Evaluating the Necessity of Language Translation Services for Emergency Dispatch Calls in Toronto*

How We Might Make Language Services More Accessible to Toronto's Diverse Communities

Emily Kim

2 February 2023

Toronto is a city known for its diverse immigrant communities, resulting in a variety of non-English first languages being spoken. This diversity can create challenges for emergency services to provide efficient support in time-sensitive situations. This report aims to examine the usage of designated language translation services by Toronto Paramedic Services' Emergency Medical Dispatchers during 911 calls from 2014 to 2021. By analyzing this data, we have discovered that Cantonese, Mandarin, and Russian are the top three languages in need of improved accessibility to specific language services during emergencies, with Spanish on an upward trend.

Introduction

Toronto is Canada's most diverse city, with nearly half of the city's population being immigrants from all around the world. Within this urban diversity are people from countries where the dominant language is not English. One important consequence of this phenomenon is the inability for these individuals to clearly express their crisis in an emergency situation, such as during a 911 call. Communication difficulties can lead to misunderstandings, and the caller may not be able to provide all the necessary information to the operator which can result in confusion and delay. Particularly in the event that they require health services, linguistic barriers can result in a life-threatening situation. Recent research studies by linguists show that during the COVID-19 global pandemic, the integration of emergency language services in the COVID-19 response helped greatly reduce the spread of the virus (Dreisbach and

*Code and data are available at https://github.com/emilykimto/Language-Services.git

Mendoza-Dreisbach 2020). Similarly, for an ethnically diverse city like Toronto, being able to accommodate those with limited English proficiency can help minimize the potential harm in emergency situations. Language interpreting services such as those offered by Toronto's Paramedic Services 911 are crucial to ensuring accurate communication, timely and effective delivery of emergency services, and improving outcomes.

As Toronto experiences a growing immigrant population, there is an increasing demand for emergency interpretation services that can accommodate the diverse linguistic backgrounds of the city's residents. This report aims to explore this need by analyzing the City of Toronto's paramedic language services data to uncover which communities have the most demand for emergency interpretation services. I will determine the top languages requiring the most cumulative sum of interpretation services duration (in minutes) between 2014-2021. Then, I will compare the top languages with highest call duration results with fluctuations over the eight year period to uncover a more recent trend in language usage.

By examining the trend of language-specific calls between 2014-2021, we can gain a deeper understanding of the city's evolving linguistic needs. This information can then inform the allocation of resources, such as staff and time, to improve the quality and accessibility of emergency services for Toronto's many ethnic communities.

Data

The data used in this report was collected by the Toronto Paramedic Services (Data 2022) from the Toronto Open Data portal (Gelfand 2020). It was last refreshed on February 2, 2022, and updates yearly. The data set contains information on the date, time, language, and duration in minutes of instances when a caller required language interpretation. Using `R` (R Core Team 2021), `tidyverse` (Wickham et al. 2019), and `dplyr` (Wickham et al. 2021), I began my analysis.

First, I grouped the data by language using `group_by` before using the function `summarize` from the `dplyr` package to calculate the sum of duration for each language presented in the data. This resulted in a new data frame with one row for each language paired with the sum of all durations for that language. With this new data frame, I renamed the "total_duration" column to "Total Duration" using the `rename()` function for further clarity. I wanted to find the top ten languages in the data set that had the highest total call duration, so I used the `top_n` function after which I arranged the data by "Total Duration" in descending order using `arrange`. This created a final table consisting of the top ten languages requiring the most language interpretation assistance during emergency calls using `knitr` (Xie 2021) and `kableExtra` (Zhu 2020).

I discovered that the based on the data from 2014-2021, out of these top ten languages, the three most used languages in emergency services based on duration were Mandarin, Cantonese, and Russian. Please refer to Table 1.

Table 1: Top 10 languages based on the total duration of calls requiring language services

Language	Duration
ARABIC	8647
CANTONESE	24102
FARSI	9181
HUNGARIAN	9292
ITALIAN	10238
MANDARIN	22884
PORTUGUESE	8994
RUSSIAN	17384
SPANISH	14638
TAMIL	11016

However, this contradicts the language composition data from the 2021 Census of Toronto (Toronto 2021), which showed the top five languages spoken at home to be Mandarin, Cantonese, Tagalog, Spanish, and Tamil. According to the Census, among the Top 15 Mother Tongue Languages in the City of Toronto {excluding Official Languages), Mandarin was ranked first, Cantonese second, and Russian tenth, making up 4.1%, 3.7%, and 1.2% of Toronto, respectively. There seems to be a significant discrepancy between the usage of Russian as a mother tongue in the census and its call duration in this analysis.

Next, to better understand the usage of different languages in emergency calls over the years, I want to create a line graph visualizing fluctuations between 2014 to 2021 using `ggplot2` (Wickham 2016). First, I used `tidyverse` (Wickham et al. 2019) and `lubridate` (Grolemund and Wickham 2011) to change the "Time Stamp" column in the data set into date-format. Afterwards, I excluded rows with missing or NA values in the "Time Stamp" column using `filter`. This resulted in a data set containing 26,225 rows of emergency instances, while the original had 27,673 rows. One important limitation of this line graph is its inability to address these rows: 1,448 calls.

I selected the top ten languages for the line graph and filtered to include only the selected languages. I calculated the cumulative duration per year for each language using `group_by` and `summarize`. I then plotted a line graph using `ggplot` with the x-axis labeled "Year" and y-axis labeled "Cumulation Duration (in minutes)". Each coloured line represents a different language from the top ten most requested languages in emergency interpretation services between the years 2014-2021. Figure 1 demonstrates that over this eight year time period, Mandarin, Cantonese, and Spanish have peaked as the top three most used language services, with Spanish overtaking Russian in call duration between 2020-2021.

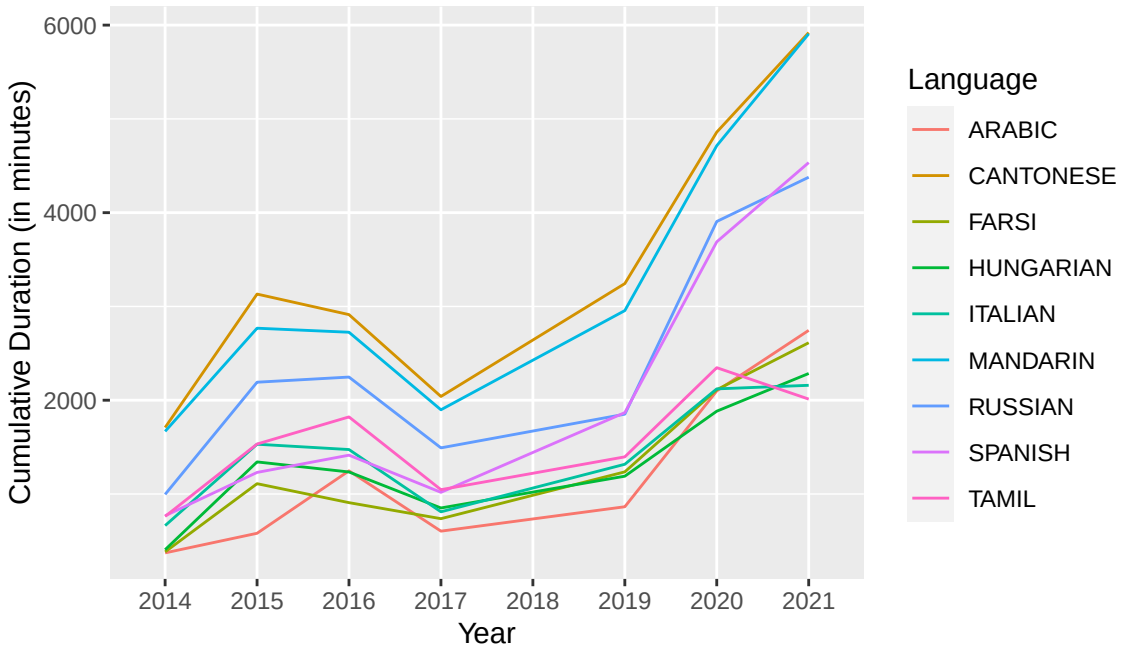

Figure 1: Cumulative Duration of Top 10 Languages Per Year Between 2014-2021

Results

My analysis of the Toronto Paramedic Services data on emergency language interpretation services from 2014-2021 revealed that the top three languages requiring language services were Cantonese, Mandarin, and Russian in terms of the total duration in minutes. The data showed that these three languages accounted for a majority of the total emergency call duration during these eight years. To put into further perspective, these three languages held a cumulative sum of 64,370 minutes or 44.7 days worth of language interpretation. This information is in contrast with the findings from the 2021 Census (Toronto 2021) investigating the language composition of Toronto, which revealed a disproportionate representation of Russian in this emergency language services data.

Figure 1 represents a line graph showing the top ten languages with the most total duration of emergency language interpretation services from 2014-2021. While some languages, such as Cantonese, Mandarin, and Russian consistently held a high demand for these services, Tamil saw a dip in more recent years. Finally, Spanish saw a spike in duration in the 2021 year, surpassing Russian in the top three languages that require the most assistance.

Discussion

Limitations

There are several limitations to this report. First, the data set was last refreshed on February 2, 2022, which may result in data that does not accurately reflect the most recent emergency call data - particularly for the 2022 year. Further, Figure 1 was unable to address 1,448 emergency calls with missing values in the "Time Stamp" column, which may impact the accuracy of the data analysis. Some ethical considerations of this report include the potential to reveak information that may be sensitive for those communities impacted by the results. This results of this report should not be used to discriminate or make assumptions regarding any particular group based on their language or ethnicity.

Moving forward

Despite these limitations and other considerations, this data highlights the language diversity in Toronto and the necessity for equitable access to emergency services. Its insights can be used to inform the development of programs and policies aimed at improving the quality and accessibility of emergency language services for Toronto's many ethnic communities. Particularly for those who speak Mandarin and Cantonese, who have been shown to consistently require additional language support in emergency situations. Additionally, it is important to note that the data analysis in this report shows Russian as the third most used language in emergency services, whereas the 2021 Census ranked Russian tenth in the list of top mother tongue languages (Toronto 2021). Moreover, my analysis indicated a recent spike increase in Spanish speakers seeking language interpretation services.

This disparity in representation as well as the sharp change in consistency in the recent 2021 year highlights the importance of considering multiple sources of data to better understand the ever-evolving linguistic needs of a diverse city. As more data is collected in the future, we can further monitor the trend of language usage over time to identify any fluctuations and create opportunities for proactive planning. This way, the city can better provide for its residents in forms of re-allocation of appropriate resources and adequate training for emergency service providers, thereby improving public safety and outcomes.

References

Data, Toronto Open. 2022. "Paramedic Services 911 Language Interpretation." https://open.toronto.ca/dataset/paramedic-services-911-language-interpretation/.

Dreisbach, Jeconiah Louis, and Sharon Mendoza-Dreisbach. 2020. "The Integration of Emergency Language Services in COVID-19 Response: A Call for the Linguistic Turn in Public Health." *Journal of Public Health* 43 (2): e248–49. https://doi.org/10.1093/pubmed/fdaa178.

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Toronto, City of. 2021. "2021 Census Backgrounder Language." https://www.toronto.ca/wp-content/uploads/2022/08/868f-2021-Census-Backgrounder-Language-FINAL.pdf.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zhu, Hao. 2020. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.