Investigating the Use of Native and Secondary Language with Data Visualization in Madagascar

Henintsoa S. Andrianarivony[1] Taratra D. Raharison[1] I. Tamby Rakotoniaina[1] Mirindra T. N. Ramarokoto[1] Noëlle Rakotondravony[2] Lane Harrison[2]

henintsoa.s.andrianarivony@gmail.com

[1]Université d'Antananarivo [2]Worcester Polytechnic Institute

Motivation

For many countries concepts and terminologies of data are often in secondary languages.

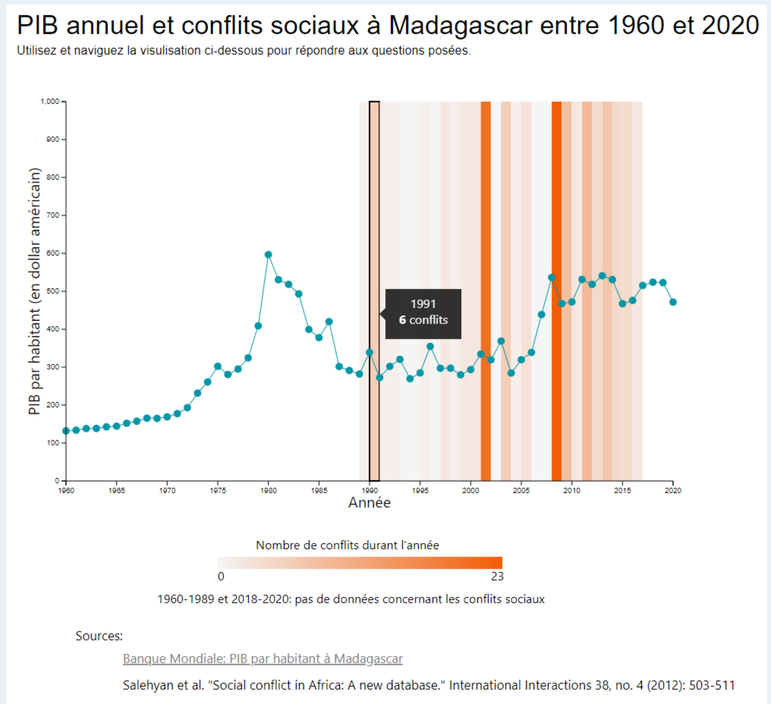
Exploring this interplay of language therefore can surface interesting aspects in the study and practice of data visualization in native language vs. secondary language.

Methodology

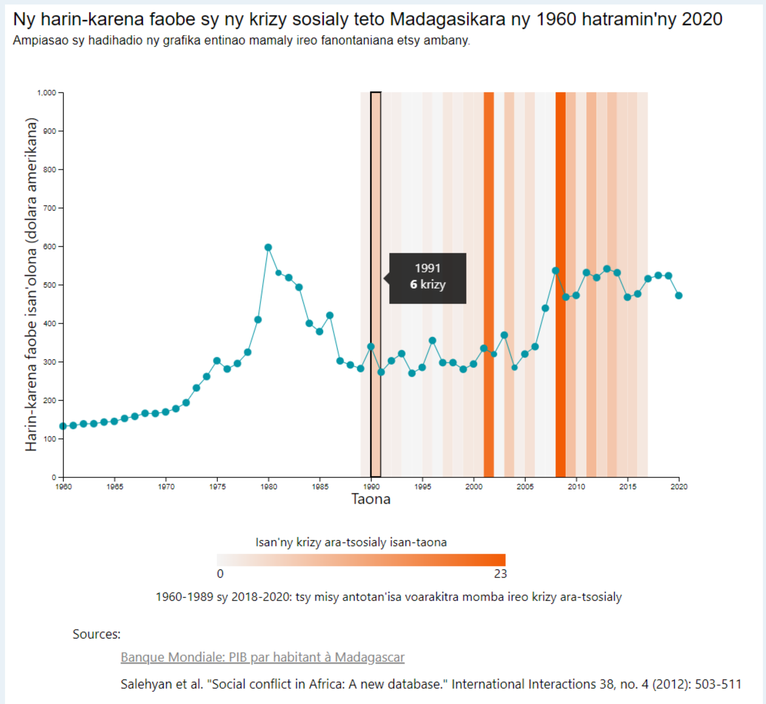
We design an online study in both **Malagasy** and **French**, with bilingual participants in Madagascar.

We visualize a dataset with the Malagasy GDP per capita using a linechart, and recorded social conflicts in Madagascar using color stripes.

Visualization in French (Secondary language)



Visualization in Malagasy (Native language)



14 participants, between subjects: 7 participants for each version.

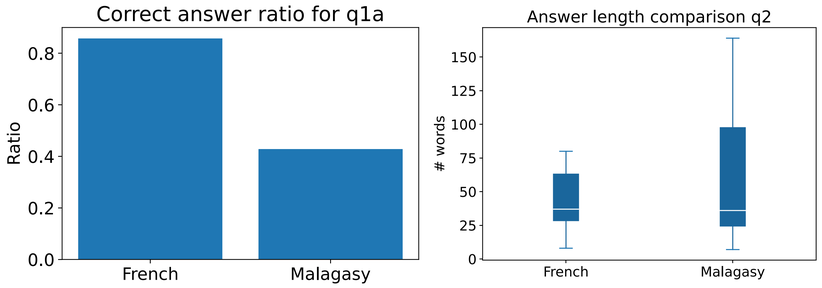
Participants are given an introduction to the study and asked to examine the interactive visualization.

Participants answer a set of understanding/insight questions afterwards.

The goal is to assess peoples' ability or difficulty in expressing and understanding the data visualization in the language of the experiment.

poster paper & additional materials

Key Results



- Concise and more accurate answers in French, e.g.:

  P4-fr: *"Suivez la courbe bleue pour le PIB par habitant et mettez le curseur sur les battons oranges pour voir le nombre de conflits durant une année." ("Follow the blue curve for GDP per capita and move the cursor over the orange bars to see the number of conflicts in a year.")*

- Extensive use of French terminology in the Malagasy condition, despite the visualization being entirely in Malagasy:

  P2-mg: *"[...] Nalefako teo ambonin'ilay **barre** faharoa farany marevaka ny **pointeur** dia nampiseho ilay isa sy ny daty mifandraika aminy." ("I moved the cursor over the second bar with darkest color to show the number and its corresponding date.")*

- Expert in survey design and data collection advised:

  P5-mg: *"You should have made tables with numerical values instead of the visualization, it would have been easier to read."*

- Most participants missed the interactivity of the visualization in both versions, instead talking about the visualization as a static chart.

Conclusion

Results suggested that in Madagascar, a country with two major languages, communicating and designing data visualizations tend to be easier for participants in **French** (language of education) than in **Malagasy** (native language).

We suspect that the difference in visualization literacy and familiarity is related in the availability of linguistic tools specific to data visualizations.

Future work might engage native-only speakers, ~80% of the population, to examine additional barriers to visualization access due to language differences.[1]

[1] https://information.tv5monde.com/info/les-malgaches-et-la-langue-francaise-je-t-aime-moi-non-plus-133105