MEASURING VARIATION IN CENTRAL PAME VOWELS

Borja Herce, Catalina Torres
University of Zurich

ABSTRACT

Central Pame (cent2145), also known by its endonym Xi'iui, is a threatened Otomanguean language spoken in central Mexico. An earlier impressionistic account described Central Pame as having an asymmetric 5-vowel system with one more front than back vowels. This study examines a set of monophthongal, oral vowels in the language. We explore the acoustic phonetic characteristics in different speakers who are bilingual (with Spanish). Results confirm that five vowel qualities can be determined acoustically. Results show three front /i, e, ɛ/, one central vowel /ɐ/ and one back vowel /o/. We find that the realisations of the front vowel /e/ differ between speakers, with young female speakers showing a more fronted vowel. The evidence suggests a vowel chain shift is taking place in front vowels.

Keywords: Otomanguean, under-documented languages, language contact

1. INTRODUCTION

Central Pame (Oto-Pamean, Otomanguean, cent2145), also known by its endonym Xi'iui, is a language spoken by around 3000 people in and around Santa María Acapulco, in the southern part of the Mexican state of San Luis Potosí. The language was first documented via the work of SIL linguists/missionaries in the 1950s [1], [2], [3], [4]. The description of the language's phonology by Gibson [5] remains the most thorough source available to date. For all its virtues, this description is of necessity dated in its methods, as it described the state of the language 70 years ago. In addition, notable recent changes have been reported in neighbouring related languages [6]. Impressionistic field work observations suggest that front Central Pame vowels may be changing as well.

1.1 Central Pame Vowels

The vowel system of Central Pame was described as an asymmetric 5-phoneme one consisting of four front /i, e, ɛ, a/, and one back vowel /o/ [5]. The instantiations and allophonic ranges of the phonemes were impressionistically described as /i/∼/ɪ/, /i̞/∼/ɛ̞/, /ɛ/∼/æ̞/, /a/, and /ʊ/∼/u/∼/o/∼/ɒ/ respectively. No phonemic distinction between different back vowels was reported despite the wide allophonic domain. Additionally, inter-speaker variation was reported regarding the quality of vowels /i/, /e/, and /ɛ/ in certain words. Additionally, /i/ was described as alternating quite freely with a diphthong /ei/. All vowels were reported to occur phonemically as either oral or nasal, as well as stressed and unstressed. According to [5], no phonemic vowel length exists in the language.

Neighbouring related languages have been described to be relatively similar to Central Pame. Some descriptions of the closest-related one, Northern Pame [7], report a 6-vowel inventory (essentially the one described for Central Pame by [5] plus an additional innovative phoneme /ə/), while others [8] find an additional phonemic distinction between /u/ and /o/. For more distantly related Chichimec, [6] also finds 7 vowels (and a contrast between /o/ and /u/) in the speech of the oldest speakers, but increasing evidence for mergers (of /ɛ/ with /e/, and /o/ with /u/) in the speech of younger speakers more heavily influenced by Spanish. Some of these disagreements and/or findings in related languages highlight uncertainties and points of controversy that also exist in Central Pame. More recent (but much less thorough) accounts of the vowel system of the language, thus, differ from the one in [5]. Accounts in [9] and [10] for example, consider that some words have an invariably /o/ or /u/-like pronunciation of the back vowel, which might point towards a (marginal) phonemic distinction between two different back vowels.

Considering observations on neighbouring languages and the lack of phonetic acoustic investigations in Central Pame, this study sets out to investigate its vowel system acoustically. Additionally, we examine whether, as reports for Chichimec [6], there is evidence for a potential merger of /e/ and /ɛ/, or other changes to front vowels.

2. METHODS

2.1 Data collection

Fieldwork was conducted in the communities of Santa María Acapulco and the neighbouring La Parada in 2022. Recordings were performed through an H1n Zoom sound recorder at a 48kHz sampling rate. They were performed indoors and in a quiet location of the respective communities. This study is based on data from 6 speakers (4f, 2m) between the ages of 18 and 31.

2.2 Materials

In consultation with native speakers of Central Pame we elaborated a list of 80 target words, which were designed to represent all possible combinations of stress-tone and vowel qualities. An elicitation paradigm was used to prompt participants to utter the selected words in carrier sentences in a naturalistic setting (a market where different objects, animals, etc. are sold). Speakers where exposed to visual stimuli consisting of images, and Spanish translations for clarity.

2.3 Data processing

Sound files were manually transcribed, and force aligned in WebMAUS, using a parameter model based on SAMPA [11]. Phonetic alignment was manually corrected in Praat [12]. A hierarchical database was build using the EMU Speech Database Management system [13]. Utterances containing disfluencies or failing to elicit the intended target word were discarded. Acoustic measures for the first and second formants were taken at vowel mid-point using the emuR [14] package in R [15]. This study examines vowel tokens in stressed syllables, from a subset of 21 target words. These were selected to avoid coarticulation effects on the vowels due to adjacent liquid, nasal, or approximant consonants. Outliers that could be linked to tracking errors were removed. The final dataset contains 339 vowel tokens.

2.4 Statistical analysis

Measurements for F1 and F2 were Lobanov normalised and these values were used in the vowel plots made with phonR [16]. We use the Pillai-Bartlett Trace (Pillai score) to measure a potential merger between vowel pairs. The Pillai score is well suited to compare between speaker differences, capture the degree of overlap or distance between vowel categories, and analyse a small number of tokens [17]. Pillai score values range from 0 to 1 whereby values closer to 0 indicate more overlap and 1 means complete separation in the F1/F2 space. To further investigate F1 and F2 we used lmerTest [18] and computed two linear mixed effects models with vowel and sex as fixed factors and participant and word as random intercepts. To examine speaker-specific differences, we carried out additional models with participant as fixed factor and word as random intercept. We also carried out pairwise comparisons using emmeans [19].

3. RESULTS

Figure 1 shows ellipses of the five Central Pame vowels for female and male speakers. The plots show some degree of overlap between /i/, /e/, and /ɛ/ while the /ɐ/ and /o/ vowels are more clearly separated in the vowel space. Table 1 summarises the Pillai scores and statistical significance obtained. The results show that for all speakers the vowels /e/ and /ɛ/ are separate, although the separation is less strong in male speakers. These results are statistically significant, thus opposing the possibility of a vowel merger between /e/ and /ɛ/. Additionally, we find that /ɛ/ and /ɐ/ are also clearly separate for all speakers, which is also confirmed by the statistical significance obtained. Although the scores of three speakers (Fe3, Fe4, M2) suggest a (near) merger between /i/ and /e/, we find that the results for speakers Fe1 and M1 show a clear (i.e. statistically significant) separation between /e/ and /i/. Results of the linear mixed effects models confirm that most vowels differ on the F1 and F2 plan. We will report on relevant pairwise comparisons to focus on the vowels of interest. Figure 2 shows the normalized F1 values for five

Central Pame vowels. Plots indicate that for the vowel /i/ speakers vary on the F1 plan, and male speakers display lower values than females.

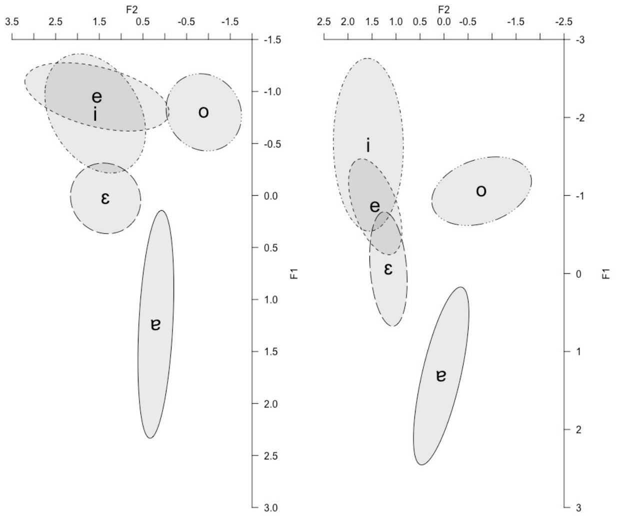

Figure 1: Lobanov normalised F1 and F2 values at midpoints of female (left) vs male (right) speakers.

Results of a model investigating F1 confirms that vowels differ. The pairwise comparisons reveal that /e/ and /ɛ/ (β = -106, p < .009) as well as /ɛ/ and /i/ (β = -120, p < .003) are separate. However, it appears that this cannot be confirmed for /i/ and /e/ (β = -14, p < .9 (n.s)). As Figure 1 shows, it appears that /i/, /e/, and /o/ are realised at a comparable height. Comparison between these vowels do not confirm a statistically significant difference.

	/e/ vs /ɛ/	/ɛ/ vs /ɐ/	/i/ vs /e/
Fe1	0.958 (p<.0001)	0.932 (p<.0001)	0.756 (p<.0008)
Fe2	0.968 (p<.0001)	0.743 (p<.0001)	0.315 (n.s)
Fe3	0.781 (p<.0001)	0.870 (p<.0001)	0.062 (n.s)
Fe4	0.778 (p<.0001)	0.865 (p<.0001)	0.152 (n.s)
M1	0.552 (p<.0001)	0.868 (p<.0001)	0.430 (p<.0001)
M2	0.387 (p<.0001)	0.934 (p<.0001)	0.078 (n.s)

Table 1: Pillai scores for speakers of different sex, statistical significance is given in parentheses.

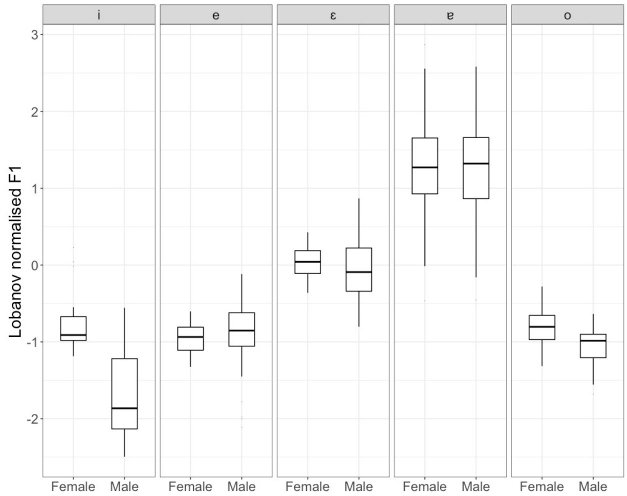

Figure 2: Boxplots of Lobanov normalised F1 values for five vowels in female (left) and male speakers (right).

Results of the model investigating F2 show a contrast between a front and a central vowel (ɛ vs a: β = 493, p < .002). Figure 3 shows the vowel /e/ displays a high degree of variability, especially for female speakers. Similarly, values for /o/ show high variance for female and male speakers. Our statistical investigation of F2 reveals that the front vowels /e/, /i/ and /ɛ/ do not differ significantly (i vs e: β = 54, p < .9 (n.s)).

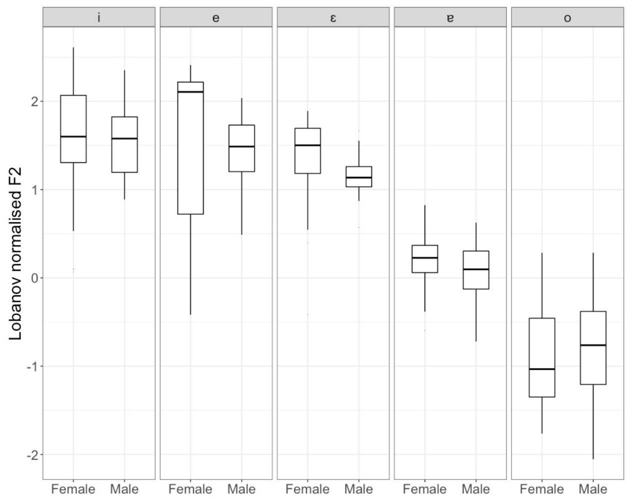

Figure 3: Boxplots of Lobanov normalised F2 values for five vowels in female (left) and male speakers (right).

Figure 4 shows the ellipses of /e/ vowel tokens for female speakers. Speakers Fe2, Fe3, and Fe4 show a high degree of fronting. Instead, speaker Fe1 realises the /e/ vowel further to the back in a less frontal position. A model that investigates speaker specific differences for F2 finds that speaker Fe1 significantly differs from all other female speakers (Fe1 vs Fe2: β = -1019, p <

.0001). We find that the two male speakers do not differ significantly. We performed the same model to investigate F1 and find a significant difference between the male speakers (M1 vs M2: $\beta = -73$, $p < .002$), although the difference is of smaller magnitude. The pairwise comparisons did not yield a speaker-specific significant difference for F1 in females.

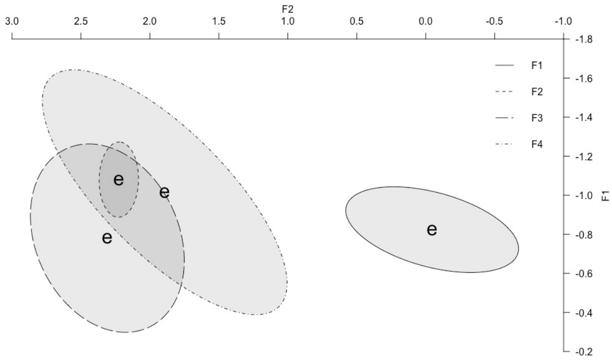

Figure 4: Lobanov normalised F1 and F2 values at midpoints of female /e/ vowels.

4. DISCUSSION

Our investigation of Central Pame confirms a five vowel inventory that can be determined acoustically. Our description of monophthongal oral vowels finds evidence for three front vowels /i, e, ɛ/, one central vowel that may be best characterised as /ɐ/, and one back vowel /o/. We find evidence that speakers vary in the realisation of front vowels, especially of the vowel /e/. In contrast to the related language Chichimec, we do not find any evidence for a vowel merger between /e/ and /ɛ/. Although male speakers seem to produce /e/ and /ɛ/ with a stronger degree of overlap, whereas female speakers clearly separate the two. Interestingly, statistical models investigating F1 and F2 find no statistically significant difference between /e/ and /i/. For the vowel /e/ we find speaker-specific differences in F2 in the female speakers and for F1 for the male speakers, showing that this vowel may be realised variably. Thus, we find that the older speakers (Fe1 31 years; M1 29 years) are different from younger speakers (18-23 years). We find increased F2 values confirming that young female speakers show a stronger degree of fronting of /e/. We speculate that this realisation could be related to some (to us unknown) sociolinguistic factors or possibly to age, since older speakers appear to display different patterns from younger speakers. The data suggests a chain shift might be taking place in the language's front vowels. Similar to the Great Vowel Shift in 15^{th} c. in English [20], front vowels appear to have become more closed, and may diphthongise when they cannot become more closed (i.e. /ɛ/>/e/, /e/>/ɪ/, /i/>/ei/) although this remains to be explored in future research with older speakers. Regarding F1 the vowel /o/ appears to be at a similar height as /e/ and /i/. However, some degree of variability is visible in the F2 values which could be indicative of a more central /ʊ/, possibly /ɵ/ or back realisation /o/. The extent to which the /o/ vowel is influenced by lexical or sociolinguistic factors goes beyond the scope of this study and remains to be explored in future work.

5. CONCLUSION

This study presents the first acoustic phonetic analysis of the Central Pame vowel system. We find evidence for a five vowel system /e, i, ɛ, ɐ, o/. Our data suggest a sound change in progress by which the front vowel /e/ is being fronted and raised. Possibly, this shift could be affecting the phonetic specification of other vowels in the inventory (/i/ and /ɛ/) which will be examined in greater detail in future work. With respect to (the) back vowel(s), although the absence of native minimal pairs suggests that the language has a single back vowel phoneme, the values for F2 are quite variable. The integration of Spanish loan words into the Pame's phonemic and phonetic system may be exerting some influence. If the original Spanish vowel distinction (i.e. /u/ vs /o/) are preserved in an increasing number of loan words (e.g. tambòl 'drum', Sp. *tambor*, krús 'cross', Sp. *cruz*), this could put additional pressure into a phonemic split in the vocabulary.

Future research will include more data from more (especially older) speakers and further examine the role of age, sex, language use, and sociolinguistic factors to account for variability in the vowel space.

7. REFERENCES

[1] Gibson, Lorna F. 1950a. Three Chichimeca texts. Unpublished manuscript. Online at https://www.sil.org/resources/archives/57422.

[2] Gibson, Lorna F. 1950b. A pedagogical grammar of Central Pame. Unpublished manuscript. Online at https://www.sil.org/resources/archives/57479.

[3] Gibson, Lorna F. 1950c. Verb paradigms in Pame. Unpublished manuscript. Online at https://www.sil.org/resources/archives/53023.

[4] Olson, Donald. 1955. Mode-Aspect-Person Inflection in Pame. (Unpublished manuscript, online at http://www.language-archives.org/item/oai:sil.org:57423)

[5] Gibson, Lorna F. 1956. Pame (Otomi) phonemics and morphophonemics. *International Journal of American Linguistics* 22, 4: 242-265.

[6] Avelino, Heriberto. 2019. The ecology of variation and change in the context of language attrition: ʔuzãʔ (Chichimec) vowels. Language Ecology 3 (1), 28–57.

[7] Berthiaume, Scott Charles. 2013. *A phonological grammar of Northern Pame*. PhD Thesis, The University of Texas at Arlington.

[8] Avelino, Heriberto. 1997. Fonología y morfofonología del Pame Norte. BA Thesis. México DF: Escuela Nacional de Antropología e Historia.

[9] Avelino, Heriberto. 1996. El sistema consonántico del Pame Central. Tercer Encuentro de Lingüística en el Noroeste. Vol I. Universidad de Sonora.

[10] Hurch, Bernhard. 2022. Pame (central) de Santa María Acapulco, Santa Catarina, San Luis Potosí. In Yolanda Lastra (Ed.) Archivo de Lenguas Indígenas de México. México DF: El Colegio de México.

[11] T. Kisler, U. Reichel, and F. Schiel. 2017. “Multilingual processing of speech via web services,” Computer Speech & Language, vol. 45, pp. 326–347.

[12] P. Boersma and D. Weenink. 2021. “Praat: doing phonetics by computer (version 6.1.52)[computer program]”.

[13] R. Winkelmann, J. Harrington, and K. Jänsch. 2017. “EMU-SDMS: Advanced speech database management and analysis in R,” Computer Speech & Language, vol. 45, pp. 392–410.

[14] R. Winkelmann, K. Jaensch, S. Cassidy, and J. Harrington. 2017. emuR: Main Package of the EMU Speech Database Management System, R package version 0.2.3.

[15] R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[16] McCloy, Daniel R. 2016. “phonR: Tools for phoneticians and phonologists.” R package version 1.0-7. http://drammock.github.io/phonR/

[17] Adank, P., Smits, R. and van Hout, R. 2004. “A comparison of vowel normalization procedures for language variation research”, *The Journal of the Acoustical Society of America*, 116(5): 3099–3107.

[18] Kuznetsova A, Brockhoff PB, Christensen RHB (2017). “lmerTest Package: Tests in Linear Mixed Effects Models.” *Journal of Statistical Software*, 82(13), 1-26. https://doi.org/10.18637/jss.v082.i13.

[19] Lenth, R .2022. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.7.4-1, https://CRAN.R-project.org/package=emmeans.

[20] Wolfe, Patricia M. 1972. *Linguistic change and the great vowel shift in English*. Univ of California Press.