Prediction on the Number of Different Language Speakers

Huaxia Zhang, Jinhui Chen, Tianyuan Liu
Department of Nuclear Science, North China Electric Power University, Beijing 102206, China

Abstract: As a cultural symbol and the communication tool, language plays an important role in human society. The prediction of different language speakers has great significance on language protection and research. According to the modern theories of difference equations, we establish a discrete model to predict the number of speakers of a language. It consists of endogenous growth (affected mainly by government's language policy) and exogenous growth (affected by languages used in schools, assimilation of cultural groups, etc.). We use logistic model to predict the former, and we develop a language influence model to quantify the latter. The prediction results show that top-10 languages in the world will remain unchanged but the order will change slightly.

Keywords: Endogenous growth; Exogenous growth; Language impact factor

1. Introduction

As a tool for transmitting information, language implements the function of human communication [1]. Affected by the government promotion, school education, etc, the number of users in different languages shows different trends. The change in language use will accelerate the process of globalization, benefit world trade. But meanwhile it can also bring some languages to the brink of extinction. Therefore, prediction of the number of language speakers is of great practical significance.

2. Model Establishment

We constructed a difference equation model to describe the number of speakers in different languages over time.

$$x_{k+1} = x_k + B_1 + B_2 \tag{13}$$

Where xk is the total number of people who speak a certain language at a time; xk+1 is the number of people who speak the language after a period of time; B1 is the endogenous growth; B2 is the exogenous growth.

2.1 Government Promotion Model

The usage of a certain language in native speaking countries is mainly promoted by the local governments. Assuming there are n countries use the language as an official language, B1 is expressed as follows:

$$B_1 = \sum_{i=1}^{n} \Omega_i \Delta p_i \tag{14}$$

Where Ωi is the proportion of population in a country using this language; Δpi is the growth of population in the country in a period of time.

We use the logistic model to predict demographic changes.

$$\Delta p = rp(1 - \frac{p}{p_m}) \tag{3}$$

Where r is population growth rate; pm is the maximum environment capacity.

2.2. Language influence model

We build a language impact evaluation system to evaluate the increase of a language.

In our evaluation system, we use the number of native speakers(f1), the number of secondary speakers(f2) and the number of population using the language(f3) with points of four, six and seven respectively to definite language impact factor f.

$$f = 4f_1 + 6f_2 + 7f_3, \tag{4}$$

in which f1=l1x, f2=(1-l1)x, f3=x. l1 is the proportion of native speakers of a certain language; x is the total number of people who speak this language.

We assume that exogenous growth B2 has positive relationship with language impact factor f .

$$B_2 = c_1 f + c_2 fx \tag{5}$$

c1f is the number of people who acquire language through study. It depends on the popularity of a language and has nothing to do with population base.

c2fx is the number of people who acquire languages through assimilation. It is related to population base.

3. Parameters Determination

3.1. Determination of endogenous growth

We only take languages with top 20 speakers and top 40 countries with the largest population into account. In the lists of countries and languages we consider, 15 languages are used officially in 31 countries.

Table 1. r, Ω, p of Different Countries[2],[3]

Language	Country	r (%)	Ω	P(million)
Mandarin Chinese	China	0.500	0.700	1405.37
English	India	1.578	0.180 0.806	1304.20 322.76
	America	0.900	0.533 0.637	182.31 102.41
	Nigeria	2.533	0.997 0.480	65.04 47.59
	the Philippines	1.873	0.660 0.569	42.54 35.87
	England	0.362	0.890	192.40
	Kenya	2.758		
	Uganda	3.300		
	Canada Pakistan	0.830		
		1.805		
Hindustani	India	1.578	0.820 0.076	1304.20 192.40
	Pakistan	1.805		

*Only a few of them are listed.

3.2. Determination of exogenous growth

We take Chinese and German as examples to demonstrate the prediction process of l1.We only find data of seven years, and we use the interpolation method to fill the missing data. We choose GM (1,1) Model to predict this proportion with known data and the complemented data.

Prediction of native speakers proportion in Chinese and Germany is shown in Fig.1.

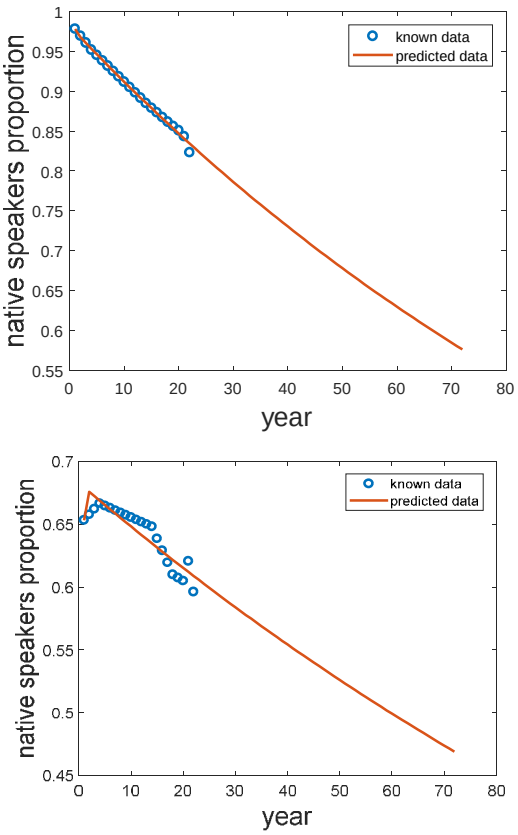


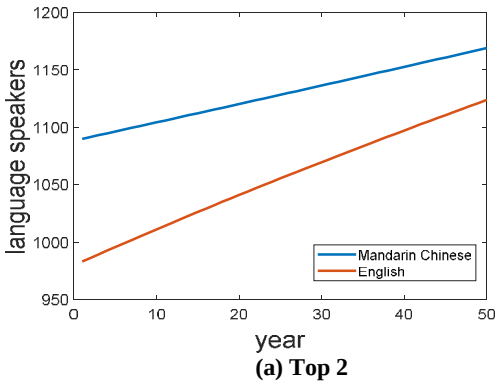
Figure 1. Prediction of Native Speakers Proportion

The existence of the coefficient c is to make the two components on the right side of Equation 5 take an equal order of magnitude.

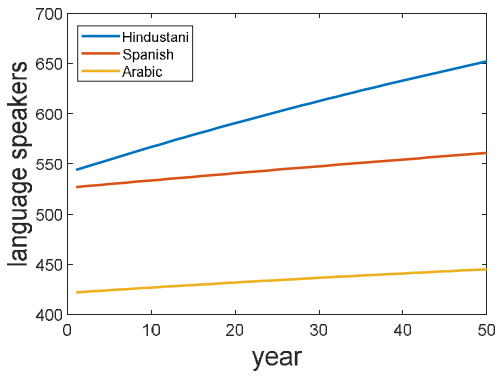
After calculating, c1 equals 4×10-4, c2 equals 1×10-8 .We assume the number of speakers in different languages have the same linear relationship with their language influence, so c1 and c2 also apply to other languages.

4. Predicting Results

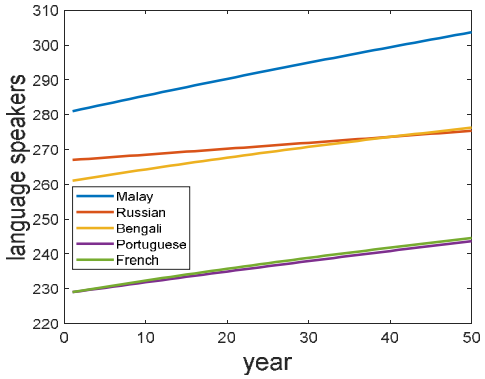
After determining the parameters of our model in Equation 1, we get the forecast curve in Matlab environment. The changing trends are shown as follows. We take the year 2017 as a starting point for forecasting. We can keep predicting the value of xk+1 with Equation 1.



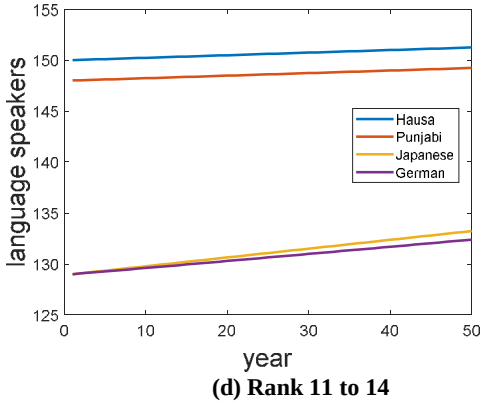
(a) Top 2



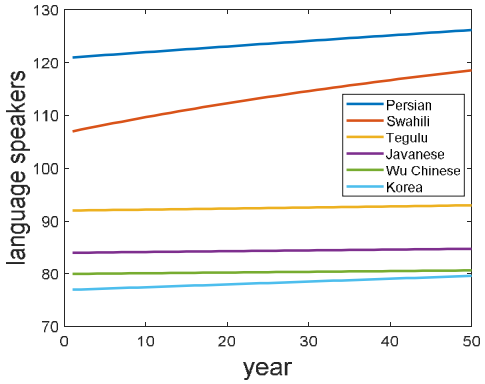
(b) Rank 3 to 5



(c) Rank 6 to 10



(d) Rank 11 to 14



(e)Rank 15 to 20

Figure 2. Prediction of Total Language Speakers

The ranking changes can be seen intuitively in Fig. 2. According to our forecast, the top-10 lists will remain unchanged in the next 50 years, but the rankings will change slightly. In the top-ten lists, Bengali will surpass Russian and rank the seventh place in about 40 years time. Portuguese and French now have the same number of speakers, but French will outperform Portuguese in the next 50 years.

5 . Conclusions

In predicting the total number of people using different languages, the development trends of dominant languages in the world are moderate and steady. There is no language in top-10 lists being replaced by another language. This is related to the stability of the world structure. This result is reasonable under the premise of a basically stable future world situation.

References

[1] Hamdy A. Taha. Operation Research: An Introduction[M]. Eighth Edition.No.14 temple street, chongwen district, Beijing: the Posts and Telecommunications Press, 2008.8.1 :393-409.

[2] Da-Yong Chang. Applications of the extent analysis method on fuzzy AHP [D]. Beijing 101149, China: Beijing Materials College, 1996.

[3] Xu Ze-shui, Liao Hu-chang. Intuitionistic fuzzy analytic hierarchy process[J].IEEE Transaction on Fuzzy Systems, 2014, 22(4):749-761.

[4] Xin Bingqing, Li Xin, Chen Xiong, Wang Chengbiao4, Liu Yun2. Study on Technical Aid Strategy of Developed Countries in the Area of Climate Change[D]. Ministry of Science and Technology, Beijing 100045, China: China Science and Technology Exchange Center, 2014.

[5] WANG Ya-ru, ZHAO Xue-yan, ZHANG Qin, LUO Li. Farmers' climate change adaptation strategies in an ecologically vulnerable alpine region: Acase of Gannan Plateau[D]. Northwest Normal University, Lanzhou 730070, China: College of Geography and Environment Science, 2016.

[6] Xin Bingqin, Li Xin, Chen Xiong, Wang Chengbiao, Liu Yun2.Study on Technical Aid Strategy of Developed Countries in the Area of Climate Change[D].Ministry of Science and Technology, Beijing 100045, China: China Science and Technology Exchange Center, 2014.

[7] ZHANG Jingjing, GUO Zhifu.Risk assessment on agricultural drought in Zhoukou City based on fuzzy theory[D].Shangqiu Normal University, Shangqiu, Henan 476000, China:Department of Environment & Planning, 2016.

[8] LI Yan-fang, ZHANG Zhong-li, Li Cheng.Toward Specialized Legislation to Address Climate Change in China[D].Beijing 100872, China: Law School, Renmin University of China, 2016.

[9] LIU Bo, TANG Wei-mu. Research and development of catastrophe risk assessment models[D]. Harbin 150080, China:Institute of Engineering Mechanics China Earthquake Administration, 2011.

[10] WU Guochao LI Quan-hai BAO Wen-yue. Realtime gray prediction application in dam deformation forecast[D].Shanghai 2000, China:College of Mapping and Geography Information, Tongji University, 2013.

CUI Chao, ZHANG Ya-dong, GUO Jin. Research on Train Wheel Diameter Calibration with Improved Gray Prediction [D].Chengdu 611756, China: School of Information Science and Technology, Southwest Jiaotong University, 2016.