



# USE OF NON-ENGLISH LANGUAGES IN THE 2020 CENSUS: INSIGHTS FROM PARADATA

Renee Ellis

Center for Behavioral Science Methods  
U.S. Census Bureau

AAPOR 2024, Portland Oregon

*Disclaimer: This presentation is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. DRB Approval Number: CBDRB-FY24-CBSM003-005*

# Motivation

- ❖ The 2020 Census was the first decennial census to extensively offer an internet mode for self-response. This included the collection of paradata for all internet self-response sessions.
- ❖ The 2020 Census internet instrument was available in 12 languages: English, Spanish, Chinese, Vietnamese, Korean, Russian, Arabic, Polish, Japanese, Portuguese, French, Haitian Creole, and Tagalog.
- ❖ Up until now most of our knowledge about translations comes from qualitative studies and we have extensive research over the last 15 years that has been used to understand difficulties with translations and navigation
- ❖ The 2020 Census provides us an opportunity to examine the use of translations in internet instruments on a scale we have not had previously
- ❖ We recently conducted a detailed paradata analysis on respondents who use the provided internet translations- that report is forthcoming
- ❖ The corpus of qualitative research on language use and the recent paradata report on language use give us an opportunity to look at the intersections between those seemingly diverse methods of knowledge collection and to discuss how paradata analysis can provide added value to survey improvement
- ❖ This presentation uses paradata analysis from the forthcoming 2020 language paradata analysis report where we focus on comparisons between the paradata and previous qualitative findings to illustrate the similarities and value added of paradata analysis

# Questions?

- ❖ How do qualitative findings compare to paradata findings?
  - Is there evidence in the paradata that shows changes made based on qualitative findings worked?
  - Are there places where the paradata shows continued issues?
- ❖ Are there areas where paradata expands our findings beyond those from qualitative research?

# Intersections between qualitative research and paradata research

	Qualitative Research	Internet paradata analysis
Definition	Naturalistic study of social meanings and processes	Study of respondent behavior during the navigation of an online instrument
Methods	Cognitive interviews, focus groups, usability testing, observations	Aggregate measures of behavior in instrument derived from parsed data like time in instrument or percentage of error triggers of clicks on help links
Data insights	Use themes or observations holistically to develop insights based on patterns	Individual measures are difficult to interpret, analysis is stronger when measures are used holistically to develop insights based on patterns
Data presentation	Data is presented thematically or as descriptive statistics. Size limitations limit the use of inferential statistics	Data is presented thematically or as descriptive statistics. Data violates usual assumptions of inferential statistics
Size	Generally small	Generally, very large

# Language research main focus

	Qualitative language research	Paradata language report
Access	Ability to find ID, mailing materials, ability to get into instrument	Use of ID, Device used for access and how translations are accessed
Navigation	Issues moving through instrument and answering questions	Issues moving through instrument and answering questions
Identification of questions where respondents have particular problems	Identified questions and the problem with those questions by language	Identified pages that were the most frequent location for break-offs, changing to a non-English language, and indicators of difficulty

# Methods

- ❖ Analyzed the paradata from about 1.6 million internet instrument sessions that selected one non-English language.
  - Included only respondents who selected instrument supported translations at least once in a non-English language
  - Excluded respondents who selected more than one non-English language
  - Analysis was at the session level
- ❖ Navigation
  - Calculated individual measures of indicators of difficulty
  - Created holistic measure that divided languages into three groups
- ❖ Access
  - Grouped languages in thirds languages for break-off and use of non-ID path
  - Based on navigation issues identified two other indicators of access
- ❖ Difficult questions
  - ❖ Identified the top pages for break-offs, changing to a non-English language
  - ❖ Identified top pages for indicators of difficulty
  - ❖ Identified pages that had the most indicators of difficulty by language

# Data

Language	N	%
All Languages	1,626,000	100%
Tagalog	16,000	1%
Arabic	20,000	1%
Polish	36,500	2%
Japanese	41,500	2%
Russian	44,500	3%
Portuguese	45,000	3%
French	46,500	3%
Haitian Creole	49,500	3%
Vietnamese	63,000	4%
Korean	85,500	5%
Chinese	157,000	10%
Spanish	1,020,000	63%

2020 internet self-response paradata. Numbers are rounded

# Difficulty with navigation

Indicators used to determine navigation difficulty

- Edit warnings
- Number of edit warnings
- Use of help
- Number of times using help
- Backing
- Number of times backed
- Timeout warnings
- Number of timeout warnings
- Use of review
- Time in instrument

## Qualitative Research

Arabic

**Russian**

**Korean**

Chinese

**Vietnamese**

**Spanish**

## Paradata research

Polish

**Russian**

Japanese

Chinese

Tagalog

Portuguese

French

**Korean**

Arabic

Haitian Creole

**Vietnamese**

**Spanish**



# Difficulty with access

## Use of ID

Japanese	91%
Polish	88%
French	88%
Haitian Creole	88%
Tagalog	86%
Russian	86%
Portuguese	86%
Korean	84%
Vietnamese	82%
Arabic	79%
Chinese	79%
Spanish	73%

Qualitative research	Paradata research break-offs	
<u>Arabic</u>	Japanese	9%
<b>Russian</b>	Portuguese	9%
<b>Korean</b>	French	9%
	Haitian Creole	9%
	Korean	9%
	Polish	11%
	Russian	12%
	Tagalog	13%
	Vietnamese	14%
<b>Chinese</b>	<u>Arabic</u>	16%
<b>Vietnamese</b>	Chinese	16%
<b>Spanish</b>	Spanish	19%

# Navigation and access patterns

Qualitative research	Navigation Difficulty	Access difficulty
Arabic	Polish	Polish
Russian	Russian	Japanese
Korean	Japanese	French
	Chinese	Haitian Creole
	Tagalog	Tagalog
	Portuguese	Russian
	French	Portuguese
	Korean	Korean
	Arabic	Arabic
Chinese	Haitian Creole	Vietnamese
Vietnamese	Vietnamese	Chinese
Spanish	Spanish	Spanish

# Question specific difficulty

	Who is in HH	Under-count	Race	Ethnicity	Relationship	County	Rent/own	Date of Birth	Login
Tagalog	◆ ▲ ●	●	◆ ▲	▲ ●	▲		◆	◆ ▲ ●	◆ ▲
Arabic	★ ◆ ▲ ●	●	★ ◆ ▲ ●	★ ◆ ▲ ●	★ ◆		◆	◆ ▲ ●	◆ ▲ ●
Polish	▲ ●	●	◆ ▲ ●	▲ ●	◆	◆	◆	◆ ▲	◆ ▲
Japanese	◆ ▲ ●		◆ ▲	▲	◆	◆	◆	◆ ▲	◆ ▲
Russian	★ ◆ ▲ ●	●	◆ ▲ ●	▲	★ ◆	★ ◆	★ ◆ ●	◆ ▲	◆ ▲
Portuguese	◆ ▲ ●	◆ ●	◆ ▲ ●	▲ ●	◆		◆	◆ ▲	◆ ▲
French	◆ ▲ ●	◆ ●	◆ ▲	▲	◆ ▲	◆	◆	◆	▲
Haitian Creole	◆ ▲ ●	◆ ●	◆ ▲ ●	◆ ▲ ●	◆		◆	◆ ▲ ●	▲
Vietnamese	★ ◆ ▲ ●	●	▲	◆ ▲ ●	★ ◆ ▲	★ ◆	★ ◆ ●	◆	◆ ▲
Korean	★ ◆ ▲ ●	●	▲	▲	★	★ ◆	★ ◆ ▲	◆	◆ ▲
Chinese	★ ◆ ▲ ●	★ ●	★ ▲	★ ◆ ▲ ●	★ ◆	◆	★ ◆ ▲	◆	◆ ▲
Spanish	★ ▲ ●	★ ●	★ ◆ ▲ ●	★ ▲ ●	★ ◆ ▲	◆	◆	●	◆ ▲ ●

★ QR ◆ PD translation ▲ PD breakoff ● PD Difficulty

# Conclusions

- ❖ Is there evidence in the paradata that shows changes made to address qualitative findings worked?
  - High use of Id's and low breakoff rates
  - Improvements in navigation by respondents using Chinese
  - Lack of issues in the county/locality question and to a lesser degree the relationship question
- ❖ Are there places where the paradata shows continued issues?
  - Spanish, and Vietnamese translations show more difficulty with access and navigation
  - Chinese translations show more difficulty with access
  - Difficulty with understanding of who is in the household and in the Race and Ethnicity questions
  - Difficulty with rent own questions in Russian, Vietnamese, Korean, Chinese
- ❖ Are there areas where paradata expands our findings beyond those from qualitative research?
  - Additional measures of access difficulty
  - Respondents using Arabic translations have more difficulty than found in qualitative research
  - Paradata identifies additional questions where respondents have difficulty

# Discussion

## Recommendations

- Optimize mobile instruments for languages that are more likely to use mobile devices
- Additional research on languages where we don't have as many years of research
- Additional research on languages with the most difficulty
- Continued research on translations with an emphasis on questions identified as more difficult

## Research extensions

- Use internal data on the other six languages to make comparisons and recommendations
- Refining analysis to better understand the patterns of how respondents change language across the survey pathway
- Investigate respondents who use browser translations



THANK YOU

QUESTIONS?

[Renee.ellis@census.gov](mailto:Renee.ellis@census.gov)

# Background- Paradata as an analysis tool

- ❖ Internet paradata contains each action of the instrument and the respondent making it a complete history of a respondent's behavior while completing the survey
- ❖ Can be used for many purposes, such as to identify usability issues and problematic questions, and to evaluate the ease-of-use of the instrument.
- ❖ Paradata is very large and complex creating challenges
  - Paradata is derived from instrument information transmitted from the respondent, this creates multiple opportunities for error in the instrument or transmission to create data issues that makes data difficult to understand and may be indistinguishable from user behavior
  - Because of the non-standard format of paradata, it violates many of the assumptions we use for weighting or statistical measures of significant difference
  - Standard paradata measures are hard to interpret
- ❖ Solutions to overcoming limitations of paradata to make it easier to communicate issues and recommend solutions
  - Improved measures to capture patterns of behavior across the survey pathway
  - Measures of significance
  - Using existing standard measures holistically to better communicate patterns and areas for improvement

# Question specific difficulty

	Who is in HH	Under-count	Race	Ethnicity	Relationship	County	Rent/own	Date of Birth	Login
Tagalog	◆ ▲ ●	●	◆ ▲	▲ ●	▲		◆	◆ ▲ ●	◆ ▲
Arabic	★ ◆ ▲ ●	●	★ ◆ ▲ ●	★ ◆ ▲ ●	★ ◆		◆	◆ ▲ ●	◆ ▲ ●
Polish	▲ ●	●	◆ ▲ ●	▲ ●	◆	◆	◆	◆ ▲	◆ ▲
Japanese	◆ ▲ ●		◆ ▲	▲	◆	◆	◆	◆ ▲	◆ ▲
Russian	★ ◆ ▲ ●	●	◆ ▲ ●	▲	★ ◆	★ ◆	★ ◆ ●	◆ ▲	◆ ▲
Portuguese	◆ ▲ ●	◆ ●	◆ ▲ ●	▲ ●	◆		◆	◆ ▲	◆ ▲
French	◆ ▲ ●	◆ ●	◆ ▲	▲	◆ ▲	◆	◆	◆	▲
Haitian Creole	◆ ▲ ●	◆ ●	◆ ▲ ●	◆ ▲ ●	◆		◆	◆ ▲ ●	▲
Vietnamese	★ ◆ ▲ ●	●	▲	◆ ▲ ●	★ ◆ ▲	★ ◆	★ ◆ ●	◆	◆ ▲
Korean	★ ◆ ▲ ●	●	▲	▲	★	★ ◆	★ ◆ ▲	◆	◆ ▲
Chinese	★ ◆ ▲ ●	★ ●	★ ▲	★ ◆ ▲ ●	★ ◆	◆	★ ◆ ▲	◆	◆ ▲
Spanish	★ ▲ ●	★ ●	★ ◆ ▲ ●	★ ▲ ●	★ ◆ ▲	◆	◆	●	◆ ▲ ●

★ QR ◆ PD translation ▲ PD breakoff ● PD Difficulty



# Additional paradata measures of access difficulty

Use of mobile devices		Use of dropdown to select language	
Haitian Creole	1%	Haitian Creole	3%
French	3%	French	7%
Tagalog	4%	Tagalog	8%
Polish	5%	Polish	8%
Japanese	8%	Portuguese	8%
Russian	9%	Vietnamese	8%
Portuguese	9%	Japanese	10%
Vietnamese	11%	Korean	11%
Korean	11%	Chinese	13%
Chinese	19%	Arabic	15%
Arabic	26%	Russian	19%
Spanish	44%	Spanish	30%