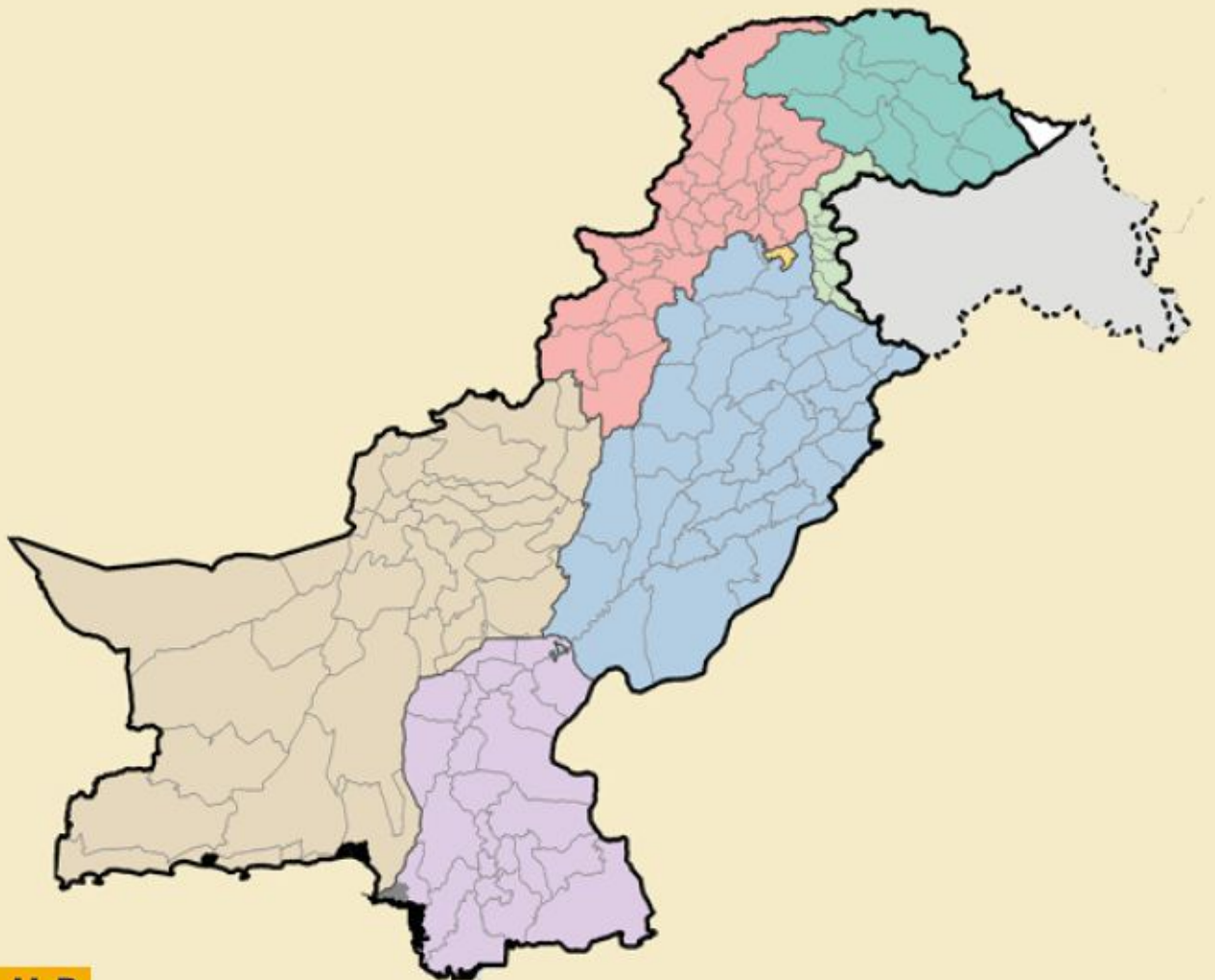


GALLUP BIG DATA ANALYSIS OF THE 2023 CENSUS | VOLUME 1

A comparative analysis of Pakistan's
mother tongues using Census data
from 1998, 2017 and 2023



Gallup Pakistan's Big Data Analysis of the 2023 Census: A Comparative Study with the 2017 and 1998 Censuses

Volume 1

Punjabi as a mother tongue on a decline, while Pashto, Saraiki, and Urdu see a rise: Gallup Pakistan's Big Data Analysis of Pakistan's Census 2023

A comparison with the 1998 and 2017 censuses reveals that Punjabi speakers decreased overall from 44% in 1998 to 37% in 2023, while Urdu speakers increased from 8% in 1998 to 9% in 2023. Pashto and Saraiki speakers were on a similar upward trajectory during these 25 years, while Sindhi speakers fluctuated slightly to 15% in 2017 but then stabilized at their original 1998 level of 14% in 2023.

Balochi speakers declined from 4% to 3% over this quarter of a century.

(Islamabad), 7th August, 2024

Gallup Pakistan, as part of its Big Data Analysis initiative, is looking at the 2023 Population and Housing Census. The complete 2023 Population and Housing Census is available [here](#).

What is the Big Data Analysis Series by Gallup:

Gallup Pakistan's Big Data series was started by Bilal I Gilani, Executive Director of Gallup Pakistan. Bilal explains the rationale of the series: *"The usual complaint from academics and policy makers is that Pakistan does not have data availability. Our experience negates that. Pakistan has lots of data, but it is not available in a usable form and not widely accessible. At Gallup we plan to bridge this gap in terms of accessibility and use of data. The Gallup Big Data series has earlier worked with data sets such as PSLM, Labor Force Survey, and Economic Survey reports as well as National Census Reports and Election Commission Data sets. The current series is using the 2023 Census (7th Population and Housing Census). We hope that these series are useful, and we welcome both feedback as well as possible collaborations as we create a public good in the form of useful data sets in Pakistan."*

For more reports and dashboards on Pakistan please visit our dedicated website: <https://www.galluppakistandigitalanalytics.com/>

What data points this current edition covers:

This series aims to present the important learnings from the 2023 Population and Housing Census for policy makers, the public, as well as for marketers in an easy and understandable way. The other reports within this series shed light on key topics like literacy ratios, population density, level of congestion in households, period of house construction, sources of energy, drinking water, and information/communication, population with disability, Pakistani citizens holding a CNIC, homeless population etc.

Key Findings

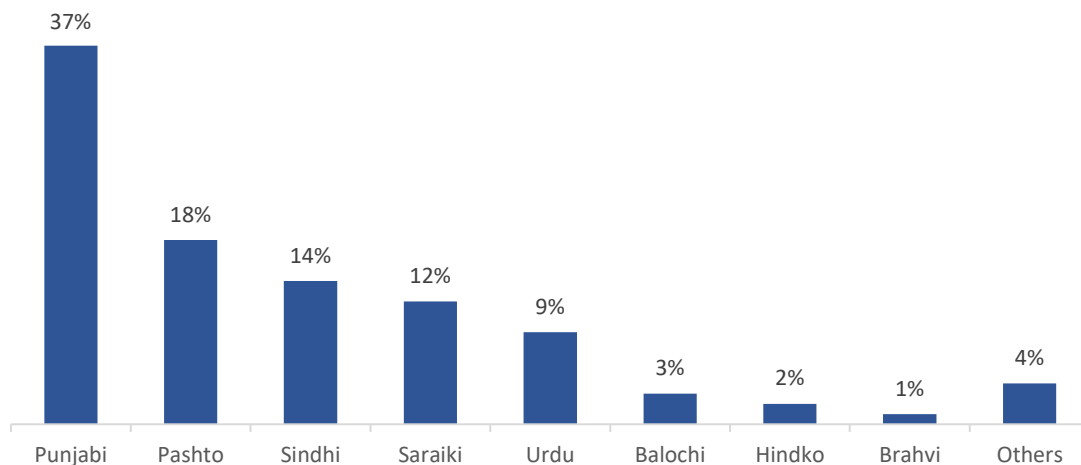
- 1. Population by Mother Tongue:** Urdu was the mother tongue of 9% of Pakistanis in 2023. Punjabi was the dominant language with the highest number of speakers (37%), followed by Pashto (18%), Sindhi (14%), Saraiki (12%), Balochi (3%), Hindko (2%), Brahvi (1%) and others (1%).
- 2. Provincial Breakdown – Comparison with 2017 Census:**
 - KPK's main regional language, Pashto, saw a 1% increase from 80% in 2017 to 81% in 2023.
 - Punjab's main regional language Punjabi has seen its speakers decline by 3%, from 70% in 2017 to 67% in 2023.
 - Sindhi speakers of Sindh experienced a 2% reduction in their proportion, from 62% in 2017 to 60% in 2023.
 - In Balochistan, from the two main regional language groups, the Balochi speakers increased from 35% in 2017 to 40% in 2023, while speakers of the Pashto language fell slightly from 35% to 34% over these 6 years.
 - In the Islamabad Capital Territory, the 3 largest regional language speakers are Punjabi, Pashto, and Urdu. Punjabi speakers decreased slightly from 52% to 51%, and Pashto speakers declined from 19% in 2017 to 18% in 2023. Meanwhile, Urdu speakers witnessed a 4% increase in their proportion, from 12% in 2017 to 16% in 2023.
- 3. Regional Language Distribution in Rural Areas – Comparison with 1998 and 2017 Censuses:** 34% of rural Pakistanis claimed Punjabi to be their mother tongue in 2023, but rural Punjabi speakers have witnessed a dramatic 9% decline over this quarter of a century. Rural Pashto speakers rose from 18% in 1998 to 22% in 2017 and 2023. Urdu speakers of rural areas have remained stable at the low rate of 2% throughout this 25-year period.
- 4. Regional Language Distribution in Urban Areas – Comparison with 1998 and 2017 Censuses:** Punjabi speakers constitute the highest proportion of urban residents at 41%, but urban Punjabi speakers have steadily declined by 7% over these 25 years. Sindhi and Pashto speakers followed similar trends; urban Pashto speakers started at 10% in 1998, increased to 12% in 2017, and then fell slightly in 2023 to 11%. Urban Urdu speakers have fluctuated during this period, starting at 20% in 1998, falling to 17% in 2017, but rising again to 21% in 2023, showing a net increase of 1% during these 25 years.

1. Population by Mother Tongue:

According to the 2023 Census results, **Urdu was the mother tongue of 9% of Pakistanis in 2023**. The Census reported on the population breakdown for the following languages: Urdu, Punjabi, Sindhi, Pushto, Balochi, Kashmiri, Saraiki, Hindko, Brahvi, Shina, Balti, Mewati, Kalasha, Kohiostani, and others. Kashmiri, Shina, Balti, Mewati, Kalasha, and Kohiostani were mother tongues of very small sub-populations, with exact numbers ranging from 7,000+ to 200,000+ people. Meanwhile, **the regional language of Punjabi was the mother tongue of almost 89 million citizens**.

The language spoken by the majority of citizens was Punjabi (37%), while the national language of Pakistan, Urdu, was the mother tongue of only 9% of Pakistanis. The domination of Punjabi is attributable to the population demographic of the province wherein the largest proportion of the Pakistani population resides in Punjab. This was followed by Pashto (18%), Sindhi (14%), Saraiki (12%), Balochi (3%), Hindko (2%), Brahvi (1%) and others (1%).

Figure 1: Population by Language (Mother Tongue) in Pakistan in 2023

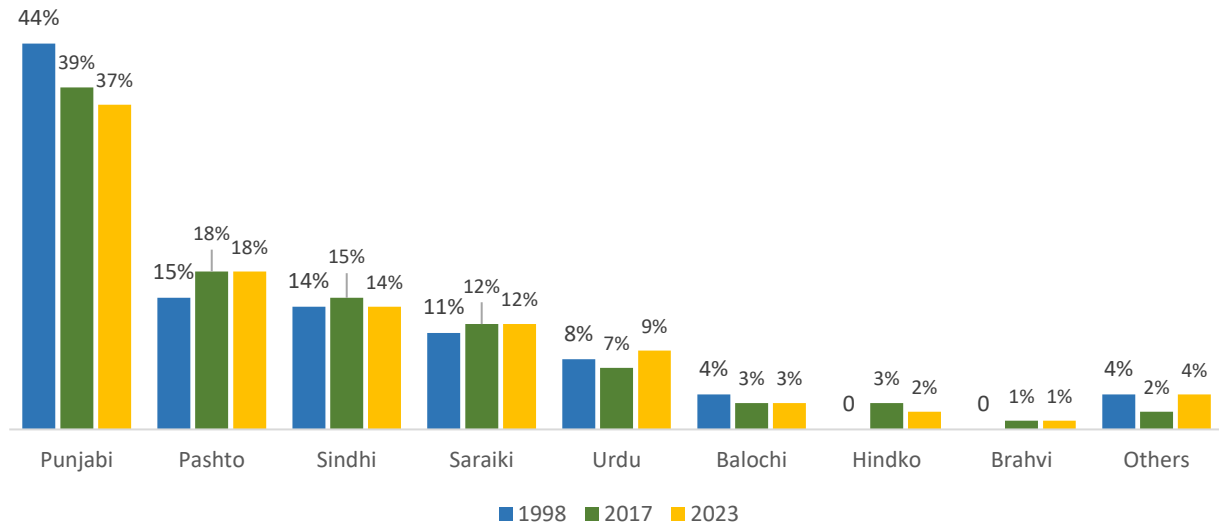


Source: 2023 Population and Housing Census, Pakistan Bureau of Statistics

1.1. Population Breakdown – Comparison with 1998 and 2017 Censuses:

A broader historical view of this data, through a comparison with the 1998 Census, highlights a decline in dominant Punjabi language speakers, from 44% in 1998 to 37% in 2023. Meanwhile, Pashto has increased from 15% to 18%, a 3% increase in this 25-year period. Saraiki has witnessed a 1% increase in these 25 years, while other languages have either remained stable or have decreased. Sindhi saw an increase in 2017, but fell back in 2023 to the same level it was in 1998, 14%. Balochi has declined from 4% in 1998 to 3% in 2017, and it has remained at this level in 2023, showing a 1% decline. Other languages, which were at 4% in 1998, fell to 2% in 2017, but climbed up to their previous level of 4% in 2023. **The official language Urdu has increased by 1% as compared to 1998, after recovering from a decline in 2017, and now stands at 9% in 2023.**

Figure 1a: Population by Language (Mother Tongue) in Pakistan – comparison of 2023 Census with 1998 and 2017 Censuses



Source: 1998, 2017, and 2023 Censuses, Pakistan Bureau of Statistics

**Hindko and Brahvi languages were not included in the 1998 census*

For a 6-year comparison with the 2017 census data, we see in Figure 1a that Punjabi speakers decreased slightly from 39% in 2017 to 37% in 2023, a reduction of 2%, while Urdu speakers increased from 7% in 2017 to 9% in 2023. Between these 6 years, Pashto speakers remained at a stable level of 18%, as did Saraiki speakers at 12%, Balochi speakers at 3%, and Brahvi speakers at 1%. However, Sindhi speakers decreased by 1 percentage point, from 15% in 2017 to 14% in 2023. Hindko speakers also reduced, from 3% in 2017 to 2% in 2023. Demographics whose mother tongue was less spoken constituted 2% in 2017, growing to 4% in 2023.

2. Provincial Breakdown – Comparison with 2017 Census

An analysis of Pakistan's total population as per the 2023 census reveals that 53% of Pakistan's population resides in Punjab, 23% in Sindh, 17% in Khyber Pakhtunkhwa, 6% in Balochistan, and 1% in Islamabad.

This shows that **the Punjabi speaking population is disproportionately less than the province's population; Punjab's share is 53% of the total population while 37% of the residents claimed Punjabi to be their mother tongue in 2023.** This highlights the diversity of population with differing mother tongues residing in the province. **A similar trend can also be seen for Sindhi** where the province has 23% of Pakistan's population but the total share of Sindhi-speaking nationals is only 14%. **Pashto speaking citizens (18%) are proportionate to the provincial share of Khyber Pakhtunkhwa (17%).**

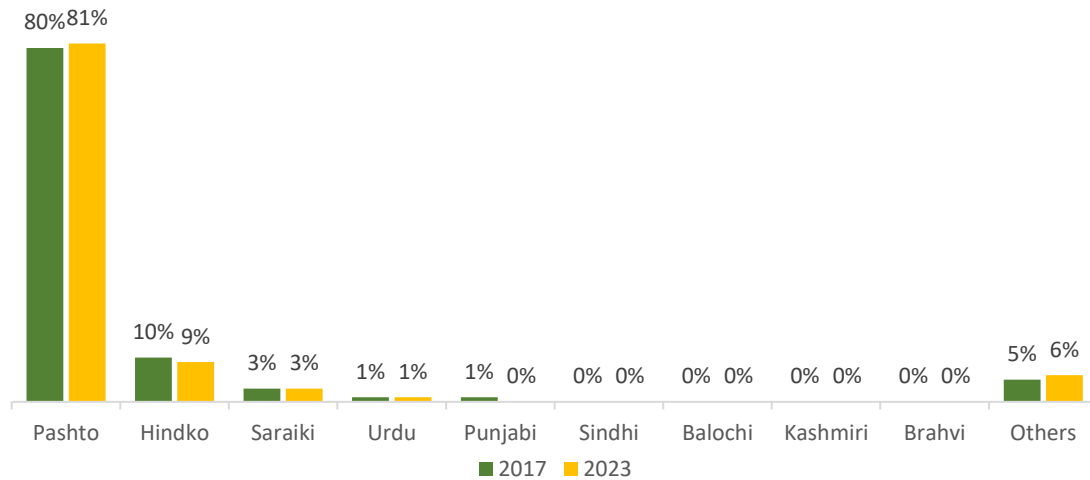
Figure 2: Provincial Breakdown of Regional Languages in 2017 and 2023

Region	Year	Urdu	Punjabi	Sindhi	Pashto	Balochi	Kashmiri	Saraiki	Hindko	Brahvi	Others	Total
Pakistan	2017	7%	39%	15%	18%	3%	0%	12%	2%	1%	3%	100%
	2023	9%	37%	14%	18%	3%	0%	12%	2%	1%	4%	100%
KPK	2017	1%	1%	0%	80%	0%	0%	3%	10%	0%	5%	100%
	2023	1%	0%	0%	81%	0%	0%	3%	9%	0%	6%	100%
Punjab	2017	5%	70%	0%	2%	1%	0%	21%	0%	0%	1%	100%
	2023	7%	67%	0%	2%	1%	0%	21%	0%	0%	2%	100%
Sindh	2017	18%	5%	62%	5%	2%	0%	2%	2%	1%	3%	100%
	2023	22%	4%	60%	5%	2%	0%	2%	1%	1%	3%	100%
Balochistan	2017	1%	1%	5%	35%	35%	0%	3%	0%	17%	3%	100%
	2023	1%	1%	4%	34%	40%	0%	2%	0%	17%	1%	100%
Federal Capital	2017	12%	52%	1%	19%	0%	2%	2%	6%	0%	6%	100%
	2023	16%	51%	1%	18%	0%	2%	2%	6%	0%	4%	100%

Source: 2017 and 2023 Population and Housing Censuses, Pakistan Bureau of Statistics

Figure 2 depicts how the language share for the entire country and the separate provinces has evolved over the past 6 years. Each province is naturally dominated by speakers of the regional languages they share their names with. **Balochistan splits its main language share between Balochi and Pashto speakers, owing to the significant Pashtun population living there and the province's shared border with Afghanistan, where one of the official languages is Pashto.**

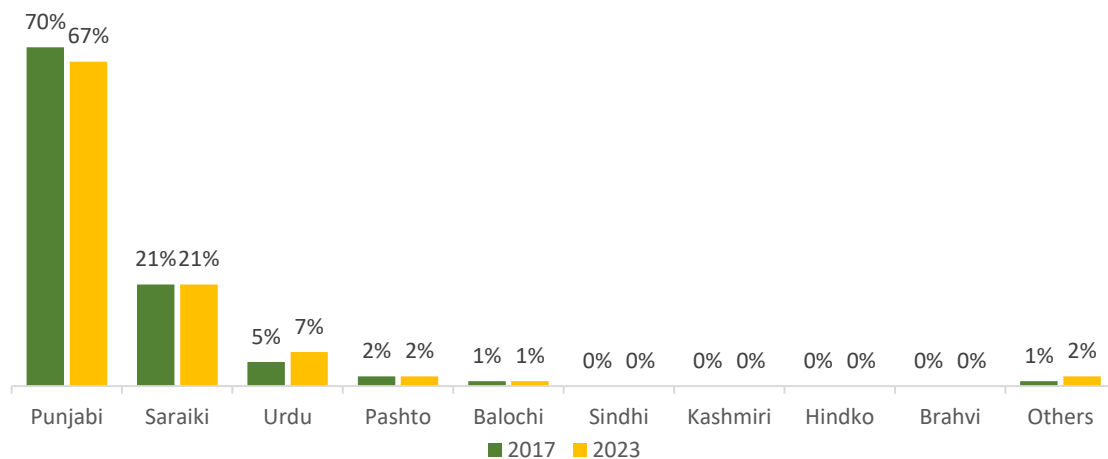
Figure 2a: KPK's language distribution – Comparison of 2017 and 2023 Censuses



Source: 2017 and 2023 Population and Housing Censuses, Pakistan Bureau of Statistics

As depicted in Figure 2a, **KPK's main regional language, Pashto, saw a 1% increase from 80% in 2017 to 81% in 2023. Meanwhile, the region's second-largest language by proportion of speakers, Hindko, witnessed a 1% decrease from 10% in 2017 to 9% in 2023.** Saraiki speakers maintained their proportional constancy at 3% from 2017 to 2023. Urdu speakers in KPK were stable at 1%, while Punjabi speakers fell from 1% to 0% during these 6 years. Sindhi, Balochi, Kashmiri, and Brahvi languages all remained at 0%, while speakers of other regional languages increased slightly from 5% in 2017 to 6% in 2023.

Figure 2b: Punjab's language distribution – Comparison of 2017 and 2023 Censuses

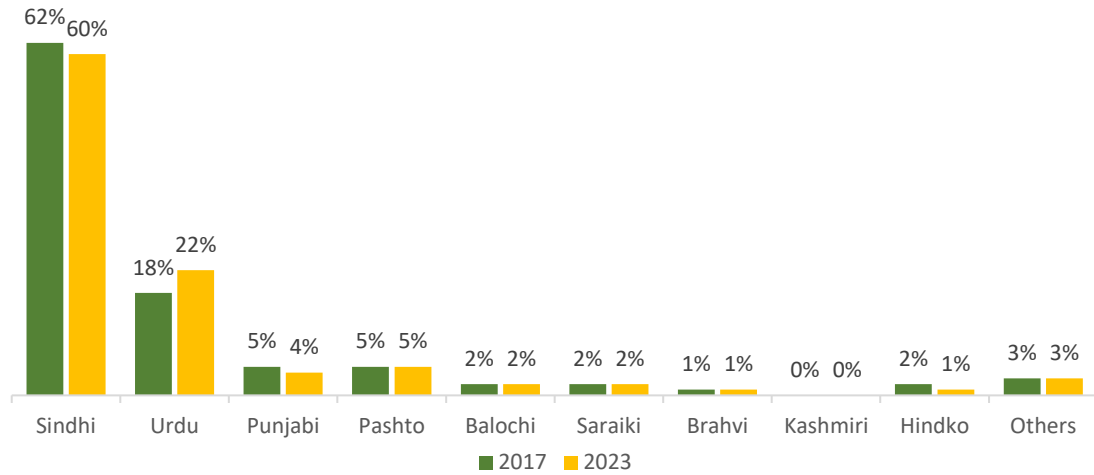


Source: 2017 and 2023 Population and Housing Censuses, Pakistan Bureau of Statistics

The Punjab province's main regional language Punjabi has seen its speakers decline by 3%, from 70% in 2017 to 67% in 2023. Saraiki speakers, who make up the second-largest language group in the Punjab province, have also been stable at 21% between these 6 years. Punjab's language distribution in Figure 2b highlights an increase in Urdu speakers during these 6 years, from 5% in 2017 to 7% in 2023. Speakers of the Pashto language have remained constant at 2%, while Balochi speakers and Sindhi speakers have not fluctuated from

1%. Meanwhile, speakers of the Kashmiri, Hindko and Brahvi languages have remained at 0% during this period. Speakers of other languages increased from 1% in 2017 to 2% in 2023.

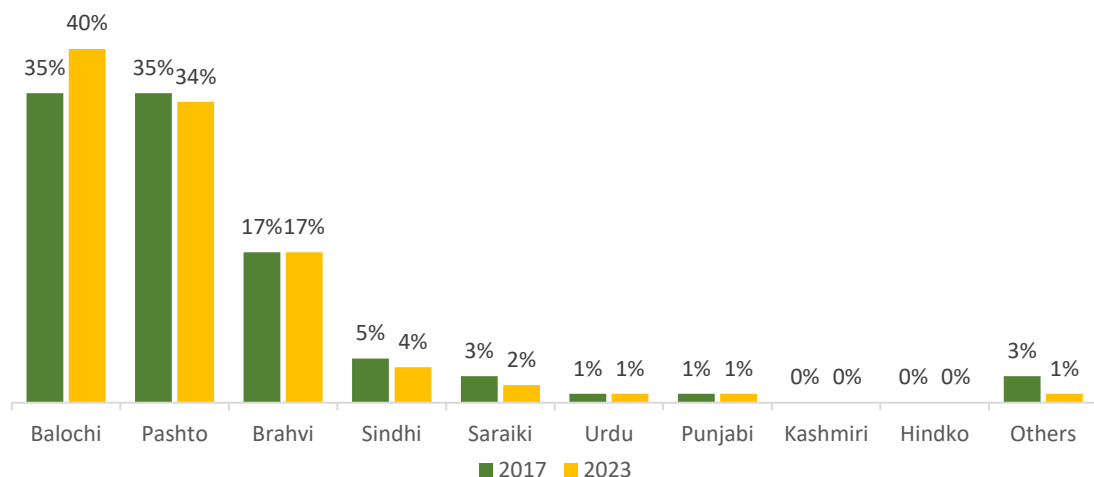
Figure 2c: Sindh's language distribution – Comparison of 2017 and 2023 Censuses



Source: 2017 and 2023 Population and Housing Censuses, Pakistan Bureau of Statistics

Sindhi speakers of Sindh experienced a 2% reduction in their proportion, from 62% in 2017 to 60% in 2023. Similar to Punjab, Sindh's provincial breakdown also underscores a rise in Urdu speakers in the province. **Urdu speakers constitute the second-largest language group in Sindh—an increase from 18% Urdu speakers in 2017 to 22% in 2023 was noted.** Punjabi speakers decreased slightly from 5% in 2017 to 4% in 2023. The proportion of speakers for the following languages remained constant during this 6-year period: Pashto at 5%, Balochi at 2%, Saraiki at 2%, Brahvi at 1%, and Kashmiri at 0%. Hindko speakers declined from 2% to 1% between 2017 and 2023. Speakers of other regional languages were also stable at 3% throughout this period.

Figure 2d: Balochistan's language distribution – Comparison of 2017 and 2023 Censuses

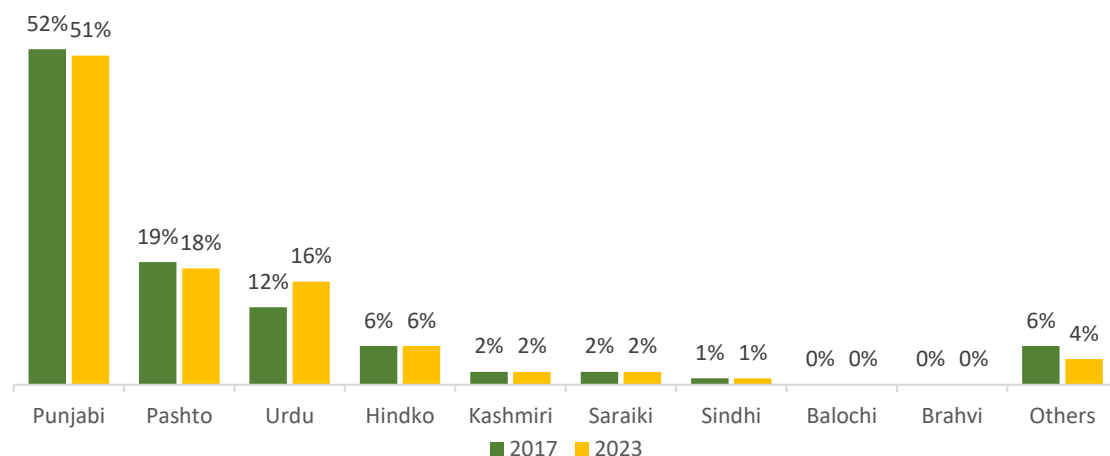


Source: 2017 and 2023 Population and Housing Censuses, Pakistan Bureau of Statistics

In Balochistan, from the two main regional language groups, the Balochi speakers increased from 35% in 2017 to 40% in 2023, while speakers of the Pashto language fell

slightly from 35% to 34% over these 6 years. As seen in Figure 2d, Brahvi speakers make up the third-largest proportion of regional language speakers in Balochistan. They remained stable at 17% during this 6-year period. Sindhi speakers of this region saw a 1% reduction as Sindhi speakers declined from 5% in 2017 to 4% in 2023. Similarly, Saraiki speakers decreased from 3% in 2017 to 2% in 2023. Urdu and Punjabi speakers both remained constant at 1%, and Kashmiri and Hindko speakers remained at 0% throughout this period. Speakers of other languages fell from 3% in 2017 to 1% in 2023.

Figure 2e: Islamabad Capital Territory's language distribution – Comparison of 2017 and 2023 Censuses



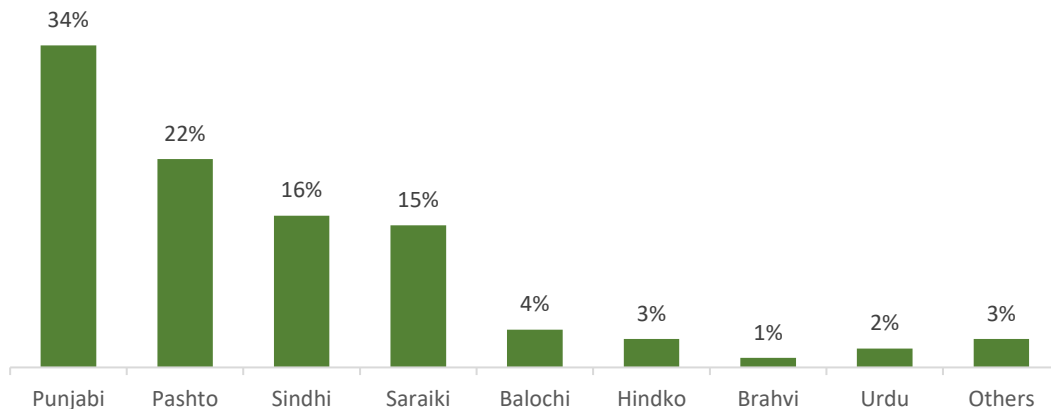
Source: 2017 and 2023 Population and Housing Censuses, Pakistan Bureau of Statistics

The division of regional language speakers in the Federal Capital of Islamabad highlighted Punjabi as the most widely spoken language, followed by Pashto and Urdu (Figure 2e). Punjabi speakers decreased slightly from 52% to 51%, and Pashto speakers experienced a similar 1% decline, from 19% in 2017 to 18% in 2023. Meanwhile, Urdu speakers witnessed an increase in their proportion from 12% in 2017 to 16% in 2023. Several language speakers remained constant in their proportions over this 6-year period: Hindko at 6%, Kashmiri at 2%, Saraiki at 2%, Sindhi at 1%, and Balochi and Brahvi both at 0%. Speakers of other languages decreased from 6% in 2017 to 4% in 2023.

3. Regional Language Distribution in Rural Areas:

As per the 2023 census, 61% of Pakistan's population reside in rural areas. Among the rural areas, **34% of rural Pakistanis claimed Punjabi to be their mother tongue**. This was followed by Pashto (22%), Sindhi (16%), Saraiki (15%), Balochi (4%), Hindko (3%), Brahvi (1%) and others (3%). **Only 2% of Pakistan's rural population were Urdu speakers.**

Figure 3: Regional Language Distribution in Rural Areas in 2023



Source: 2023 Population and Housing Census, Pakistan Bureau of Statistics

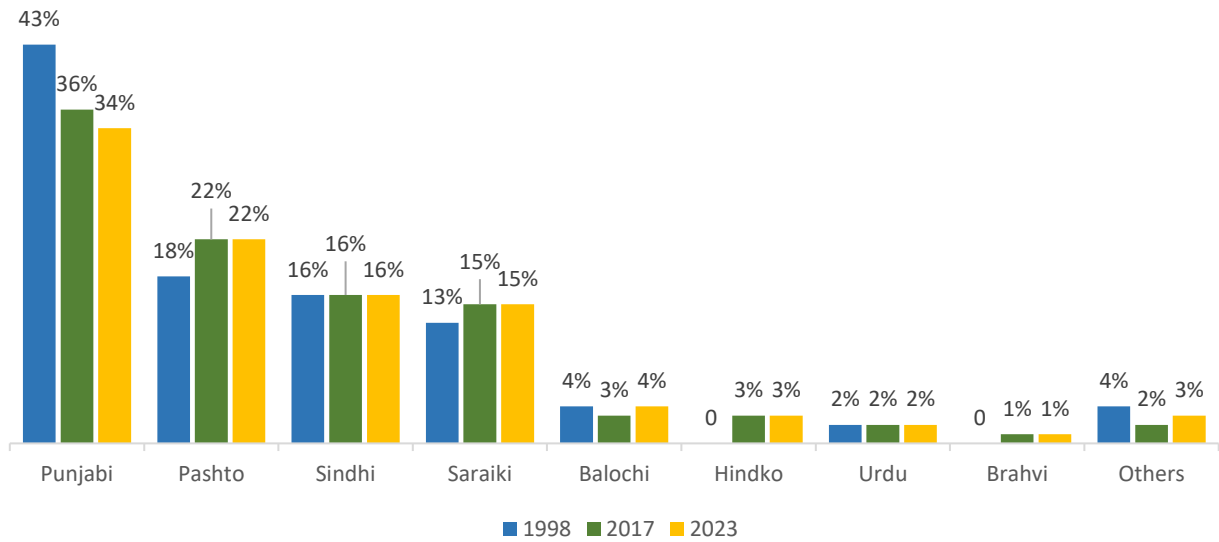
3.1. Regional Language Distribution in Rural Areas – Comparison with 1998 and 2017 Censuses:

The comparison with the 1998 Census for rural area data reveals a trend that is mostly similar to the one seen in Figure 1, except for large differentiations in Urdu. As the official language, Urdu has been more popular among urban areas as compared to the rural areas. Figure 3a visualizes this trend, showing how **Urdu speakers have remained stable in rural areas at the low rate of 2% throughout this 25-year period.**

Meanwhile, **Punjabi has witnessed a dramatic 9% decline in rural speakers over this quarter of a century**, steadily decreasing over the years to 34% in 2023. **Rural Pashto speakers rose from 18% in 1998 to 22% in 2017, and remained at this new higher level in 2023.** Rural residents who claim Sindhi as their mother tongue were anchored at 16% throughout these 25 years, without experiencing any rises or falls.

Saraiki rural speakers increased from 13% in 1998 to 15% in 2017, and persisted at this level in 2023 as well. Balochi also regained its foothold in rural areas in 2023, returning to 4%, the same level as in 1998, after experiencing a 1% dip in 2017.

Figure 3a: Regional Language Distribution in Rural Areas Comparison with 1998 and 2017 Censuses



Source: 1998, 2017, and 2023 Censuses, Pakistan Bureau of Statistics

**Hindko and Brahvi languages were not included in the 1998 census*

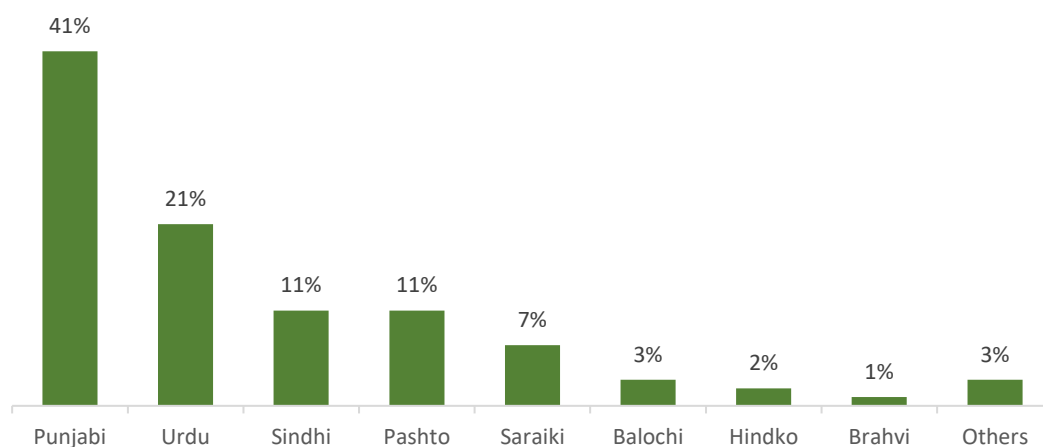
Hindko and Brahvi speakers were not recorded in the 1998 census, but they experienced no changes between 2017 and 2023, remaining at the same consistent levels of 3% and 1%, respectively. Rural residents whose mother tongues were other languages were 4% in 1998, which decreased to 2% in 2017, and then increased to 3% in 2023.

4. Regional Language Distribution in Urban Areas:

According to the 2023 Census, 39% of Pakistan's population lives in urban areas. **Punjabi speakers constitute the highest proportion of these urban residents, at 41%.** This is followed by Urdu speakers at 21%, Sindhi and Pashto both with 11%, Saraiki at 7%, Balochi at 3%, Hindko with 2%, and Brahvi speakers being 1%. Speakers of other regional languages constitute 3%.

The proportion of Urdu speakers who reside in urban areas (21%) are significantly higher than those who reside in rural areas (2%).

Figure 4: Regional Language Distribution in Urban Areas in 2023



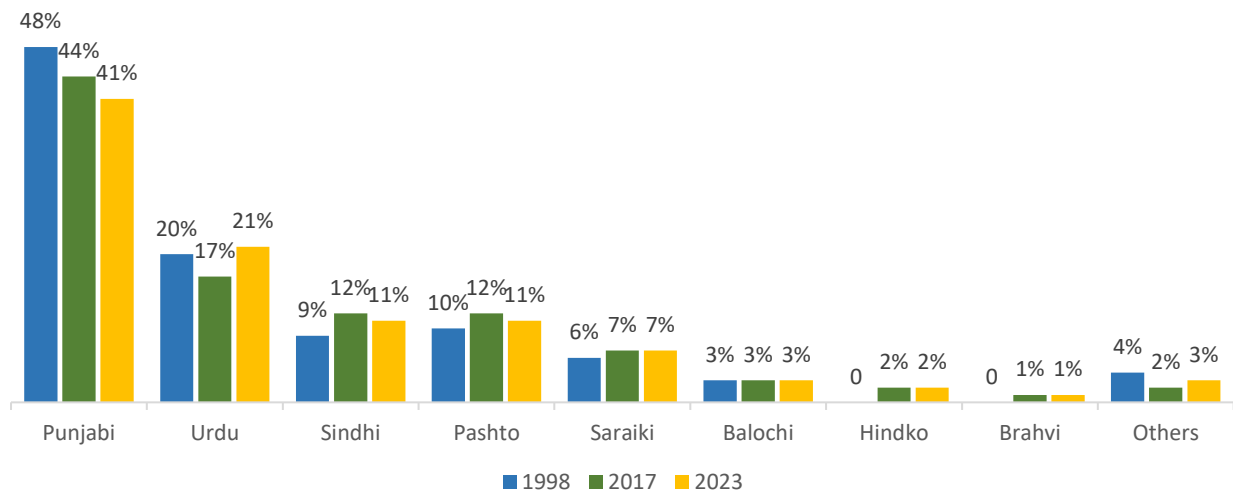
Source: 2023 Population and Housing Census, Pakistan Bureau of Statistics

4.1. Regional Language Distribution in Urban Areas – Comparison with 1998 and 2017 Censuses:

Figure 4a illustrates the trend in the population of speakers of different languages residing in urban areas for the years 1998, 2017, and 2023. **As the largest regional language, Punjabi has declined in terms of its urban speakers as well as rural speakers, falling steadily over this quarter century by a total of 7% in urban areas. Urban Urdu speakers have fluctuated during this period, starting at 20% in 1998, falling to 17% in 2017, but rising again to 21% in 2023, thus witnessing a net increase of 1% between 1998 and 2023.**

Urban Sindhi speakers have increased from 9% in 1998 to 12% in 2017, halting at the slightly reduced rate of 11% in 2023. While they have increased by a net 2% over 25 years, Sindhi speakers have declined during the 6-year period between 2017 and 2023. Pashto speakers have followed a similar pattern, starting at 10% in 1998, increasing to 12% in 2017, and dipping slightly in 2023 to 11%.

Figure 4a: Regional Language Distribution in Urban Areas, Comparison with 1998 and 2017 Censuses



Source: 1998, 2017, and 2023 Censuses, Pakistan Bureau of Statistics

**Hindko and Brahvi languages were not included in the 1998 census*

The proportion of Saraiki speakers in urban areas increased from 6% in 1998 to 7% in 2017 and has remained at 7% in 2023. Balochi speakers residing in urban areas remained consistently at 3% throughout this 25-year period. Hindko and Brahvi speakers were not recorded in the 1998 census, but their numbers remained stable between 2017 and 2023, at 2% and 1%, respectively. Speakers of other languages who reside in urban areas were the highest in 4% in 1998, falling to 2% in 2017, and rising to 3% in 2023.

Wish to know more about 2023 Census?

The 7th Population and Housing Census of Pakistan started self-enumeration from February 20th and field operations from March 1st, 2023, with results delivered by April 30th, 2023. It was Pakistan's first digital census under the vision of "Digital Pakistan". The technology architecture for this first-ever digital census was developed after extensive research and learning from other countries that have digitized their censuses, and tailored accordingly to Pakistan's specific conditions and available resources. The digitization of the census makes the process more inclusive, accessible, and user-friendly. This international standard census has provided timely data for policy planning and enabled real-time monitoring for the provision of quality data.

You can find more information on the national census and complete tables on www.pbs.gov.pk

You can find Gallup Pakistan's 2017 Census Report on Mother Tongues in Pakistan [here](#).

2017 Census Series

Volume 1: Male Literacy Ratio: <https://www.gallup.com.pk/post/32297>

Volume 2: Female Literacy Ratio: <https://www.gallup.com.pk/post/32361>

Volume 3: Population by Administrative Units: <https://www.gallup.com.pk/post/32389>

Volume 4: Level of Congestion: <https://www.gallup.com.pk/post/32423>

Volume 5: Sources of Energy: <https://www.gallup.com.pk/post/32452>

Volume 6: Sources of Drinking Water: <https://www.gallup.com.pk/post/32471>

Volume 7: Sources of Information/Communication: <https://www.gallup.com.pk/post/32496>

Volume 8: Owned Housing Units by Period of Construction:
<https://www.gallup.com.pk/post/32525>

Volume 9: Construction Material Used for Outer Walls:
<https://www.gallup.com.pk/post/32564>

Volume 10: Population with Disability: <https://www.gallup.com.pk/post/32625>

Volume 11: Pakistani Citizens Holding Computerized National Identity Cards:
<https://www.gallup.com.pk/post/32720>

Volume 12: Number of Persons Living Abroad: <https://www.gallup.com.pk/post/32720>

Volume 13: Homeless Population: <https://gallup.com.pk/post/32888>

Volume 14: Mother Tongue: <https://gallup.com.pk/post/33151>

Wish to Contribute to Gallup Big Data Analysis?

Gallup Pakistan is looking for collaboration with researchers to expand both the Big Data Analysis project as well as this study of national level data. If you have any ideas, please write to Bilal I Gilani, Project Director of the Big Data Analysis Series at Gallup Pakistan.

Methodology:

The data for this analysis is based on the “7th Population and Housing Census” by the Pakistan Bureau of Statistics, Government of Pakistan.

The 2023 Census, or the “7th Population and Housing Census”, by the Pakistan Bureau of Statistics can be accessed at: <https://www.pbs.gov.pk/digital-census/detailed-results>

The 2017 Census, or the “6th Population and Housing Census 2017”, by the Pakistan Bureau of Statistics can be accessed at: <https://www.pbs.gov.pk/content/final-results-census-2017>

The 1998 Census data on ‘Population by Mother Tongue’ by the Pakistan Bureau of Statistics can be accessed at: <https://www.pbs.gov.pk/node/97>

For more survey data, see out website <http://www.gallup.com.pk/> or contact Gallup Pakistan at caf@gallup.com.pk

Disclaimer: Gallup Pakistan has not conducted the survey work. We are only analyzing the publicly disclosed data and reports provided by the Government of Pakistan. This is a public service activity done on individual initiative of Gallup Pakistan and not in any way endorsed, commissioned, or supported by the Government of Pakistan.

Disclaimer: Gallup Pakistan is not related to Gallup Inc. headquartered in Washington D.C. USA. We require that our surveys be credited fully as Gallup Pakistan (not Gallup or Gallup Poll). We disclaim any responsibility for surveys pertaining to Pakistani public opinion except those carried out by Gallup Pakistan, the Pakistani affiliate of Gallup International Association. For details on Gallup International Association see website: www.gallup-international.com

Disclaimer: The views and inferences expressed in the article are that of the author themselves and Gallup Pakistan does not take any responsibility in this regard. This series, and many such initiatives, are internally funded by Gallup Pakistan and Gilani Research Foundation. No outside country or local funding has been received for this current activity.



Disclaimer: Gallup Pakistan is not related to Gallup Inc. headquartered in Washington D.C. USA. We require that our surveys be credited fully as Gallup Pakistan (not Gallup or Gallup Poll). We disclaim any responsibility for surveys pertaining to Pakistani public opinion except those carried out by Gallup Pakistan, the Pakistani affiliate of Gallup International Association. For details on Gallup International Association see website: www.gallup-international.com

Contact Details:

Islamabad : +92 51 2655630

Email: isb@gallup.com.pk

www.gallup.com.pk