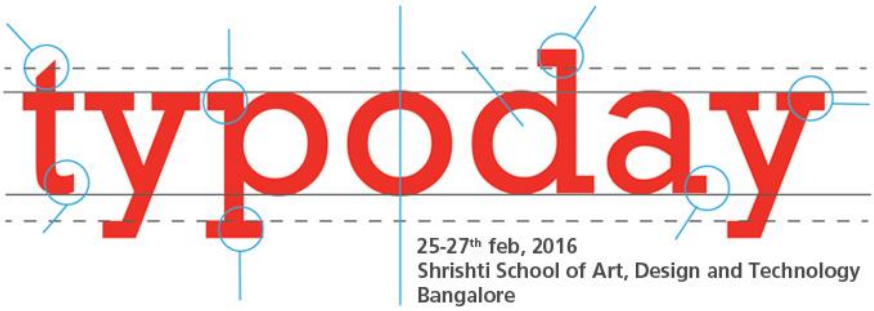


Typography and Education
http://www.typoday.in

"Idital"- Digitization and revival of tribal script, Sora Sompeng

Sony Salma Priyadarshini, Indian Institute of Technology Bombay,India, ssalmap@gmail.com,

Abstract: Sora Sompeng is the script of Sauras, a tribal group predominant in Odisha and Andhra Pradesh. Saura mostly uses Odia and Telugu to write their language. This has pushed Sora to the verge of extinction. Recently, a growth of interest in Sora Sompeng shows that there is a scope for reviving and digitizing the script.

This paper describes the initial efforts to analyze and understand Sora Sompeng in order to make a suitable typeface for the script. It discusses the various aspects of the script, its visual features and the design decisions taken towards achieving this.

Key words: *Sora Sompeng, Sauras, Tribal Script, Digitization, Type design, Monolinear*

1. Introduction

Tribals are people who have sustained their way of life for many generations, living in communities that are largely self-sufficient, and are clearly different from the mainstream and dominant society. In India, tribals (also called adivasis) account for about 8.6 percent of the population — majority of them residing in six states: Madhya Pradesh, Chhattisgarh, Maharashtra, Orissa, Jharkhand and Gujarat.

These tribals are commonly perceived as a fragment of our past, rather than a participant in the present and the future. They are viewed as backward, pre-literate people — frozen in time and incapable of change. This viewpoint does not reflect the truth as tribal societies are constantly evolving and adapting, like the mainstream society. They are

important, beyond the exotic nature of their dresses and their dances, for the vast amount of accumulated knowledge they carry. This knowledge is not only related to the forest, its flora and fauna; but also their history, myths and stories, society, medicine and livelihood. For the preservation of this knowledge, it is important for it to be documented by a proper script that captures the nuances of the language. Realizing this, many tribals have started creating new scripts for their traditional languages. This recent phenomenon can be seen in many examples such as; Warang Chiti, the script of the Ho tribe (who live near the borders of Jharkhand and Odisha) is around 50 years old; Tolong Siki, script of the Oraon Tribe (who are spread throughout the eastern part of India) is around 25 years old. More recently, in 2015 Mishmi tribes of Arunachal Pradesh — Idu, Kman and Taraon — with the help of linguist Roger Blench are developing a new writing system built on the Roman alphabet but expanded by characters and diacritics [1].

Most tribals live in the remote pockets of the forest but there are few communities who live very close to urban centers. The latter are especially economically dependent on these centers for their livelihood, are influenced by them in their language and culture. The remote tribals mostly retain their traditional ways of living. As we progress and connect more and more communities, it is important to protect the sensitive tribal cultures and languages before they get eroded by the dominant culture. Various efforts have been effective to achieve this, one of which is making sure that their language can have an effective script.

2.2 Sauras and their script, Sora Sompeng

The Sauras (alternative names and spellings include Saora, Sora, Savara and Sabara) are a tribe from Southern Odisha, north coastal Andhra Pradesh in India. They are also inhabitant the hills of Jharkhand, Madhya Pradesh and Maharashtra [2]. At home they speak their mother tongue Saura which is an Austro-Asiatic language.

Sauras have used other non-native dominant scripts before developing their own script. For instance in Odihsa, the Odia script was used while communicating with others. In 1936, Mangei Gomango, a prominent Saura scholar, created the script for Sauras. He broke down the language into 24 letters and 10 digits. This script was named Sora Sompeng and was dedicated to Akshara-Bramha(Alphabet deity of Sauras) [3].

Later on an organization named Matarvanam Vignyan Prachar Ashram(MVPA) established a block printing press in Dombasara, Gunupur, Odisha. It printed a limited number of books and other documents in past. But since then, reduced exposure and minimal involvement of the state in promoting the script has resulted in demand for the press dying down and its eventual closure. Revival efforts for the script were blocked by the unavailability of printing machines. With the advent of new technologies, more possibilities have opened up. Digital printers are accessible today and personal printers are commonplace. The

digitization of the script allows for publication of educational, cultural and other documents without the need for specialized printing presses.

The “Idital” project was started by us to create a font for the Sora script in response to the increased availability of technology within the Saura people. “Idital” is the main deity of the Sauras to whom their traditional mural paintings (“*Italons*”) are dedicated. The project was started with the help and support of experts of tribal language and activists with the aim of creating and popularizing a font for the Saura language.

2. Classification

Sora language belongs to the class of Munda family. Gregory Anderson classifies Sora under the South Munda Branch [4]. Sora speakers are concentrated mainly in the districts of Ganjam, Gajapati and Rayagada in Southern Odisha. Around 300,000 people speak in Sora. It is mostly an oral language and is not written or read extensively. Most texts in the recent past were written using Oriya and Telugu scripts. Though thriving in some areas, it is considered to be threatened overall.

2.1 The Character Set

Sora is composed of 25 letters and 10 digits [Figure 1]. There 18 consonants and 7 vowels in the alphabet. The letters in Sora Sompeng are named after the deities in the Sora pantheon. It is an Abiguda (alphasyllabary) with each consonant has an inherent vowel /a/. Separate letters for vowels are used to change or mute the inherent vowel.

The following are the consonants, vowels and digits of Sora script-

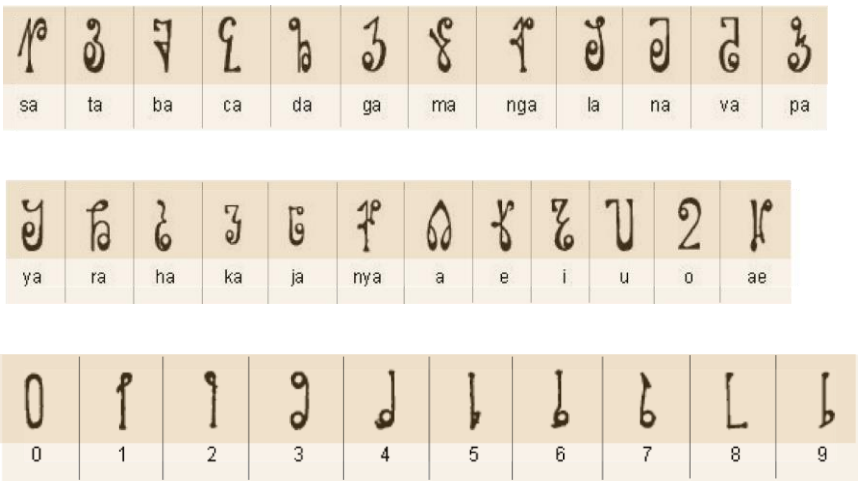


Figure.1 Character and Digit set

3 Process

3.1 Collection of Reference Material

Availability of authentic reference materials was one of initial challenges of the project. Looking at the letterpress books printed in Sora Script in the past, one can get a glimpse of the immense drive and enthusiasm to promote the script. This has sadly waned down in present times. Most of the books printed at Dombsara printing press have been lost in the course of time. An almost worn-out letterpress school book was used as the primary reference [Figure 2]. The primitives for the font were derived from the modulated font found in the book.

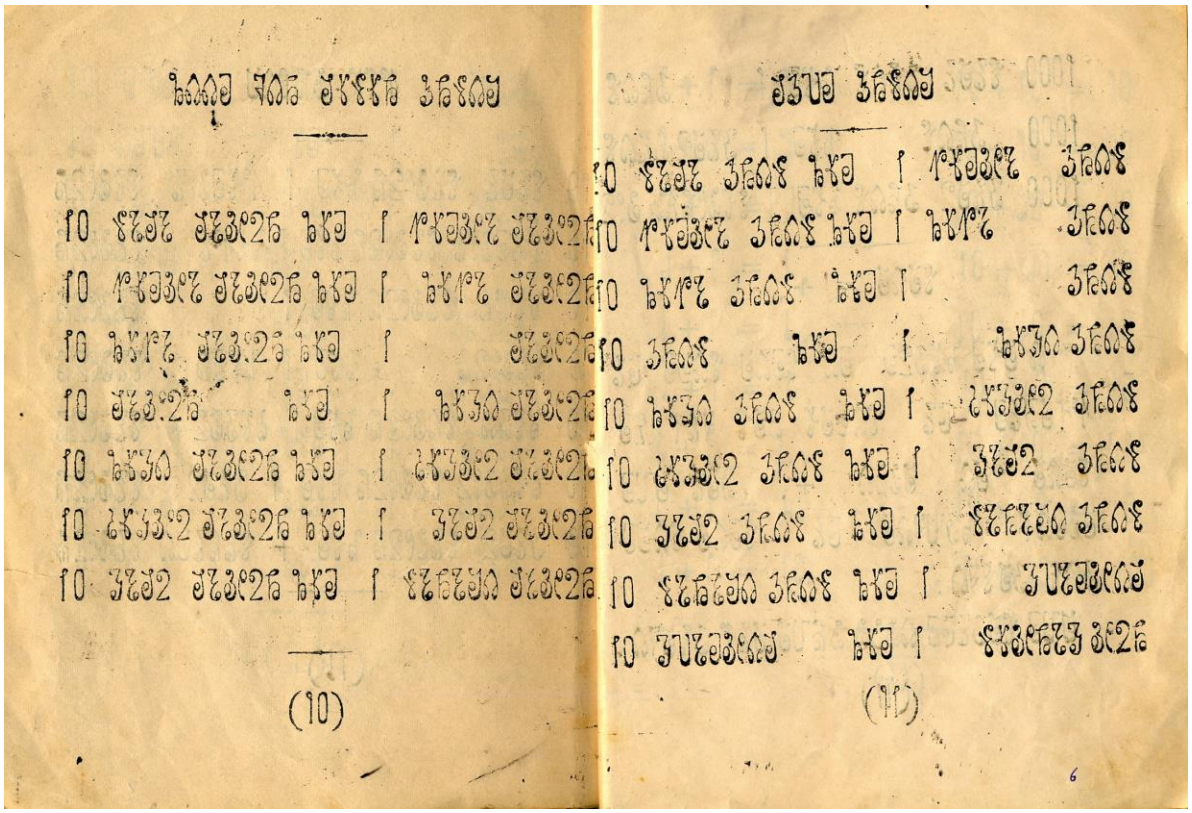


Figure.2 Page from reference book

3.1 Categorization

Characters were categorized into sets based on the similarities of their visual structures. This grouping helped in maintaining harmony across the characters sharing similar visual grammar.

The beauty and challenge in taking a letterpress book as reference is that each bed is manually crafted and composed. From choosing the typeface, picking out letters, kerning, inking the press, inking the letter bed, applying pressure to transfer the print onto a paper etc.is done manually. Manual process has inconsistencies unlike digital printing. We found many such inconsistencies in our reference book. In order to accommodate these inconsistencies multiple references for each character were considered. Having multiple interpretations of a character made the study, intense and extensive. Several characteristics from multiple references of the font were studied; for example the

extension of curve, the movement of bowl, the length of stem. Studying the characteristics from these references made the process of making the font easier.

We looked into two kinds of categorization, one using morphological elements within a letter (prominent parts of the characters) and the other by considering the entire character.

3.2.1 Morphological categorization

Morphological categorization is based on prominent graphical elements shared by characters [Figure 3]. Few of the identified elements are vertical bar, slants, knots and loops, with loops being the most frequent. This was done to understand how few elements can give a sense of agreement between many characters. Even though the same element might have different behavior across characters, figuring out how an element behaves helps us understand characters that share similarities.

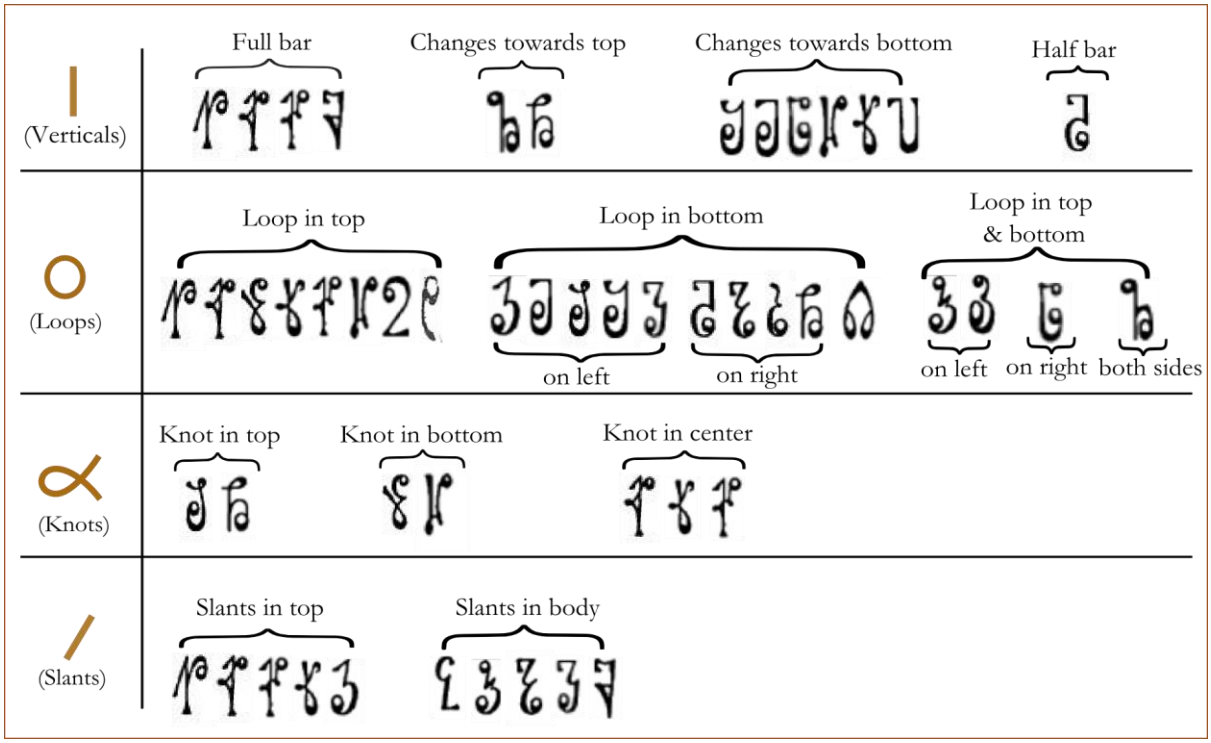


Figure.3 Morphological Categorization

3.2.2 Character wise categorization

This categorization was based on visual characteristics of entire character rather than the elements it contains [See figure 4]. Since element wise grouping considers only a specific detail, a lot of characters can be put together in a group, although the overall structure might not be similar. Character wise grouping considers morphological similarities but more importantly the similarities of the entire character with other characters.

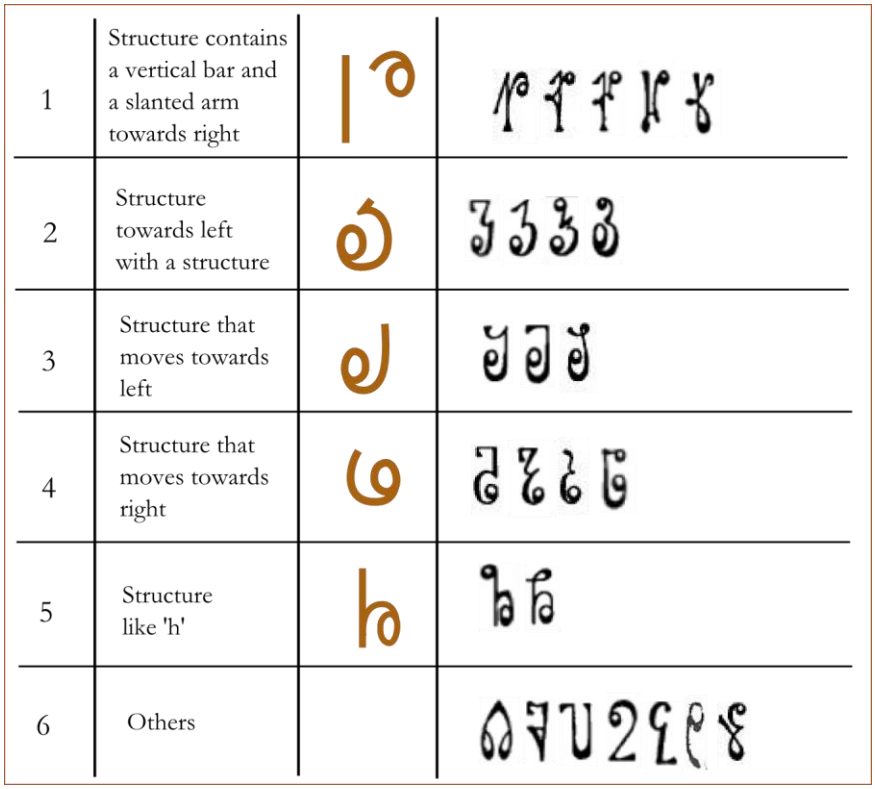


Figure.4 Character-wise categorization

3.3 Tools

The font of the reference material was written with a crow quill. The influence of Telugu can be seen in bottom heavy nature of the glyphs. Observations such as these helped us understand the local influences on the script.

Since this was the first attempt in reinterpreting the font, deciding an appropriate tool to capture the characteristics poses a challenge to the type designer. It is difficult to capture all the niceties the characters and restrict it to a single tool. Choosing the tool is an important decision that a designer takes. We explored with several tools for varied interpretations of the font [Figure 5]. Tools such as boru, flat tipped brush, round-tipped marker etc. were used.

Drawing only the skeletal structure of the characters and adding weight around it seemed to be the most appropriate solution to start out. It was during the tool exploration stage, we decided to forego all potential tool influences and design a mono-linear font. This eliminated the additional tool characteristics which might have come into play otherwise. It moved towards a direction with simplicity in forms and possibly helped us obtain the basic structural details of the original script.

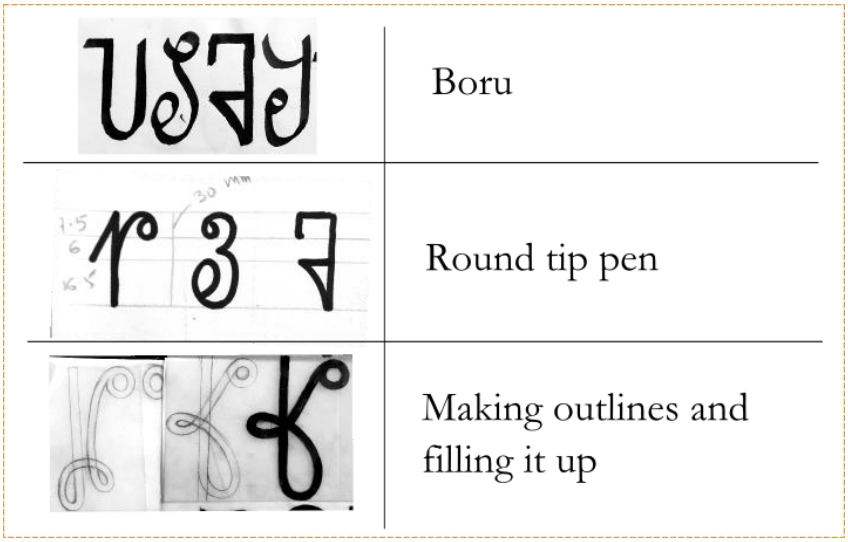


Figure.5 Explorations in different tools

3.3 Grid

A grid in typography is the framework that allows for arranging the font within a certain amount of space. It helps in establishing relationship between different elements within a font.

The Sora script does not have ascenders and descenders. Hence combinations of characters are placed side by side and sit within the same horizontal space. This makes the character height the essential factor in designing the grid. A grid based on the font height was designed for building a proportional skeletal structure of the characters.

Each character was measured and the font height was taken from the characters with vertical bar [Figure 6]. The total height of the character was 8mm as measured from the reference book. The heights at which most characters display distinct movements were measured and averaged. This was mapped into 10 unit height grid. The grid was divided into 3 parts; the top part was in 2.5 units, the middle part was 2 units and the bottom part was 5.5 units [Figure 6]. Hence the division of height in ratio of 2.5:2:5.5. All the characters and digits were made to fit into this grid. Although little flexibility was allowed, the character's stroke movements were appropriately restricted in this division of space. We also measured the height to stroke ratio and tried to maintain the same during the font making process.

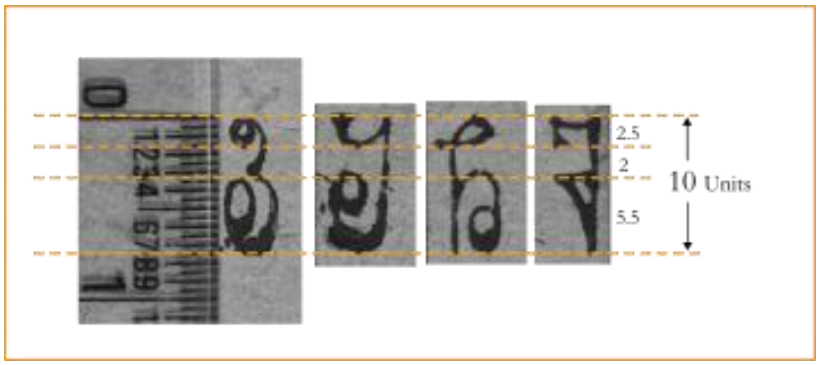


Figure.6 Grid design, the scale in the figure is in mm (left most)

3.4 Anatomical Analysis of Sora

There can be many interpretations of an element in a character. Each curve, knot, line and even the slightest projection of a character's element can open up possibilities for a font designer. We looked into the various ways in which we could detail out the elements. Although this can be an ongoing process we limited ourselves to a few due to constraint time.

A character is made up of many elements — curves, knots, lines etc. These elements were individually detailed out keeping in mind that it should also be in harmony with the other elements of the character.

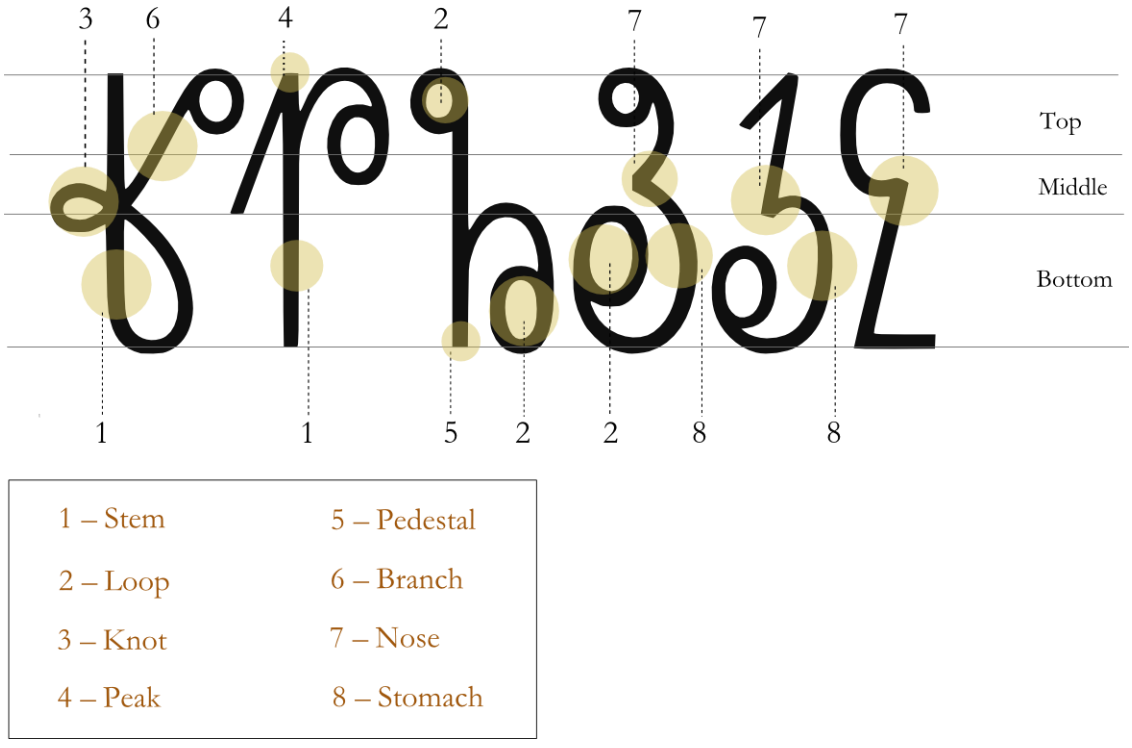


Figure.7 Anatomy of Sora Sompeng

3.4.1 Stem

Stem is an element which has a prominent vertical bar. Some characters such as Ba [Figure 8] have the stem throughout the body height, while Ra [Figure 8] and Na [Figure 8] have the stems transforming either at the top or the lower part of the character.

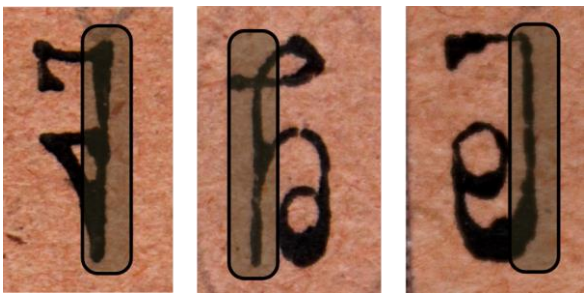


Figure.8 Stem in ‘ba’, ‘ra’ and ‘na’

Body of the stem has a fixed width but in characters where the stem had to transform to a curve, scaling was done to achieve visual balance. In character ‘na’ [Figure 9], the stem moves out towards to the left to be a curve in the lower half. Again scaling was done in this zone for a smooth transition from stem to a curve.

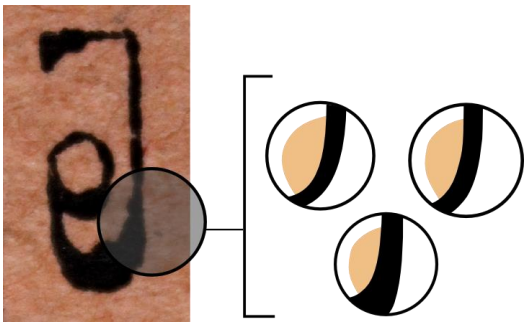


Figure.9 Scaling character ‘na’

3.4.2 Loop

Loop is an element in which a stroke of a letter overlaps (or touches) over another stroke but not cross-over; usually to create small counter spaces [5]. The character ‘sa’ [Figure 10] has a loop hanging on its branch. The counter space of the loop can be a circle, an oval, tear drop shaped and so on. A visually heavy loop stresses the branch where it hangs. The loop should have a balanced weight.

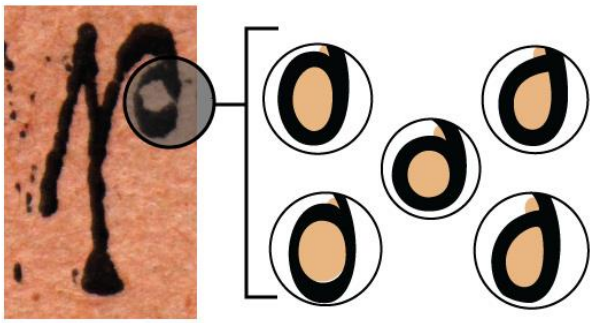


Figure.10 Loop in character ‘sa’

3.4.3 Knot

Knot can be seen as visual extension of a loop. A knot is an element in which one stroke overlaps over another and traverses ahead [5] such as in letter ‘ra’ [Figure 11]. Various knot explorations were explored; such as counter spaces of different shapes and axis. The part of the knot which jutted out had to be aligned to be in harmony with the rest of the knot.

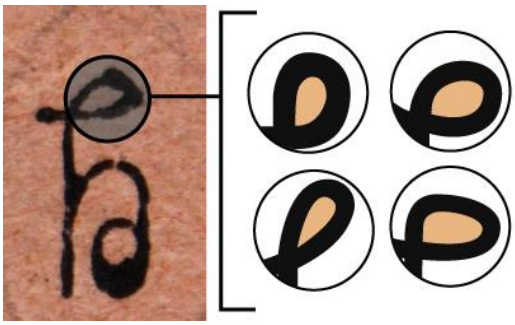


Figure.11 Knot in character 'ra'

3.4.4 Peak

Peak is joinery element which can be seen as a pointed top when two strokes meet.

The peak in the character ‘sa’ [Figure 12] is culmination or the point of meet of two elements- a slant and a vertical bar. Various kinds of peaks were explored — square, chamfered, rounded etc.

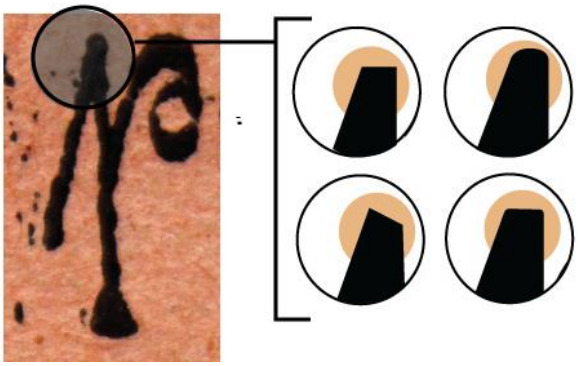


Figure.12 Peak in character ‘sa’

3.4.5 Pedestal

Pedestal is an element which can be seen as a bottom most part of a vertical bar.

The pedestal in the reference was flared up towards the end such as in character ‘sa’ [Figure 13]. Few variations of such terminals were explored such as — square, triangular/ flared up etc.

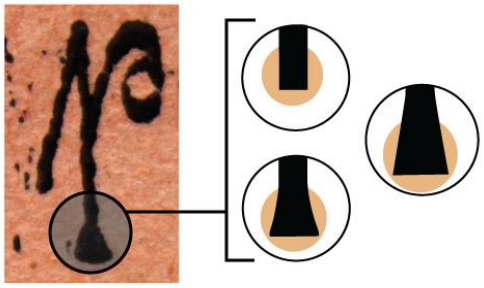


Figure.13 Pedestal in character ‘sa’

3.4.6 Branch and Branching Distance

Branch is an element which can be seen as arm extending out of the vertical bar.

The joinery of the branch with the vertical bar can have a smooth (seamless) transition or it can be cornered (non-seamless)[Figure 14]. Scaling was done at the point where the branch moves out of the stem.

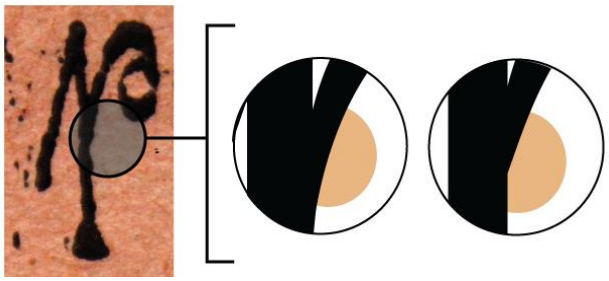


Figure.14 Branch in character ‘sa’

The distance of the branching element from the vertical bar [Figure 15]. Since vertical bar is the central element, the other elements need to be balanced around it. This distance is also crucial to determine the counter spaces formed in between these elements.

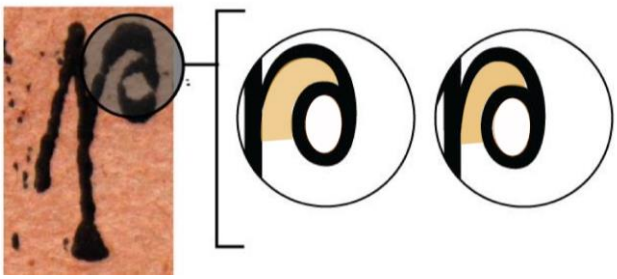


Figure.15 Branching Distance in ‘sa’

3.4.7 Nose

The nose is an in-body (not at the top or bottom but lies in the middle of the grid) joinery element which is made when two strokes join to make an acute-angle [Figure 16]. Various explorations of the nose included rounded tip, sharp/cornered tip, chamfered tip etc.

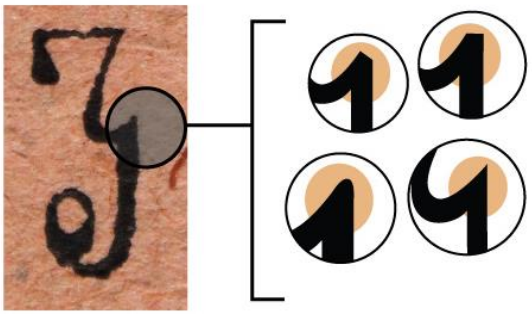


Figure.16 Nose in character ‘ka’

3.4.8 Stomach

The stomach is an element which starts with a kink (usually a nose) at the middle of the body, bulges out and is followed by a loop at the end[Figure 17]. Various explorations such as the distance between the loop and the bulge (counter space), scaling, the inward bent of the loop etc. were considered.

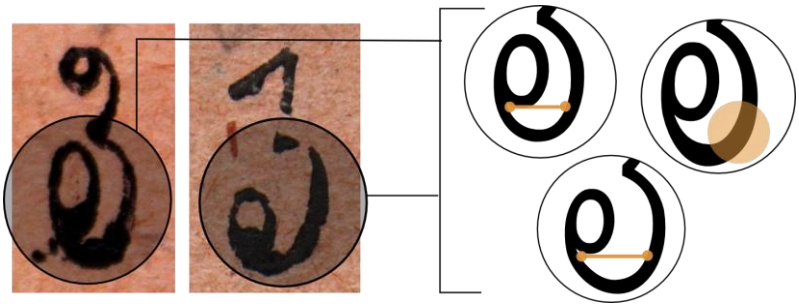


Figure.17 Loop in letter ‘ta’ and ‘ga’

3.4.9 Corner

A corner is the angle between two elements. We considered the angle from 0—90 degrees [See Fig 18]. In this case a vertical bar and a slant decide how far or close the elements can be from the visual center of the character. The angle was to be visually corrected such that it does not affect the overall tension in the character

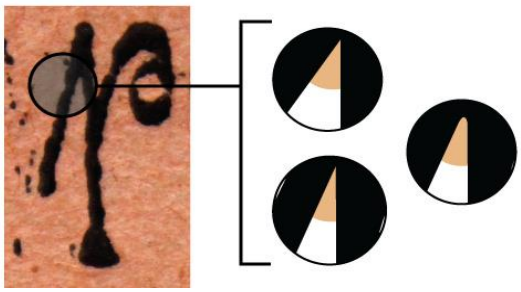


Figure.18 Corner in character ‘sa’

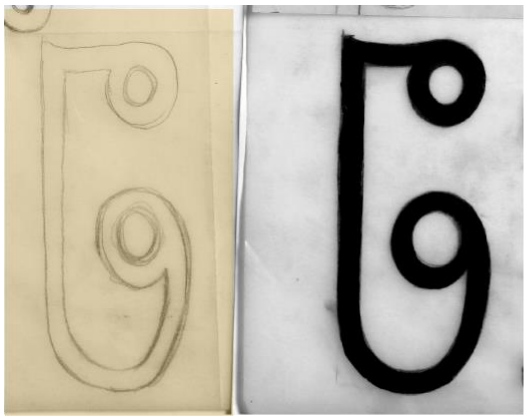
Out of all the anatomical explorations only a few were taken forward and finalized. For example sharp peaks were excluded. Since there are so many round elements in the font, a sharp element stands apart. Hence this was not going with the visual language of other elements. Such decisions allowed maintaining harmony across the characters.

3.5 Hand rendering

The characters were hand-rendered in mono-linear style. Stroke modulation was given so that the characters appear visually balanced. Horizontal strokes need to be lighter, for instance. The stroke modulation was not only applied for horizontals and verticals but also applied for loops, knots and other kinds of strokes. For instance few characters such as the ‘nya’ [Figure 19] had multiple loop elements. When these loops come close they can cause uneven gray. Such characters needed more attention because the each loop and knot has to be balanced individually to distribute weight across the character. This in turn helps in creating a balance when combined to make a word, a sentence or a page and achieve a uniform gray.

These details had to synchronize in their visual qualities with other characters. After a series of iterations, the stroke width was decided based on the efficiency in visibility and readability across an acceptable range of font sizes.

Figure.19 Letter ‘nya’ with multiple loops



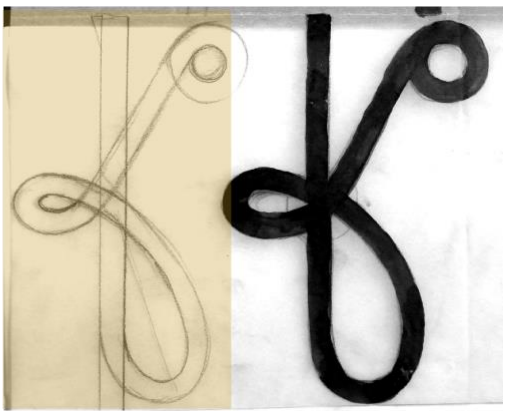


Figure.20 Hand rendered letters ‘ja’ and ‘e’

3.6 Digitization

After hand rendering and we went to digitalize the fonts. First the sheets were scanned to get high-resolution images. These images were edited to high-contrast and the most appropriate sketch attempts were chosen. These high-contrast, high-resolution cropped out letters were transferred to Illustrator to live-trace the images and create vector outlines.

The outlines were copied and pasted into in FontLab Studio within a height of 1000 pts. Some type samples were printed; this was carefully examined, re-checked and re-drawn a few times.

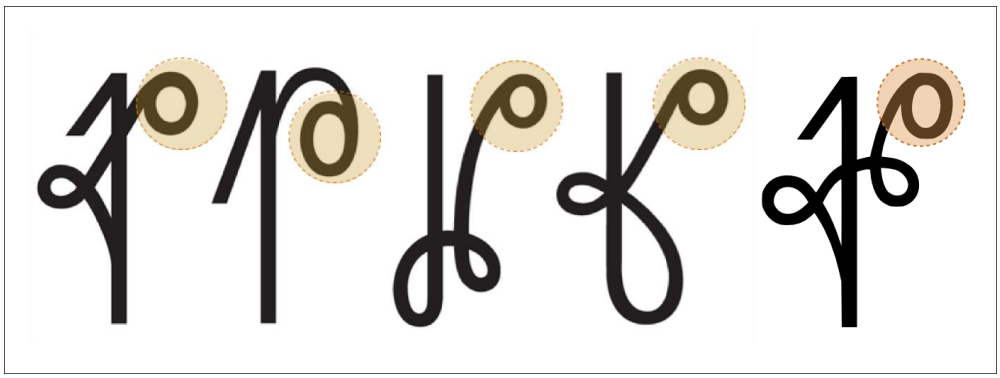


Figure.21 A set of characters with similar features

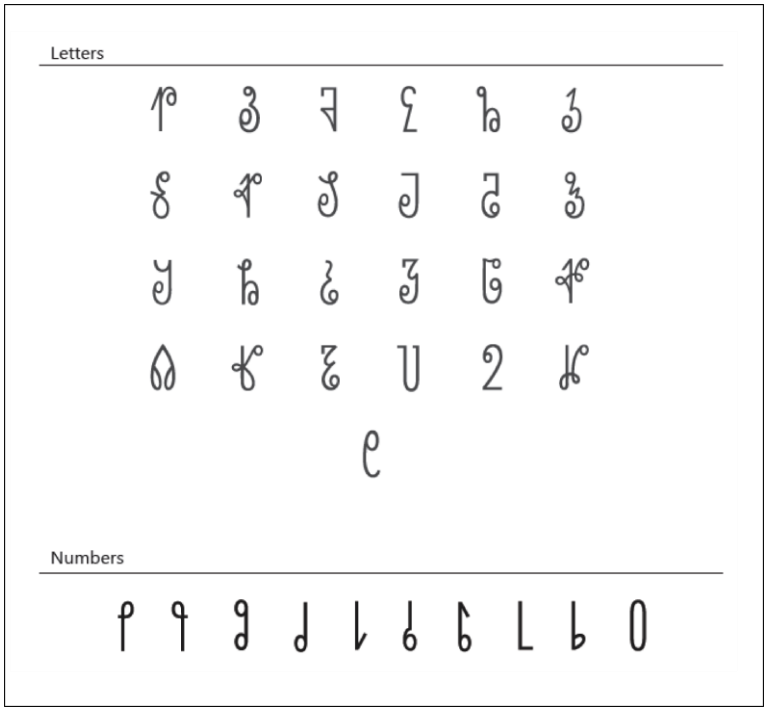


Figure.22 Digitized set of 25 characters and 10 digits

3.6 Short-comings and further improvements

The fonts were magnified for further visual corrections [See Figure 23]. The font still needs more refinement in its detailing. This is an iterative process. Few of issues which were identified for correction can be seen in Figure 24. Some of the issues included resolving and correcting the joinery to prevent uneven gray, non-matching joineries in the same character,

Figure .23 Magnifying the font for visual corrections

a b c d e

Figure.24 Issues identified for further corrections; a) letter ‘sa’ having a really long branch; b) letter ‘ba’ having a blocky joinery; c) letter ‘nga’ having a blocky joinery; d) letter ‘i’ having different detailing for similar looking joinery; e) letter ‘nya’ being unbalanced by the multiple loops. The loops need to be balanced to attain some amount of rhythm between the elements.

We evaluated our font from both — users without typographic knowledge and typography expert and got positive response. This included voluntary feedback and design ideas to improve the font further. Some of the expert review suggestions included decrease in the font height which would substantially reduce the vertical stress in the font. The future scope lies in performing expert evaluation and incorporating necessary changes to enhance the readability of the font.

3.7 Example

Figure.25 A children's letter book

6. Conclusion

A language can be most appropriately written in its own script. The Sora Sompeng typeface would facilitate future activities involving the native script which in turn would enable the tribe to take forward their literary heritage in both formal and informal spaces. Our project is an attempt to keep alive their identity as a tribe by contributing to the movement of documenting endangered languages and cultures.

Initiatives have been taken to promote the script by introducing primary education in parallel scripts i.e.in both Odia and Sora. Text books, calenders, informal documents and other artifacts of daily use have already been made by the community members. The response and enthusiasm has made this project into a collaborative community project. Like Sora, similar efforts can be extended to other scripts. With the technological advances, more scripts can be incorporated and co-exist under the typographic umbrella.

Acknowledgement

Prof. Girish Dalvi, Indian Institute of Technology, Bombay, girish.dalvi@iitb.ac.in

References

[1] http://www.thehindu.com/todays-paper/tp-national/tp-otherstates/in-danger-of-extinction/article1818261.ece. thehindu.com. 2007-03-31. [As seen on 12-01-2016]

[2] https://en.wikipedia.org/wiki/Sora_people [As seen on 02-01-2016]

[3] Web site: Michael. Everson. Michael Everson. Proposal for encoding the Sora Sompeng script in the UCS. pdf. 2009-06-08. International Organization for Standardization. Working Group Document

[4] Anderson, Gregory D.S. (1999). "A new classification of the Munda languages: Evidence from comparative verb morphology." Paper presented at 209th meeting of the American Oriental Society, Baltimore, MD [As seen on 01-01-2016]

[5] Girish Vinod Dalvi, (2010) "Conceptual model for Devanagari typefaces" [As seen on 8-02-2016]