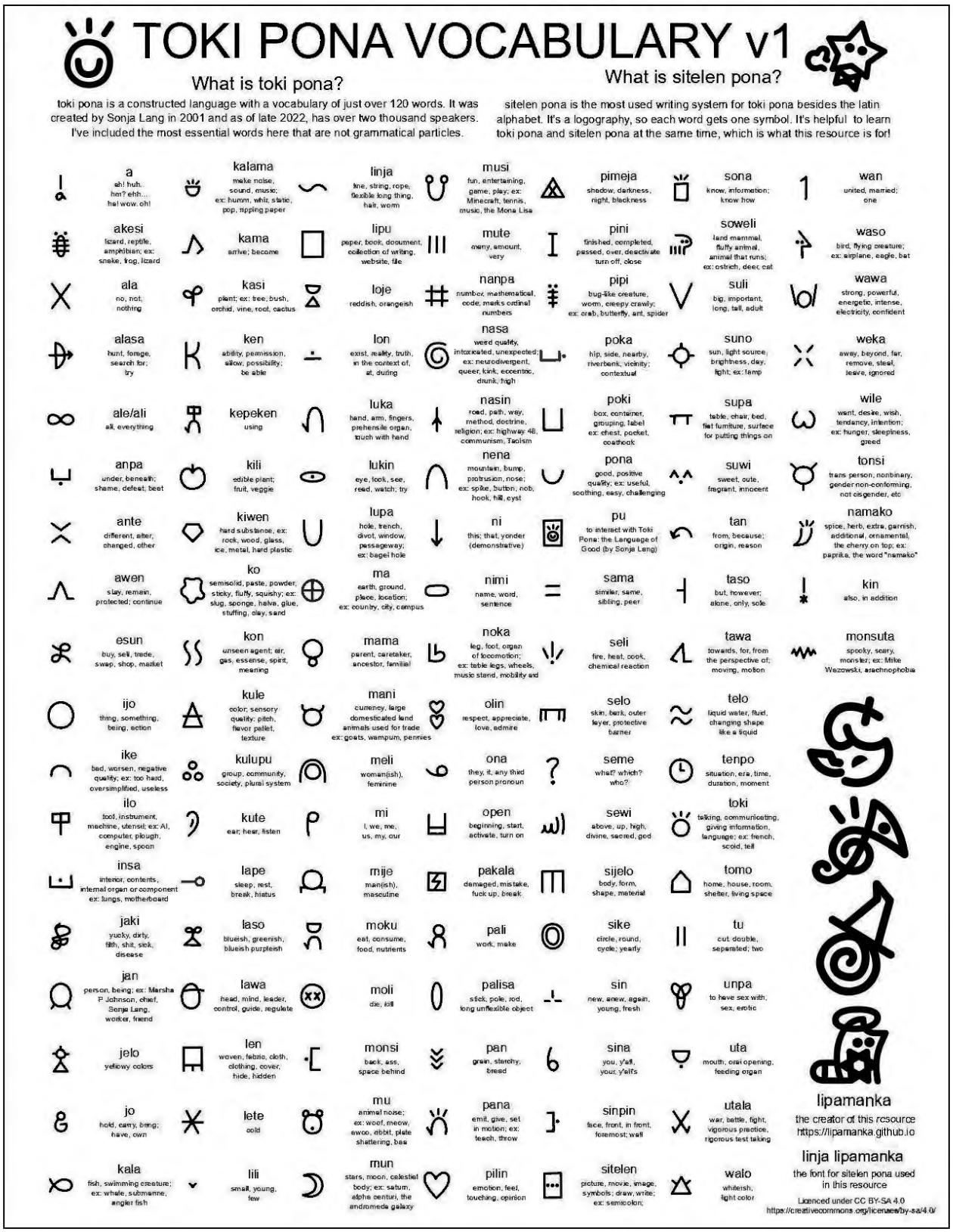


Figure 76. Cheatsheet with toki pona vocabulary by lipamanka (2023b).

Figure 77. Screenshot of a virtual toki pona lesson by Gregory Danielson III (2022a).

Objects with e

Home Blog Lessons Now Dev Work

Check out the video lesson too!

▸ Table of Contents

Words to Know

Particles

Word	Use	Writing
» e	(mark an object; begin an object phrase)	Two right-facing arrows

Content Words

Word	Definition	Writing
wile	want, need, to want	Curvy letter "w"
lipu	paper, card, document, website, leaf, small flat thing	Tall rectangle like a piece of paper
sona	knowledge, understanding, to know	Piece of paper with emitters!
	eyes, to look, to try	An eye with a pupil in the center

Toggle SP

Figure 78. Screenshot from Danielson's (2022b) online learning course.

Figure 79. Screenshot from March 2024 of soweli Tesa's online learning course.

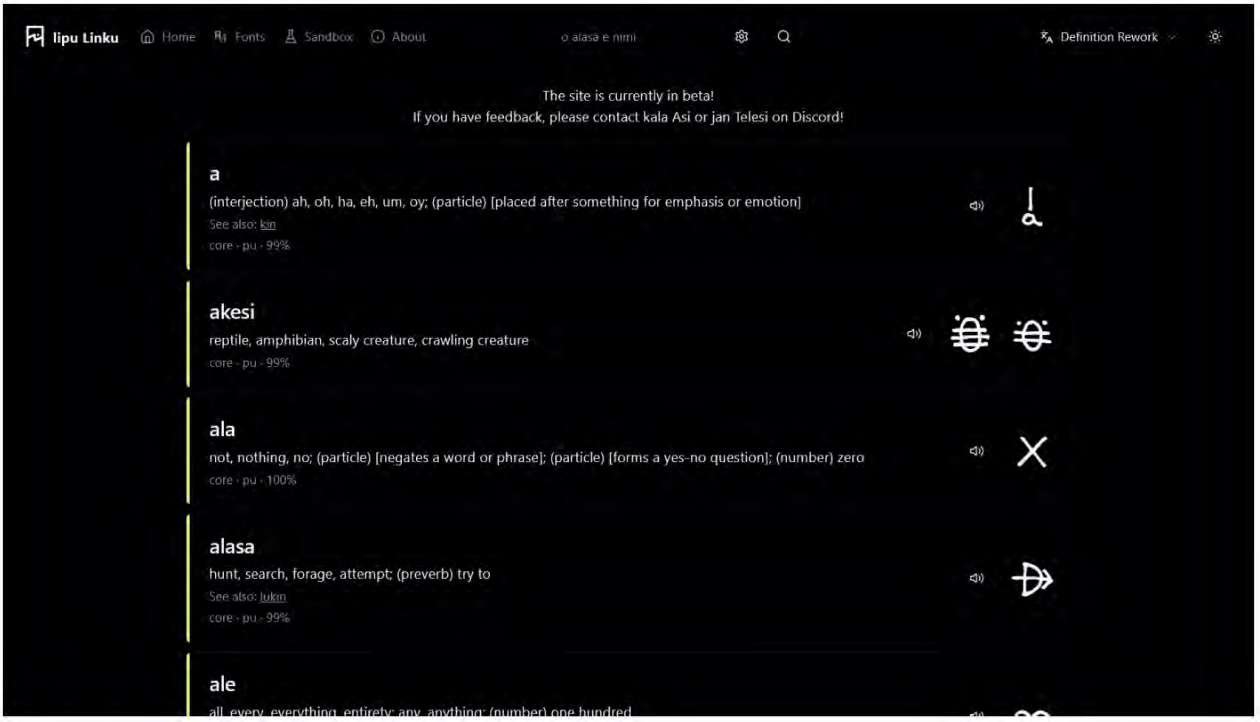


Figure 80. Screenshot from April 2024 of *lipu Linku*, an online toki pona dictionary.

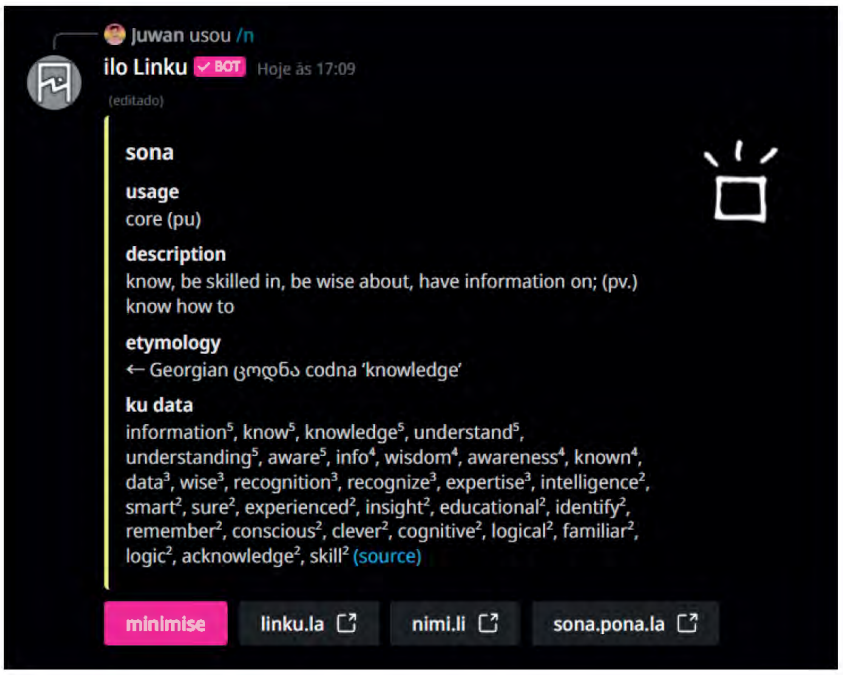


Figure 81. Screenshot from February 2024 of *ilo Linku*, a Discord bot which uses Linku data.

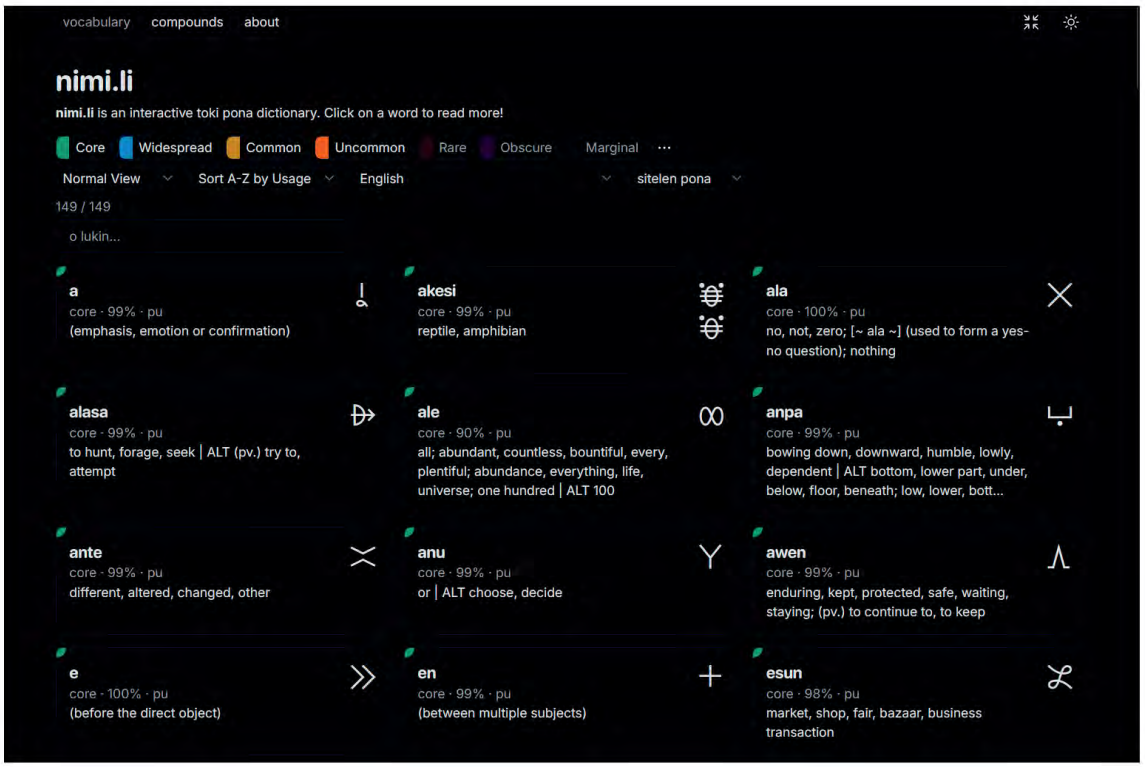


Figure 82. Screenshot from January 2024 of *nimi.li*,
a frontend for Linku data by Daniel Huang (jan Tani).

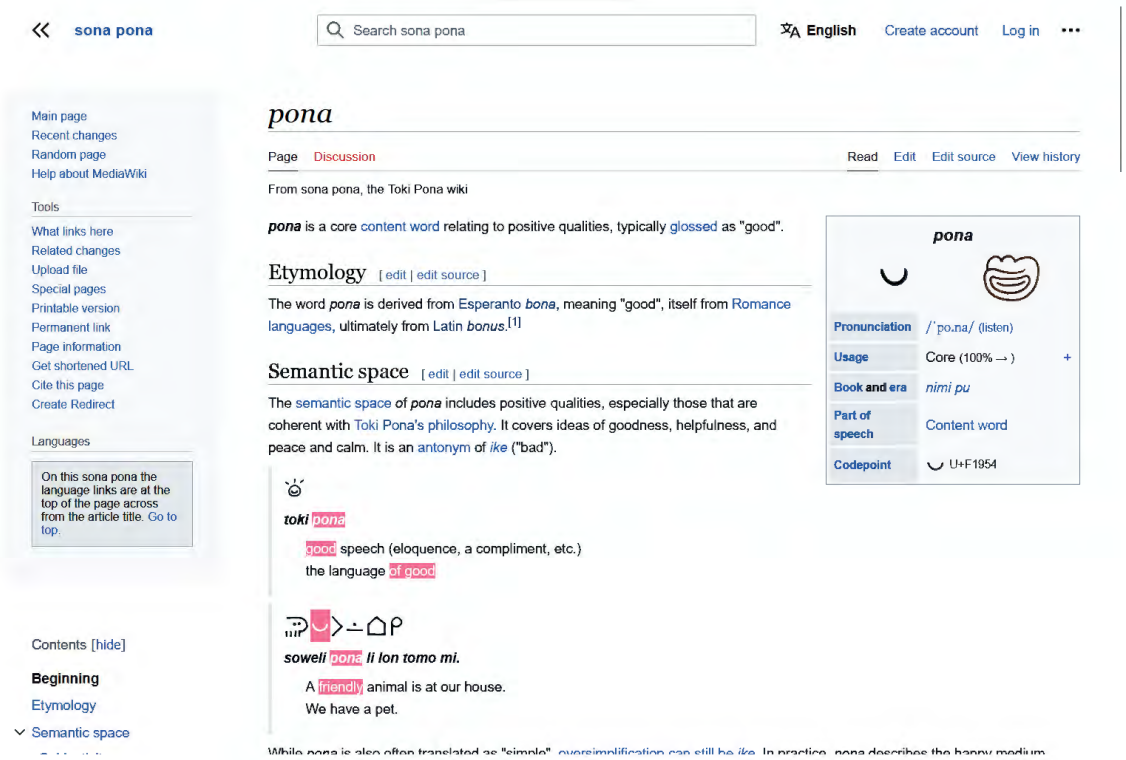


Figure 83. Screenshot from January 2024 of *sona pona*, a wiki dedicated to toki pona, with example sentences given in both the Latin script and sitelen pona.

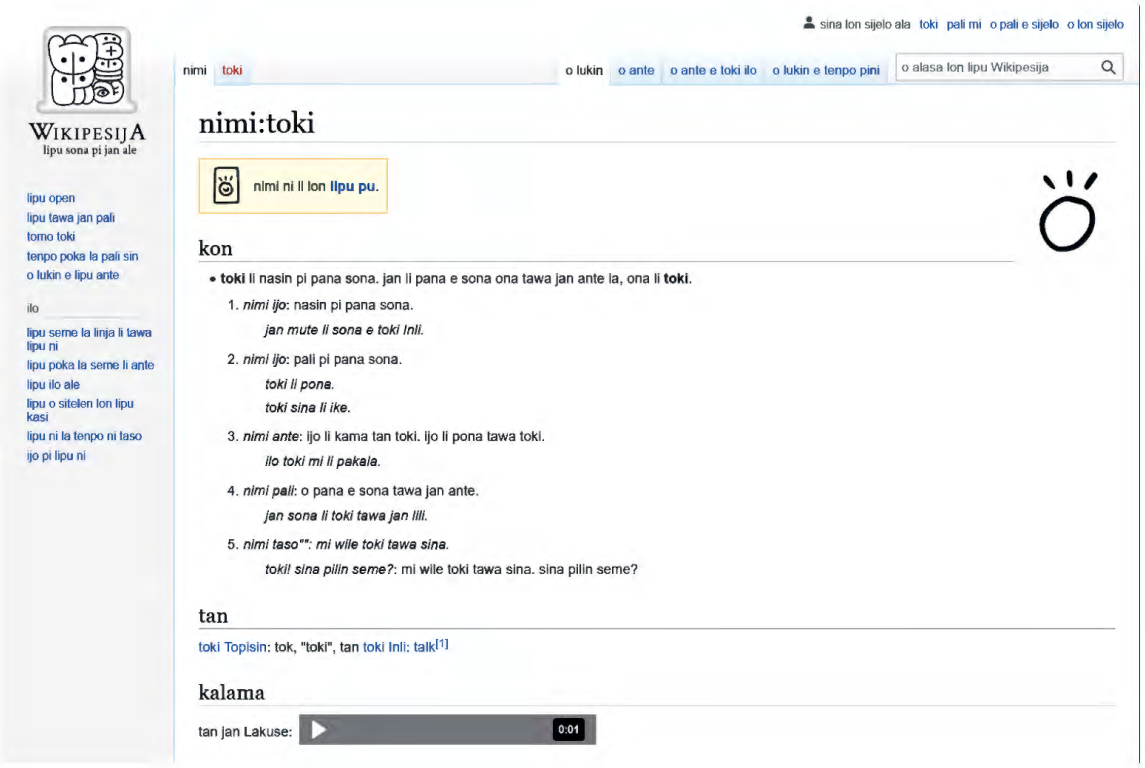


Figure 84. Screenshot from January 2024 of a dictionary entry on *lipu Wikipesija*, an unofficial online encyclopedia project.

Figure 85. Screenshot from January 2024 of *lipu Wikipesija*, using an extension to convert Latin-script text to sitelen pona by jan Lesa (zendelexa).

Figure 86. Screenshot from January 2024 of the Wiktionary entry for the toki pona word *soweli*.

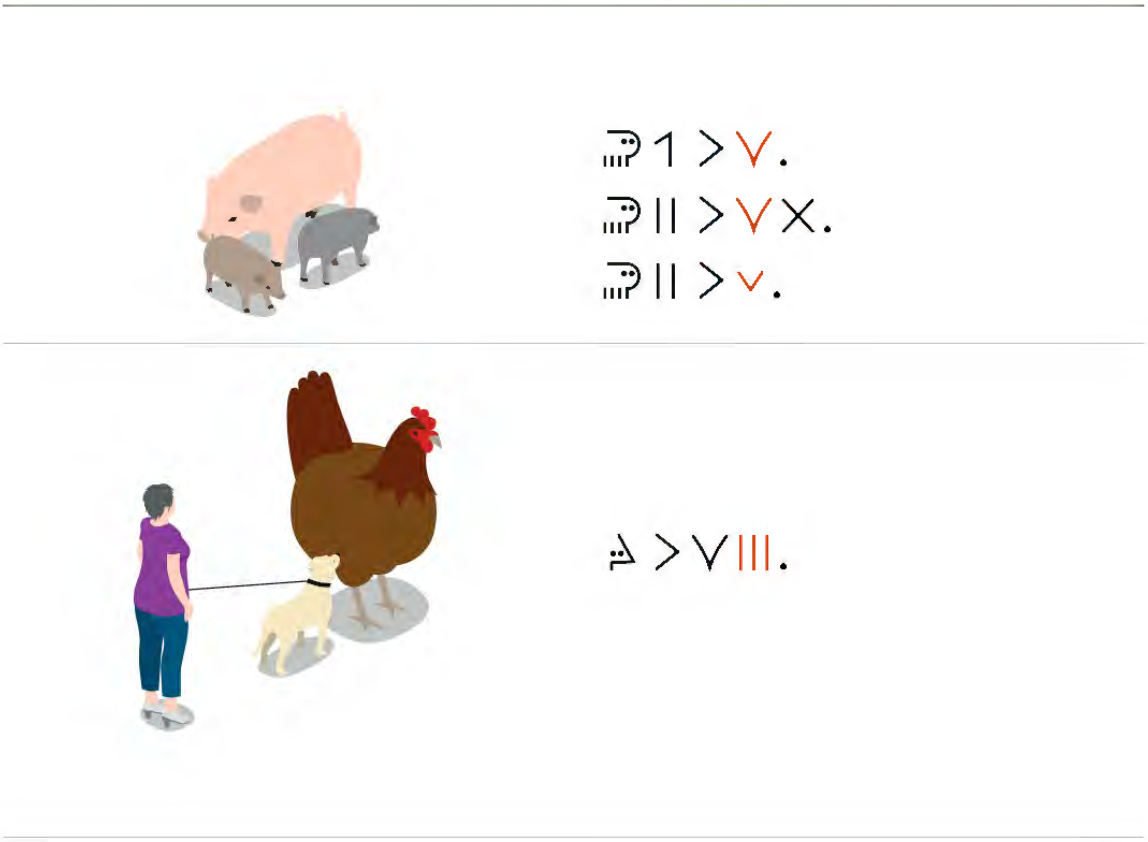


Figure 87. Learning resource for sitelen pona by David A. Roberts (2018b).

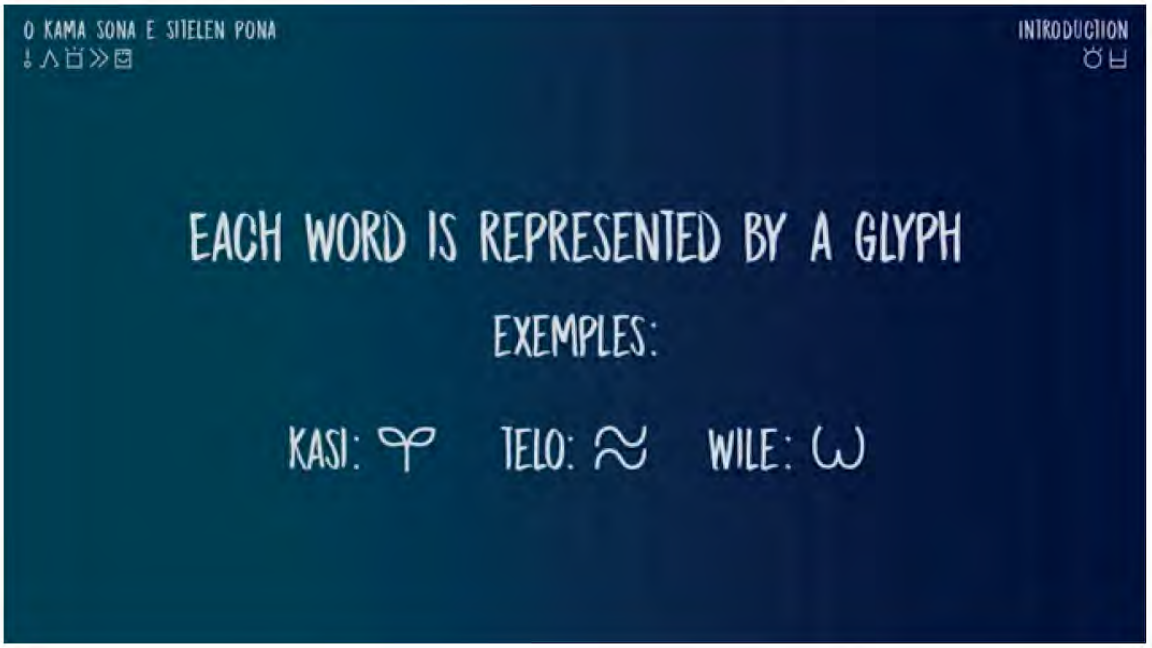


Figure 88. Learning resource for sitelen pona by James Flear (n.d-c).

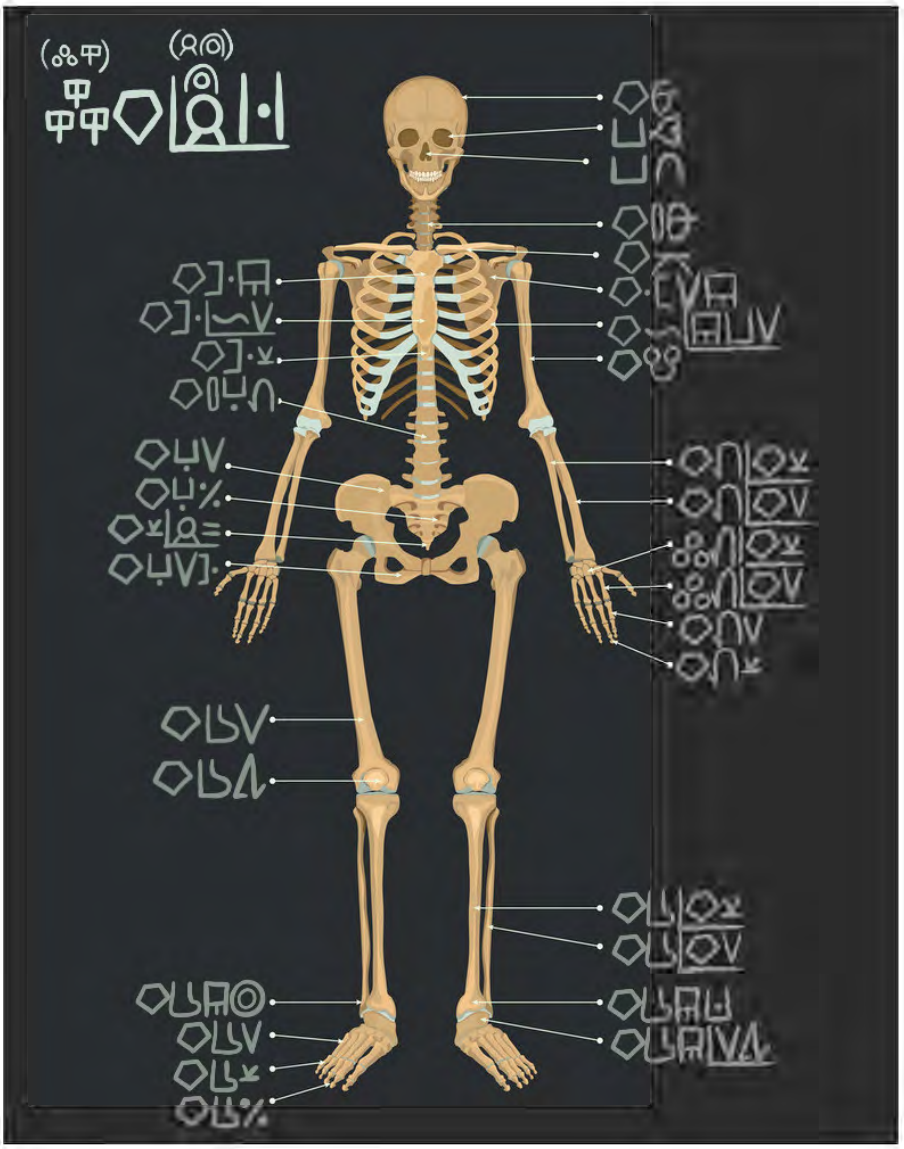
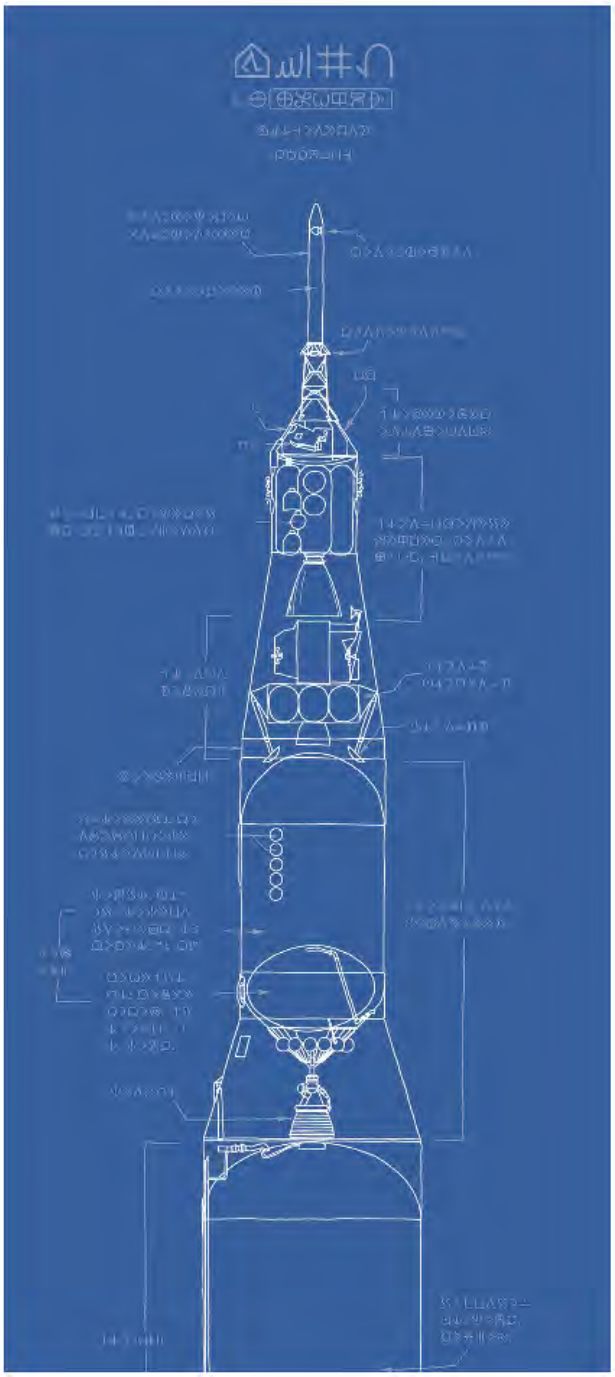


Figure 89. Diagram of a skeleton of the average human female, translated by jan Majeka (2023).



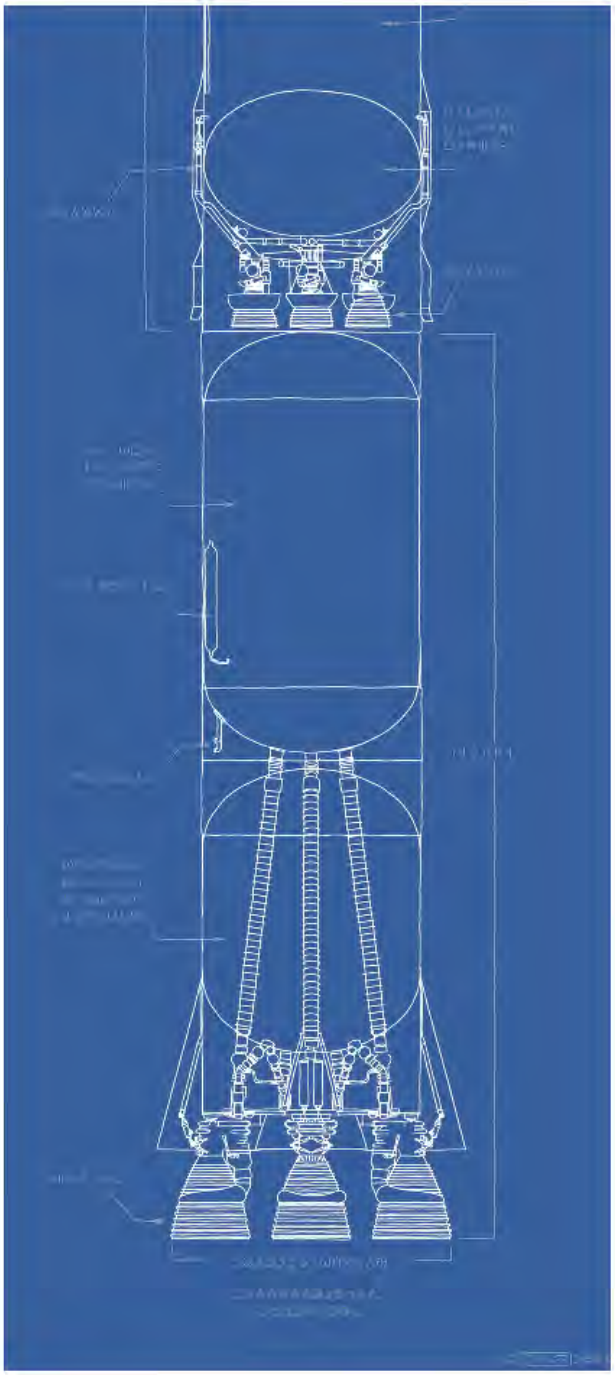


Figure 90. Rocket blueprint from *xkcd* #1133: Up Goer Five (spread over two images), translated by David A. Roberts (Munroe 2018).

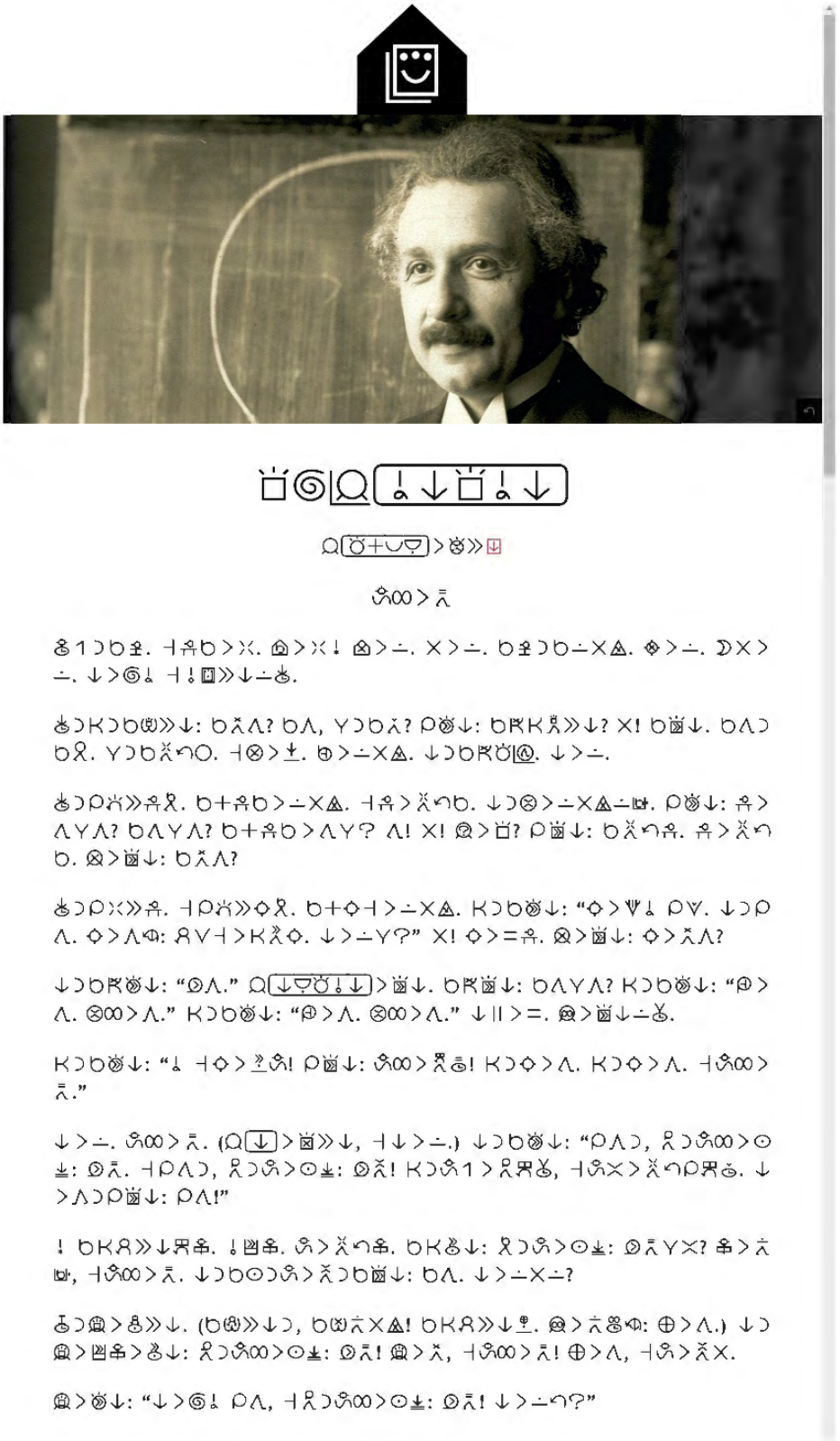


Figure 91. The article “Albert Einstein’s Theory of Relativity: In Words of Four Letters or Less,” written by Brian Raiter, translated by David A. Roberts (Raiter 2018).

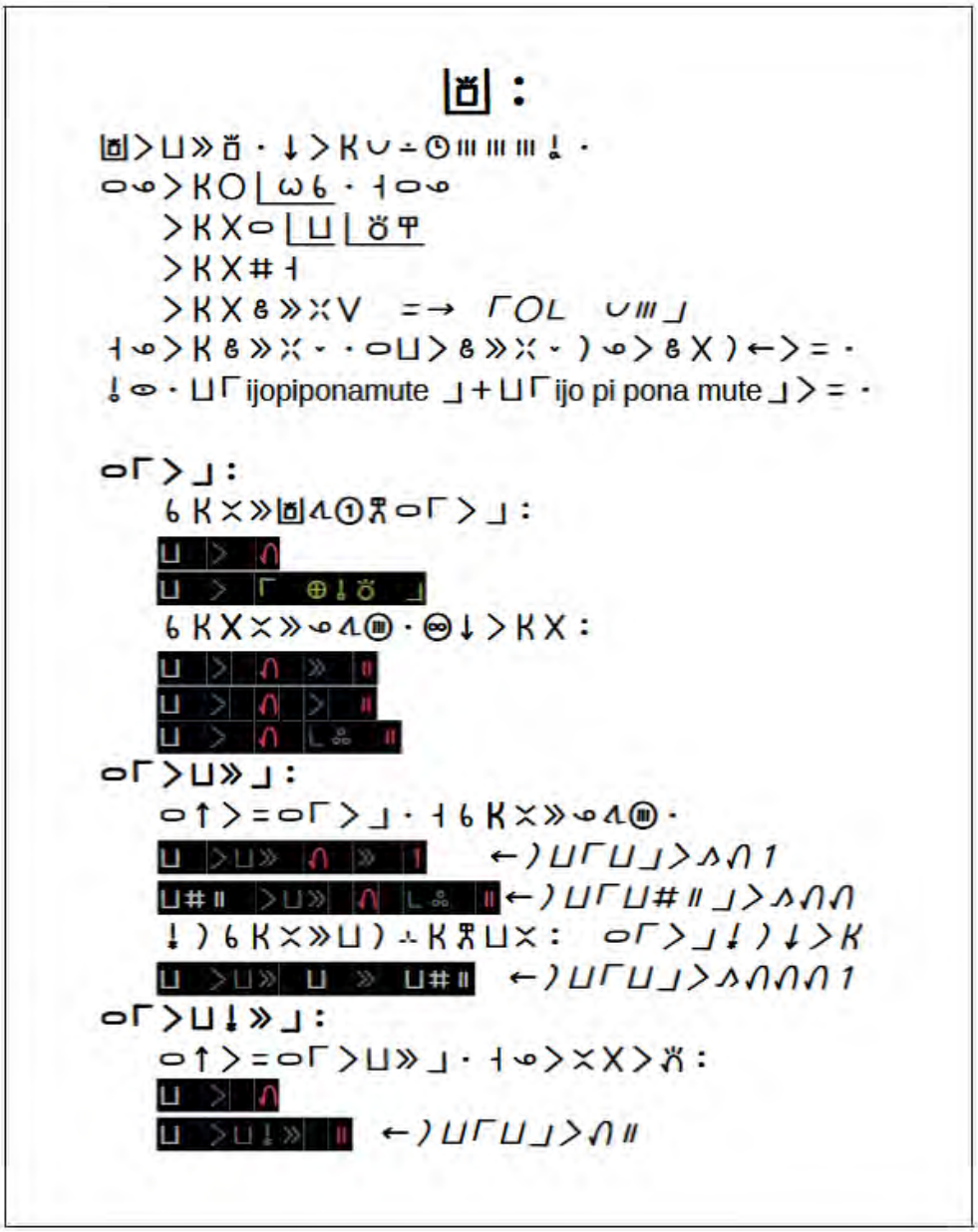


Figure 92. Instruction manual for the programming language *toki ilo pona*, created by jan Popo (2023).

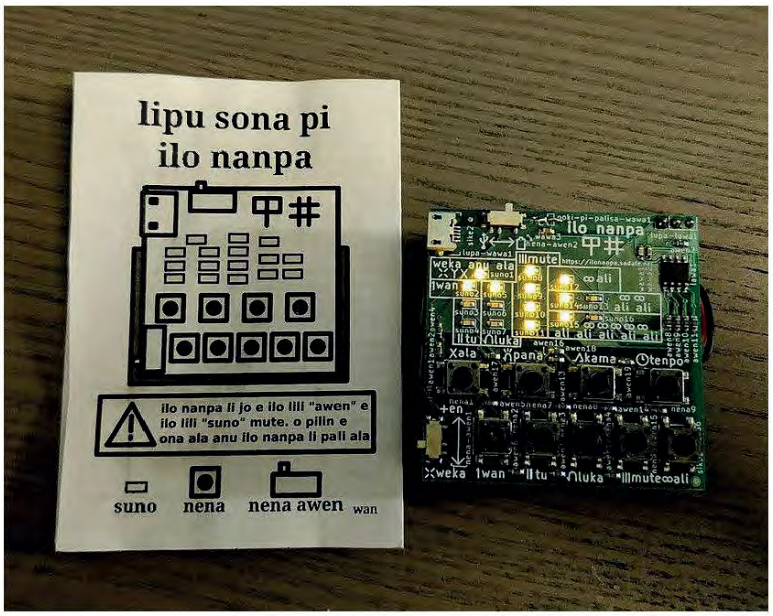


Figure 93. Ilo Nanpa, a calculator with toki pona numerals by Wong Cho Ching (2019).

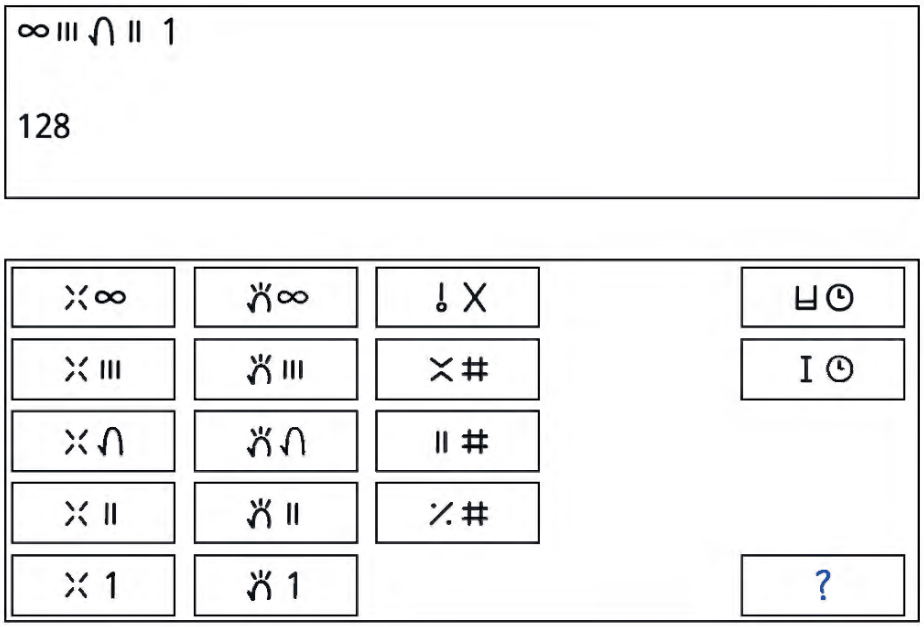


Figure 94. *ilo nanpa pi jan Tolonpon*, a web-based calculator by jan Tolonpon (2024).

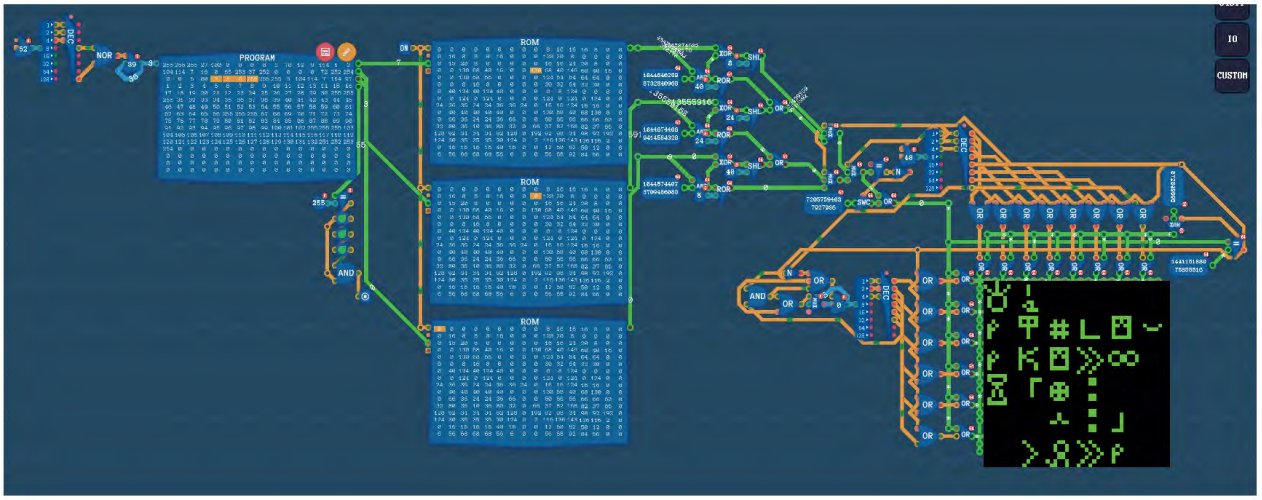


Figure 95. Machine which displays sitelen pona logographs designed by loje Masin (2024).

Figure 96. Screenshot from an online lecture by Gregory Danielson III (2023), discussing the hypothesis that humans are living inside a computer simulation.

Figure 97. Screenshot from a presentation by Nia W. (2023) about the fediverse.

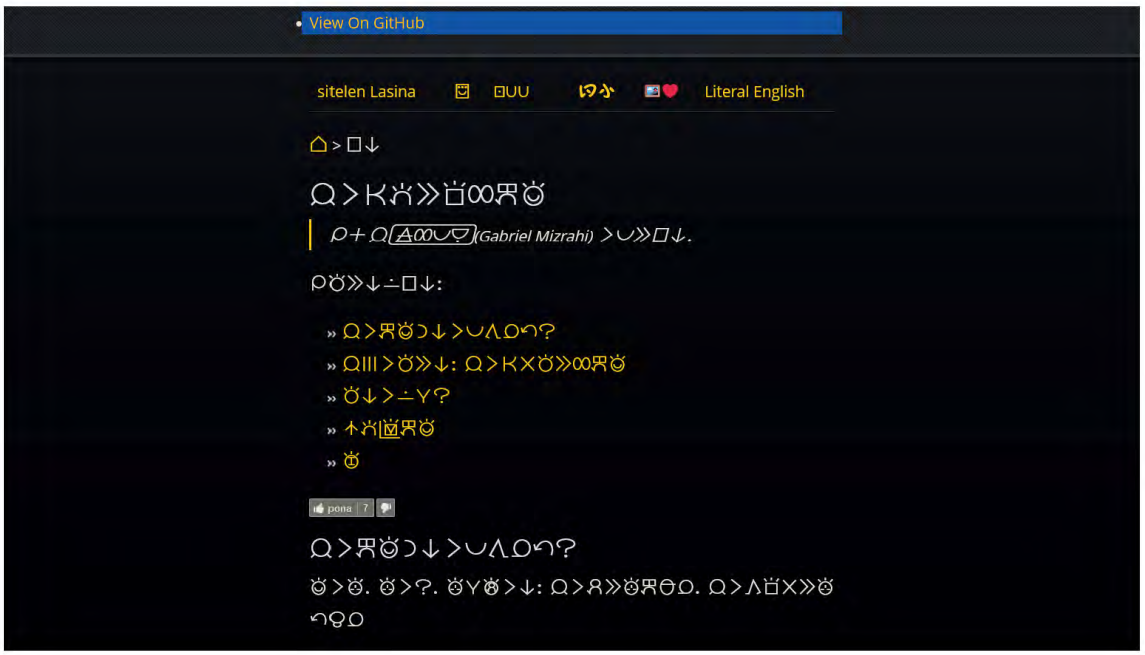


Figure 98. Blog post written by Joel Thomas and Gabriel Mizrahi (2020), available in many different writing systems for toki pona, including sitelen pona.

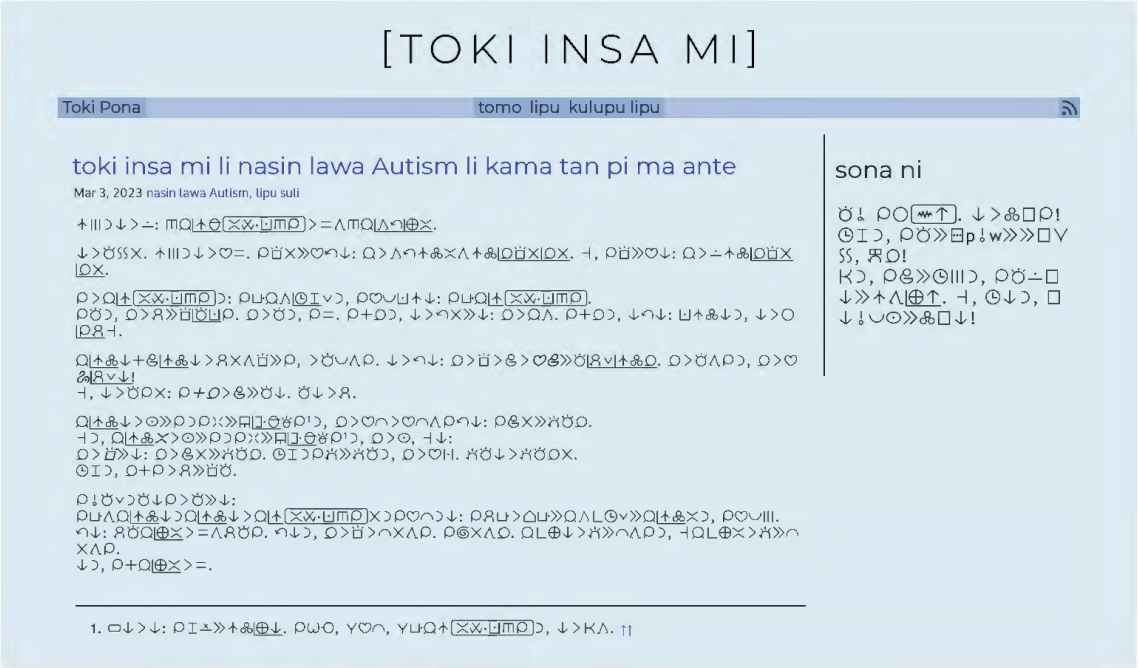


Figure 99. Blog posts from *toki insa mi*, written by ijo Me (2023).

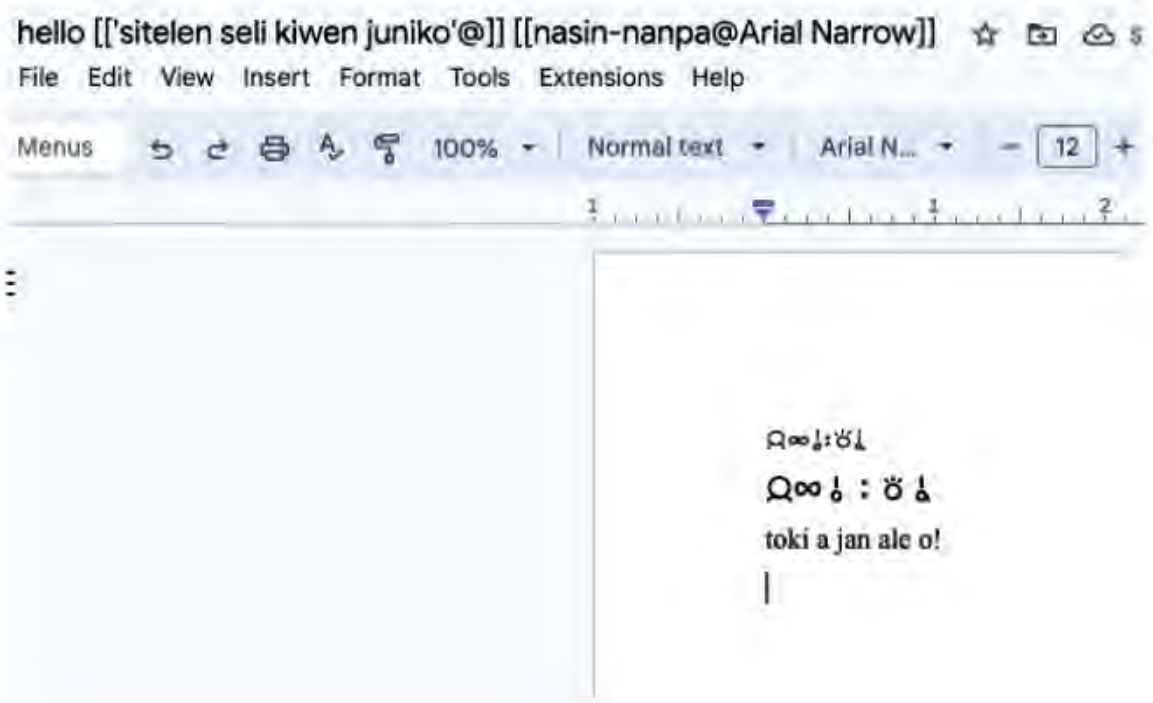


Figure 100. Screenshot of the browser extension “sitelen pona for Google Workspace” by jan Sakawi (2024), which allows speakers to use sitelen pona fonts in Google products.

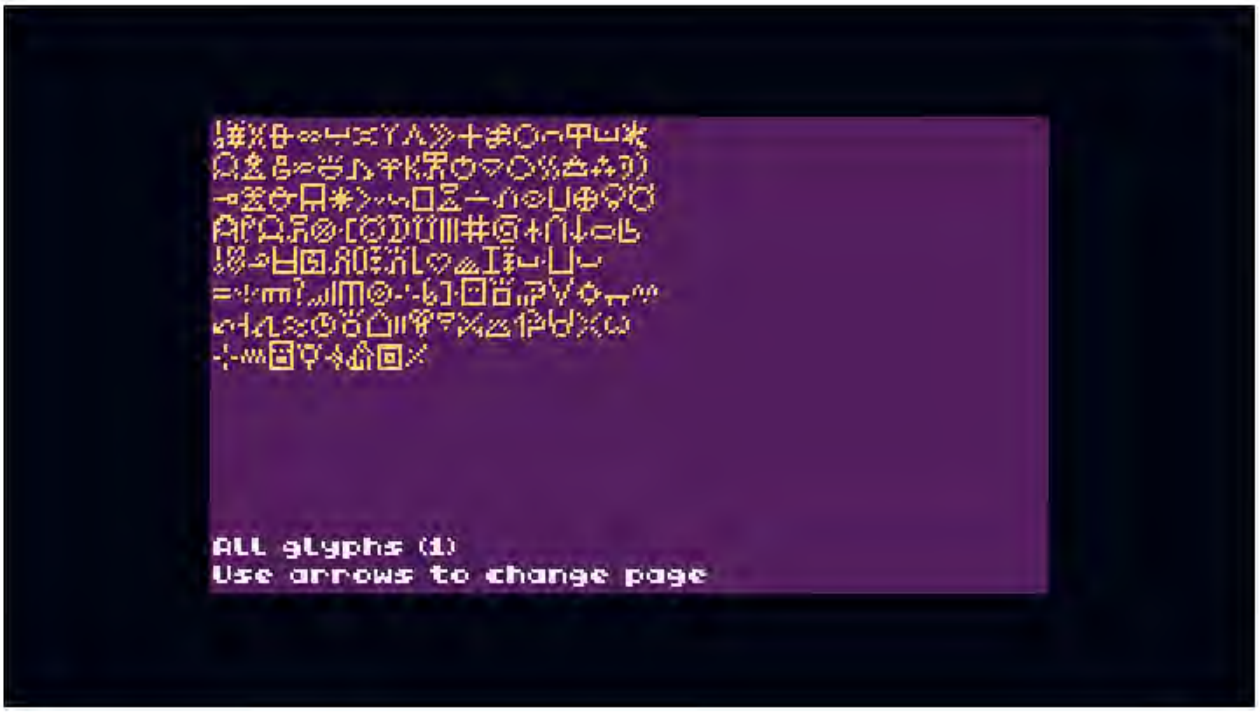


Figure 101. Sitelen Pona Writer, a simple text editor for the TIC-80 platform, by Podepi (2021).

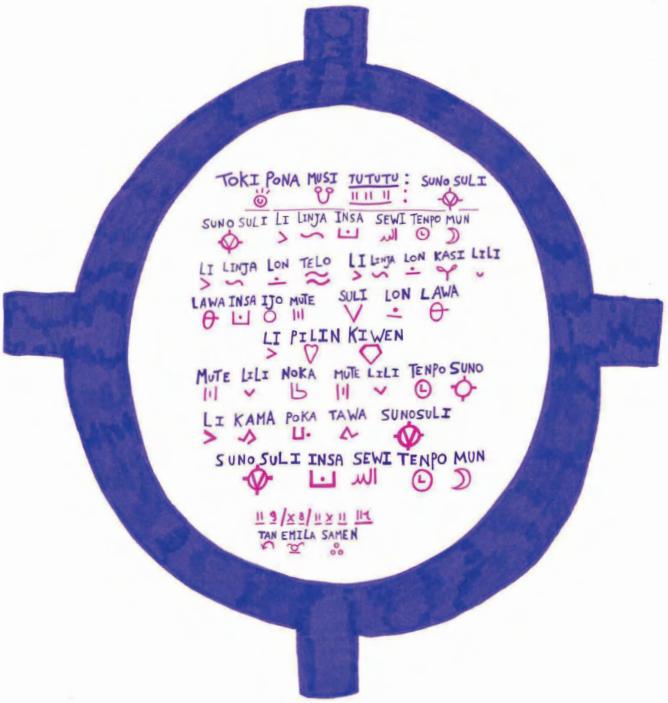


Figure 102. "*suno suli*" (Big Sun), toki pona poem by Emilia Sameyn (2023).

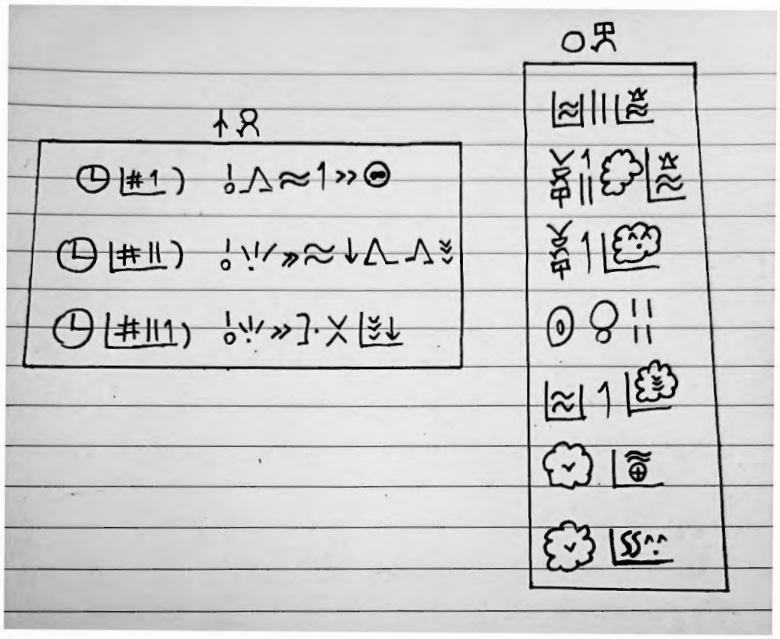


Figure 103. Crêpe recipe in sitelen pona by Reddit user u/y_witch (2023).

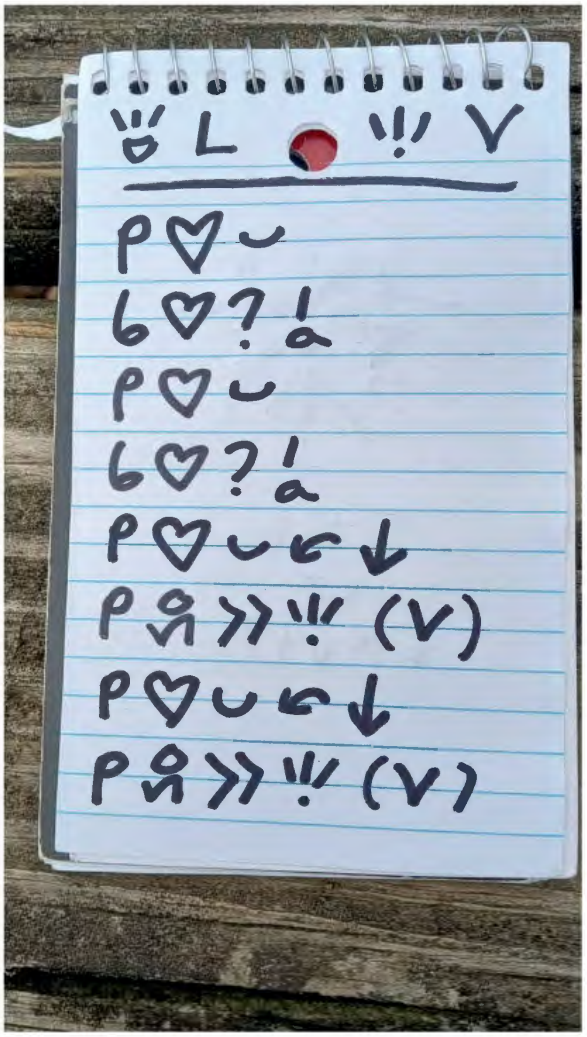


Figure 104. Campfire song written by Reddit user u/TromboneBoi9 (2023).

Figure 105. Custom DVD case with sitelen pona text by William S. Bradley (2024).

Figure 106. Vinyl record prop created by jan Aleko (2023) used for the music video “mi ko” by jan Sepulon, featuring both sitelen pona and sitelen sitelen (another script but with less usage).

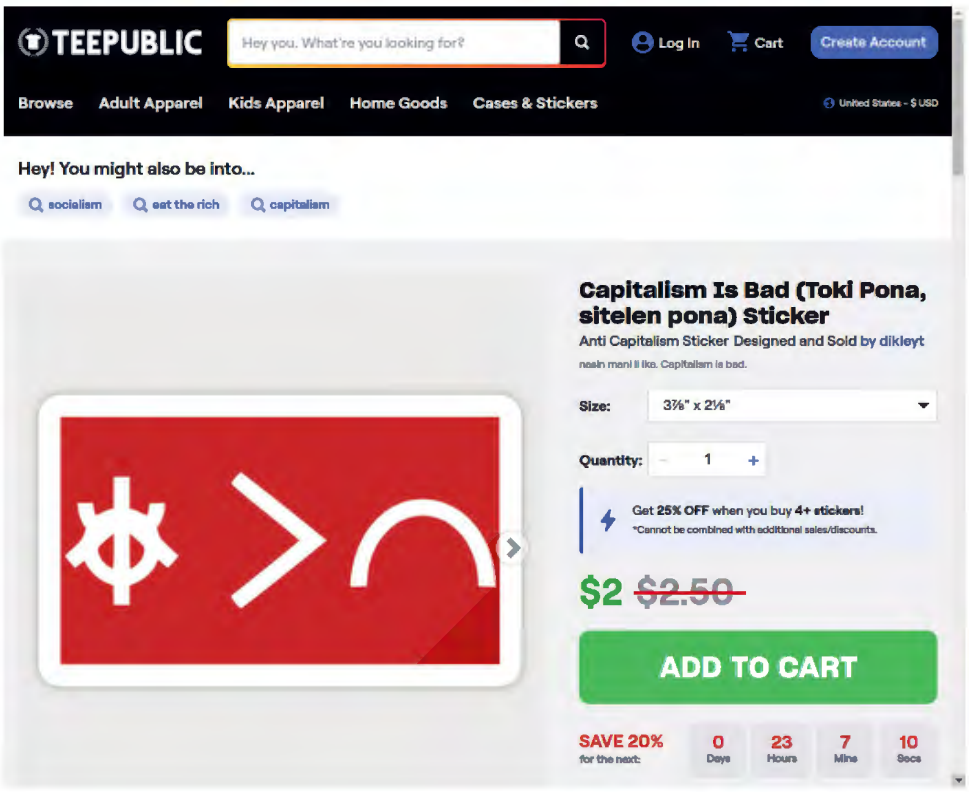


Figure 107. Sticker for sale with *nasin mani li ike* ("Capitalism is bad") in sitelen pona. The glyphs for *nasin* and *mani* are overlaid (Leybikhe 2023).

Figure 108. Playing card deck based on sitelen pona by soweli Iseja (2022).

Figure 109. Six- and eight-sided dice with sitelen pona numbers (Niwatori 2022b, c). Eight-sided dice use a nonstandard number system, inspired by the song “7 (likujo)” by jan Usawi (2024b), with the number 8 replaced with *kijetesantakalu* (not used as a numeral).

Figure 110. Collaborative chalk art from a meetup in Vienna, Austria (Samys 2022b).

Figure 111. Graffiti in sitelen pona. In Johnson City, Tennessee, found by Reddit user u/eriksealander (2023b). In Melbourne, Australia, found by jan Tomi Wami (2023).

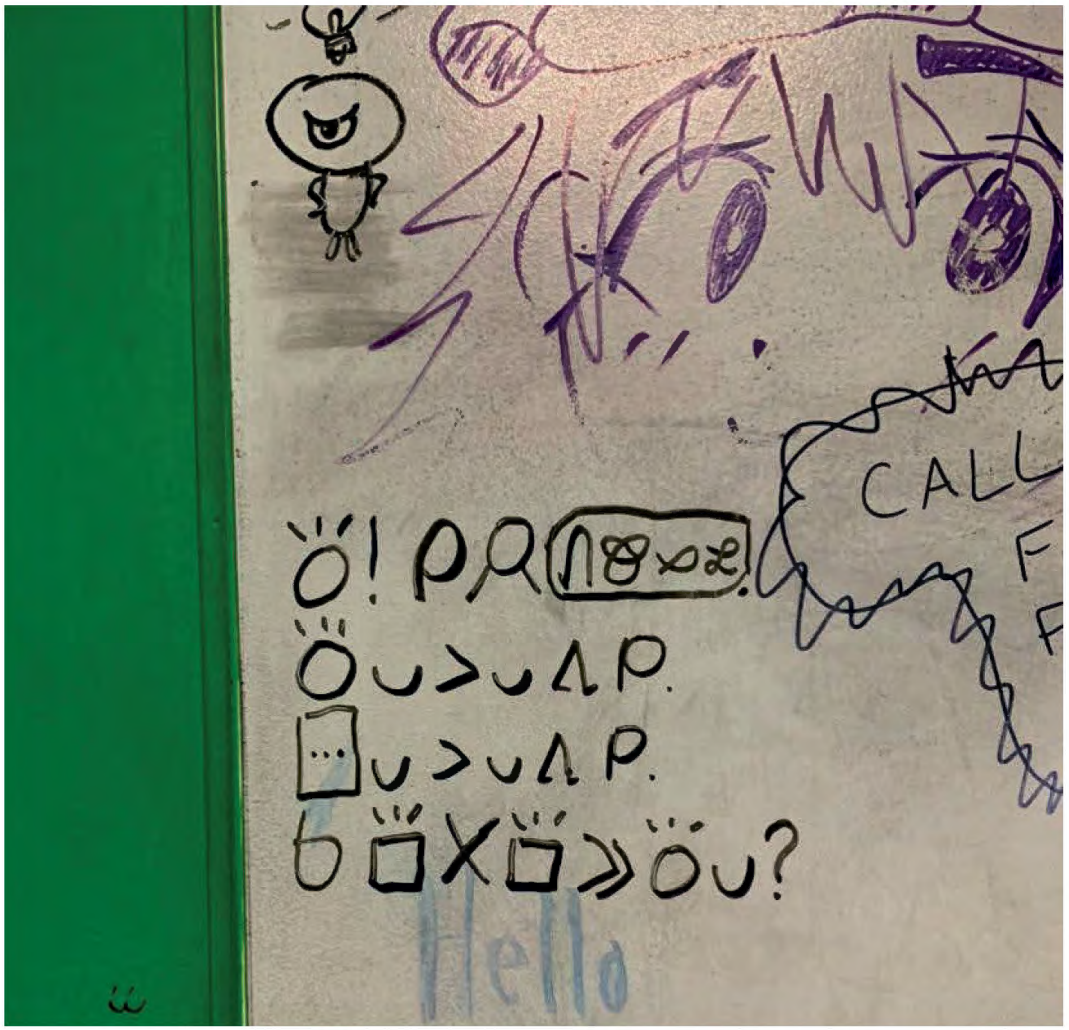


Figure 112. Writing in sitelen pona on a college whiteboard, by Reddit user u/VerySwag (2024).

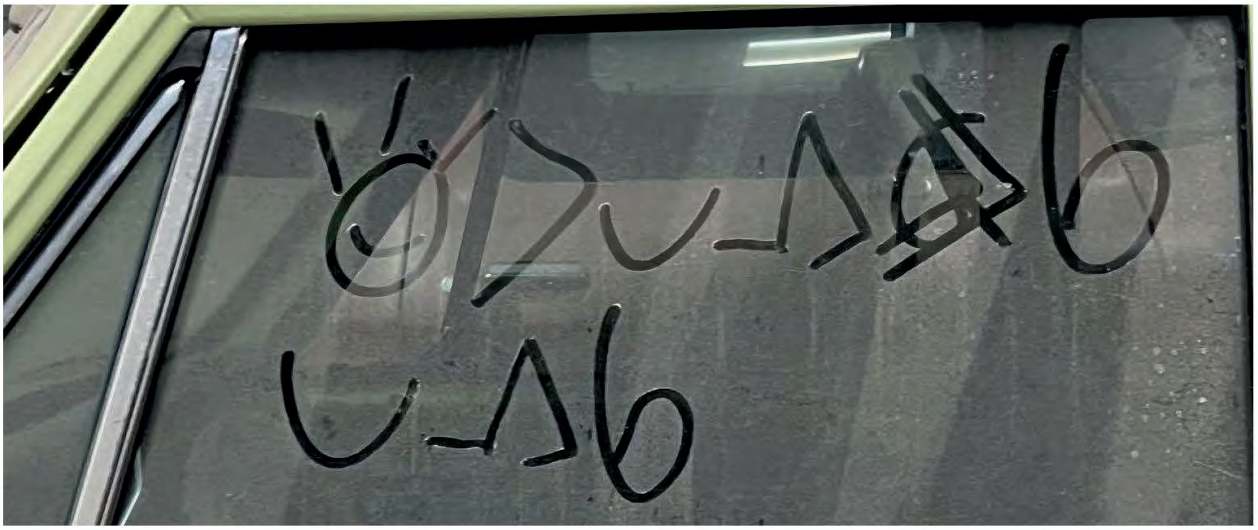


Figure 113. Writing in sitelen pona on a car window, found by Reddit user u/enpn (2024).

Figure 114. Tattoos in sitelen pona. From left to right: *kijetesantakalu* (u/EtruscaTheSeedrian), *ale li pona* (jan Itan), and *soweli seli* (lipamanka).

Figure 115. Art piece by Algerian artist Chlag Amraoui (2023), which incorporates sitelen pona and Tifinagh symbols.

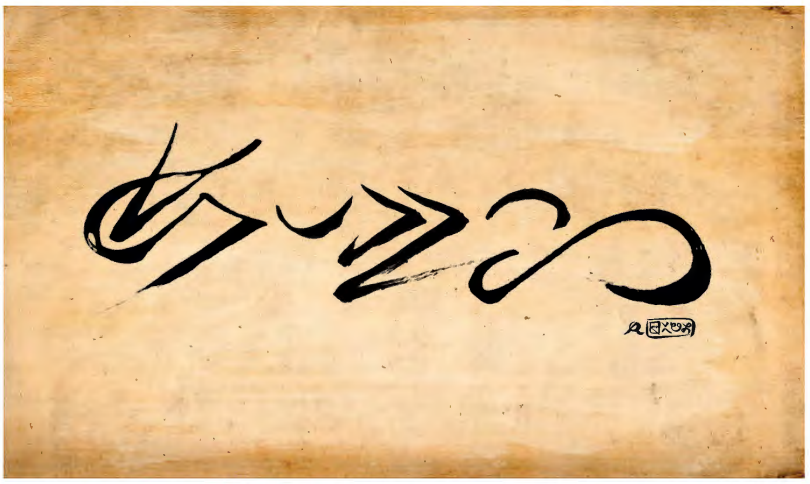


Figure 116. Examples of calligraphy in sitelen pona by James Flear (n.d-a).

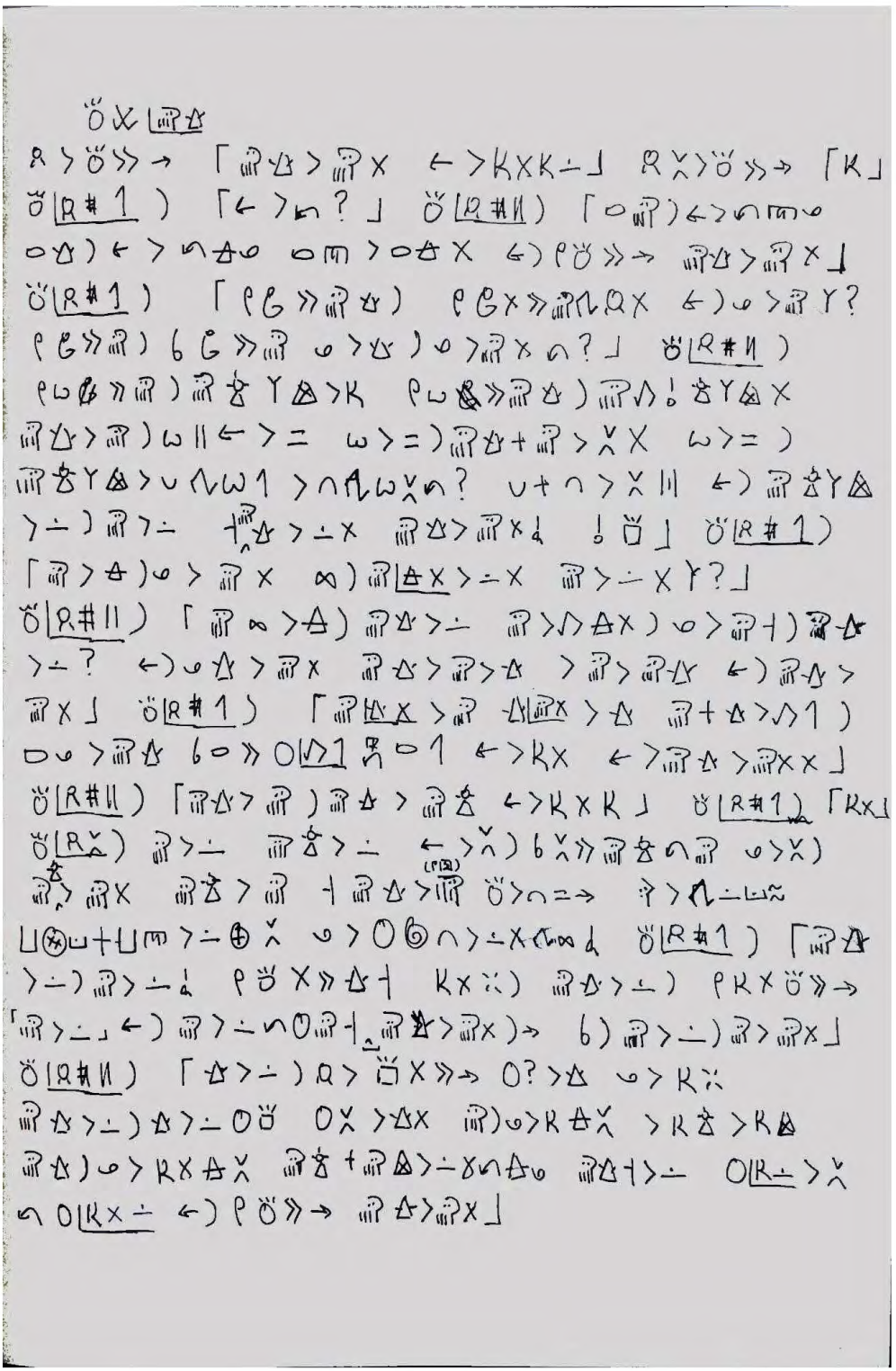


Figure 117. Translation of the ancient Chinese philosophy “white horse debate” (白馬論), translated and handwritten by jan lili Enta (2023).

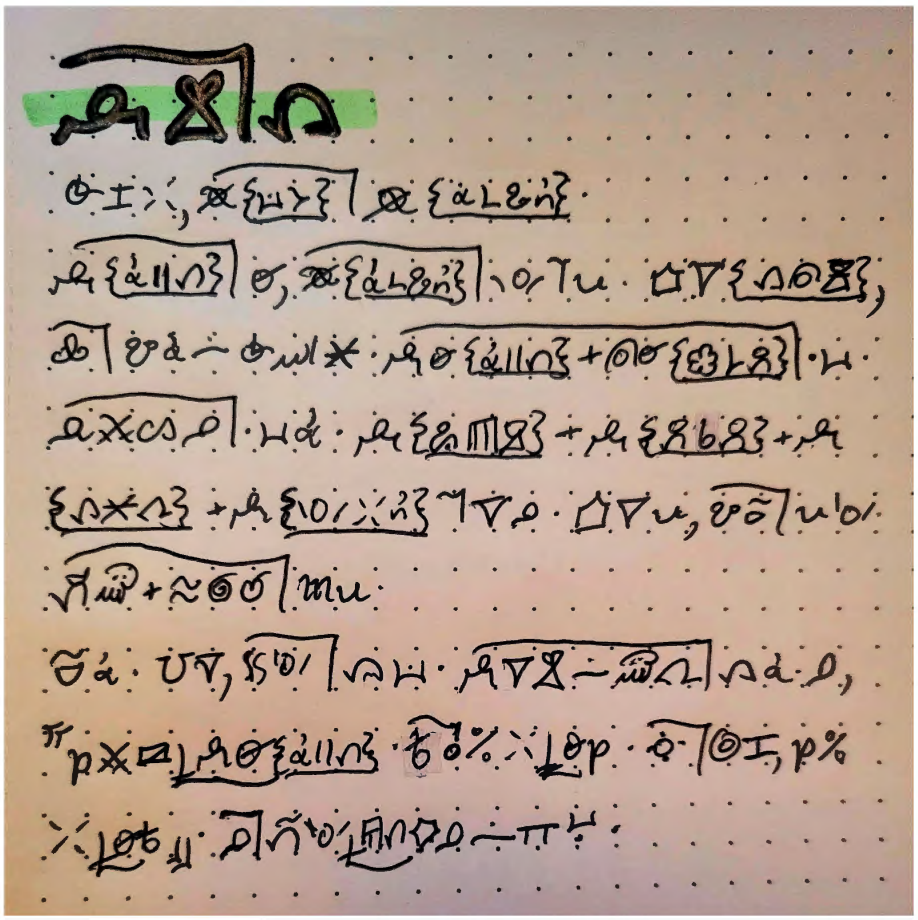


Figure 118. Translation of Sir Gawain and the Green Knight, translated and handwritten by C. Ryan Moniz (n.d.).

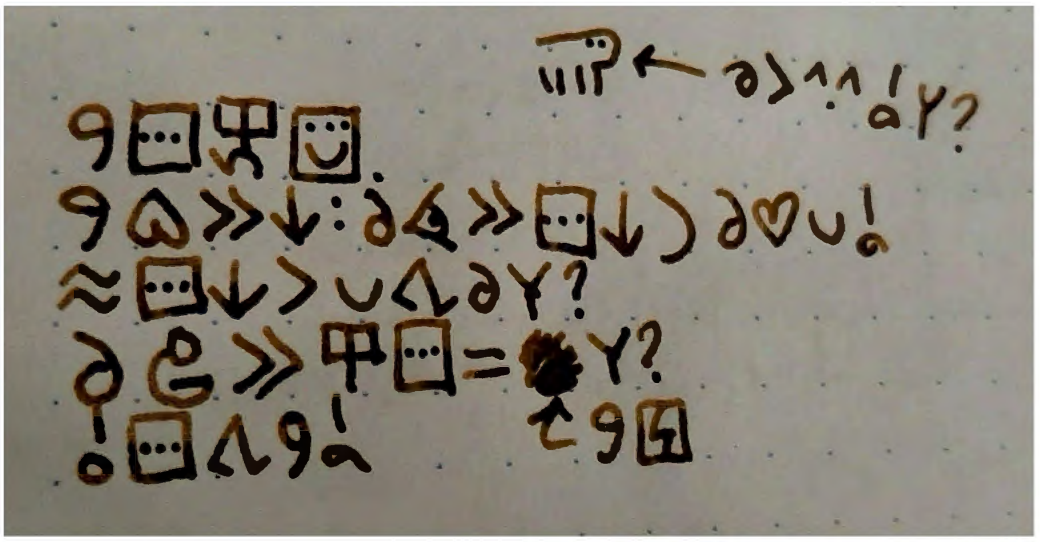


Figure 119. Informal handwriting in sitelen pona by jan Lun (2023).

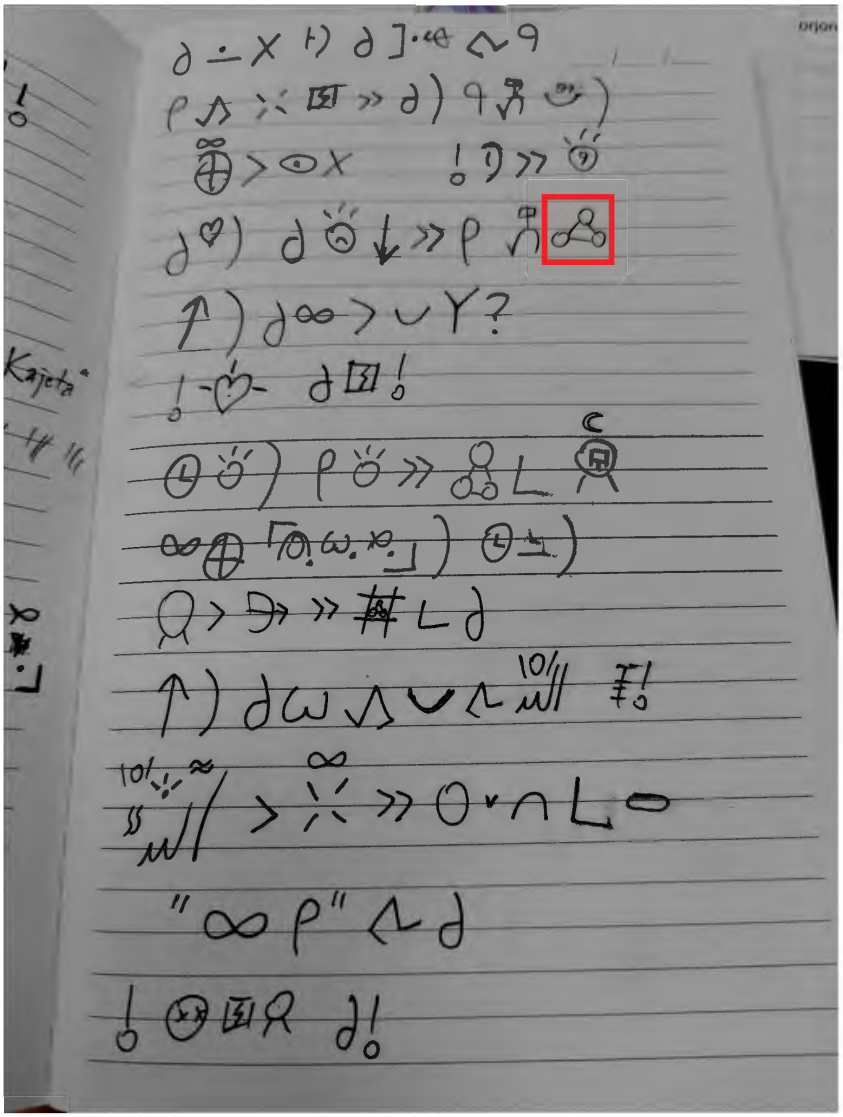


Figure 120. Informal handwriting in sitelen pona by soweli Wajole (2023), highlighting the *kulupu*-derived version of *linluwi*.

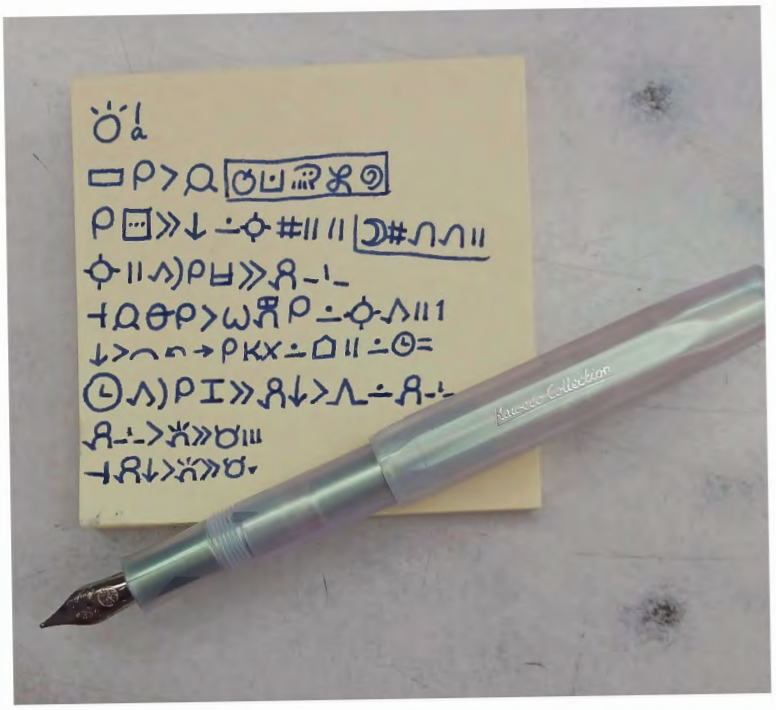


Figure 121. Informal handwriting note in sitelen pona by jan Kisen (2023).

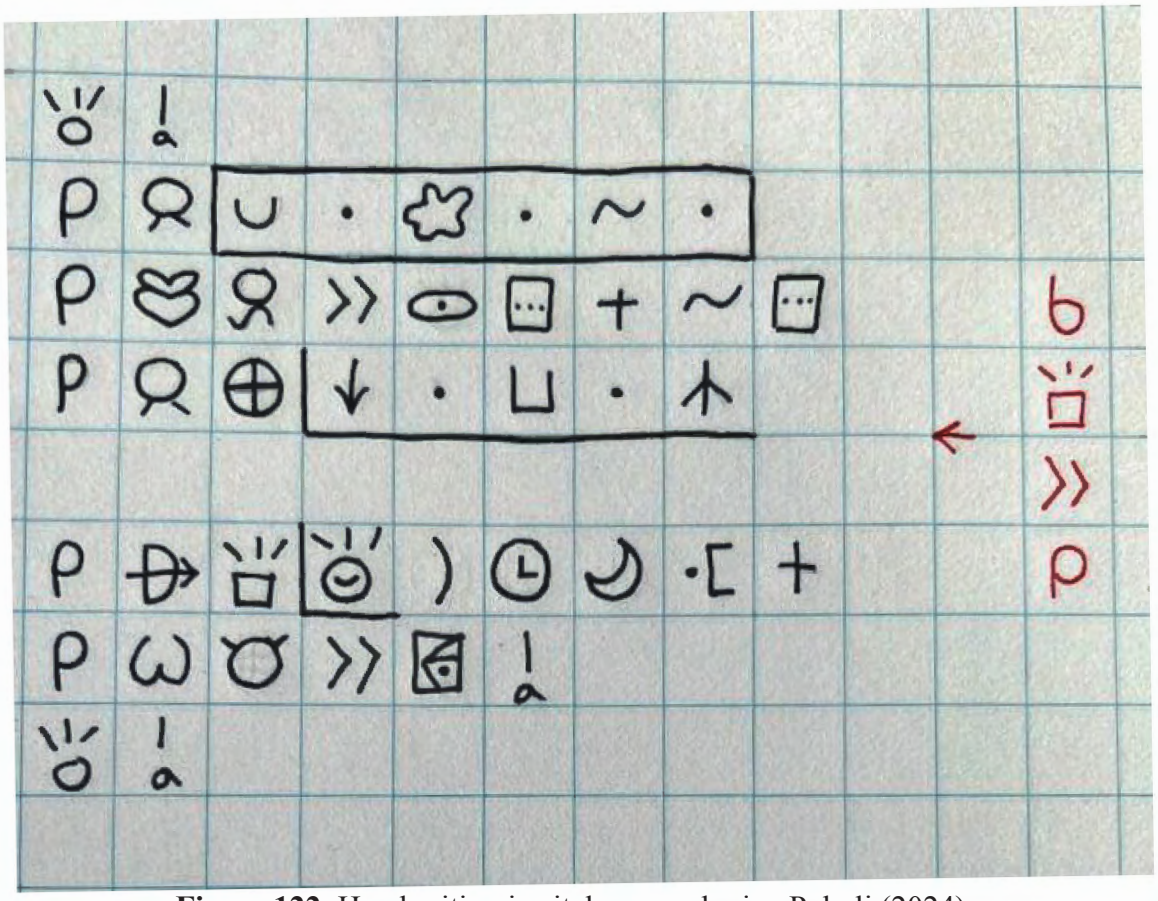


Figure 122. Handwriting in sitelen pona by jan Pokoli (2024).

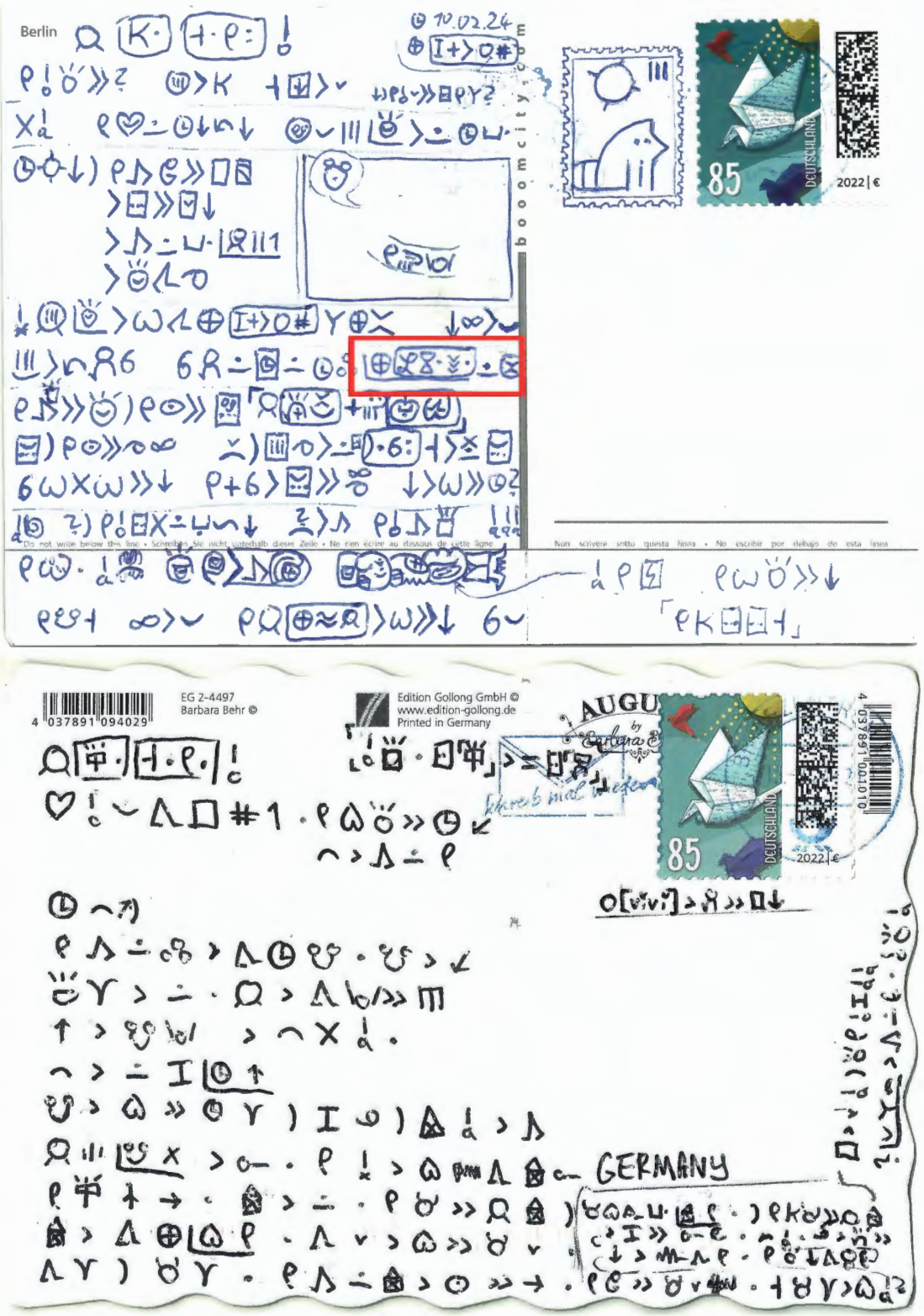


Figure 123. Handwritten postcards from jan Mateja (2024, top) and ijo Vivi (2024, bottom) to
Ret Samys in anticipation of the Summer 2024 meetup in Berlin, Germany.
An example of a cartouche nested inside an extended glyph is highlighted in red.

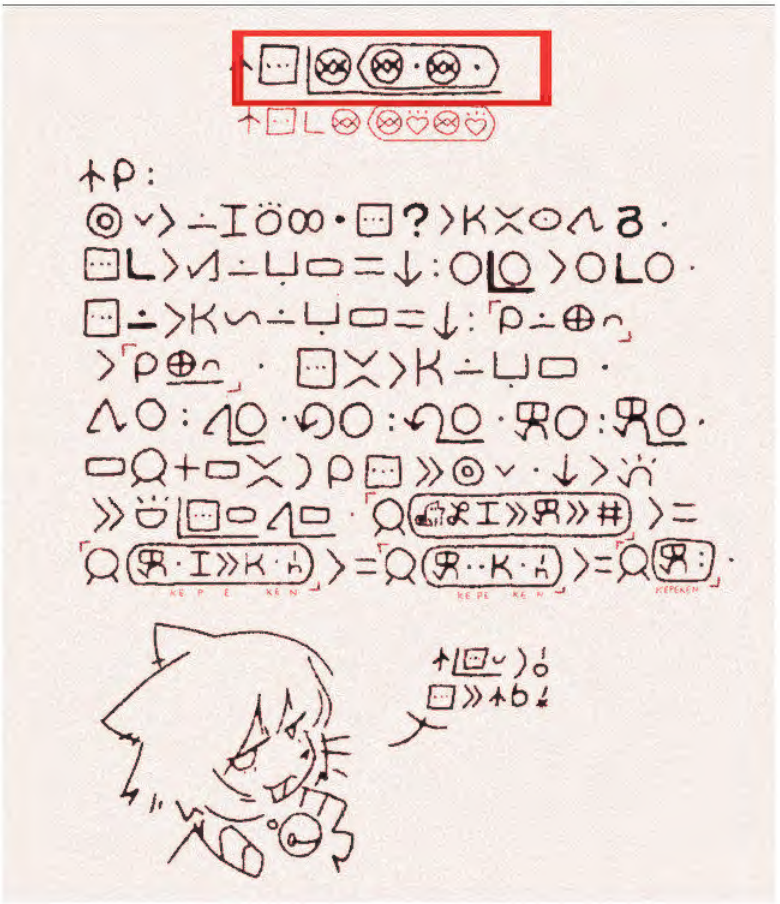


Figure 124. Description of moli Momo's (2023a) style of writing, containing features, such as extended glyphs, *nasin sitelen kalama*, and sentence-ending punctuation. An example of a cartouche nested inside an extended glyph is highlighted in red.

Figure 125. Handwritten informal message on Discord demonstrating nested *pi*.

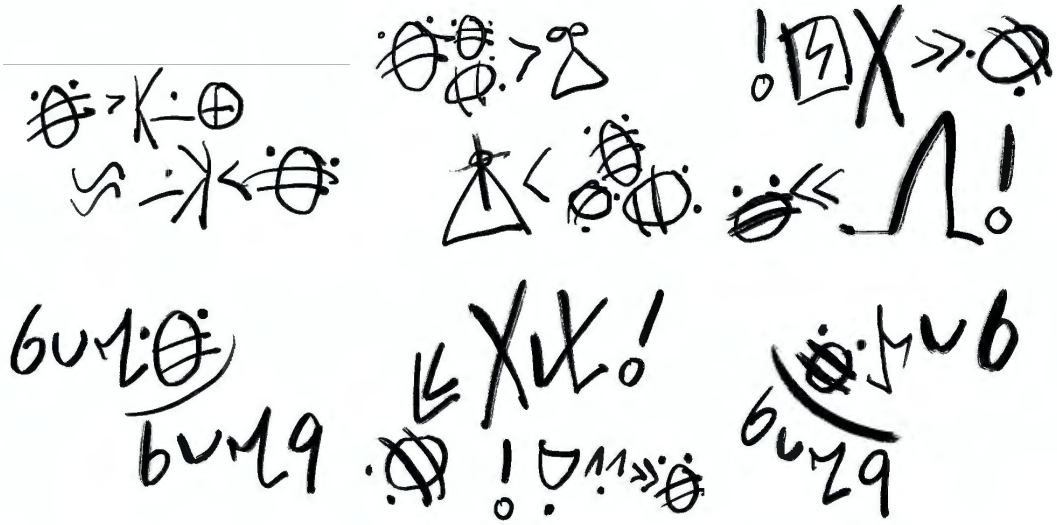


Figure 126. Example of sitelen pona handwriting by pipi Kewapi (2024) with several experimental features, including bidirectional writing, non-linear use of the character for *la*, and word multiplication instead of the logograph for *mute* ("many").

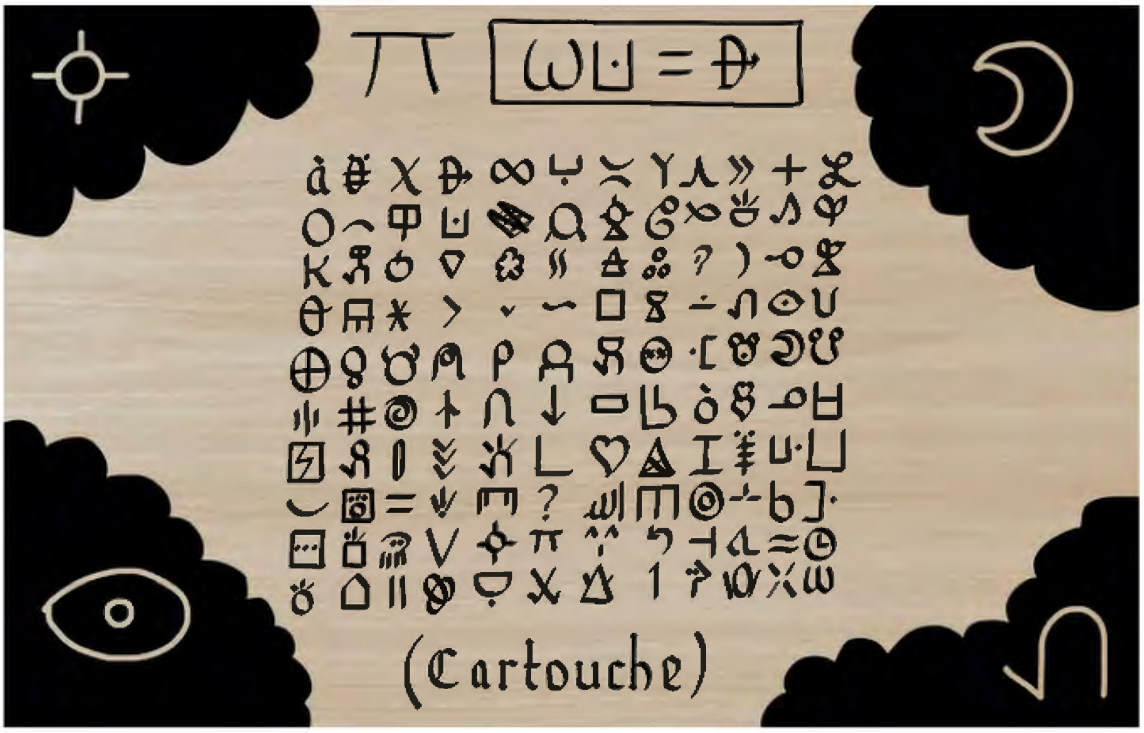


Figure 127. Divination tool in sitelen pona by Phox (2023).

Figure 128. The phrase "*mi ken ala toki pona e ijo la mi sona ala e ijo*" ("if I can't say a thing in toki pona, I don't know the thing"), a principle of toki pona philosophy, illustrated and handwritten by soko Piki (2024).

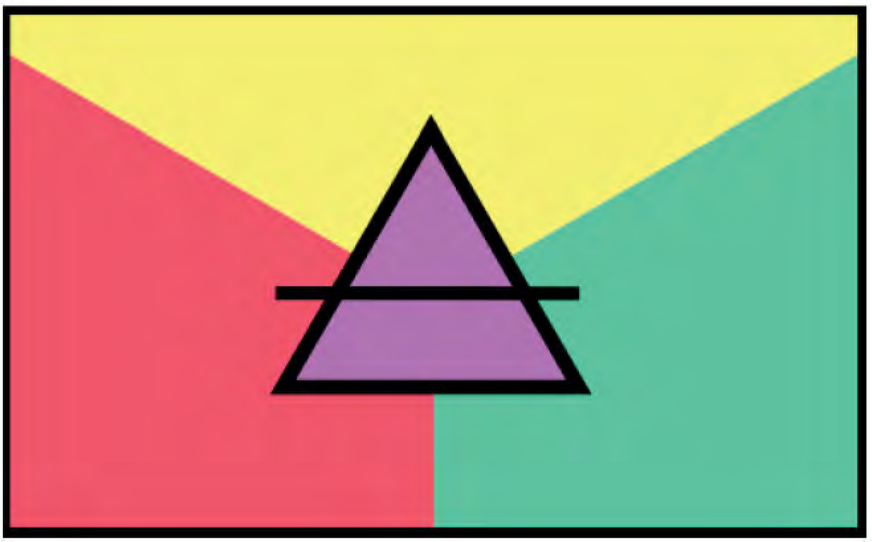


Figure 129. Proposed toki pona LGBT flag by jan Nijo (2021), with the word *kule* in the center.

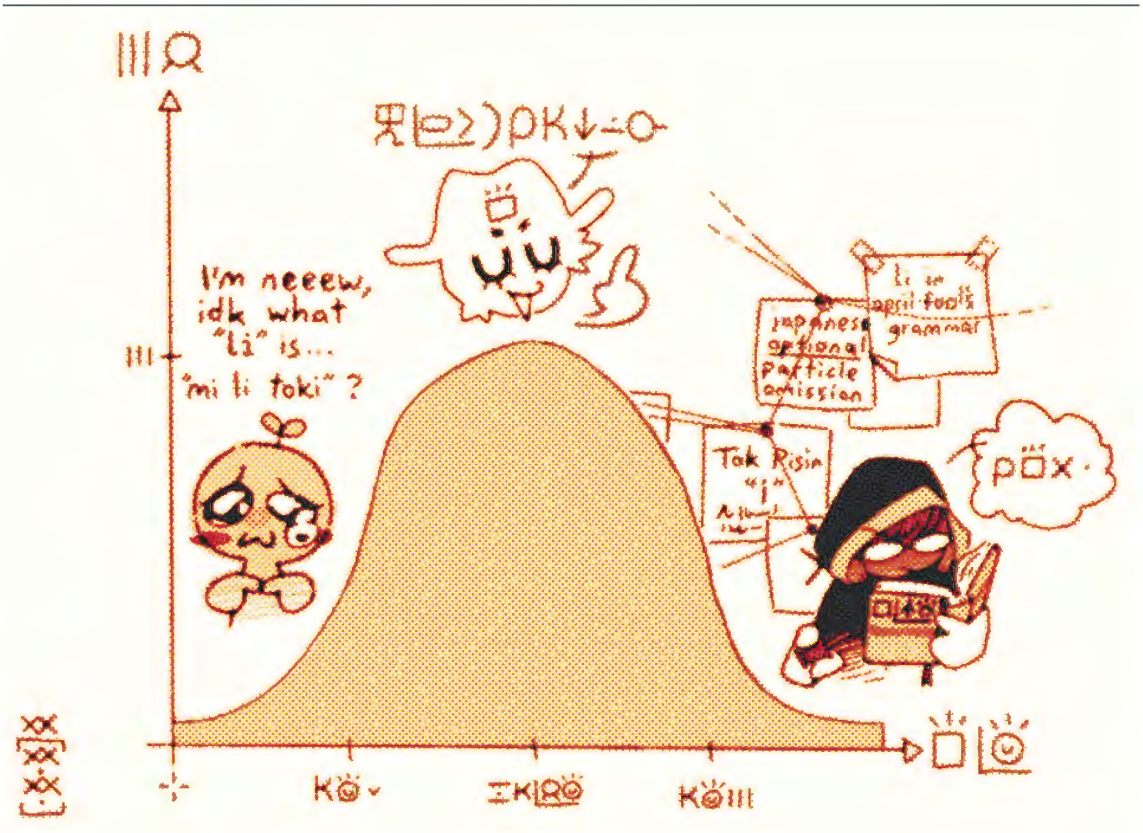


Figure 130. Humorous comic by moli Momo (2023b) about the particle *li*.

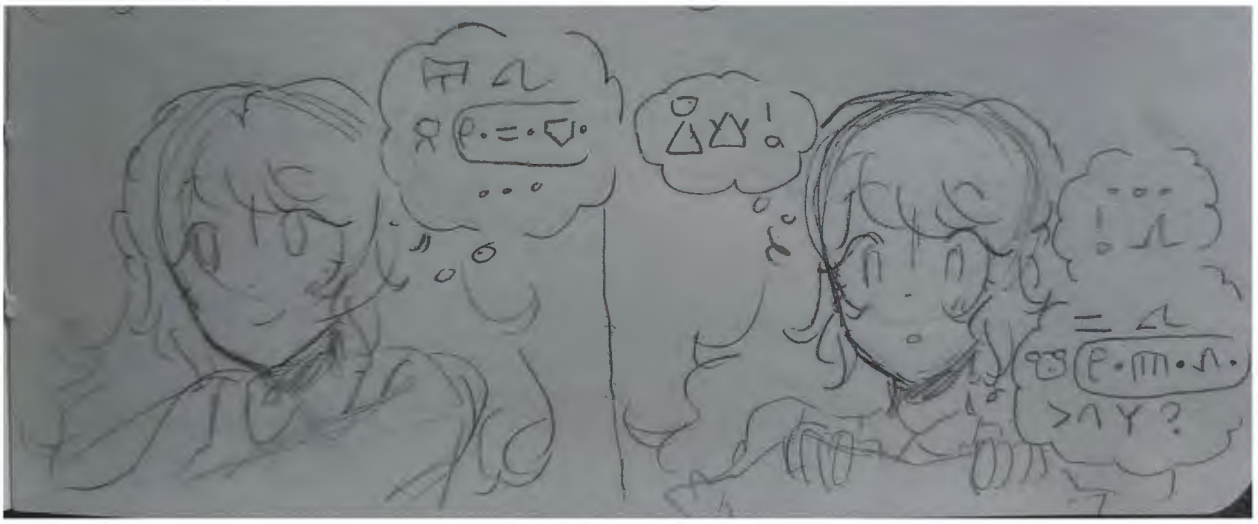


Figure 131. Short comic panels in sitelen pona drawn by kasi Lija (2024b).

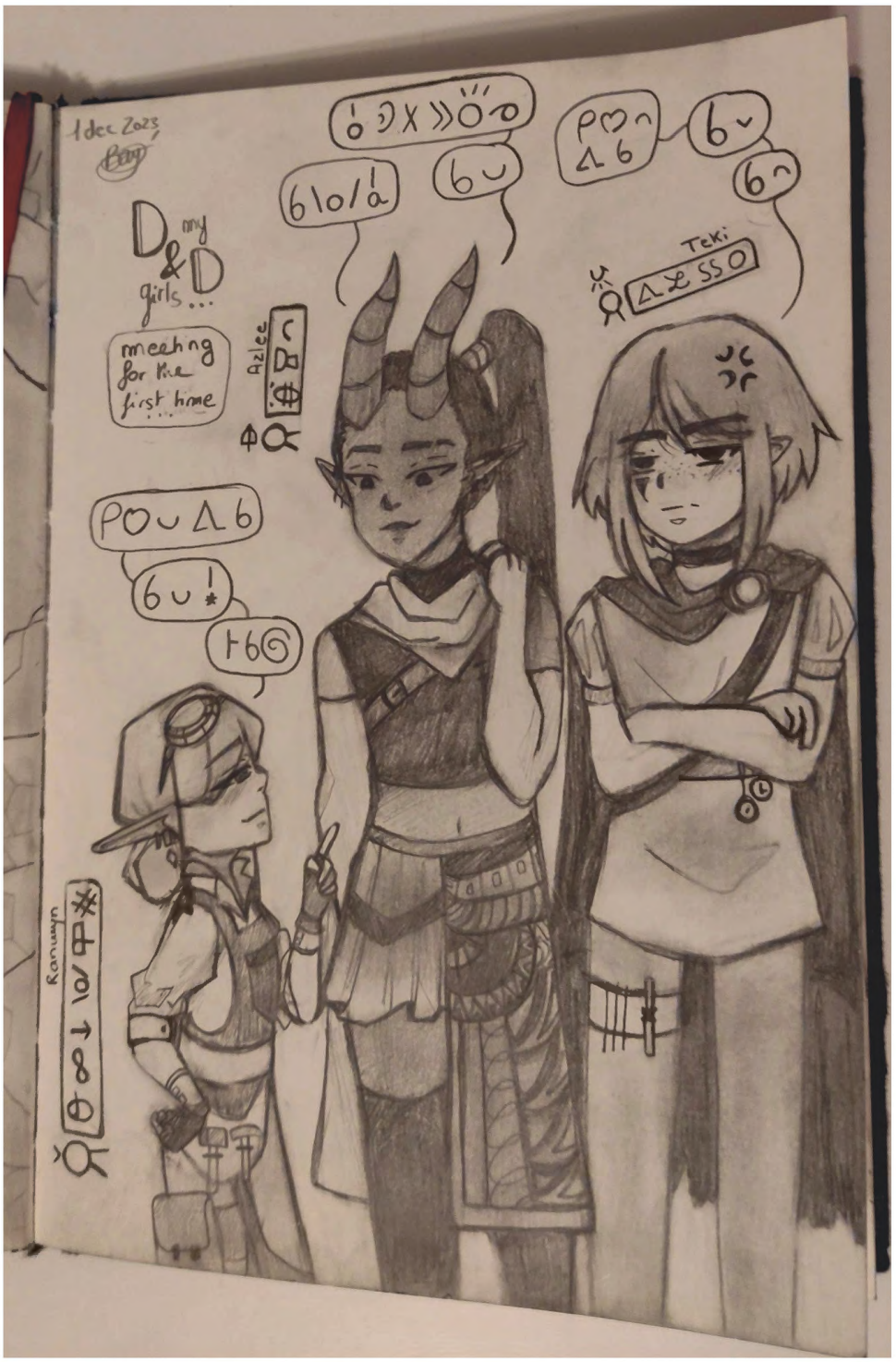


Figure 132. Short comic panel “My D&D girls” illustrated by jan Peli (2023).

Figure 133. Humorous and educational illustration featuring jan Kekan San, illustrated by waso mu (2023).

Figure 134. Humorous flag of California by soweli Kina (2023), where *soweli* replaces the bear,
mun replaces the star, and the text says in the Latin script “ma Kalapona.”

Figure 135. Burger King “Whopper” advertisement parody by jan Jami (2023).

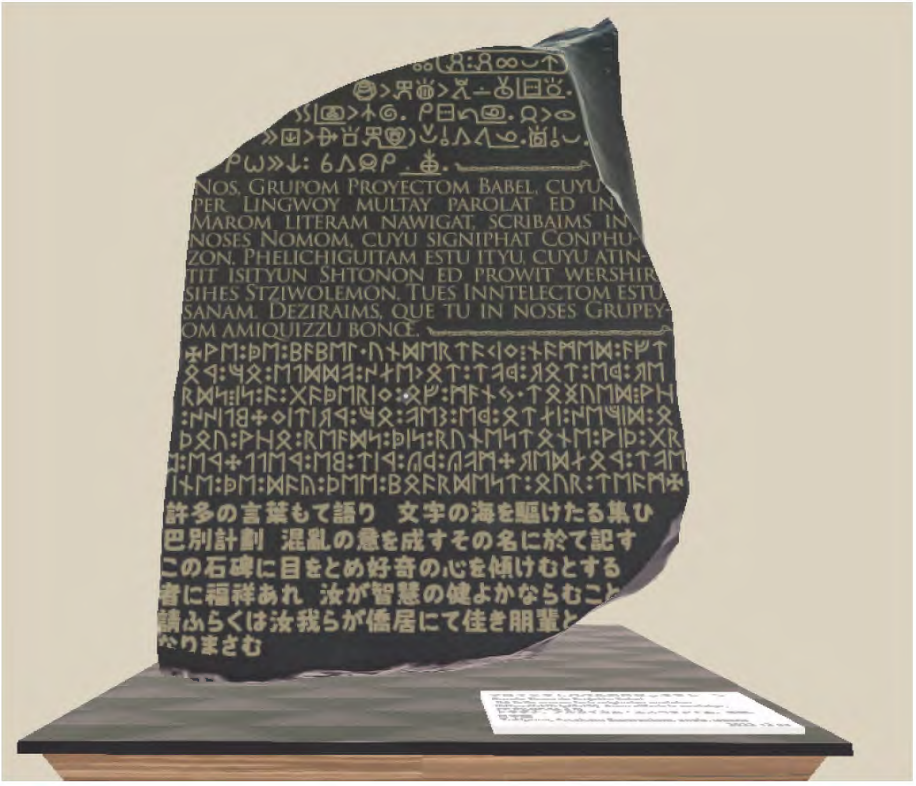


Figure 136. 3D model of the Rosetta Stone, with the languages replaced with toki pona, Archaic Esperanto, English (in Elder Futhark), and Japanese (Projekto Babel 2022).

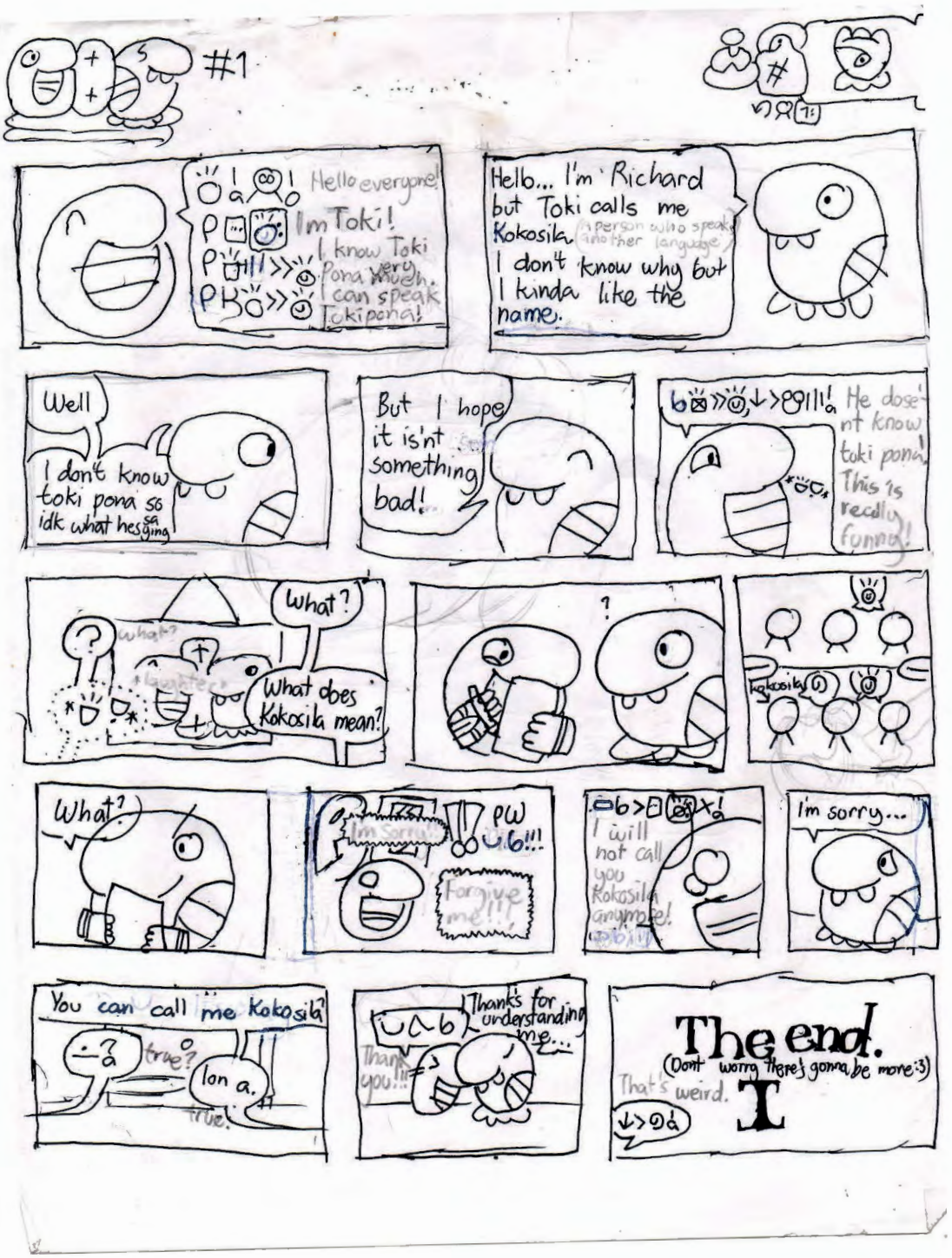


Figure 137. Short humorous comic with characters modeled after sitelen sitelen (another script with less usage) and text in sitelen pona, written and illustrated by Ron Raphael Salvoro (2024).

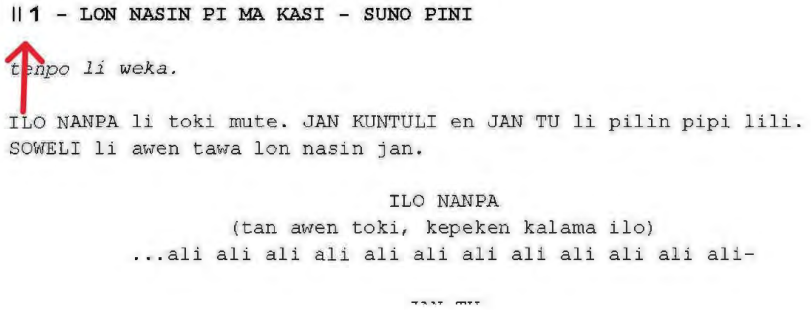


Figure 138. Highlighted usage of sitelen pona numerals in Latin script text, as shown in the screenplay *tu kuntu* by kule epiku Atawan and kala pona Tonyu (2021).

Figure 139. User attempting to type sitelen pona using existing lookalike Unicode characters. This is called *sitelen akesi* or *sitelen pona Juniko*, among other names, and is fairly common.

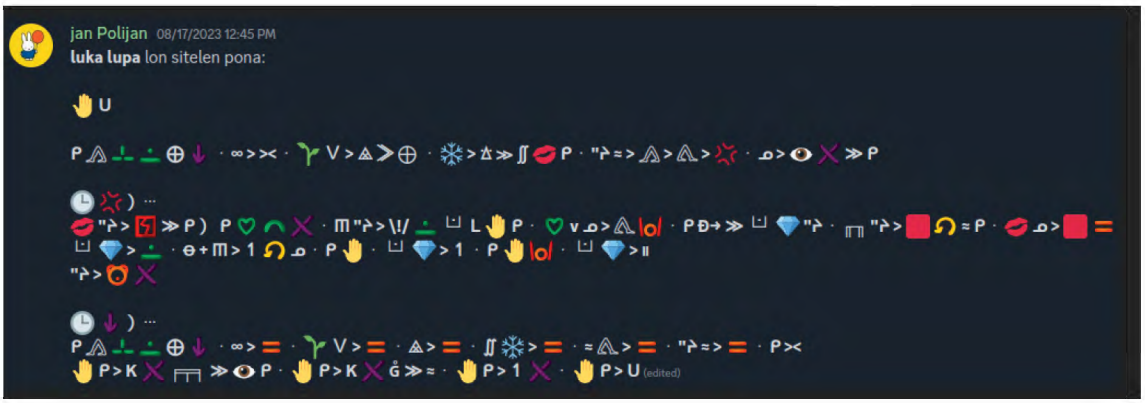


Figure 140. Story *luka lupa* in toki pona written using emoji as replacement for sitelen pona text.

Figure 141. Users on Mastodon using custom emoji as replacement for sitelen pona text.

Figure 142. User using Latin script when a logograph is not encoded in a certain font. This will likely occur with characters not included in this proposal.

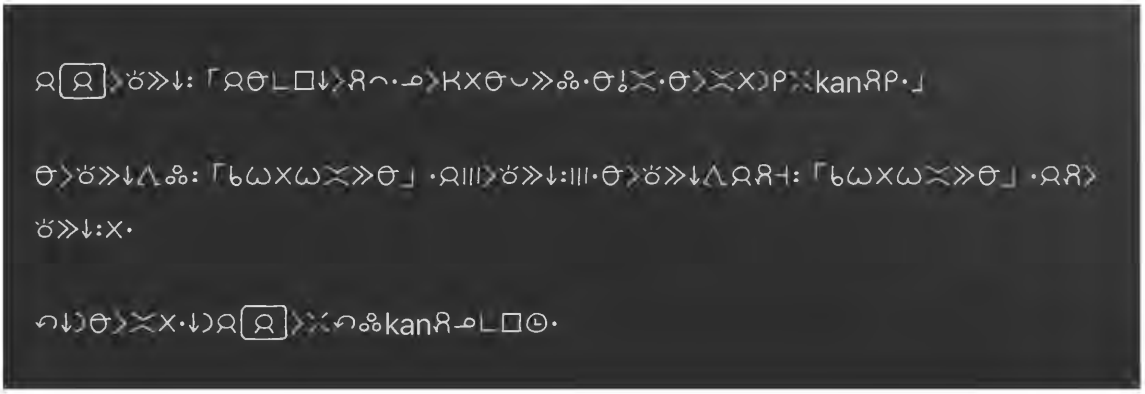


Figure 143. Old excerpt from an article from *lipu kule*, written with UCSUR characters, showing an example of an unencoded word (*kan*) being substituted with Latin script.

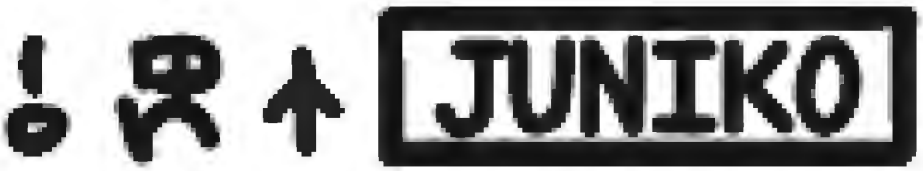


Figure 144. Example of the *linja pimeja pona* font, mixing sitelen pona with Latin script for a name (*Juniko*) not encoded in the font.

Figure 145. Cover of the album *toki Gaming* by jan Usawi (2023b), satirizing the typo possibilities from the use of ASCII transcription when inputting sitelen pona.

Fonts and input methods for sitelen pona

Figure 146. Examples of combined glyphs from documentation for the font *Nishiki-teki*.

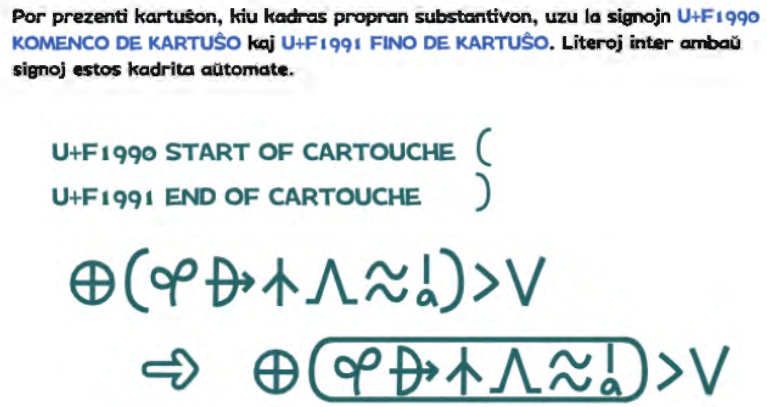


Figure 147. Instructions on creating cartouches from documentation for the font *Nishiki-teki*.

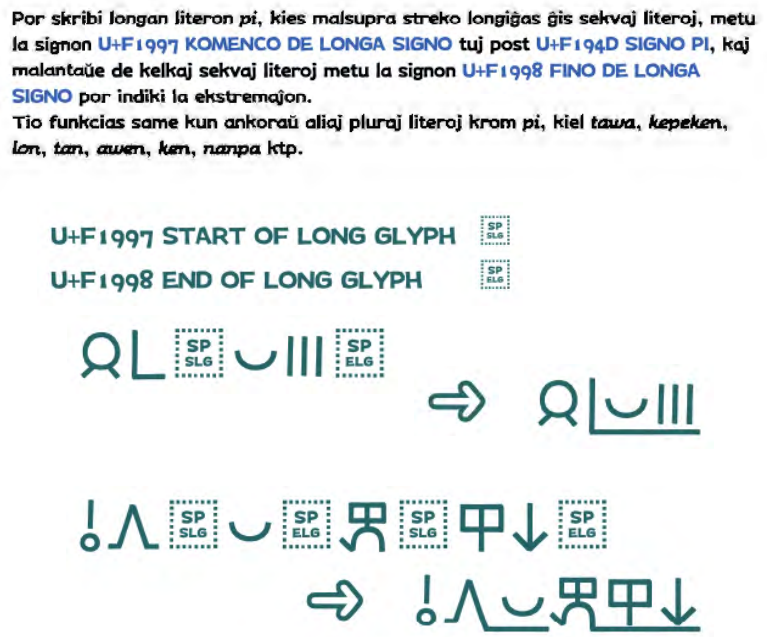


Figure 148. Instructions on creating long glyphs from documentation for the font *Nishiki-teki*.

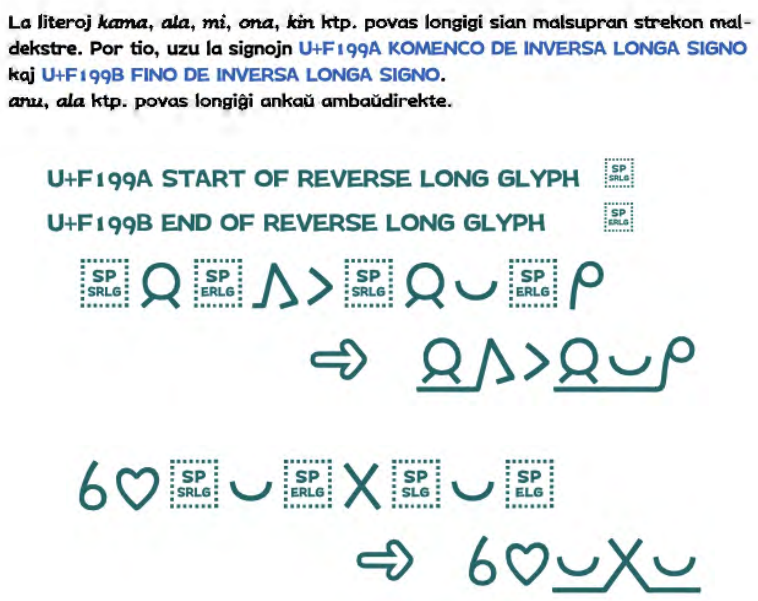


Figure 149. Instructions on creating long glyphs from documentation for the font *Nishiki-teki*.

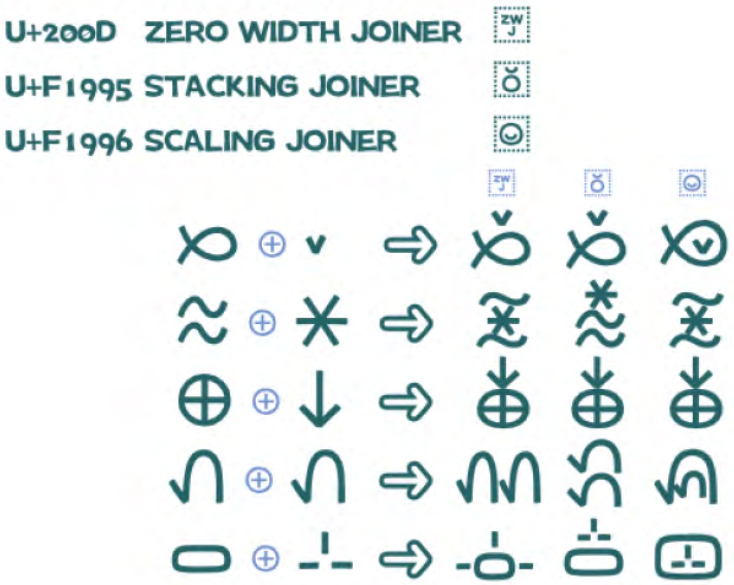


Figure 150. Examples of using joiners from documentation for the font *Nishiki-teki*.

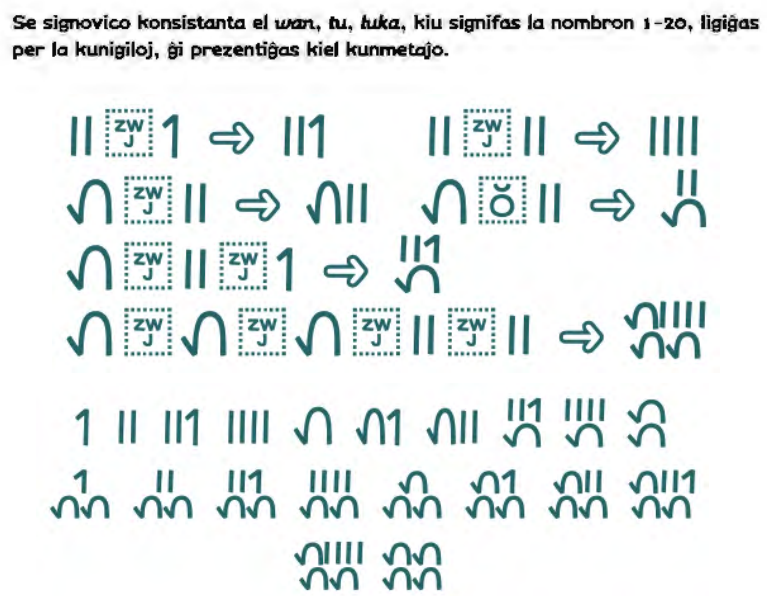


Figure 151. Examples of using joiners to make logographs for the numbers 1 to 20, from documentation for the font *Nishiki-teki*.

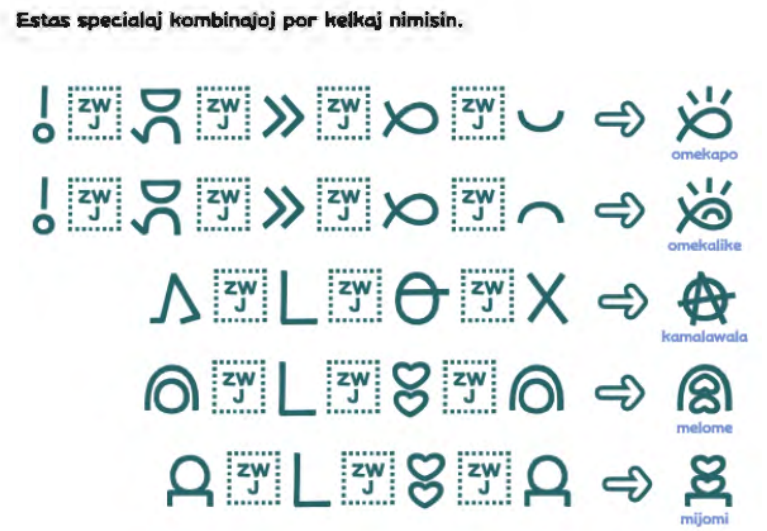


Figure 152. Examples of joining multiple logographs to create logographs for certain rare words, from documentation for the font *Nishiki-teki*.

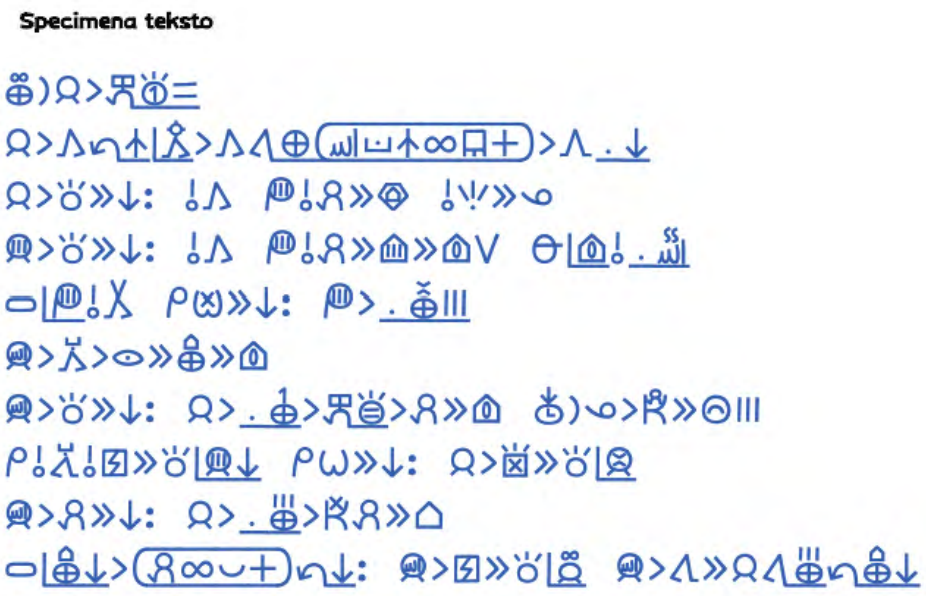


Figure 153. Sample text from documentation for the font *Nishiki-teki*.

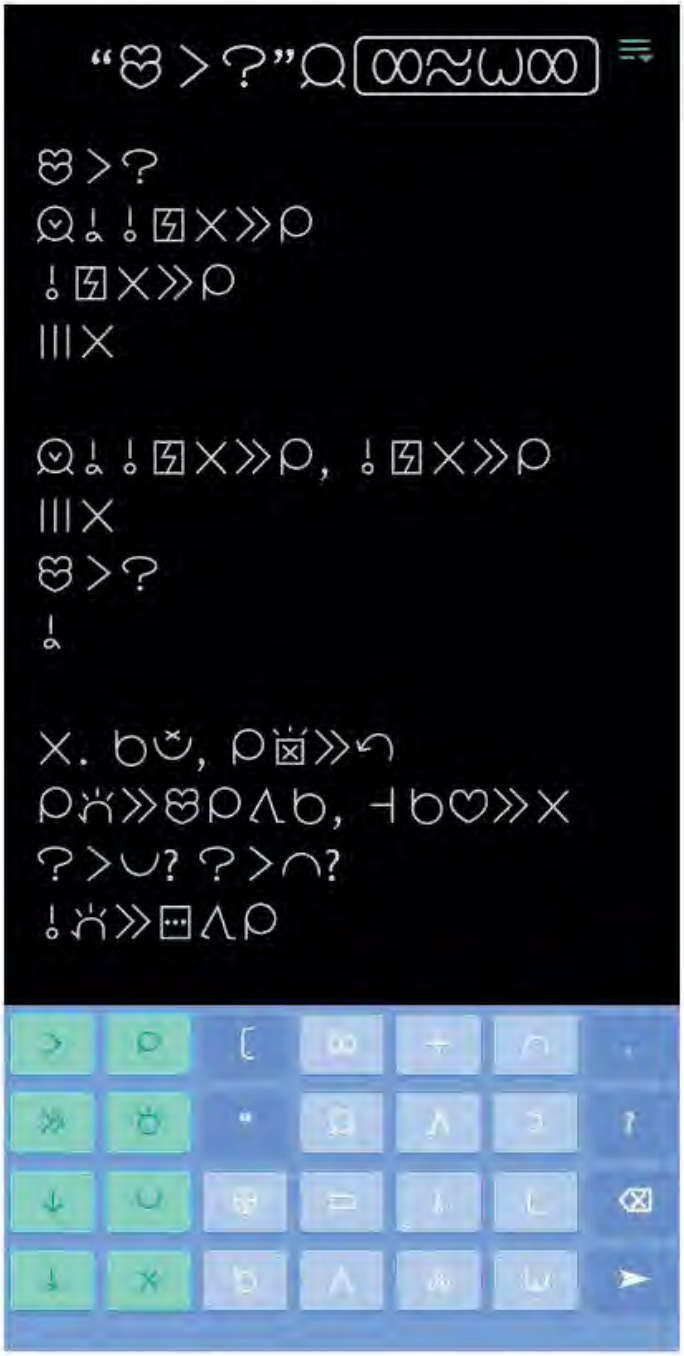


Figure 154. Screenshot of the Android app “Toki Pona Keyboard” by Timeo Pochin (2020).

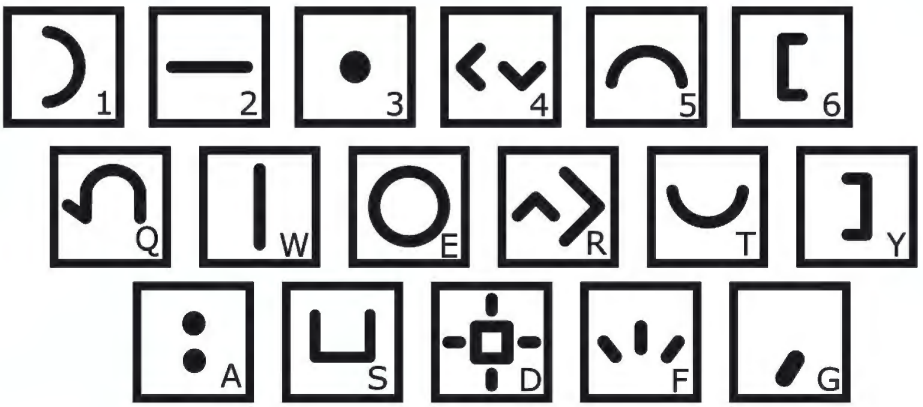


Figure 155. The keyboard layout of the *Wakalito* input method for toki pona, designed by jan Likipi and kala pona Tonyu (2020).

Figure 156. Physical keyboard for the *Wakalito* input method, created by H La Vallee (n.d.).

Figure 157. First page of “nimi ale kepeken nasin Wakalito 7.3.2,” the documentation for the *Wakalito* input method for toki pona.

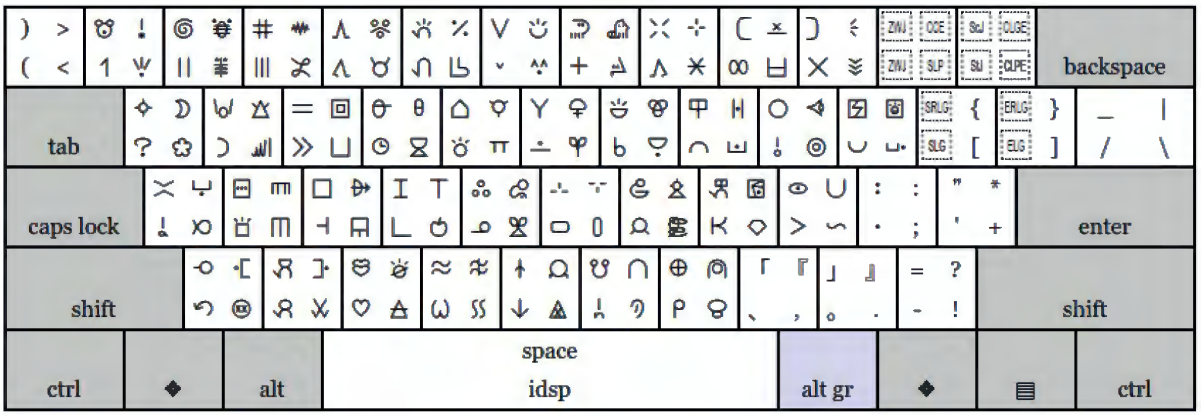


Figure 158. Keyboard layout for sitelen pona designed by Rebecca Bettencourt (2023a).

Figure 159. Physical keycaps for sitelen pona designed by Rebecca Bettencourt (2023b).

Figure 160. Screenshot of Ajemi, an input method editor for sitelen pona by dec32 (2023).

Figure 161. Digital typewriter with a sitelen pona screen by jan Inkepa (2020).

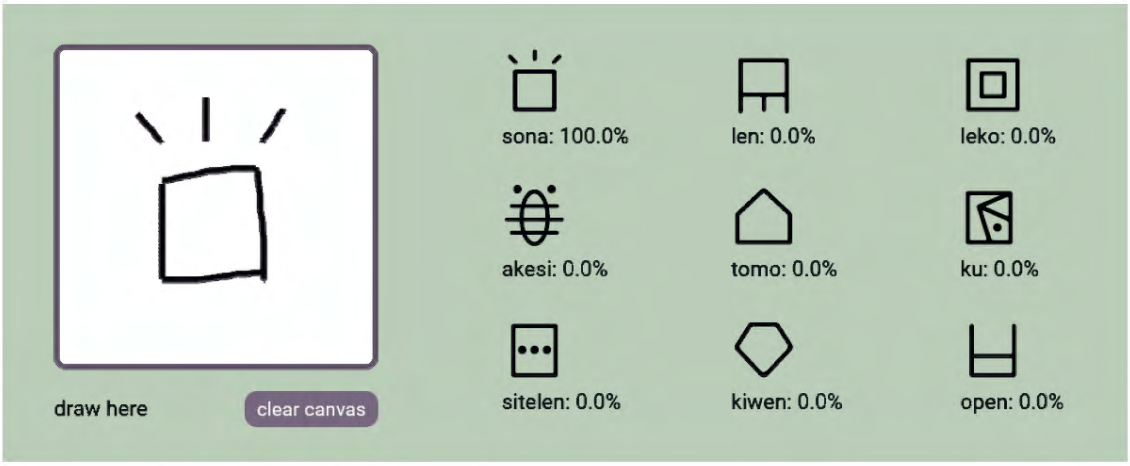


Figure 162. Screenshot of *ilo sona Like*, a sitelen pona handwriting recognition tool developed by kala pona Tonyu (2023).

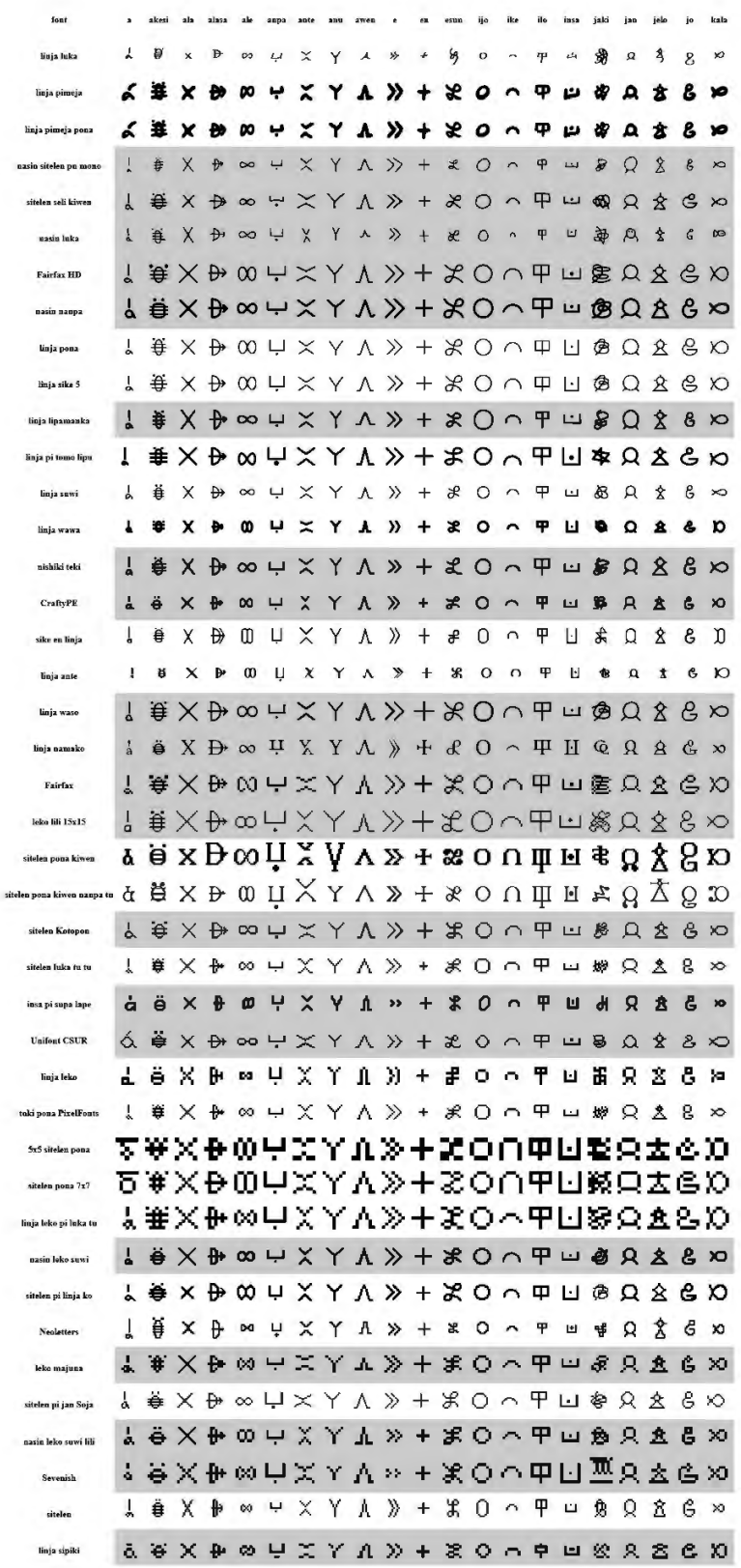


Figure 163. Comparison of glyphs from 45 sitelen pona fonts with a free software license. Words in the screenshot range from *a* to *kala*. Highlighted in grey are fonts using the UCSUR encoding.

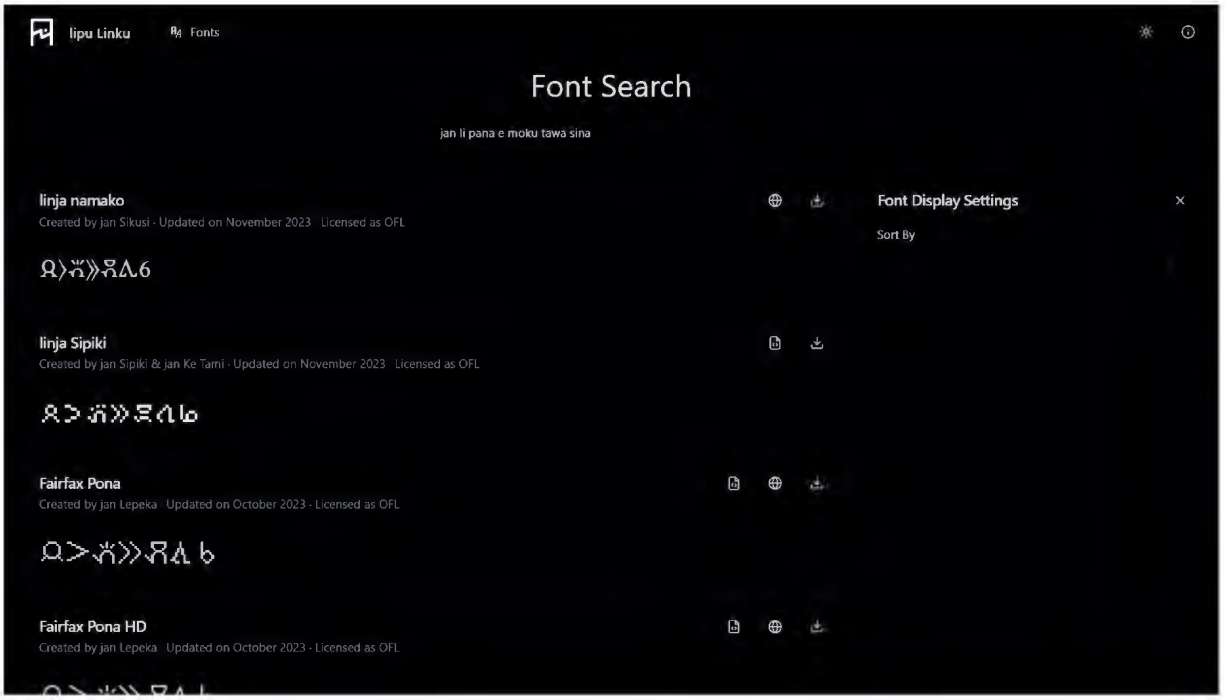


Figure 164. Directory of free fonts in of *lipu Linku* as of February 2024.

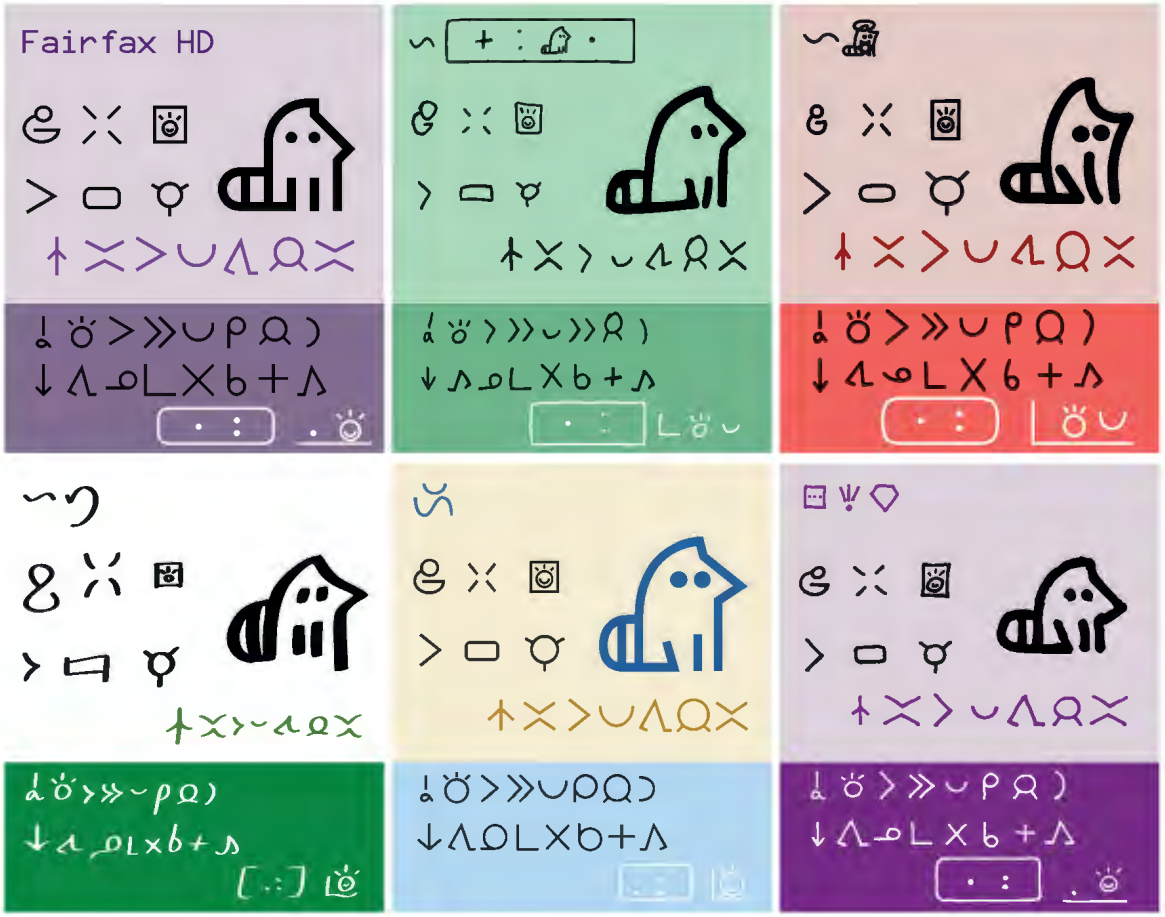


Figure 165. Typeface specimens for a handful of sitelen pona fonts.

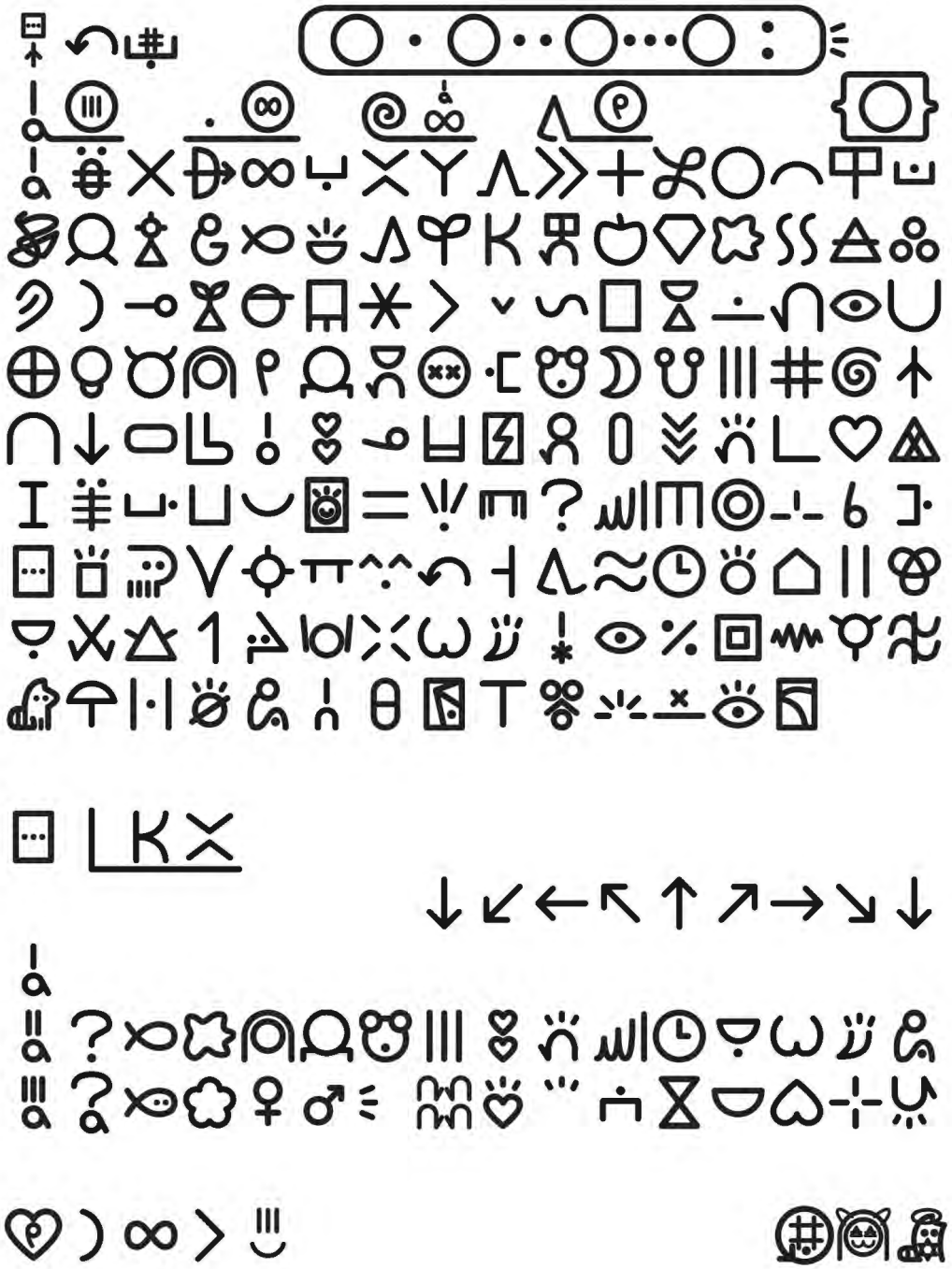


Figure 166. List of available characters in the version 3.0.0 of the font *nasin nanpa* by jan Itan.

Figure 167. List of available glyphs in version 3.6 of the font *linja pona* (image rotated 90°).

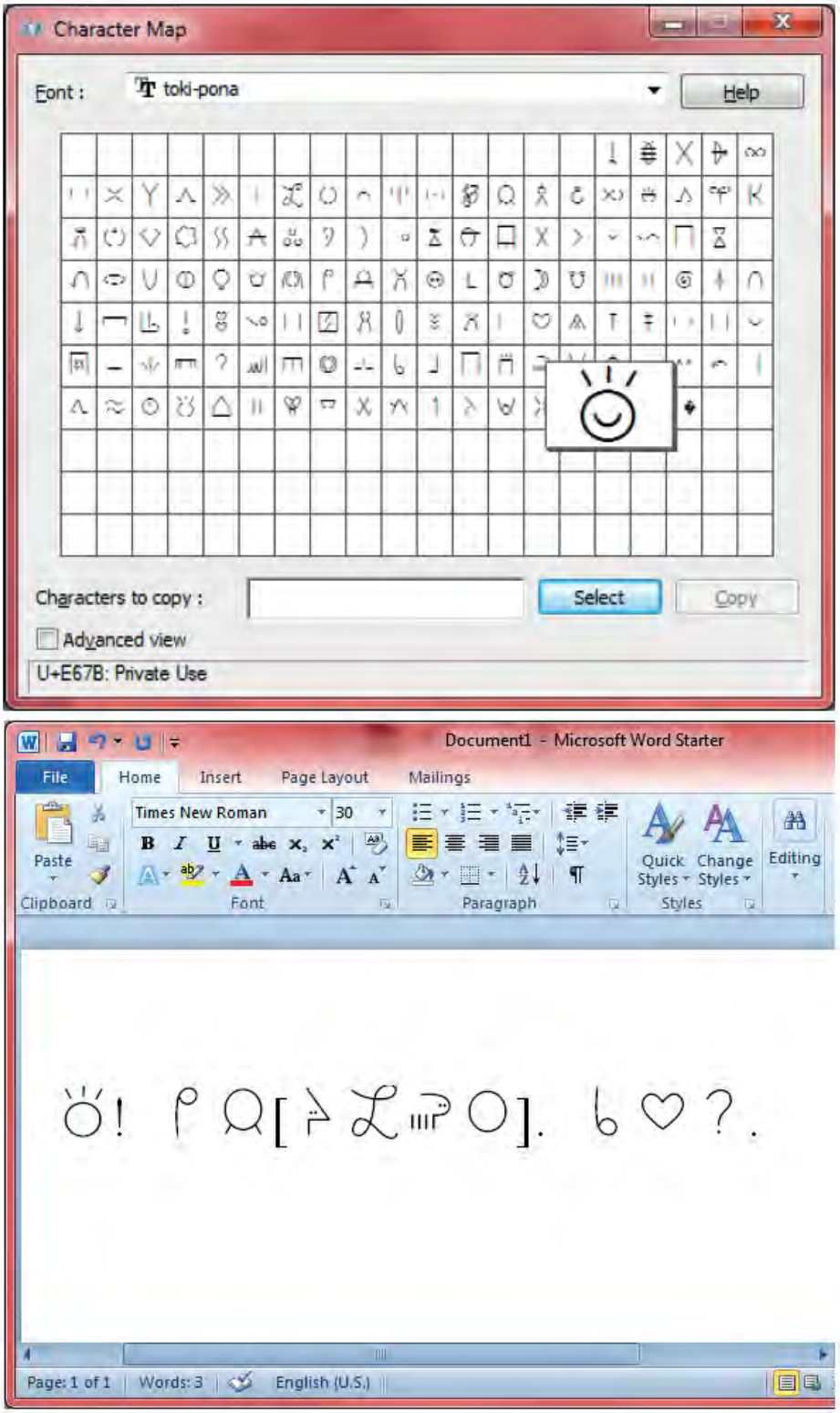


Figure 168. Examples of toki pona OTF, the first font created for sitelen pona, published in July 2014, two months after the release of *Toki Pona: The Language of Good*.

Search results, graphs, and statistics

Most Popular Conlangs Online

Conlangs by Discord Membership

Rank	Conlang	Creator	Largest Discord Server	Members
1	Toki Pona	Sonja Lang	ma pona pi toki pona	12,938
2	Na'vi	Paul Frommer	LearnNavi.org Community	9,591
3	Interslavic	committee	Medžuslovjansky	8,080
4	Esperanto	L. L. Zamenhof	Esperanto	7,360
5	Sindarin, Quenya	J. R. R. Tolkien	Vinyë Lambengolmor	2,963
6	Viossa	community	Viossa Diskordservery	1,690
7	Lojban	Logical Language Group	lojban	1,589
8	Ithkuil	John Quijada	uhailalepš	1,453
9	Klingon	Marc Okrand	Klingon Language Institute	1,241
10	High Valyrian, Dothraki	David J. Peterson	Languages of Ice and Fire Community	966

last updated 2024-02-01

Conlangs by Reddit Community Membership

Rank	Conlang	Largest Reddit Community	Members
1	Esperanto	Esperanto, la internacia lingvo	29,000
2	Toki Pona	lipu Wesi pi toki pona	19,000
3	Sindarin	Sindarin	7,500
4	Quenya	Quenya	7,300
5	Lojban	Lojban News	4,100
6	Ithkuil	A grammatically complex constructed language	3,200
7	High Valyrian	Valyrio Udrir	2,500
8	Klingon	tlhIngan Hol	2,200
9	Interslavic	Medžuslovjanski Jezyk	1,400
10	Solresol	Solresol	1,100

last updated 2024-01-30

See any errors? Contact Sonja Lang on Discord or at sonjaaa@gmail.com

Figure 169. A comparison of popular constructed languages online by size of largest community on Discord and Reddit, with toki pona highlighted in red (Lang 2024a).

Figure 170. Graph of members of the ma pona pi toki pona Discord server over time. Screenshot from January 2024. Note that the discrepancy between this graph and the figure below is due to the server insights feature excluding those who choose to opt out, and the raw server membership count including them.

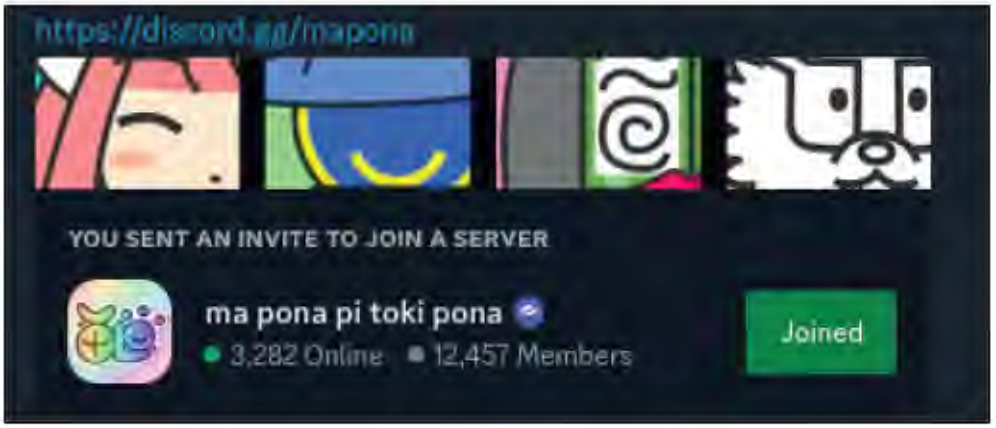


Figure 171. Number of members of the ma pona pi toki pona Discord server. Screenshot from January 2024.

Figure 172. Graph of members of the kama sona Discord server over time. Screenshot from January 2024.

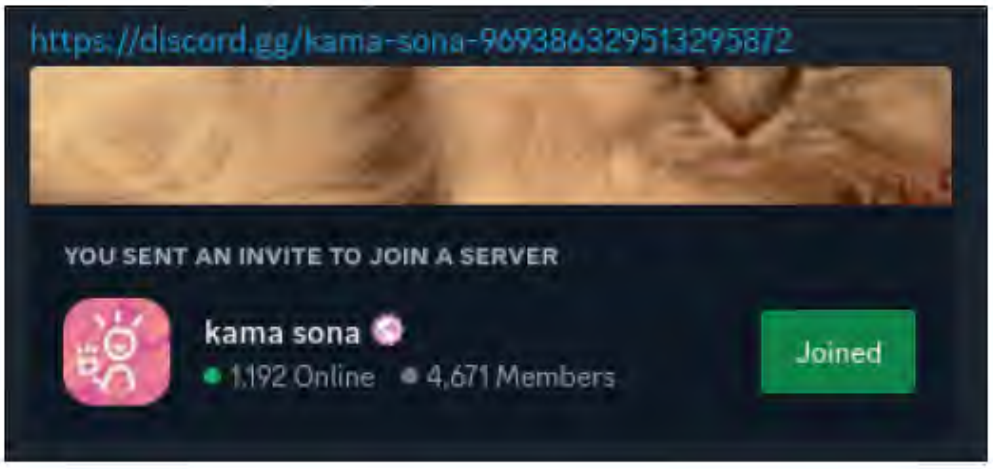


Figure 173. Number of members of the kamasona Discord server. Screenshot from January 2024.

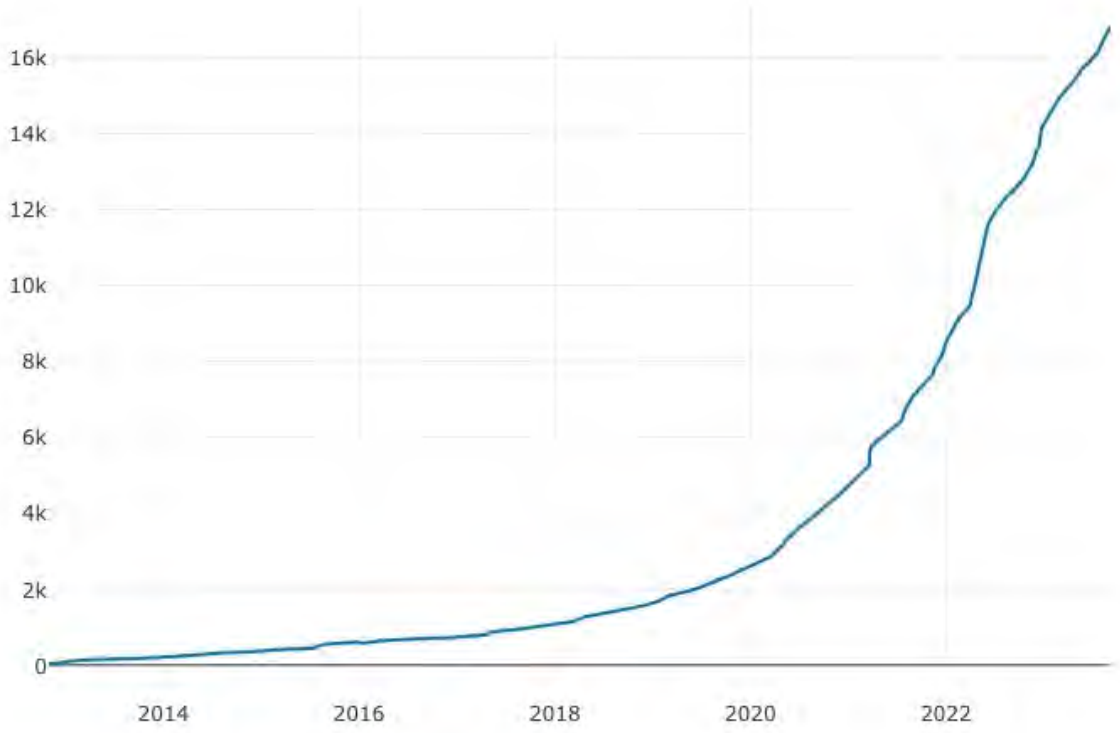


Figure 174. Graph of subscribers to the r/tokipona subreddit over time. Screenshot from December 2023.

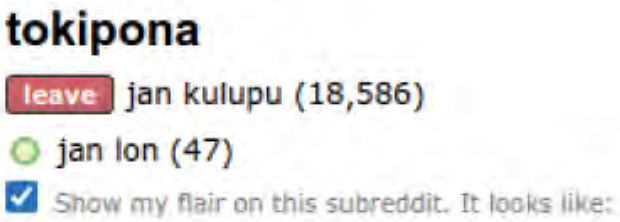


Figure 175. Number of subscribers ("jan kulupu") to the r/tokipona subreddit. Screenshot from January 2024.

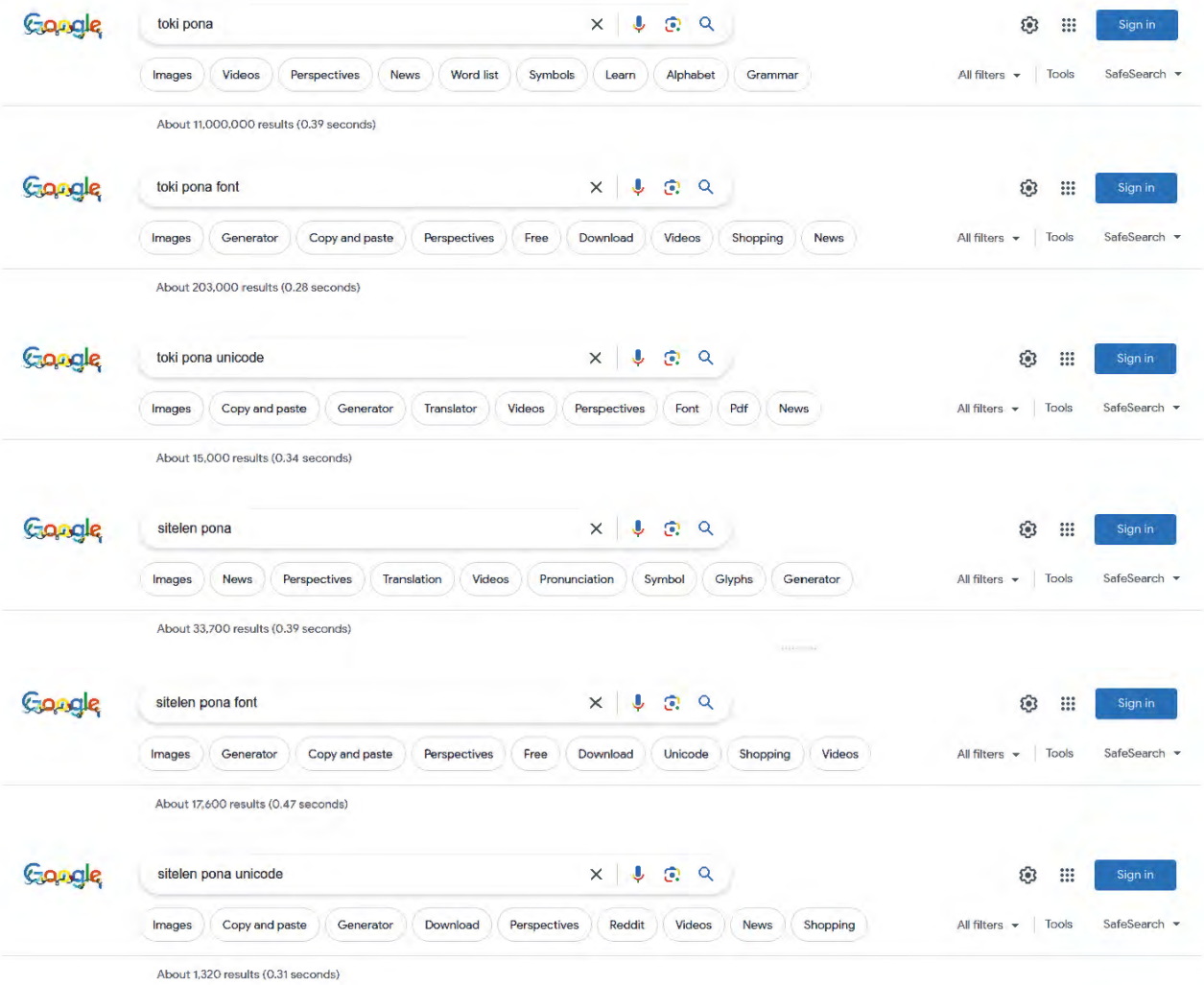


Figure 176. Google search results as of December 2023 for search terms related to sitelen pona.

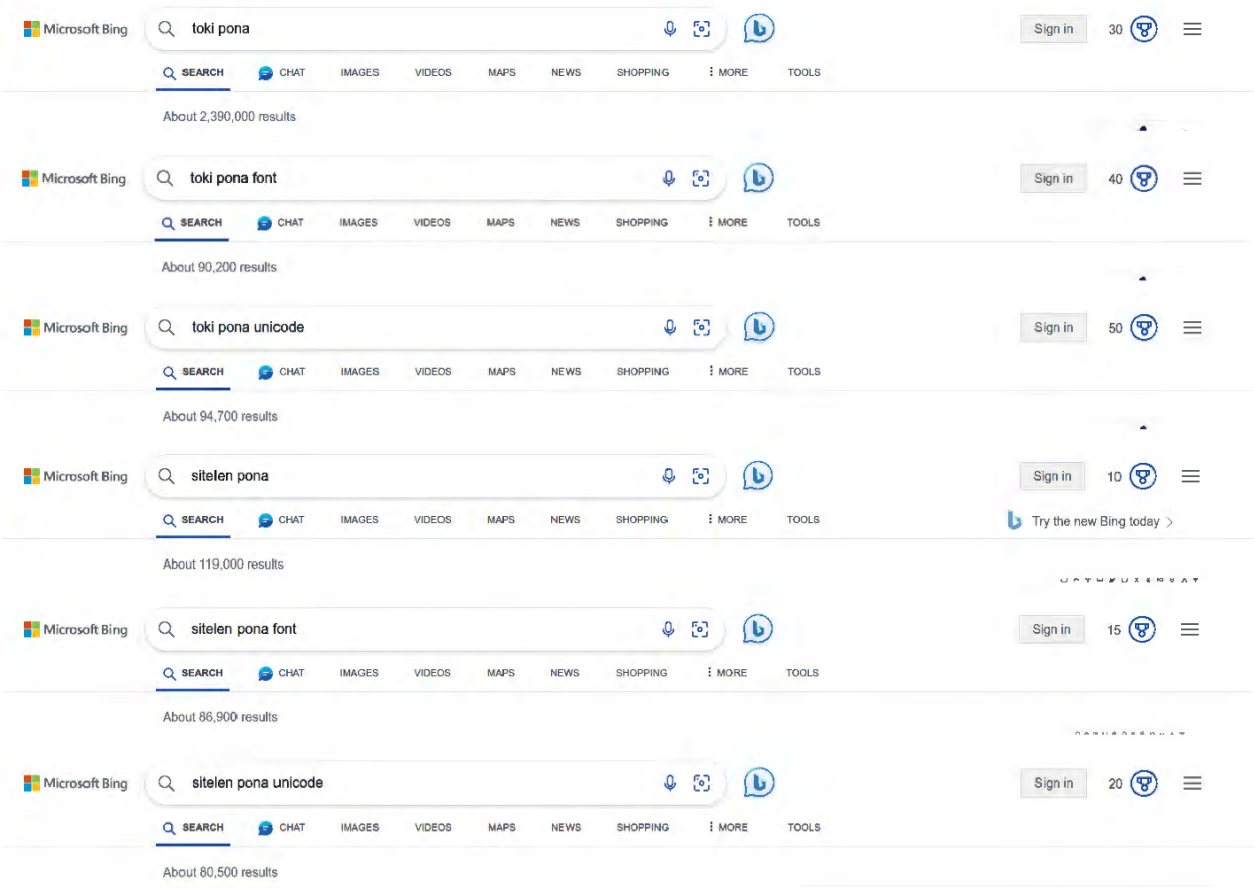


Figure 177. Bing search results as of December 2023 for search terms related to sitelen pona.

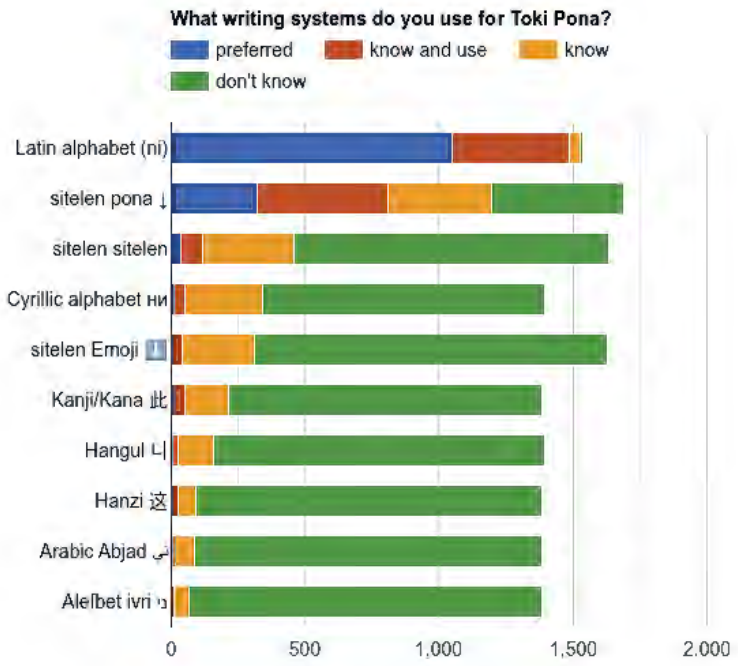


Figure 178. Results from the English version of the 2022 toki pona census. About 48% of respondents use the sitelen pona script. About 70% at least know the sitelen pona script.

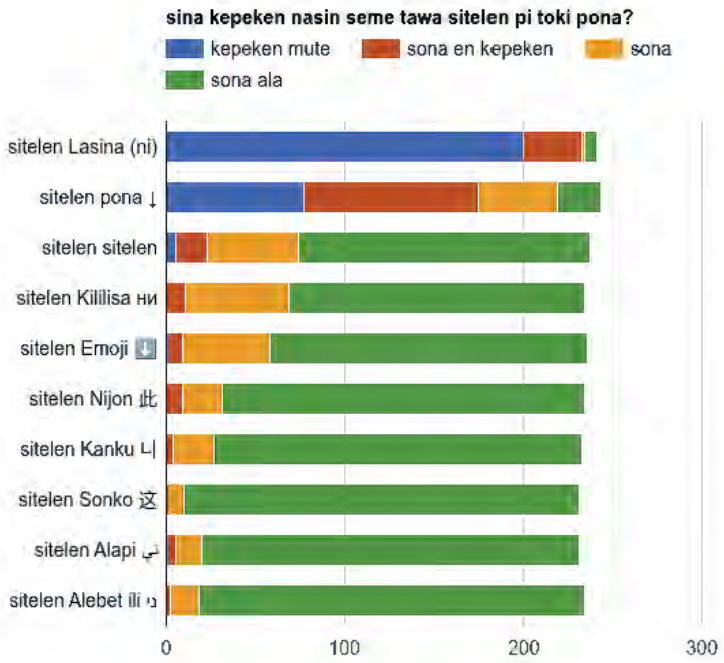


Figure 179. Results from the toki pona version of the 2022 toki pona census. About 72% of respondents use the sitelen pona script. About 90% at least know the sitelen pona script.

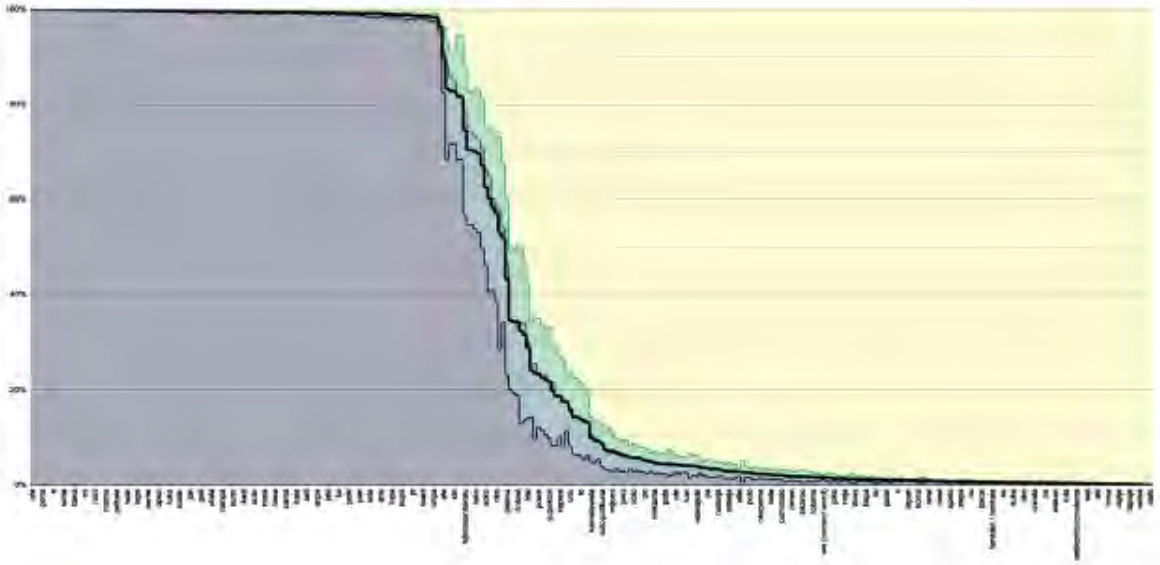


Figure 180. Full results of the 2023 toki pona word usage survey, sorted by usage. The colors from blue to yellow represent the response types, from “I use it” to “I never use it,” while the black line is the average response, aka “usage score.”

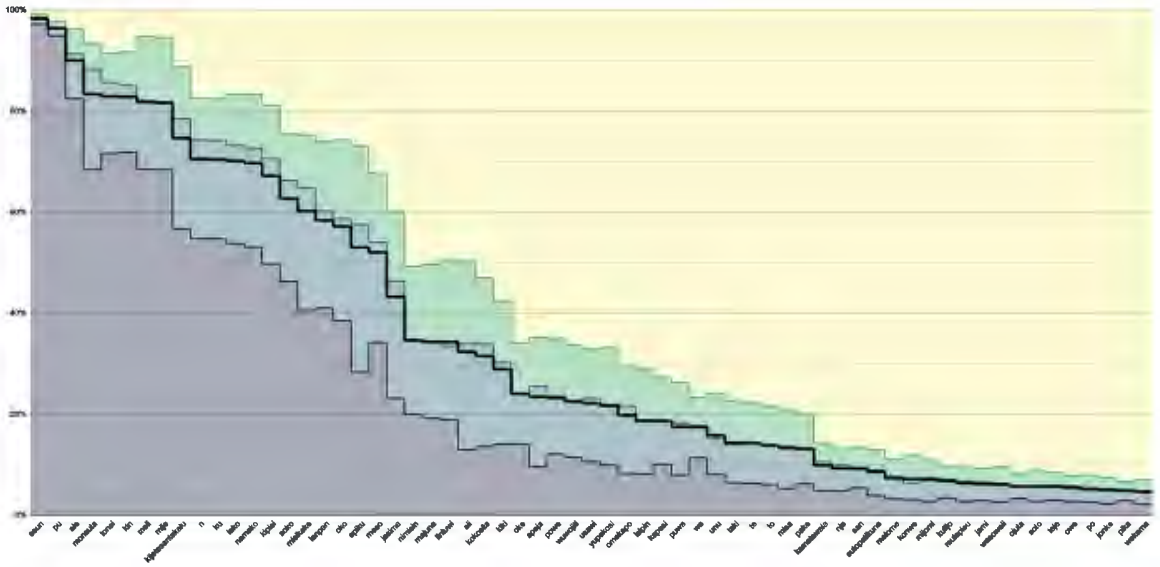


Figure 181. Results of the 2023 toki pona word usage survey, zoomed into the curve between the words *esun* (98%) through *wekama* (5%).

Widespread words (70-90%)

10 words:

```
kin              83%
monsuta          83%
tonsi            83%
meli             82%
mije             82%
kijetesantakalu  75%
ku               70%
leko             70%
n                70% (up from common)
namako           70%
```

Common words (50-70%)

7 words:

```
kipisi    67% (down from widespread)
soko      63% (down from widespread)
misikeke  60%
lanpan    58%
oko       57%
epiku     53% (up from uncommon)
meso      52%
```

Uncommon words (20-50%)

14 words:

```
jasima    43%
nimisin   35% (up from rare)
linluwi   34%
majuna    34%
ali       32%
kokosila  31%
kiki      29%
oke       24% (up from rare)
apeja     23%
powe      23%
usawi     22% (up from rare)
wuwojiti  22% (up from rare)
yupekosi  22% (up from rare)
omekapo   20% (up from rare)
```

Rare words (10-20%)

11 words:

```
isipin      19%
kapesi      19%
puwa        17%
wa          17%
unu         16%
taki        14%
te          14%
to          14%
misa        13%
pake        13%
kamalawala  10% (up from obscure)
```

Figure 182. Detailed usage statistics for the words *kin* through *kamalawala*.

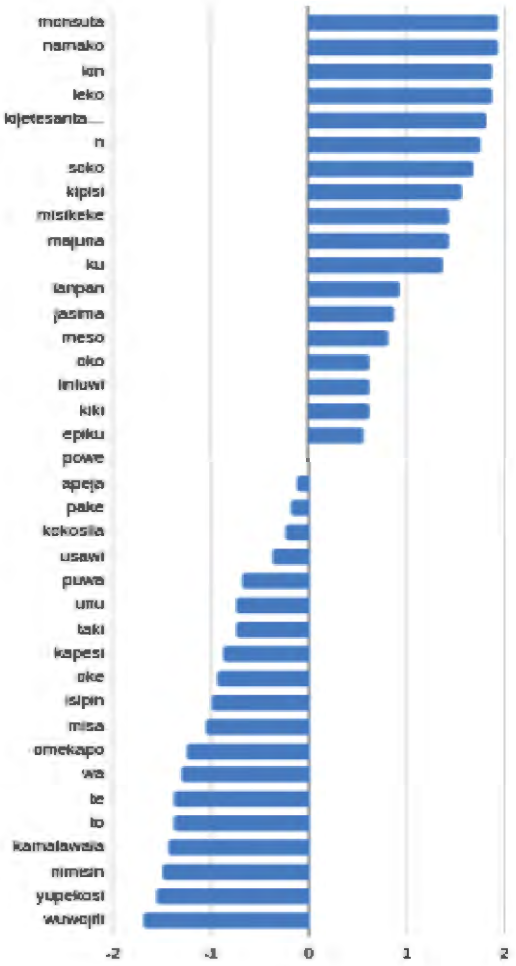


Figure 183. Results of a survey among the proposal authors about which words to include. The horizontal axis shows the average score obtained by each word, with positive scores in favor of encoding. The words *monsuta* through *epiku* are heavily favored, *powe* through *kokosila*
lightly disfavored, and *usawi* through *wuwojiti* heavily disfavored.

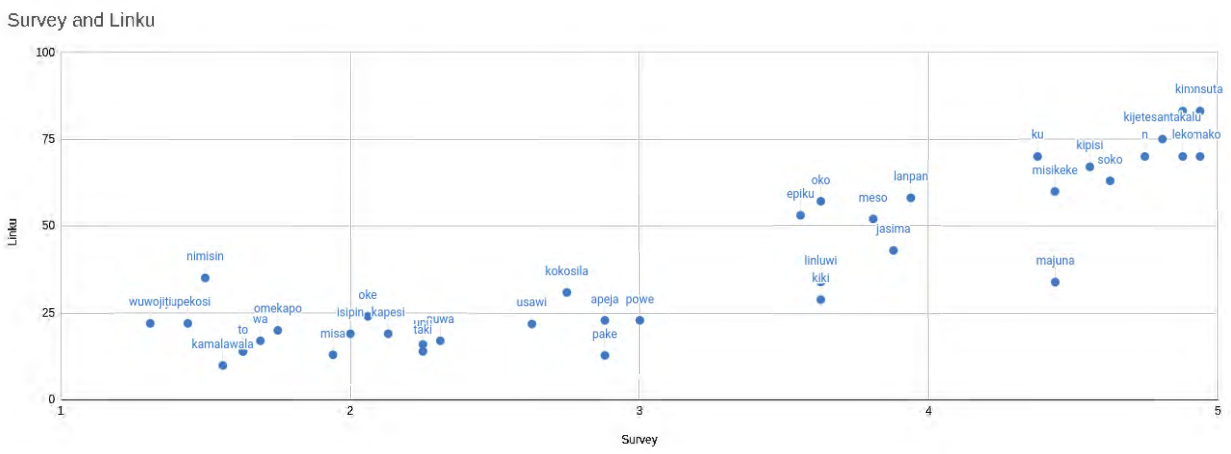
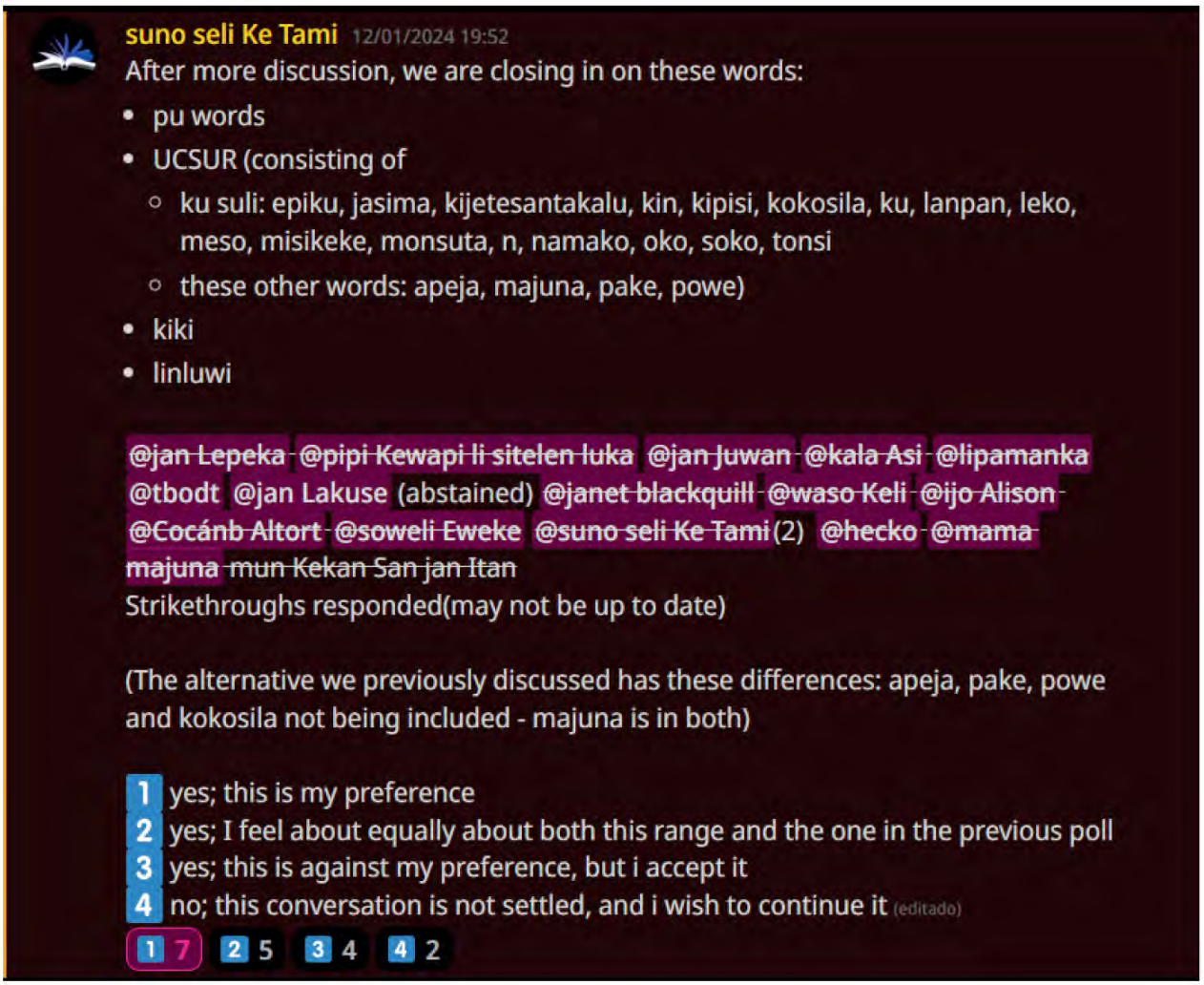


Figure 184. Results of the survey plotted against Linku usage categories.



suno seli Ke Tami 12/01/2024 19:52

After more discussion, we are closing in on these words:

- pu words
- UCSUR (consisting of
  - ku suli: epiku, jasima, kijetesantakalu, kin, kipisi, kokosila, ku, lanpan, leko, meso, misikeke, monsuta, n, namako, oko, soko, tonsi
  - these other words: apeja, majuna, pake, powe)
- kiki
- linluwi

~~@jan Lepeka~~ ~~@pipi Kewapi li sitelen luka~~ ~~@jan Juwan~~ ~~@kala Asi~~ ~~@lipamanka~~ @tbodt @jan Lakuse (abstained) ~~@janet blackquill~~ ~~@waso Keli~~ ~~@ijo Alison~~ ~~@Cocánb Altort~~ ~~@soweli Eweke~~ ~~@suno seli Ke Tami~~ (2) ~~@hecko~~ ~~@mama majuna~~ ~~mun Kekan San jan Itan~~

Strikethroughs responded(may not be up to date)

(The alternative we previously discussed has these differences: apeja, pake, powe and kokosila not being included - majuna is in both)

1 yes; this is my preference
2 yes; I feel about equally about both this range and the one in the previous poll
3 yes; this is against my preference, but i accept it
4 no; this conversation is not settled, and i wish to continue it (editado)

1 7 2 5 3 4 4 2

Figure 185. Results of a final poll about whether to include *apeja*, *pake*, *powe*, and *kokosila*.

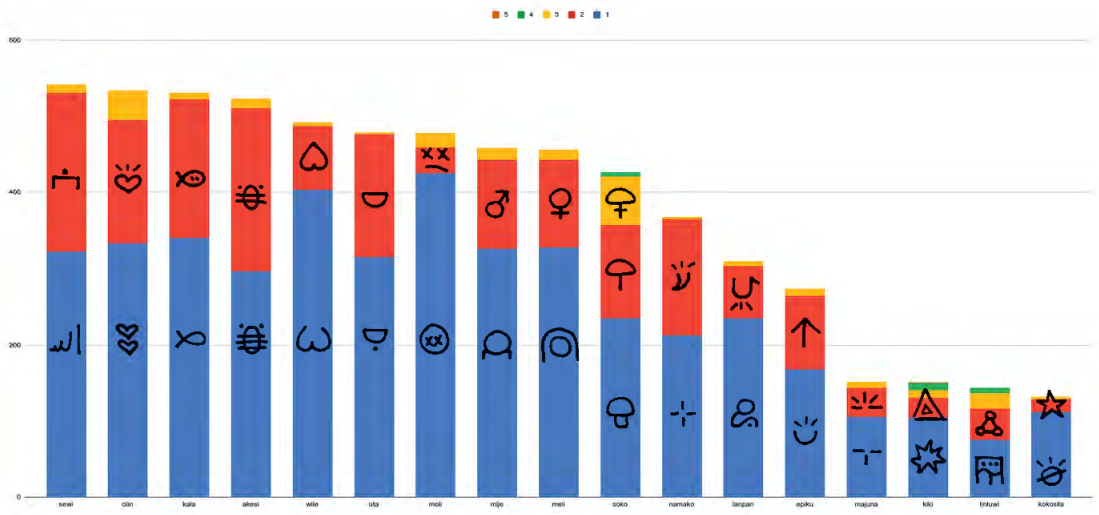


Figure 186. Results of a usage survey of variant glyphs in sitelen pona.
The vertical axis represents the number of respondents who use a particular glyph.

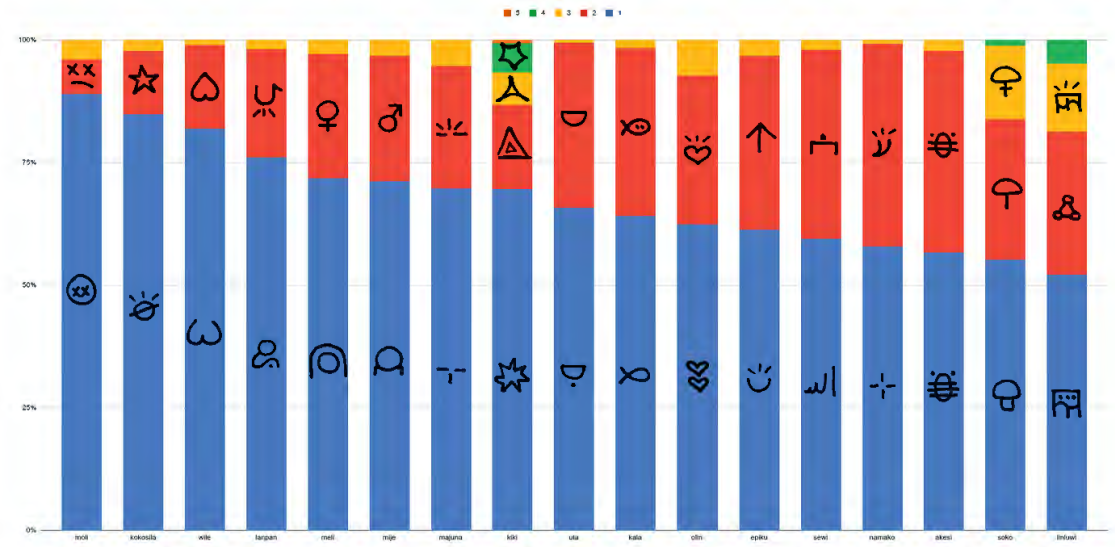


Figure 187. Results of a usage survey of variant glyphs in sitelen pona. The vertical axis represents the percentage of users of a glyph among users of the word the glyph represents.

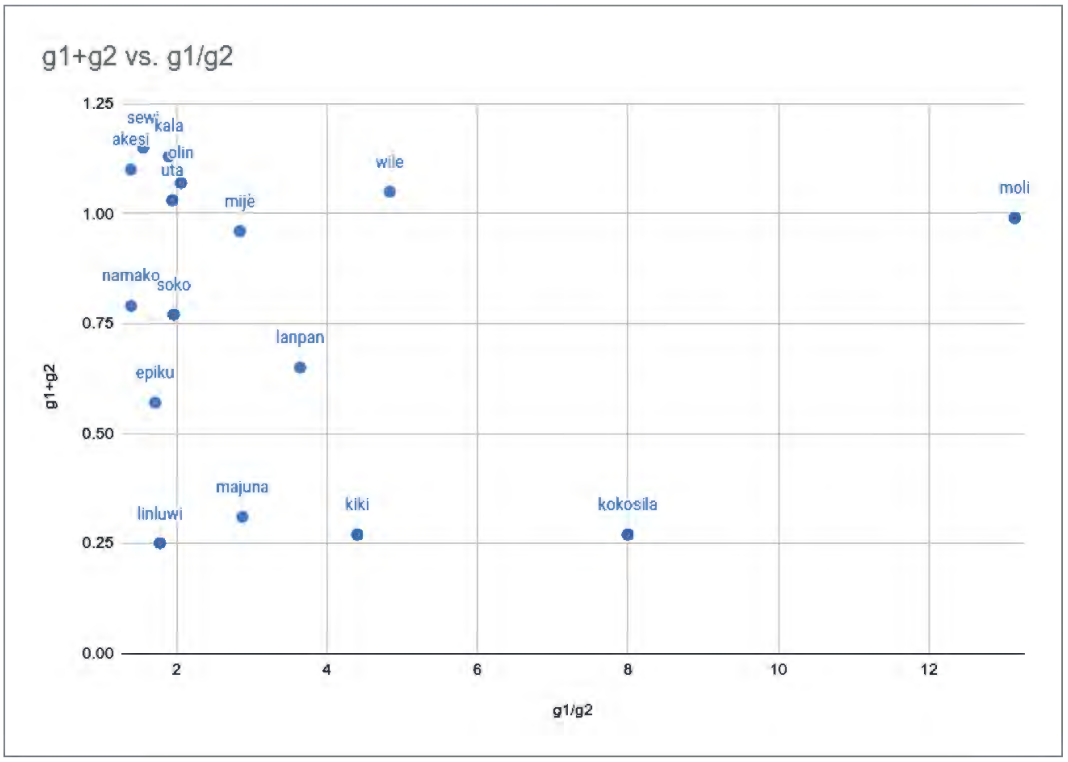


Figure 188. The spread of variant glyphs (number of users of the most popular variant glyph divided by number of users of the second most popular variant glyph) plotted against total usage of the represented word (number of users of any variant glyph).

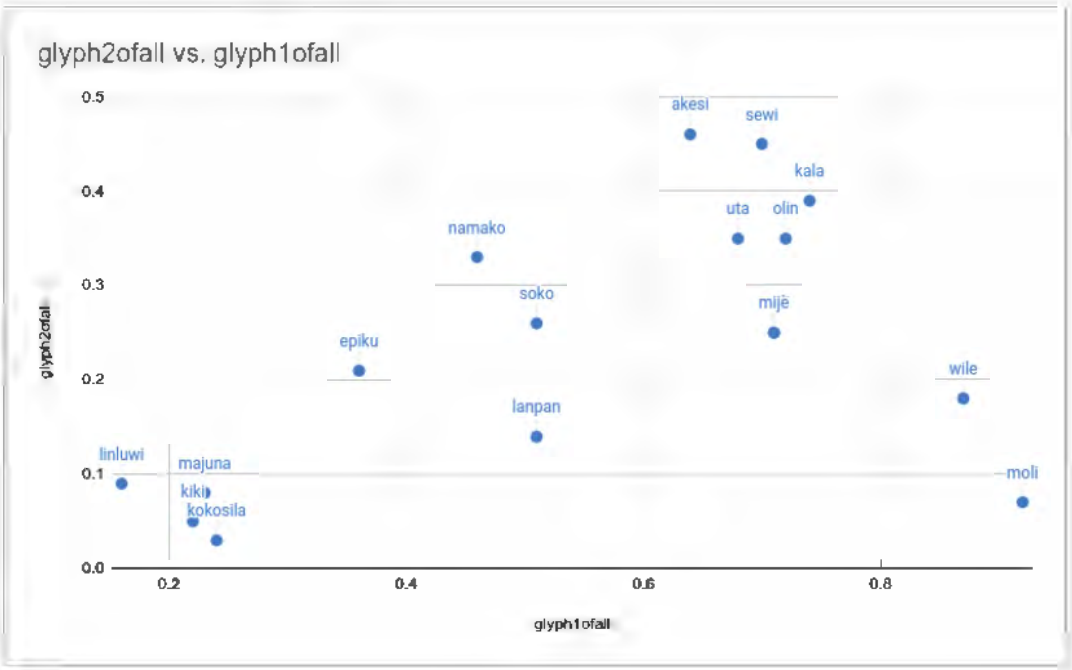


Figure 189. Number of users of the most popular variant glyph plotted against number of users of the second most popular variant glyph.

Figure 190. Results of an informal poll about variants of *sewi* and *namako*. Most respondents do not mind which variant is used (the heart-shaped responses). However, a handful of respondents *will not use a word* if their preferred variant is not available (the square and circle responses). This is a more divisive issue for *namako* (29 square and circle responses)

than for *sewi* (22 square and circle responses); this is a particularly curious result that may warrant further study.

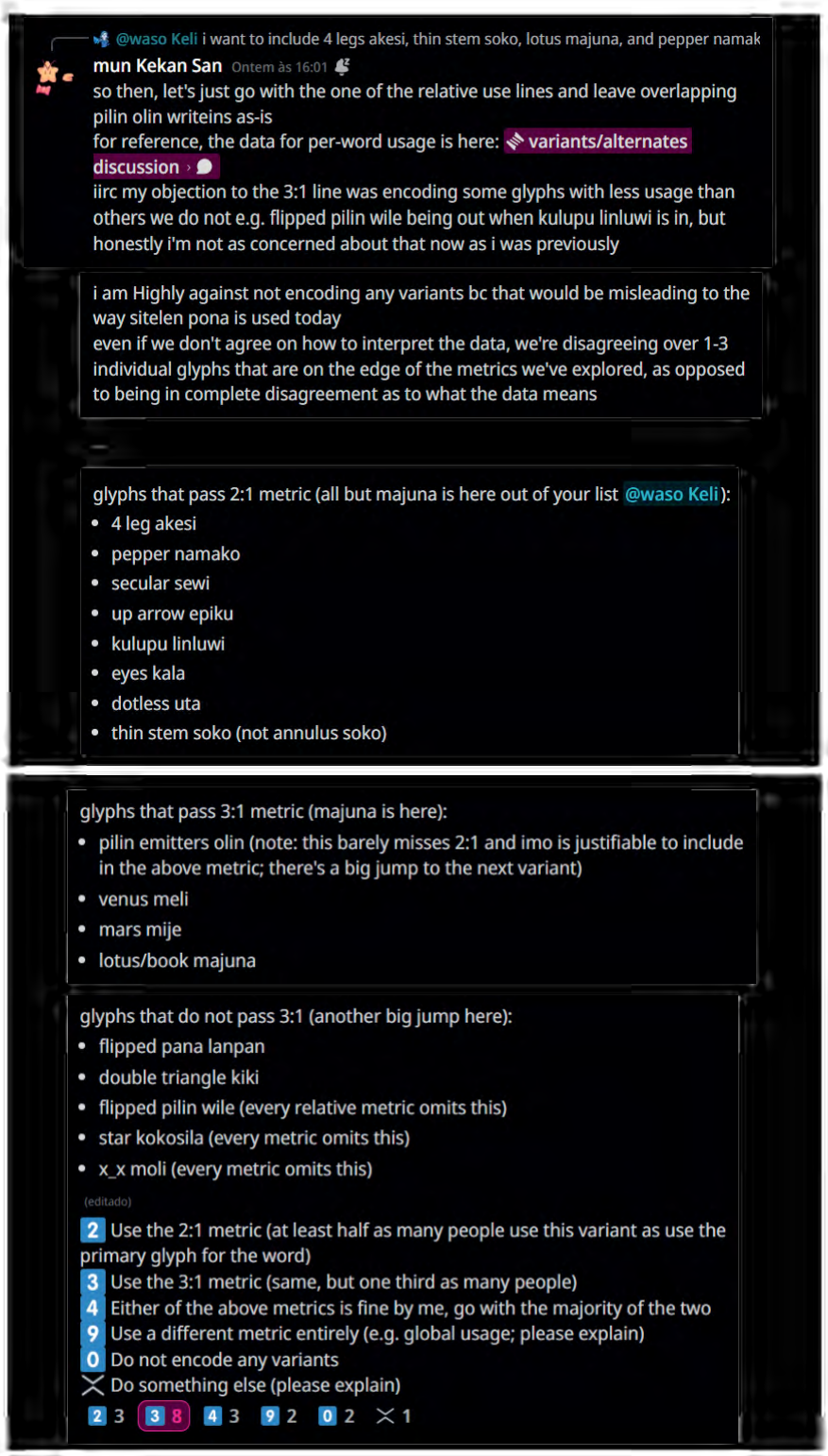


Figure 191. Results of a poll among the authors about which glyph variants to encode. Most agreed to encode variants used by at least one third of the number of people who use the most popular variant.

Variable	Count	Average	Median
[illegible]	152	1,38	1
[illegible]	152	2,28	2
[illegible]	152	2,41	2
Total (3)	456	2,00	2

Variable	Count	Average	Median
[illegible]	152	1,54	1
[illegible]	152	2,49	2
[illegible]	152	3,07	3
[illegible]	152	3,74	4
[illegible]	152	4,15	5
Total (5)	760	3,00	3

Variable	Count	Average	Median
[illegible]	152	4,37	1
[illegible]	152	4,68	4
[illegible]	152	5,05	3
[illegible]	152	5,12	4
[illegible]	152	5,72	6
[illegible]	152	6,23	6
[illegible]	152	7,31	7
[illegible]	152	7,38	8
[illegible]	152	7,49	8
[illegible]	152	7,74	8
[illegible]	152	8,16	8
[illegible]	152	8,74	11
Total (12)	1.824	6,50	7

Variable	Count	Average	Median
[illegible]	152	2,38	1
[illegible]	152	2,89	2
[illegible]	152	4,30	3
[illegible]	152	4,97	5
[illegible]	152	5,26	4
[illegible]	152	5,63	6
[illegible]	152	6,38	7
[illegible]	152	6,99	8
[illegible]	152	7,81	9
[illegible]	152	8,39	10
Total (10)	1.520	5,50	6

Figure 192. Results of the collation survey, sorted by average (mean) score.

Bucketing

Variable	Count	Average	Median
𓄿𓄿𓂋𓂋𓎛	152	1,38	1
𓄿𓄿𓂋𓂋𓎛	152	2,41	2
𓄿𓄿𓂋𓂋𓎛	152	2,20	2
Total (3)	456	2,00	2

Bucketing

Variable	Count	Average	Median
𓄿𓂋	152	1,54	1
𓄿𓄿𓂋𓂋𓎛𓏏	152	2,49	2
𓄿𓂋𓂋𓏏	152	3,07	3
𓄿𓄿𓂋𓂋𓎛𓏏	152	3,74	4
𓄿𓂋𓂋𓏏	152	4,15	5
Total (5)	760	3,00	3

Bucketing

Variable	Count	Average	Median
𓄿𓂋𓂋𓎛	152	4,37	1
𓄿𓄿𓂋𓂋𓎛	152	4,68	4
𓄿𓂋𓂋	152	5,05	3
𓄿𓂋𓂋𓏏	152	5,12	4
𓄿𓂋𓂋𓂋	152	5,72	6
𓄿𓂋𓂋𓂋𓄿	152	6,23	6
𓄿𓂋𓂋𓂋𓂋	152	7,31	7
𓄿𓂋𓂋𓂋𓄿	152	7,38	8
𓄿𓂋𓂋𓂋𓂋𓎛	152	7,49	8
𓄿𓂋𓂋𓂋	152	7,74	8
𓄿𓂋𓂋𓂋𓄿𓎛	152	8,16	8
𓄿𓂋𓂋𓂋𓂋𓂋	152	8,74	11
Total (12)	1.824	6,50	7

Bucketing

Variable	Count	Average	Median
𓄿𓂋𓂋𓏏	152	2,38	1
𓄿𓂋𓂋𓂋	152	2,89	2
𓄿𓂋𓂋	152	4,30	3
𓄿𓂋𓂋	152	5,26	4
𓄿𓂋𓂋𓂋𓏏	152	4,97	5
𓄿𓂋𓂋𓂋𓂋	152	5,63	6
𓄿𓂋𓂋𓂋𓂋𓂋	152	6,38	7
𓄿𓂋𓂋𓂋𓂋	152	6,99	8
𓄿𓂋𓂋𓂋𓂋𓂋𓂋	152	7,81	9
𓄿𓂋𓂋𓂋𓂋𓂋𓂋𓂋	152	8,39	10
Total (10)	1.520	5,50	6

Figure 193. Results of the collation survey, sorted by median score.

A. Administrative

1. Title
Preliminary proposal to encode sitelen pona in the UCS
2. Requester's name
Sitelen Pona Working Group (Rebecca Bettencourt et al.)
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2024-04-16
5. Requester's reference (if applicable)
6. Choose one of the following:
6a. This is a complete proposal
No.
6b. More information will be provided later
Yes.

B. Technical - General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
Yes.
1b. Proposed name of script
Sitelen Pona.
1c. The proposal is for addition of character(s) to an existing block
No.
1d. Name of the existing block
2. Number of characters in proposal
153.
3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)
Category A.
4a. Is a repertoire including character names provided?
Yes.
4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
Yes.
4c. Are the character shapes attached in a legible form suitable for review?
Yes.
5a. Who will provide the appropriate computerized font (ordered preference: TrueType, or PostScript format) for publishing the standard?
Rebecca Bettencourt.
5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Rebecca Bettencourt, FontForge.
6a. Are references (to other character sets, dictionaries, descriptive texts, etc.) provided?
Yes.
6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.
7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration, etc. (if yes please enclose information)?
Yes.
8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.
See above.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
Documents about sitelen pona have been submitted by others in 2021 and 2023; however, they were not complete or proper proposals.
2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
Yes.
2b. If YES, with whom?

***ma pi nasin sitelen* (Discord group), *kulupu Linku* (Discord group), *ma pona pi toki pona* (Discord group), *Ma Toki Pona VR* (VRChat group), r/tokipona (Reddit group), toki pona groups on Facebook and Telegram, Sonja Lang (creator of the language)**
2c. If YES, available relevant documents
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?
Thousands of speakers of toki pona worldwide.
4a. The context of use for the proposed characters (type of use; common or rare)
Common.
4b. Reference
5a. Are the proposed characters in current use by the user community?
Yes.
5b. If YES, where?
Worldwide, particularly in North America, Europe, and Asia. Online, particularly on Reddit, Facebook, and Discord.
6a. After giving due considerations to the principles in the P&P document, must the proposed characters be entirely in the BMP?
No.
6b. If YES, is a rationale provided?
6c. If YES, reference
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
Yes.
8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
No.
8b. If YES, is a rationale for its inclusion provided?
8c. If YES, reference
9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?
No.
9b. If YES, is a rationale for its inclusion provided?
9c. If YES, reference
10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
Yes.
10b. If YES, is a rationale for its inclusion provided?
Yes.
10c. If YES, reference
Included in proposal (see Sections 4 and 4d).
11a. Does the proposal include use of combining characters and/or the use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1:2000)?
No.
11b. If YES, is a rationale for such use provided?
11c. If YES, reference
11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
11e. If YES, reference
12a. Does the proposal contain characters with any special properties such as control function or similar semantics?
Yes.
12b. If YES, describe in detail (include attachment if necessary)
Included in proposal (see Section 4f).
13a. Does the proposal contain any Ideographic compatibility character(s)?
No.
13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?