

Indigenous Identity and Mother Tongue: A Linguistic Diagnosis of University Students in Puebla, Mexico

Abelardo Romero-Fernández¹, Fabio Morandín-Ahuerma¹, Laura Villanueva-Méndez¹,
Yonatan Puón Castro², Verónica Sánchez Hernández², Annel Sapianni Palestina-Aguilar¹

¹Complejo Regional Nororiental, Benemérita Universidad Autónoma de Puebla, Puebla, Mexico

²Facultad de Lenguas, Benemérita Universidad Autónoma de Puebla, Puebla, Mexico

Email: abelardo.romero@correo.buap.mx

How to cite this paper: Romero-Fernández, A., Morandín-Ahuerma, F., Villanueva-Méndez, L., Puón Castro, Y., Sánchez Hernández, V., & Palestina-Aguilar, A. S. (2026). Indigenous Identity and Mother Tongue: A Linguistic Diagnosis of University Students in Puebla, Mexico. *Advances in Anthropology*, 16, 167-184. <https://doi.org/10.4236/aa.2026.164009>

Received: September 18, 2026

Accepted: September 25, 2026

Published: September 28, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This study aimed to describe the gap between Indigenous self-identification and reported mother tongue among undergraduate students of the Northeastern Regional Complex (CRN) of the Benemérita Universidad Autónoma de Puebla, Mexico, explore the distribution of this gap across campuses, and make visible the discrimination, institutional ideologies, and perceived language shift surrounding it. A questionnaire was administered online through Microsoft Forms to 747 students across four campuses; 721 gave consent and constituted the sample for analysis, which was conducted using descriptive statistics given the small size of the Indigenous-language-speaking subgroups, focusing on the three campuses with sufficient sample size: Teziutlán, Libres, and Zacapoaxtla. Among students who self-identified as Indigenous (n = 74), 78.4% (58 of 74) reported Spanish, rather than an Indigenous language, as their mother tongue, a gap that varied notably across campuses and coexisted with a linguistic hierarchy positioning English above Indigenous languages in participants' repertoire, with subtle forms of linguistic discrimination not always recognized as such, and with a perception widely shared among Indigenous-language speakers that Spanish is actively replacing their mother tongues. The findings offer an empirical diagnosis that could guide institutional efforts toward linguistic inclusion at the CRN-BUAP, and represent the first systematic diagnosis of this gap among university students in the Sierra Nororiental of Puebla, a region absent from existing literature on Indigenous higher-education populations in Mexico.

Keywords

Indigenous Self-Identification, Mother Tongue, Language Shift, Higher Education, Puebla, Mexico

1. Introduction

Indigenous language shift is a phenomenon of global scale that is systematically linked to economic development and to the spread and expansion of dominant languages (Amano et al., 2014). Within this landscape, one of its less visible, yet increasingly documented, manifestations is the gap between Indigenous ethnic self-identification and actual Indigenous language use. That is, people who identify as Indigenous but do not speak the language of their people, and Indigenous-language speakers who do not always self-identify as such (Martínez et al., 2026).

Across Latin America, this gap has been consistently documented in different countries. Martínez et al. (2026), drawing on data collected in Bolivia, Guatemala, Mexico, and Peru, countries that together concentrate more than two-thirds of the Indigenous population in the Americas, found that Indigenous language proficiency predicts Indigenous self-identification, but not consistently: there are people who self-identify as Indigenous without speaking the language, and Indigenous-language speakers who do not self-identify as such.

In Mexico, this gap is considerably wide. According to the 2020 Population and Housing Census of the National Institute of Statistics and Geography (INEGI), of the 23.2 million people who self-identified as Indigenous, only 7.1 million (30.8%) spoke an Indigenous language, while 16.1 million (69.2%) did not. This gap has been documented at the national level as a phenomenon that requires treating self-ascription and language use as explicitly distinct analytical variables, which are neither interchangeable nor equivalent for public policy design (Ibarra-López et al., 2020).

The Northeastern Regional Complex of the BUAP, with campuses in Teziutlán, Libres, and Zacapoaxtla, was created following the reform of the Organic Statute approved by the University Council in December 2016, which recognized Regional Complexes as multi-campus academic units aimed at deconcentrating and territorially organizing the institution's academic activities (Benemérita Universidad Autónoma de Puebla, 2016). To date, the complex offers 13 undergraduate programs across its four campuses. In Teziutlán: Nursing, Dentistry, Physiotherapy, Medicina General y Comunitaria, Clinical Nutrition, Psychology, and Pharmaceutical Chemistry (the latter incorporated in the 2025-2026 cycle); in Libres: Business Administration, Public Accounting, and Law; in Zacapoaxtla: Business Administration and Public Accounting; and in Cuetzalan del Progreso: Gestión Territorial e Identidad Biocultural. The Teziutlán campus also offers an Associate Degree in Imaging Technology (Benemérita Universidad Autónoma de Puebla, 2025). The complex is part of a regionalization model in which eight out of every ten first-year BUAP students come from a regional complex (Benemérita Universidad Autónoma de Puebla, 2024).

Teziutlán and the Sierra Nororiental of Puebla constitute a region of high cultural diversity: according to the 2020 INEGI Census, the municipality registers 103,246 inhabitants, of whom at least 6,297 speak Náhuatl, 175 speak Totonac, and 21 speak Maya (Instituto Nacional de Estadística y Geografía, 2020), lan-

guages that reflect the historical presence of Indigenous communities in the area. The economically active population (31,296 people) is concentrated mainly in maquila and manufacturing activities, commerce (close to 10,000 establishments), the textile and furniture industry, mining (the Autlán ferromanganese plant, operating since 1970), and hydroelectric power generation (the Atexcaco Power Plant, operating since 2013) (Instituto Nacional de Estadística y Geografía, 2020).

In this context, the creation of the Northeastern Regional Complex takes on particular relevance: regions such as the Sierra Nororiental of Puebla have been identified in the literature as “educational deserts”, a term designating geographic areas historically lacking quality public higher-education provision (González Reyes et al., 2022; Mejía-Pérez & Worthman, 2017). Research on university decentralization in Mexico shows that, in municipalities where the only option for accessing higher education is a regional campus of an autonomous university, the population tends to depend on primary economic activities, presents lower prior educational attainment, and faces higher dropout rates (Mejía-Pérez & Worthman, 2017). In this sense, academic units such as the Northeastern Regional Complex not only expand access to higher education in a historically underserved region, but also become a key point of contact for addressing the wellbeing and mental health of a student population that, in many cases, is the first generation in their family to reach this educational level.

The persistence of Indigenous ethnic identity, even when the Indigenous language may stop being transmitted as a mother tongue, has recently been noted in the sociolinguistic literature. Martínez et al. (2026) found, across four Latin American countries including Mexico, that although Indigenous language proficiency can be related to Indigenous self-identification, this relationship is not determinative: some people self-identify as Indigenous without necessarily mastering the ancestral language. Similarly, Muehlmann (2008) found that among young Cucapá people in Baja California, Mexico, Indigenous identity is located in the awareness of a shared history of colonization and marginalization, rather than solely in language fluency. Likewise, Shulist (2016) describes how, in multilingual communities, the process of ethnic “graduated authenticity” rests on dimensions that can extend beyond linguistic competence. Güemez (2026) found that approximately 16% of people in Mexico changed their Indigenous self-identification at least once over a ten-year period, suggesting that the gap observed in specific populations reflects a dynamic social phenomenon.

The hierarchy among Indigenous languages, Spanish, and English in Latin American higher education has been characterized as a structural, rather than incidental, phenomenon. Roux and Soler Millán (2024) found that in Colombia, although official frameworks exist to protect and preserve native languages, these coexist with “unilingualism” and “unibilingualism” ideologies that in practice favor and privilege Spanish and English, structurally sidelining Indigenous languages. In the Mexican context, Despaigne (2019) noted that it is easier to find English courses than Indigenous-language courses in educational institutions, ev-

idence of how English has become successfully embedded in the country's institutional linguistic ecosystem.

Territorial heterogeneity within a single region can generate marked differences in Indigenous language vitality, even among geographically close localities. [Cannuto \(2017\)](#), through a comparative study of two P'urhépecha communities in Michoacán—Ichán and Tacuro—showed how intergenerational transmission at the family level favors the continuity of widespread everyday Indigenous language use in both localities. In a context closer to the present study, [Gomashie \(2022\)](#) reported that, in Santiago Tlaxco, Puebla, bilingual youth mainly use Náhuatl only with their grandparents, while with other interlocutors a mix of Spanish and both languages predominates.

According to [Ng \(2007\)](#), linguistic discrimination can be difficult to identify and visualize because it often occurs through subtle, normalized forms embedded in everyday school routines. For [Kvietok Dueñas \(2021\)](#), this dynamic is present in educational settings and relates to students ceasing to use Indigenous languages in common spaces, such as classrooms, and being frequently labeled by teachers as “language deniers”, without considering that this decision may stem from pressures within earlier socialization processes.

According to [Cashman \(2009\)](#), language shift is defined and measured in sociolinguistics through multiple complementary indicators rather than a single metric, including what the author identifies as “language panic” versus “language pride” as forces in constant tension and competition. In the Mexican context, the “ecology of pressures” model, developed by [Terborg \(2006\)](#) specifically to explain the shift of Indigenous languages toward Spanish, was applied by [Gomashie and Terborg \(2025\)](#) to the Náhuatl-speaking community of Santiago Tlaxco, Puebla.

Universities play a dual role in relation to Indigenous languages: they can act as agents of revitalization and visibility when they build intercultural spaces and recognize Indigenous knowledge, but they can also reproduce monolingual pressure dynamics when inclusion does not go beyond declarative statements. In Mexico, this tension has driven the creation of specific intercultural and Indigenous universities ([Navarrete-Cazales & Alcántara-Santuario, 2015](#)). However, most higher-education institutions in the country, including those where Indigenous enrollment is concentrated, as is the case of the CRN-BUAP, are not intercultural universities.

In Latin America, research on Indigenous language shift and the identity-language gap has been studied unevenly across a limited number of cases, particularly in Chile, Bolivia, and urban Andean communities, which recent reviews have explicitly identified as a limitation of the field ([Quiroga Curín et al., 2025](#)). This geographic concentration, together with the scarcity of comparative evidence on university populations in Mexico, positions a diagnosis such as that of the North-eastern Regional Complex of the BUAP as a valuable contribution to a geographic and population gap recognized by the discipline itself.

In response to this gap, the objective of this study was to describe the gap be-

tween Indigenous self-identification and reported mother tongue among the student body of the Northeastern Regional Complex (CRN) of the Benemérita Universidad Autónoma de Puebla (BUAP), explore its distribution by campus, and document three elements of the context in which this gap is embedded: experiences of linguistic discrimination, perceptions of institutional linguistic hierarchies, and the perceived shift toward Spanish among Indigenous-language speakers.

2. Method

2.1. Design

A cross-sectional observational study was conducted (Levin, 2006), through the administration of a structured questionnaire with closed and open-ended questions, administered at a single point in time to the population of interest. This type of instrument constitutes a widely used tool for sociolinguistic data collection (Dollinger, 2012).

For the purposes of this study, mother tongue was operationally defined as the first language learned in childhood, at home or within the community, based on direct self-report (“What is your mother tongue: the first language you learned in childhood, at home or in your community?”), following the operational approach to mother tongue used in self-report language questionnaires such as the Language and Social Background Questionnaire (Anderson et al., 2018). An Indigenous-language speaker was defined as any participant who reported an Indigenous language (Náhuatl, Totonac, or another Indigenous language of the region) as their mother tongue under this definition. The “gap” examined throughout this study refers specifically to the discrepancy between Indigenous ethnic self-identification (a separate, single-choice item) and this mother-tongue variable; it does not refer to general language proficiency or current language use, which were measured through separate instrument items (Section 2.3).

2.2. Participants

Undergraduate students of the Northeastern Regional Complex (CRN) of the Benemérita Universidad Autónoma de Puebla (BUAP)—comprising the campuses of Cuetzalan, Libres, Teziutlán, and Zacapoaxtla—were invited to participate through in-person outreach in classrooms, with voluntary and anonymous participation.

A total of 747 people accessed the questionnaire. Of these, 26 (3.5%) explicitly declined to participate by answering “No, I do not wish to participate” at the initial consent question. The remaining 721 people (96.5%) gave their consent and completed the questionnaire, constituting the analytic sample for this study. The distribution across campuses was uneven: Teziutlán ($n = 602$; 83.5%), Libres ($n = 96$; 13.3%), Zacapoaxtla ($n = 22$; 3.1%), and Cuetzalan ($n = 1$; 0.1%). Comparative analyses focus on the three campuses with analytically sufficient representation: Teziutlán, Libres, and Zacapoaxtla. The sample spanned 13 academic programs,

concentrated mainly in the early semesters (2nd and 4th) and among students who identify as women.

2.3. Instrument

The “Mapa Lingüístico CRN-BUAP” questionnaire was applied, designed for this study based on self-report language questionnaires widely used and validated in research on bilingualism and language use (Anderson et al., 2018; Marian et al., 2007). The instrument focused on two reference languages, mother tongue and current language of greatest use, following the logic of instruments such as the Language and Social Background Questionnaire (LSBQ) by Anderson et al. (2018).

The instrument included, among others, the following closed-ended items analyzed in this study: Indigenous self-identification (“Do you self-identify as Indigenous?”; response options: Yes/No), mother tongue (single-choice: Spanish/Indigenous language [specify]/English/Other), current linguistic repertoire (multiple-choice checklist of languages understood, spoken, read, and written), experiences of linguistic discrimination (multiple-choice, context-specific items for schooling before university and within the CRN-BUAP), agreement with statements on institutional linguistic hierarchies (five-point Likert scale: Totally disagree to Totally agree), and, for Indigenous-language speakers, perceived replacement of the mother tongue by Spanish (single-choice: Yes, to a great extent/Yes, to some extent/No/Not sure). Participants who selected more than one language in checklist items were coded as reporting each selected language; incomplete responses to a given item were treated as missing data for that item and excluded from the corresponding analysis, without imputation. The questionnaire also included open-ended items (e.g., inviting participants to describe experiences of discrimination or to explain their perception of language replacement in their own words); these open-ended responses are not analyzed in the present study, which reports findings from the closed-ended items only, and are reserved for a separate qualitative analysis.

2.4. Procedure

Outreach for the instrument was carried out in person in classrooms across the CRN’s four campuses. The questionnaire was administered exclusively online, through Microsoft Forms, during May 2026. Participation was voluntary, anonymous, and confidential.

2.5. Data Analysis

Given the size and composition of the sample, the analysis was conducted using descriptive statistics: frequencies and percentages for categorical variables, and means/response distributions for Likert-type scales (Mursa et al., 2025). No statistical inference tests were applied, since the subgroups of interest did not meet the minimum size recommended for reliable use.

The questionnaire included two mother-tongue items: an initial direct question (used as the primary variable throughout this study, per Section 2.1) and a closing confirmation question with identical wording. Of the 31 participants who reported an Indigenous language as their mother tongue on at least one of these two items, 12 (38.7%) did so consistently on both, while 19 (61.3%) gave a different response between the initial and confirmation items. This inconsistency is discussed as a substantive finding in Section 4 rather than treated as measurement error. All analyses in this study use the initial mother-tongue item as the single analytic denominator ($n = 20$ Indigenous-language mother-tongue speakers overall, corresponding to 16 self-identified Indigenous participants and 4 who did not self-identify as such); the confirmation item was used exclusively for this consistency check and not to redefine subgroup membership elsewhere in the study.

2.6. Ethical Considerations

Participants were presented with informed consent, following principles of respectful research that is culturally appropriate for Indigenous peoples and ethnic groups (Chaparro et al., 2025). Participation was voluntary, anonymous, and confidential.

3. Results

3.1. The Gap between Indigenous Self-Identification and Mother Tongue

Of all participants ($n = 721$), 74 (10.3%) self-identified as Indigenous. Cross-referencing this variable with reported mother tongue showed that, of these 74 cases, 58 (78.4%) indicated Spanish as their mother tongue, while only 16 (21.6%) reported an Indigenous language (Table 1). Among those who did not self-identify as Indigenous ($n = 647$), 642 (99.2%) reported Spanish as their mother tongue and 4 (0.6%) an Indigenous language.

Table 1. Indigenous self-identification by reported mother tongue ($n = 721$).

Mother tongue	Self-identifies as Indigenous ($n = 74$)	Does not self-identify ($n = 647$)
Spanish	58 (78.4%)	642 (99.2%)
Indigenous language	16 (21.6%)	4 (0.6%)
English	0 (0.0%)	1 (0.2%)

Source: the authors.

Regarding the Indigenous people to which participants belonged, among those who self-identified as Indigenous ($n = 74$), 75.7% (56/74) identified as Nahua, 18.9% (14/74) as Totonac, 4.1% (3/74) as Mixtec, and 1.4% (1/74) as Triqui. Among those who reported an Indigenous language as their mother tongue ($n = 20$, using the initial mother-tongue item as the analytic denominator; see Section 2.5), 65.0% (13/20) corresponded to Náhuatl and 35.0% (7/20) to Totonac (Table 2). Regarding gender, the distribution within both subgroups was similar to that

of the overall sample: 71.6% (53/74) of those self-identified as Indigenous and 75.0% (15/20) of those reporting an Indigenous mother tongue identified as women.

Table 2. Ethnic and gender profile of the Indigenous subpopulation.

Variable	Category	n	%
Indigenous people (self-identification, n = 74)	Nahua	56	75.7
	Totonac	14	18.9
	Mixtec	3	4.1
	Triqui	1	1.4
Gender (self-identification, n = 74)	Woman	53	71.6
	Man	20	27.0
	Other	1	1.4
Specific Indigenous mother tongue (n = 20)	Náhuatl	13	65.0
	Totonac	7	35.0
Gender (Indigenous mother tongue, n = 20)	Woman	15	75.0
	Man	5	25.0

Source: the authors.

3.2. Linguistic Repertoire

Overall, 66.0% (476/721) of participants reported speaking a single language, 30.4% (219/721) two languages, and 3.6% (26/721) three or more languages (**Table 3**). Regarding the specific languages mentioned by participants, Spanish was reported by 87.5% (631/721) and English by 32.7% (236/721), making the latter the second most frequent language in the sample's repertoire, well above any Indigenous language. Among Indigenous languages, Náhuatl was the most frequently reported (6.0%; 43/721), followed by Totonac (1.5%; 11/721) and, in a single case, Triqui (0.1%; 1/721). Other foreign languages were reported less frequently: Portuguese (1.0%; 7/721), Italian (0.8%; 6/721), Korean (0.7%; 5/721), French (0.6%; 4/721), German (0.4%; 3/721), Japanese (0.3%; 2/721), and, in one case each, Chinese and Thai (0.1%; 1/721 each).

Table 3. Self-reported linguistic repertoire (n = 721).

Number of languages	n	%
1 (monolingual)	476	66.0
2 (bilingual)	219	30.4
3	18	2.5
4	4	0.6
5 or more	4	0.6

Source: the authors.

3.3. Distribution of the Gap by Campus

The proportion of Indigenous self-identification varied notably across the three campuses analyzed: Teziutlán (54/602; 9.0%), Libres (10/96; 10.4%), and Zacapoaxtla (10/22; 45.5%). Examining, within each campus, the proportion of self-identified Indigenous participants who reported an Indigenous mother tongue, differentiated patterns were found: in Teziutlán, 11 of 54 (20.4%); in Libres, 0 of 10 (0.0%); and in Zacapoaxtla, 5 of 10 (50.0%). In this context, the proportion of students who reported having experienced or observed some form of linguistic discrimination within the CRN-BUAP also varied by campus: 52 of 602 (8.6%) in Teziutlán, 6 of 96 (6.2%) in Libres, and 3 of 22 (13.6%) in Zacapoaxtla (**Table 4**).

Table 4. Indigenous self-identification, Indigenous mother tongue, and linguistic discrimination, by campus.

Campus	Indigenous self-identification	Indigenous mother tongue	Linguistic discrimination
Teziutlán	54/602 (9.0%)	11/54 (20.4%)	52/602 (8.6%)
Libres	10/96 (10.4%)	0/10 (0.0%)	6/96 (6.2%)
Zacapoaxtla	10/22 (45.5%)	5/10 (50.0%)	3/22 (13.6%)

Source: the authors.

3.4. Perceptions of Institutional Linguistic Ideologies and Hierarchies

Participants (n = 721) were asked about their level of agreement with four statements related to the status of Indigenous languages at the CRN-BUAP. 80.9% (583/721) agreed or strongly agreed that the institution has respected the cultural and linguistic diversity of the student body. However, 47.0% (339/721) also agreed or strongly agreed that Spanish is perceived as the only “appropriate” language in most academic and administrative spaces and processes. Likewise, 80.9% (583/721) considered that valuing and making Indigenous languages visible should be normalized, and 79.8% (575/721) supported the institution offering specific support for Indigenous-language-speaking students (**Table 5**).

Table 5. Level of agreement with statements on institutional linguistic ideologies and hierarchies (n = 721).

Statement	% agree or strongly agree
The CRN-BUAP respects the cultural and linguistic diversity of the student body	80.9
Spanish is expected as the only “appropriate” language in academic and administrative spaces	47.0
Valuing and making Indigenous languages visible should be normal	80.9
The CRN-BUAP should offer specific support for Indigenous-language-speaking students	79.8

Source: the authors.

3.5. Experiences of Linguistic Discrimination

While the previous section presented linguistic discrimination broken down by campus, this section describes the same variable for the sample as a whole (n = 721), distinguishing when it occurred and the type of experience reported. 5.7%

(41/721) of participants reported having experienced discrimination related to language during their schooling prior to university. Within the CRN-BUAP, 1.7% (12/721) reported having experienced it personally and 6.8% (49/721) having observed it in others; 16.1% (116/721) indicated they were not sure, and 75.4% (544/721) reported not having experienced or observed it (**Table 6**).

Table 6. Reported experiences of linguistic discrimination (n = 721).

Context	Response	n	%
Prior schooling (before university)	Yes	41	5.7
	No	680	94.3
CRN-BUAP	Experienced personally	12	1.7
	Observed in others	49	6.8
	Not sure	116	16.1
	No	544	75.4

Source: the authors.

3.6. Perceived Replacement of Mother Tongue by Spanish

Among participants who reported an Indigenous language as their mother tongue (n = 20), 15 (75.0%) responded to this item. Of these, 53.3% (8/15) perceived that Spanish has replaced their mother tongue “to a great extent”, and 13.3% (2/15) “to some extent”; overall, 66.7% (10/15) reported some degree of replacement. 33.3% (5/15) indicated they perceive no replacement, and none selected “not sure” (**Table 7**).

Table 7. Perceived replacement of mother tongue by Spanish, among Indigenous-language speakers who responded to this item (n = 15 of 20).

Response	n	%
Yes, to a great extent	8	53.3
Yes, to some extent	2	13.3
No	5	33.3
Not sure	0	0.0

Source: the authors.

4. Discussion

4.1. The Gap between Indigenous Self-Identification and Mother Tongue

This finding can be considered within the broader national context, with an important conceptual caveat. The national-level INEGI (2020) figure of 69.2% refers to the proportion of self-identified Indigenous people who do not speak an Indigenous language—a measure of current language ability—whereas the 78.4% figure reported here refers specifically to mother tongue, a related but conceptually distinct construct (see Section 2.1). These two indicators are therefore not directly

comparable, and this study's finding should not be interpreted as indicating a larger gap than the national one; rather, both figures independently point in the same direction—a substantial disconnect between Indigenous self-identification and the ancestral language, whether measured as current speaking ability (national level) or as mother tongue (this study). One possible explanation for the gap observed in the present sample may lie in the university context: entering higher education entails, for many Indigenous students, going through schooling trajectories conducted mostly in Spanish, which may have intensified language shift relative to the general population of the country. This finding aligns with the idea that ethnic identity can be maintained and sustained through pathways other than linguistic competence (Martínez et al., 2026; Muehlmann, 2008; Shulist, 2016), and with the evidence that Indigenous self-identification in Mexico is a trait that changes over time and across different contexts (Güemez, 2026). From this perspective, this gap could be regarded as evidence that, among the student community of the CRN-BUAP, ethnic belonging is defined predominantly through bonds that are not limited to language and may be associated with family ancestry, community memory, or identification with an ethnic group.

The inconsistency observed between the initial and confirmation mother-tongue items, in which 19 of 31 participants (61.3%) who reported an Indigenous language on at least one of the two items answered differently between them (Section 2.5), is consistent with this broader pattern of fluidity rather than representing a simple measurement artifact. If Indigenous self-ascription itself is a dynamic, context-dependent trait (Güemez, 2026), it is plausible that reporting an Indigenous language as one's mother tongue is similarly sensitive to how the question is framed, to the point in the questionnaire at which it is asked, or to momentary shifts in how participants position themselves linguistically. This possibility does not undermine the validity of the present findings, which rely on a single, consistently applied item as the analytic denominator (Section 2.5), but it does suggest that mother-tongue self-report among this population may be less stable than typically assumed, and that this instability is itself a phenomenon worth investigating in future research, ideally through repeated measurement over time.

4.2. Linguistic Repertoire and Institutional Hierarchy

That English is far more present than any Indigenous language in the self-reported linguistic repertoire of the participating student community is a relevant finding. It is important to distinguish this repertoire measure from the institutional-norm items reported in Section 3.4, which reflect students' perceptions of language expectations at the CRN-BUAP rather than their own language repertoire. Considered together, these two distinct measures suggest, though do not by themselves constitute direct proof of, a linguistic hierarchy consistent with that documented in higher-education institutions elsewhere in Latin America (Roux & Soler Millán, 2024) and specifically in Mexico (Despaigne, 2019). In this regard, Despaigne notes that in Mexico it is easier and more common to find English courses than Indig-

enous-language courses within educational institutions themselves, which is consistent with the repertoire data collected in this study. This does not imply that English is directly displacing Indigenous languages—both occupy different positions within the pattern observed—but the combined repertoire and institutional-norm findings are consistent with symbolic prestige and the availability of formal instruction systematically favoring English over Náhuatl, Totonac, or any other language native to the region. However, it is worth noting that the Medicina General y Comunitaria program is the only one at the CRN-BUAP that offers Náhuatl (two courses) and Totonac (two courses) courses as part of students' basic training, with the aim that they can communicate adequately with the communities where they might practice.

4.3. Regional Heterogeneity of the Gap

Among the three campuses, Zacapoaxtla concentrated both the highest proportion of Indigenous self-identification (45.5%) and the highest proportion of Indigenous mother tongue (50.0%), while Libres, with lower self-identification (10.4%), recorded no Indigenous mother-tongue retention whatsoever (0.0%). Teziutlán, however, introduced a relevant particularity: despite having the lowest Indigenous self-identification of the three campuses (9.0%), its mother-tongue retention (20.4%) is higher than that of Libres, indicating that both indicators reflect different community contexts in each locality and do not move proportionally across campuses, despite belonging to the same region. This heterogeneity among nearby localities has been documented in other contexts in Mexico, in which intergenerational transmission at the family level explains why nearby communities can diverge in the vitality of their Indigenous language beyond mere geographic proximity (Canuto, 2017). The case of Santiago Tlaxco, in the state of Puebla, offers a relevant parallel: there, Náhuatl remains vigorous among older generations while younger people increasingly use Spanish (Gomashie, 2022), a pattern of generational shift that could also be occurring in Zacapoaxtla, although the data from this study do not allow this to be confirmed directly. It is worth emphasizing that, given the small size of the sample subgroups, these findings should be treated as initial and exploratory, and can serve as a starting point for in-depth research with greater coverage across the three campuses.

4.4. The Paradox between Institutional Discourse and Practice

This result reproduces, almost literally, the tension between discourse and practice that Roux and Soler Millán (2024) identify in Colombian language policies: a tension in which a widely endorsed discourse of respect for diversity coexists, within the same institutional space, with a practical expectation of exclusive use of the dominant language, even though this reality is not consistent with the discourse of formal policies. The fact that students perceive that the CRN-BUAP respects cultural diversity while, at the same time, Spanish is expected as the only “appropriate” language in most academic and administrative spaces and processes

does not necessarily indicate a contradiction, but rather that different dimensions of the phenomenon may coexist: a declarative dimension, visible in official discourse and in general attitudes toward linguistic diversity, and a pragmatic dimension, present in day-to-day experiences of which language is functional and valid in academic and administrative settings. This pattern takes on particular significance when one considers that the CRN-BUAP is an educational institution without a specific intercultural origin, whose academic offerings and structure were not designed around Indigenous linguistic diversity—that is, it is not an intercultural university per se (Navarrete-Cazales & Alcántara-Santuário, 2015). In this context, the gap between attitude and practice identified in this study could reflect a structural condition of Mexican universities that do not identify as intercultural yet also receive and serve Indigenous populations without having been conceived specifically for that purpose. Nonetheless, making this phenomenon visible can be a starting point for institutional discourse on cultural diversity to align more closely with the reality experienced day to day in academic and administrative spaces and procedures.

4.5. Normalized Linguistic Discrimination

The group that responded “not sure” (16.1%) deserves its own analysis, distinct from a mere information gap. Ng (2007) described how linguistic discrimination can become normalized and embedded to the point of being woven into everyday life such that, at some point, it stops being perceived as a clear or explicit discriminatory act, even when it has a real effect on Indigenous-language speakers. From this perspective, the uncertainty reported by one-sixth of the sample’s participants may reflect difficulty recognizing discriminatory experiences when they occur in subtle forms, rather than their absence, since such experiences can be very subtle and may take the form of a simple comment about accent, an instinctive correction of pronunciation, or a joke that becomes normalized among peers.

This type of subtle linguistic discrimination can lead to the disuse of an Indigenous language in common spaces such as the university, causing speakers of these languages to be regarded as “language deniers” on the assumption that they lack interest or feel ashamed of their origin, without recognizing that this disuse could be related to prior social pressure, consistent with Kvietok Dueñas’s (2021) findings on how silence or disuse of an Indigenous language in the classroom can be misinterpreted. If some of this dynamic is also at play at the CRN-BUAP, the 16.1% who reported uncertainty could represent, in practice, a rate of normalized discrimination higher than the 8.5% who report it explicitly.

4.6. Perceived Replacement and Active Language Shift

That 66.7% (10 of the 15 respondents who answered this item) of those who speak an Indigenous language as their mother tongue perceive that Spanish has replaced their language is, perhaps, the study’s most eloquent finding, because it captures language shift as a conscious experience of the speakers themselves

rather than as an external figure calculated by researchers. [Cashman \(2009\)](#) distinguishes between “language panic” and “language pride” as competing subjective forces within a shift process; the proportion found here suggests that, for the majority of Indigenous-language speakers at the CRN-BUAP who responded to this item, the former outweighs the latter. This pattern coincides with what [Gomashie and Terborg \(2025\)](#) found in Santiago Tlaxco, Puebla, applying the “ecology of pressures” model developed specifically for the shift of Mexican Indigenous languages toward Spanish ([Terborg, 2006](#)): Náhuatl remains strongly used among older generations but loses ground to Spanish among young people. That a similar pattern emerges at the CRN-BUAP, with a different instrument and population, supports the idea that this reflects a broader regional dynamic that also runs through the Sierra Norte de Puebla, rather than an isolated particularity of a single community.

Taken together, these findings describe the following linguistic landscape for Indigenous languages at the CRN-BUAP: Indigenous identity persists strongly among the student body even as the corresponding mother tongue has begun to fall into disuse or has already been lost; this loss is unevenly concentrated across campuses and is not uniform across the region; there is an institutional linguistic hierarchy that favors Spanish and, increasingly, English; forms of linguistic discrimination exist that are not always recognized as such, yet are consciously experienced by those who still speak the Indigenous language, and which could lead to a process of active replacement. This pattern responds directly to the research gap identified in the literature on Indigenous university students in Mexico, concentrated so far in a limited number of institutions and regions ([Quiroga Curín et al., 2025](#)), and contributes empirical evidence situated in a region of the Sierra Nororiental of Puebla that had until now remained absent from that body of knowledge.

5. Conclusion

This study characterized, for the first time systematically, the gap between Indigenous self-identification and mother tongue among undergraduate students of the Northeastern Regional Complex of the BUAP. The central finding is compelling: nearly eight in ten students who self-identify as Indigenous do not have an Indigenous language as their mother tongue, a gap even more pronounced than that reported at the national level for Mexico. This gap is not evenly distributed across the territory but instead varies markedly across campuses; it coexists with an institutional linguistic hierarchy in which English far exceeds any Indigenous language in the student body’s repertoire; it is sustained by experiences of linguistic discrimination that are not always recognized as such; and it is actively perceived, by the majority of those who still speak an Indigenous language, as a process of replacement by Spanish.

The study has limitations that should be clearly acknowledged. Its cross-sectional design prevents observing how these patterns evolve over time. The sub-

groups of Indigenous-language speakers, although sufficient for descriptive analysis, are small, so the corresponding findings should be interpreted as exploratory signals rather than precise estimates. Territorial coverage was also uneven: the sample was concentrated mainly at the Teziutlán campus, and the Cuetzalan campus was not included in the comparative analysis due to its limited enrollment. Finally, since the data are self-reported, some degree of social desirability bias cannot be ruled out in responses related to discrimination and institutional ideologies.

Despite these limitations, the study contributes empirical evidence situated in a region of the country so far absent from the literature on Indigenous university students in Mexico, and offers a reference diagnosis that can guide institutional actions for linguistic inclusion at the CRN-BUAP. Future research could deepen these findings through interviews or focus groups with Indigenous-language speakers, expand coverage across campuses with lower sample representation, and conduct longitudinal follow-up on the shift patterns documented here.

Funding

This research was funded by the Vicerrectoría de Investigación y Estudios de Posgrado (VIEP), Benemérita Universidad Autónoma de Puebla, Project ID 00742, Folio 0505-PV.

Institutional Review Board Statement

This study was conducted in accordance with the Declaration of Helsinki and was reviewed and approved by the Vicerrectoría de Investigación y Estudios de Posgrado (VIEP), Benemérita Universidad Autónoma de Puebla (Project ID 00742, Folio 0505-PV, approved on 27 March 2026).

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study. Participation was voluntary, anonymous, and confidential; consent was requested explicitly at the start of the questionnaire, and those who did not consent did not proceed further.

Data Availability Statement

The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy considerations related to the identifiability of a small subpopulation of Indigenous-language speakers.

Acknowledgements

The authors thank the Benemérita Universidad Autónoma de Puebla, VIEP, and the student community of the Complejo Regional Nororiental who participated in this study. During the preparation of this manuscript, the authors used Claude

(Sonnet 5, Anthropic) to translate the manuscript into English and to review its grammar and language accuracy. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Author Contributions

Conceptualization, A.R.F.; methodology, A.R.F. and F.M.-A.; validation, Y.P.C. and V.S.H.; formal analysis, F.M.-A.; investigation, L.V.M. and A.S.P.A.; data curation, L.V.M. and A.S.P.A.; writing—original draft preparation, A.R.F.; writing—review and editing, A.R.F., F.M.-A., L.V.M., Y.P.C., V.S.H. and A.S.P.A.; supervision, A.R.F.; project administration, A.R.F.; funding acquisition, A.R.F. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Amano, T., Sandel, B., Eager, H., Bulteau, E., Svenning, J., Dalsgaard, B. et al. (2014). Global Distribution and Drivers of Language Extinction Risk. *Proceedings of the Royal Society B: Biological Sciences*, 281, Article ID: 20141574. <https://doi.org/10.1098/rspb.2014.1574>
- Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The Language and Social Background Questionnaire: Assessing Degree of Bilingualism in a Diverse Population. *Behavior Research Methods*, 50, 250–263. <https://doi.org/10.3758/s13428-017-0867-9>
- Benemérita Universidad Autónoma de Puebla (2016). *Reformas al Estatuto Orgánico de la BUAP: creación de los Complejos Regionales*. Gaceta Universidad BUAP. https://repositorio.buap.mx/rdci/public/inf_public/2016/0/suplemento_g209_2_.pdf
- Benemérita Universidad Autónoma de Puebla (2024). *La BUAP celebra 25 años de regionalización*. Boletines BUAP.
- Benemérita Universidad Autónoma de Puebla (2025). *Anuario Estadístico 2024–2025*. https://repositorio.buap.mx/rdocencia/public/inf_public/2025/0/Anuario-Estadistico-2024-2025.pdf
- Canuto, F. (2017). Mantenimiento de una lengua minorizada: El caso del purépecha en Ichán y en Tacuro, Michoacán (México). *Onomázein Revista de lingüística filología y traducción*, NE III, 77–96. <https://doi.org/10.7764/onomazein.amerindias.05>
- Cashman, H. R. (2009). The Dynamics of Spanish Maintenance and Shift in Arizona: Ethnolinguistic Vitality, Language Panic and Language Pride. *Spanish in Context*, 6, 43–68. <https://doi.org/10.1075/sic.6.1.04cas>
- Chaparro, O. L., Garzón, L., Ciuoderis, K., Vargas, A., Kendall, N., Nieto, J. V. et al. (2025). Ethical Guidelines and Best Practices for Health Research with Indigenous Peoples and Ethnic Groups in Colombia: Insights from Experience. *Ethics & Behavior*, 36, 183–198. <https://doi.org/10.1080/10508422.2025.2517153>
- Despaigne, C. (2019). EFL Teachers' Perspectives on the Role of English in Two Mexican Private Universities. *Profile: Issues in Teachers' Professional Development*, 21, 43–57. <https://doi.org/10.15446/profile.v21n1.69239>
- Dollinger, S. (2012). The Written Questionnaire as a Sociolinguistic Data Gathering Tool:

- Testing Its Validity. *Journal of English Linguistics*, 40, 74-110.
<https://doi.org/10.1177/0075424211414808>
- Gomashie, G. A. (2022). Bilingual Youth's Language Choices and Attitudes towards Nahuatl in Santiago Tlaxco, Mexico. *Journal of Multilingual and Multicultural Development*, 43, 182-193. <https://doi.org/10.1080/01434632.2020.1800716>
- Gomashie, G. A., & Terborg, R. (2025). Theory of Ecology of Pressures as a Tool for Classifying Language Shift in Bilingual Communities. *Open Linguistics*, 11, Article ID: 20250065. <https://doi.org/10.1515/opli-2025-0065>
- González Reyes, H., Mejía-Pérez, G., & González-Callejas, J. L. (2022). La distribución territorial de las Universidades para el Bienestar Benito Juárez García: Entre desiertos y espejismos educativos. *Revista Educación Superior y Sociedad (ESS)*, 34, 123-146. <https://doi.org/10.54674/ess.v34i2.625>
- Güémez, B. (2026). Ethnic Fluidity as Classification Propensities: Culture, Socioeconomic Status, and Changes in Indigenous Self-Identification in Mexico. *Sociology of Race and Ethnicity*. <https://doi.org/10.1177/23326492261432226>
- Ibarra-López, I., Calleros, H., & Cortés-Moreno, J. D. (2020). Estimaciones, identidad y la relación entre uso de la lengua y autoadscripción. *Papeles de Población*, 26, 185-211. <https://doi.org/10.22185/24487147.2020.103.07>
- Instituto Nacional de Estadística y Geografía (2020). *Censo de Población y Vivienda 2020, Cuestionario Ampliado*. INEGI.
- Kvietok Dueñas, F. (2021). 'Llegando a Secundaria Les Ha Dado Amnesia...ya No Quieren Hablar': Indigenous Speakerhood Socialization and the Creation of Language Deniers in Quechua Education. *Linguistics and Education*, 61, Article ID: 100888. <https://doi.org/10.1016/j.linged.2020.100888>
- Levin, K. A. (2006). Study Design III: Cross-Sectional Studies. *Evidence-Based Dentistry*, 7, 24-25. <https://doi.org/10.1038/sj.ebd.6400375>
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 940-967. [https://doi.org/10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067))
- Martínez, C., Telles, E., & Torche, F. (2026). From Language to Self-Identification: Indigenous Classification in the Americas. *Social Forces*, 104, 1630-1651. <https://doi.org/10.1093/sf/soaf152>
- Mejía-Pérez, G., & Worthman, S. (2017). La geografía de las oportunidades. El caso de las sedes de las universidades autónomas en municipios con poca oferta de educación superior. *Revista Iberoamericana de Educación Superior*, 8, 25-48. <https://doi.org/10.22201/iisue.20072872e.2017.23.244>
- Muehlmann, S. (2008). "Spread Your Ass Cheeks": And Other Things That Should Not Be Said in Indigenous Languages. *American Ethnologist*, 35, 34-48. <https://doi.org/10.1111/j.1548-1425.2008.00004.x>
- Mursa, R. A., Patterson, C., McErlean, G., & Halcomb, E. (2025). How Many Is Enough? Justifying Sample Size in Descriptive Quantitative Research. *Nurse Researcher*, 33, 35-40. <https://doi.org/10.7748/nr.2025.e1958>
- Navarrete-Cazales, Z., & Alcántara-Santuario, A. (2015). Universidades interculturales e indígenas en México: Desafíos académicos e institucionales. *Revista Lusófona de Educação*, 31, 145-160.
- Ng, S. H. (2007). Language-Based Discrimination: Blatant and Subtle Forms. *Journal of Language and Social Psychology*, 26, 106-122.

<https://doi.org/10.1177/0261927x07300074>

Quiroga Curín, J., Mayo, S., & Becerra Lubies, R. D. (2025). Language Revitalization and the Persistence of Mapudungun: A Review of Critical and Territory-Based Perspectives. *Íkala, Revista de Lenguaje y Cultura*, 30.

Roux, G., & Soler Millán, G. C. (2024). *Unilingualism* and *Unibilingualism* in Colombia. *Applied Linguistics*, 45, 348-363. <https://doi.org/10.1093/applin/amad030>

Shulist, S. (2016). “Graduated Authenticity”: Multilingualism, Revitalization, and Identity in the Northwest Amazon. *Language & Communication*, 47, 112-123. <https://doi.org/10.1016/j.langcom.2015.04.001>

Terborg, R. (2006). The “Ecology of Pressures” in the Shift of an Indigenous Language to Spanish: The Presentation of a Model. *Forum: Qualitative Social Research*, 7, Article 39. <https://doi.org/10.17169/fqs-7.4.167>