

Declination of Otomi language using cellular automata to determine this linguistics behavior

Claudia Rosas¹, Alberto Ochoa² & Marcela Herrera¹

¹ Computer Science department, Universidad Autónoma de Querétaro,

² Juarez City University, Mexico

Abstract. Many causes exist for the extinction of a language, in this case scenario we talk about Spanish colonization of the Mexican territory, more specifically the central area where the Otomi civilization lived. Nowadays we can observe a palpable decrease of population which first language is Otomi. But why is it happening? what is the impact? The model of the automata simulating the situation may give a prediction of when this important language will disappear completely or decrease significantly since today there are only 327.319 active speakers. Otomi language has an important amount of information concerning to the area, history or culture, which may get lost as its speakers disappear. The software intends to achieve an accurate forecast of its declination so this knowledge can be preserved. The acquisition of updated data is from Instituto Nacional de Lenguas Indigenas (INALI) who oversees recollecting statistics in the regions of interest and its information can be accessed by anyone who requires it. It is governmental approved information so the software can be trust in results. The simulation made by the cellular automata intends to highlight the key factors affecting the conduct and time in the progression of Otomi declination.

Keywords: Cellular automata, language declination, computational simulation, Otomi.

1 Introduction

About 97% of the world's population speaks approximately 250 languages, which represents only about 4% of the world's languages; furthermore, only 3% of the world's population speaks around 5700 languages. Thus, almost all linguistic diversity of the world is guarded by a very small number of people. It is very likely that by the end of the twenty-first century approximately 5 thousand languages of the world will disappear to be replaced by dominant languages (DL). [1]. When people talk about Latin or Classical Greek they always refer as to a dead language while they refer as living languages to those that are still spoken as the first language by people in the present. There are many factors that cause the death or declination of a language such as: linguistic substitution or change of language, linguistic evolution or the disappearance of the speakers caused by violence or epidemics. The fate of languages is not always death but change. The natural course of a language is to change, often new words are added because new things are created, verbs or objects need to be identified so the language has to change to satisfy these needs. Nevertheless, in the Otomi's scenario there are two main factors: change and violence. Otomi population was murdered or christianized by Franciscan monks after the battle of *La noche triste*; the Otomies became one of the best allies of the Spaniards against the Nahuas but after the battle where they were defeated and most of them died subsequently the Spaniards decided to transform them politically and socially. Today Otomi language presents a series of characteristics that represent

the experimented changes, for example there are numbers only from 1 to 10 and the following ones are the same as they are in Spanish. Also, there are terms that cannot be translated, they just adapt to the language; for example the most evident ones are scanner, software or hardware. Such words are added to the language and change it, there are other examples in Mexico like aguacate (avocado), molcajete which come from Nahuatl and became a part of the Spanish used day-to-day. This tendency goes over in every language that pretends to evolve and adapt to survive. Every each one of the indigenous languages in Mexico contributes a great intellectual value to the linguistic diversity in the world, and Otomi is one with the greater number of speakers. It is very important to preserve these languages and avoid their total disappearance, since at least 14 of them are in danger of extinction. The number of speaker's decreases due to education system, social pressure and the fact that it is not transmitted to new generations, then it reaches the point where only adults and seniors practice it. In Mexico, when they talk about social pressure as a factor that affects the language, they refer to the stigmatization of indigenous languages that currently still exist; they are considered as a burden, as though they're one of the reasons why the country is not developing properly. Since independence period, the government chose to make Spanish the official language so the country could be homogenized, establishing it in education and legal systems without considering the use of indigenous languages and how it would affect the population. In Mexico, when they talk about social pressure as a factor that affects the language, they refer to the stigmatization of indigenous languages that currently still exist, they are considered as a burden, as though they're one of the reasons why the country is not developing properly. Since independence period, the government chose to make Spanish the official language so the country could be homogenized, establishing it in education and legal systems without considering the use of indigenous languages and how it would affect the population. Indigenous languages have prevailed in Mexico by speaking it without any grammar established. It was until recently that the Otomi grammar has been developed and delivered by the secretary of public education Emilio Chuayffet Chemor in the Día Internacional de la Lengua Materna (International day of First Language). Language shapes you. Language represents the people, ideologies, tradition, sometimes even religion. Having the opportunity to predict its behavior and save at least some of this richness is the main goal of these cellular automata.

2 Related Work

In 2009 the Journal of Artificial Societies and Social Simulation (JASSS) published an article where the shift of Catalan language in Spain was simulated, showing interesting results in the cellular automata outcome. "Predicting the future of threatened languages can be a useful way to determine the use of language policies to reverse a language shift. As the example of Catalan in Valencia has illustrated, modeling the language shift and carrying out simulations based on cellular automata can highlight some relevant factors in the speech behavior of individuals regarding the SL, which allows the future of the SL to be forecast in a given sociolinguistic"[2]

a. Importance and behavior of the language

To the question of why we should worry about preserving languages, linguist Christopher Mosley responds: "Because each language is a mental universe structured in a unique way in its kind, with associations, metaphors, and ways of thinking, vocabulary, grammar and phonetic system. All these elements work together in a structure that, because it is extremely fragile, can disappear forever with great ease." [1]

Also Dr. Miguel León-Portilla states: “There are, of course, people who consider that the death of these (indigenous) languages is inevitable and that, in addition, there is no reason to be mourn the loss because linguistic unification is highly desirable. In contrast with such an attitude, there are others who think that the disappearance of any language impoverishes humanity.[1] All the languages in which any women and men learned to think, love and pray deserve to be respected as part of their human rights. And this applies to all Amerindian languages and to all those in the world that are spoken.” [1]. If we talk about two or more languages in a community, a hierarchical structure is always adopted, with one becoming the dominant language (DL, in this case Spanish) and the other the subordinate one (SL, in this case Otomi). Both languages can live within such a hierarchy for long periods of time, but changes such as social, political and economic events can break the balance. In such cases, the pressure on the speakers of the subordinate language produces a change in their speech behavior. The speakers of the subordinate language may notice that their language has lost value relative to the dominant language. They may decide that it is no longer useful and stop speaking it in all domains of use.[2] As mentioned before, the natural process in language is to change, nevertheless in the process languages may die or suffer from a drastic decrease of its use or its impact because of several reasons and in the case of Otomi there are three main reasons considered:

- Speakers disappearance due to economic resources or death of speakers
- The speakers choose a different language as they first, in this circumstance: Spanish over Otomi.
- Pressure on the community created by the external influences, which produces an impoverishment of the structure of the language. As explained, in Mexico the indigenous languages may be cause of shame o considered a burden for the education or governmental systems.
- Media like radio and television do not consider these languages for its content or the geographical isolation of the speakers.

This research takes these aspects into build the cellular automata that will model the conduct of the Otomi in Queretaro region.

b. Queretaro region

In view of the regional study in Querétaro, we start from the location of the Otomi communities (ñāñho or ñāñha) as one of the groups with greater numeric and / or historical weight of the entity. [3] Otomi is spoken as showed in the next map according with INALI official information in 2015

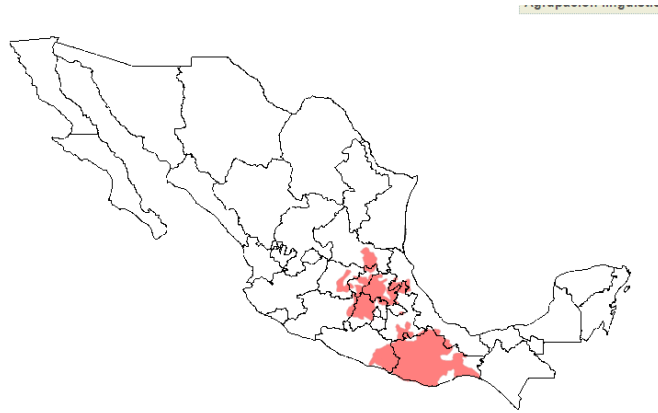


Fig. 1 Distribution of Otomi in Mexico

Specifically in Queretaro, located in the center region of Mexico, the distribution is as shown:



Fig. 2 Otomi speakers distribution in Queretaro

c. cellular automata (CA)

A cellular automaton is a decentralized computing model providing an excellent platform for performing complex computations with the help of only local information (Stavros, Christos, Gerasimos, & Evi, 2012). In the Queretaro region, we are talking about a community that uses two languages, Spanish as dominant language (DL) and Otomi as subordinate language (SL). In the model proposed at this point we talk about the strength of SL, the social pressure to encourage DL and the number of neighbors engaged with DL:

- State 0 or monolingual state: The person only speaks the DL.
- State 1 or bilingual state with preference for the DL: The person usually speaks the SL, but also speaks the DL, depending on the communication setting but external factors ease the person to transmit the DL to his or her children.
- State 2: bilingual state with preference for the SL: The person usually speaks the SL, but also speaks the DL.

The next diagram explains graphically the behavior of the language modeled in this research:

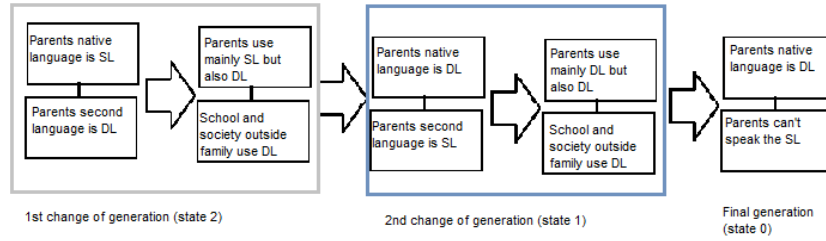


Figure 3 Transition model

Based on the work Base on the work of the JASSS, the CA follows the next procedure: The transition rule determines the future state in $t + 1$ of a given cell, which has a given state in t . The new state is determined by the sum of the neighborhood values, including the cell target. The sum can be a value between 0 and 18.

There are three thresholds:

- S_a : a sum value below the threshold produces a sharp transition, meaning that state 2 changes sharply to state 0. The established threshold for S_a is 3.
- S_b : a sum value below the threshold produces a transition from a higher-value state to a lower-value state, but a sum value above the threshold produces a transition from a lower-value state to a higher-value state. . The established threshold for S_a is 9.
- S_c : a sum value above the threshold produces a transition from a lower-value state to a higher-value state. The established threshold for S_a is 14.

The threshold values should be $S_a < S_b < S_c$. and indicate the individual's level of engagement with the SL. When there is a greater level of engagement, the individual needs a lower threshold value to move up to a higher-value state. So the individual increases his/her usage and eventually the transmission of the SL with only a minimal number of current neighbors using the SL. On the contrary, when there is a lower level of engagement, the individual needs a higher threshold value to move up to a higher-value state. Thus, the individual decreases his/her usage and eventually the transmission of the SL if there are not a large number of current neighbors using the SL. [2]

To state	0	1	2
0	$\sum neighbors \leq S_b$	$\sum neighbors > S_b$	----
1	$\sum neighbors < S_b$	$S_b \leq \sum neighbors \leq S_c$	$\sum neighbors > S_c$
2	$\sum neighbors \leq S_a$	$S_b < \sum neighbors < S_b$	$\sum neighbors \geq S_b$

Table 1 Transition Rules

d. simulation

The initial states were:

Language	State of the CA
Spanish	0
Spanish and Otomi	1
Otomi	2

Table 2 Cells state rules

The community of this model is a discrete two-dimensional torus- shaped world. The world contains 150x165, each cell contains 10 individuals, based on INALI statistics of 2015 where there are 24,736 Otomi speakers in Queretaro and 11.7% speaks only Otomi and 88.3% are bilingual (Spanish and Otomi). At each unit of time, a cell can only be classified in one of the three possible language states (0, 1 or 2), indicating the individual's strength in the use of the SL. Each cell has eight adjacent neighbors on the side and the vertex (a Moore neighborhood with a radius of 1) and the sum of neighbor values indicates the social pressure on the individual to use the DL or the SL. A low sum value means an individual has few opportunities to interact with his/her neighbors using the SL, but if the sum value increases, the individual's opportunities to interact using the SL also increase. [2]

Population over 3 years old that speaks Otomi in Queretaro in 2015				
Total	Do not speak spanish		Speak spanish	
24,736	Total	%	Total	%
	2895	11.7	21,841.8	88.3

Table 3 Information extracted from INALI Poblacion de 3 años y más hablante de alguna lengua indigena por agrupación lingüística según entidad federativa

3 Results

The simulation started with the numbers above; black's cells stand for people who only speaks Spanish (in a geographical context where people tends to speak Otomi). Blue cells mean bilingual people and white cells people who only speak Otomi.

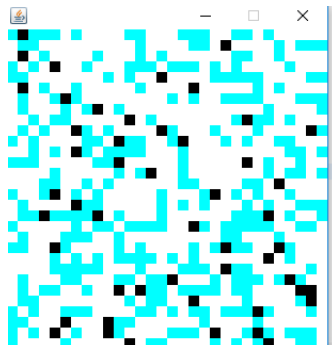


Figure 4 Initial population

After of 5 iterations, where two iteration represent a human generation, the automata stabilizes in only black cells, meaning that after a little more of two human generation people will only speak Spanish in the Queretaro regions. We can say that the grandparents will know Otomi as well as their children but perhaps their grandchildren will only know a few words and unable to transmit them to next generations.

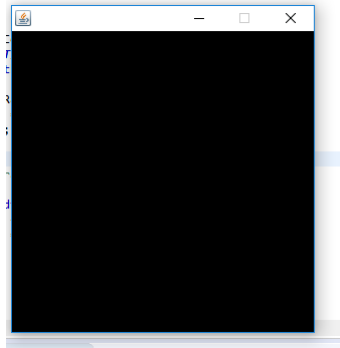


Fig. 5 Final population

Conclusions

If the speakers know a language but don't use it anymore, that language has become extinct regardless of their knowledge. Being able to predict the future of Otomi language can be useful to help determine the actions to inverse the language declination. The CA developed in this research helps to make visible graphically and numerally the shift of Otomi language in Queretaro region. However, there are other aspects, usually social aspects that cannot be measurable or are taken in consideration in this work, like the vulnerability of population due to geographic distribution or number of alphabetized people that can transmit language with a background of education and mortality rate.

We propose a visualization based on contextual changes based on paradigmatic changes of Otomi speakers and how they should adapt their knowledge of language to a modern context where there are not enough words to define many concepts such as changing aspects of technology, such as it is possible can see in the following figure.



Figure 6. Transitional model of Otomi speakers on the time.

References

1. López Sánchez Javier, Arellano Martínez Alejandra, Cárdenas Demay Almandina, Flores Castro Marisol, 2015, Mexico. *Lenguas indígenas nacionales en riesgo de desaparición*.
2. Herrando, F. B., S. F., D. A., M.-A. E., V. R.-S., & Marcos. (2009). Forecasting a Language Shift Based on Cellular Automata. *Journal of Artificial Societies and Social Simulation* , 5.
- 3.S. A., C. K., G. K., & E. P. (2012). Cellular Automata: Simulations Using Matlab . *The Sixth International Conference on Digital Society*, (pp. 63-68). Rion, Greece.
4. Sarmiento, B. U. (2003). Las regiones indígenas de Querétaro. *Identidad memoria histórica . Seminario 2003* (p. 17). CDI.