Language Decline and Competition

Demetrios Lambrou

Table of Contents

Figures

Equations

Introduction

Linguistic competition is a phenomenon that occurs when two or more languages are spoken in the same geographical region and compete for speakers and social status. In this competition, individuals may choose to learn one language over another based on factors such as the social prestige or economic opportunities associated with a particular language. As a result, one language may become dominant over the others, leading to a decrease in the number of speakers of minority languages. This phenomenon is known as language shift or language loss. Linguistic competition can also have effects on language structure and grammar. When speakers of different languages come into contact, they may borrow words, phrases, or grammatical features from each other's languages, leading to the creation of new dialects or creoles. Additionally, the pressure to maintain distinct linguistic identities can also lead to linguistic purism, where speakers consciously avoid using loanwords or grammatical features from other languages to maintain the purity of their own language. Overall, linguistic competition is a complex phenomenon that involves social, economic, and linguistic factors, and can have far-reaching effects on language use, structure, and identity.

Linguistic decline refers to the situation when a language gradually loses its speakers and may eventually become extinct. A good example of linguistic competition and decline is the case of Quebec French. Historically, Quebecois French faced challenges due to the dominant position of English in North America. In the past, the use of French was suppressed, and Francophones were often discriminated against in areas such as education and employment. As a result, many Francophones experienced linguistic assimilation and shifted to English, resulting in a decline in the use and status of Quebecois French. In more recent times, the decline of Quebecois French has been attributed to several factors, including the influence of media, immigration, and the increasing use of English in daily life. The widespread use of English-language media, including television, movies, and music, has resulted in an increasing number of Quebecois using English words and phrases in their speech. Immigration has also had an impact, as many newcomers to Quebec are not Francophone and do not speak French as their first language. Additionally, the increasing use of English in daily life, particularly in business and technology, has led to a decline in the use of French in certain domains. This has had a significant impact on the status of Quebecois French, as its speakers have become increasingly concerned about its future. In response, efforts have been made to promote and protect the language, including language laws and education policies aimed at preserving and revitalizing the language. Such examples include Bill 101 and the newly passed Bill 96. The government has also sought to limit non-francophone immigration to the province to decrease the pool of allophone speakers.

Linguistic decline can be modeled using differential equations. This paper will look at the competition between two undefined languages using three models: the Abrams-Strogatz Model, the Castello Model, and the Mira Model.

Abrams-Strogatz Model

In their 2003 article "Modelling the dynamics of language death", Strogatz and Abrams used data on the number of speakers from endangered languages in 42 regions of Peru, Scotland, Wales, Ireland, and Alsace-Lorraine to come up with the following set of differential equations:

Equation 1 - Abrams-Strogatz Model

$$\frac{dx}{dt} = yP_{yx}(x,s) - xP_{xy}(x,s)$$

$$\frac{dy}{dt} = xP_{xy}(x,s) - yP_{yx}(x,s)$$

where x and y represent the fraction of the total population that speak the two languages. Note that both x and y must sum to 1 since they are complementary. P_{xy} is the probability that a speaker of language x would switch to language y. Likewise, P_{yx} is the probability that a speaker of language y would switch to language x. They are represented by the following equations:

Equation 2 - Abrams-Strogatz Probabilities

$$P_{xy}(x,s) = (1-s)(1-x)^a$$

$$P_{yx}(x,s) = sx^a$$

where s represents the prestige/attractiveness of a language and a is an exponent that was found to be roughly constant (1.31 ± 0.25) across the groups examined by Abrams and Strogatz. It is important to note that Sutantawibul *et al* in their 2018 paper found values for a as high as 1.5 and as low as 0.84 while fitting the model to data not used by Abrams and Strogatz. For the sake of simplicity, the section will keep the value for a found by Abrams and Strogatz. The model's main flaw is that is does not account for bilingualism. As such, the model predicts that there will always be a language that dominates and a language that goes extinct when two are in competition with each other. Likewise, the s parameter is incredibly subjective, with even the slightest change yielding different results. Below are figures with different parameters for *s*. The fraction of speakers of the two languages was kept constant and the default value of *a* was used. The time *t* is in years. 50 years of modelling is shown on the graphs.

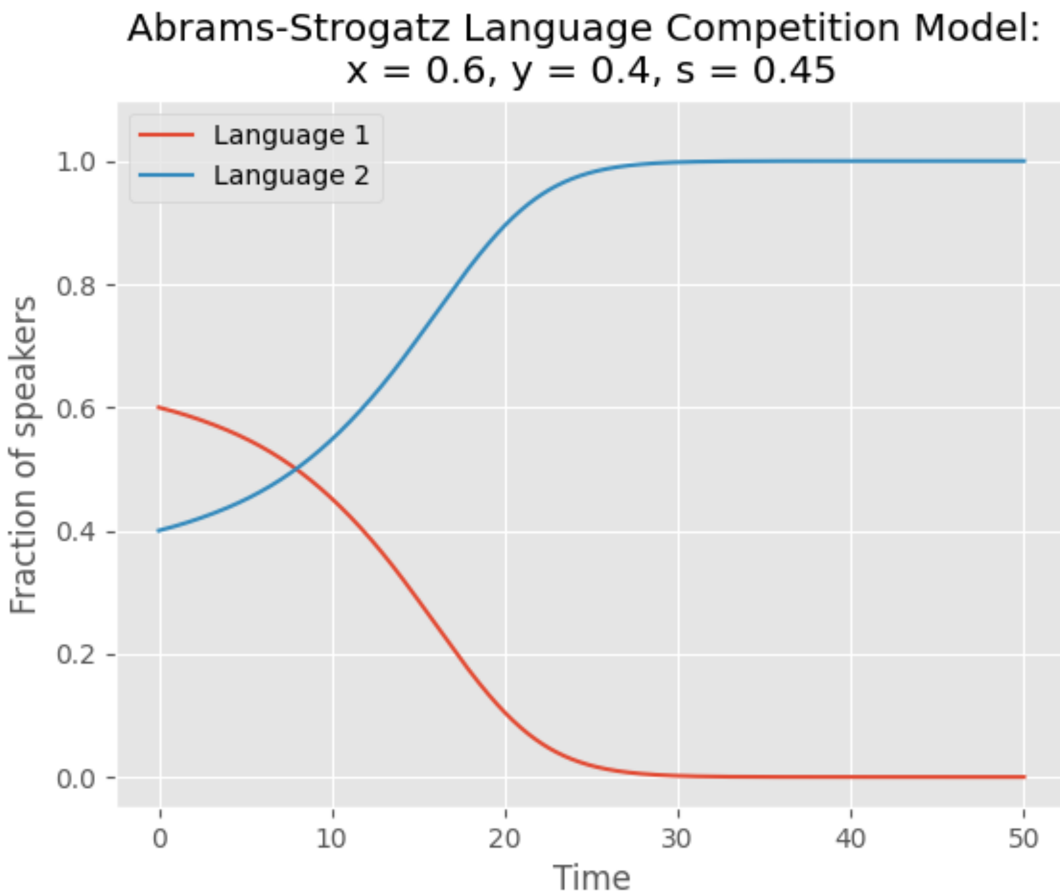

Figure 1 - Abrams-Strogatz 1

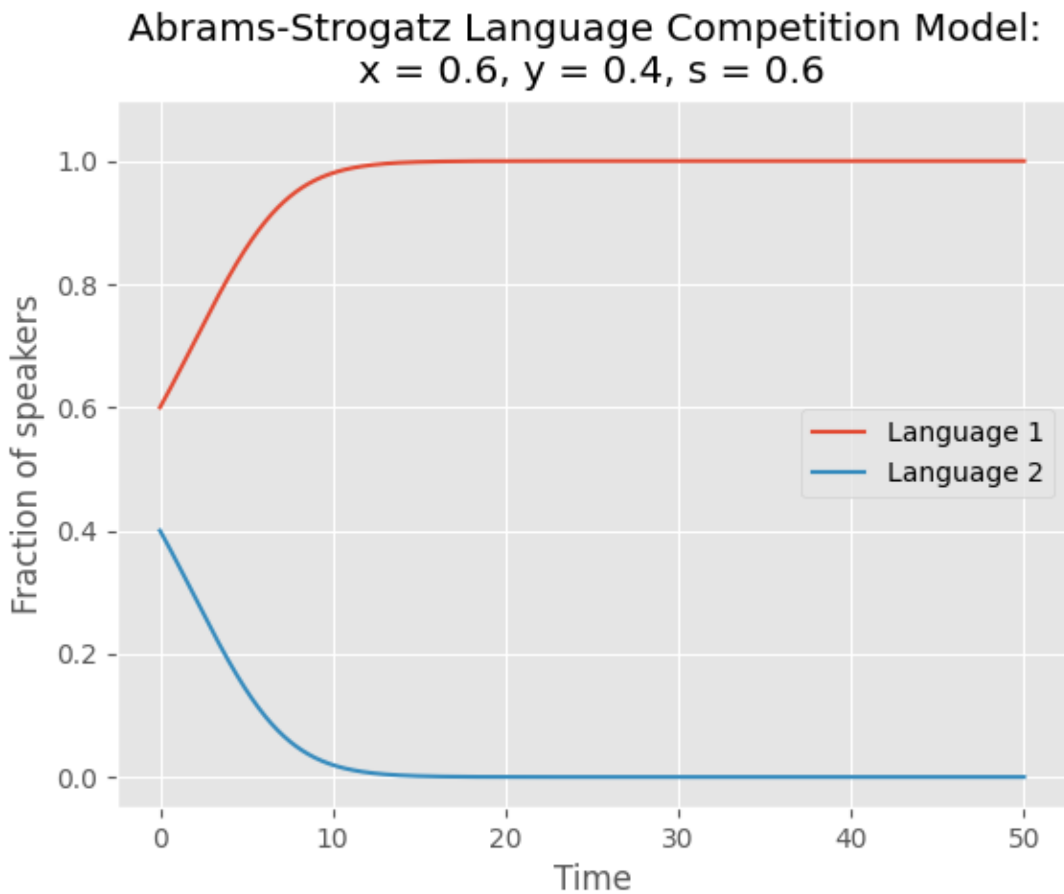

Figure 2 - Abrams-Strogatz 2

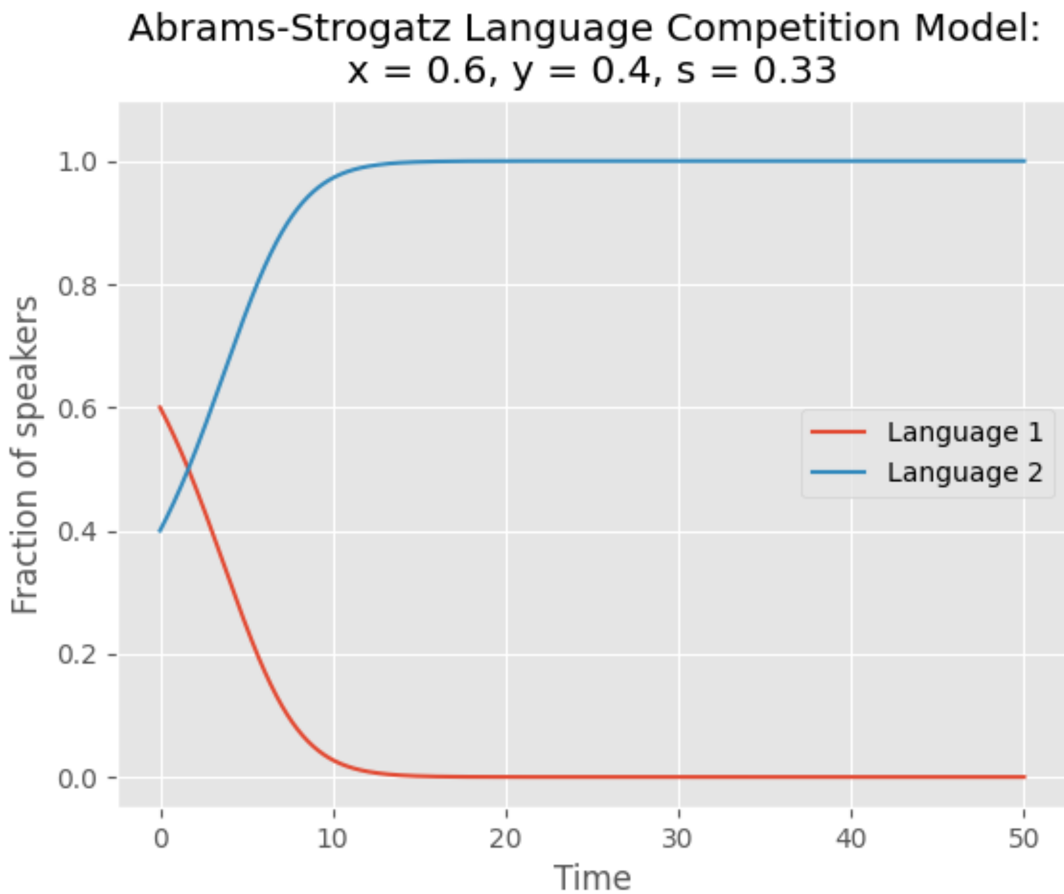

Figure 3 - Abrams-Strogatz 3

Castello Model

Castello, *et al's* model is an extension of the Abrams-Strogatz model. Whereas the Abrams-Strogatz model did not account for bilingualism, Castello, *et al's* does. The Castello model is represented by the following differential equations:

Equation 3 - Castello Model

$$\frac{dx}{dt} = yP_{yx} + bP_{bx} - x(P_{xy} + P_{xb})$$

$$\frac{dy}{dt} = xP_{xy} + bP_{by} - y(P_{yx} + P_{yb})$$

$$\frac{db}{dt} = xP_{xb} + yP_{yb} - b(P_{bx} + P_{by})$$

where the new variable *b* represents the fraction of bilingual speakers in the population. The probabilities in the equation represent the same as those in the Abrams-Strogatz model yet with the variable *b* included. For example, P_{xb} represents the probability that a speaker of language *x* would become bilingual. The *s* and *a* constants represent the same thing as in the previous model. Likewise, P_{xy} and P_{yx} are governed by the same equation as the previous model and represent the probabilities of switching from language *x* to *y* and vice-versa. The newly introduced probabilities are governed by the following equations:

Equation 4 - Castello Probabilities

$$P_{xb} = (1 - s)xy^a$$

$$P_{yb} = sxy^a$$

$$P_{bx} = s(1 - x - y)(1 - y)^a$$

$$P_{by} = (1 - s)(1 - x - y)(1 - x)^a$$

Below are figures with different parameters for *s*. The fraction of bilinguals and speakers of the two languages was kept constant and the value of *a* found by Abrams and Strogatz was used. The time *t* is in years. 50 years of modelling is shown on the graphs.

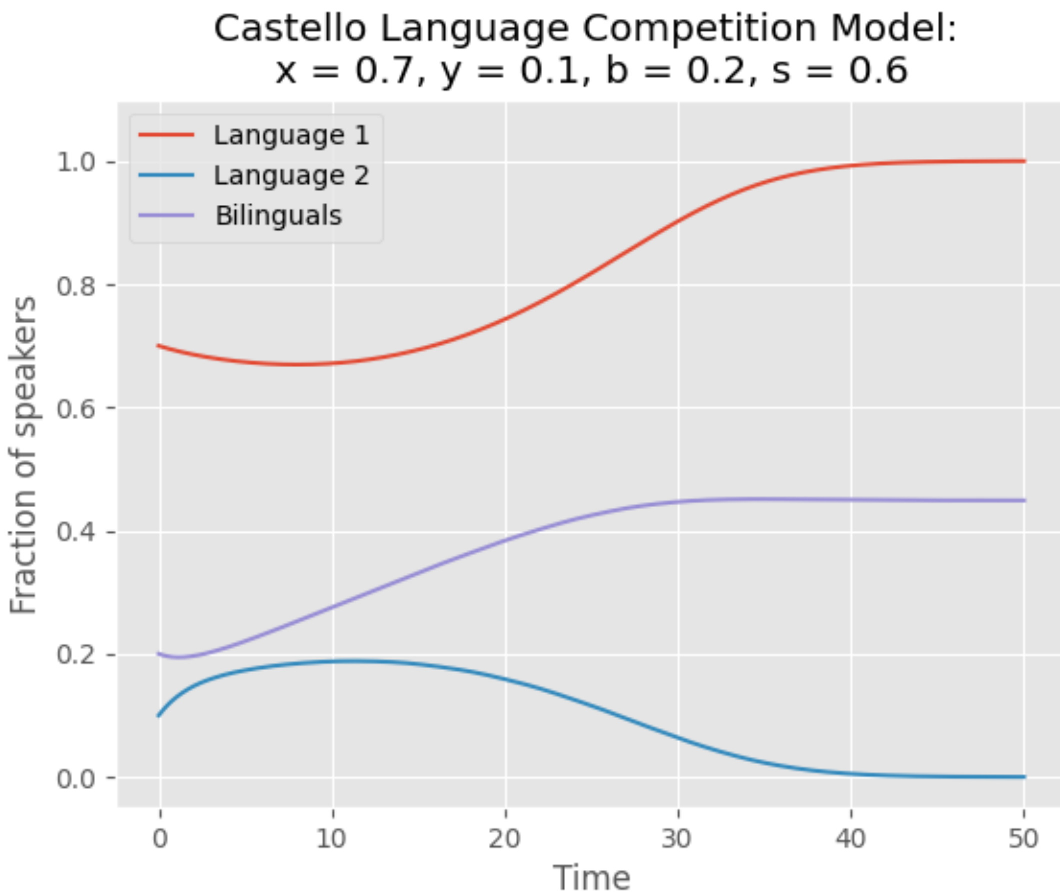

Figure 4 - Castello 1

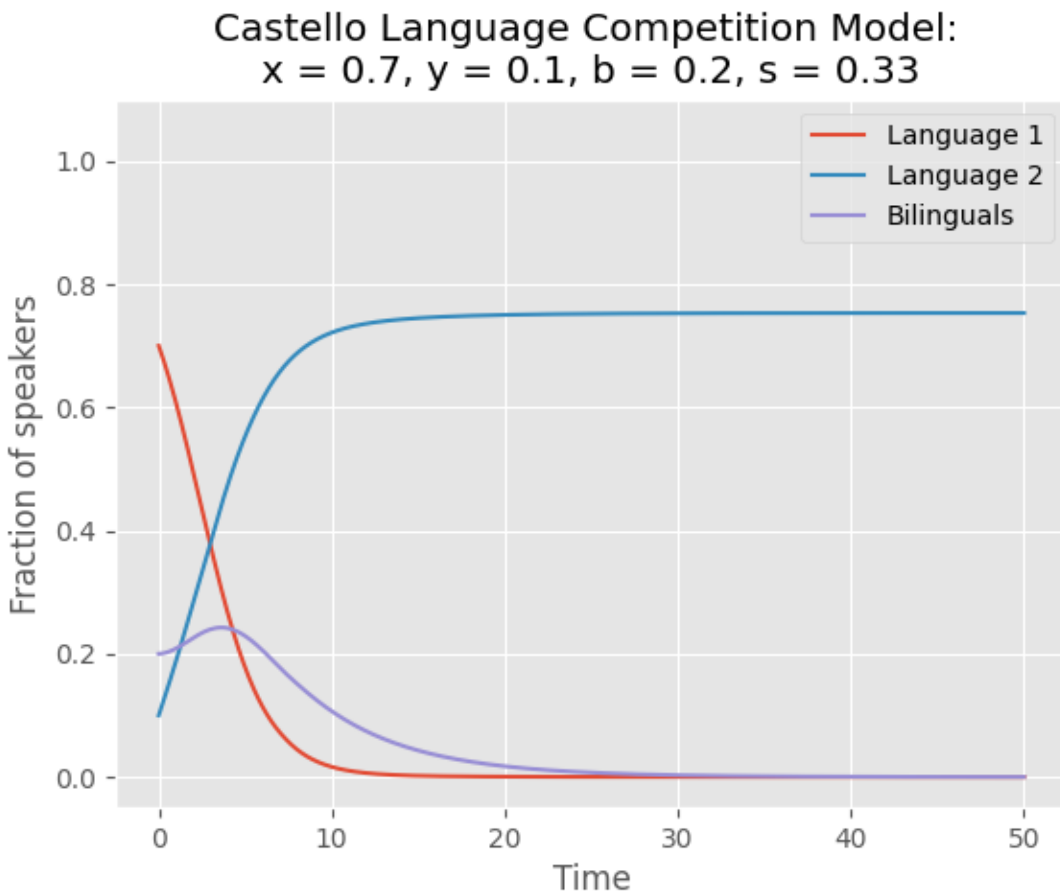

Figure 5 - Castello 2

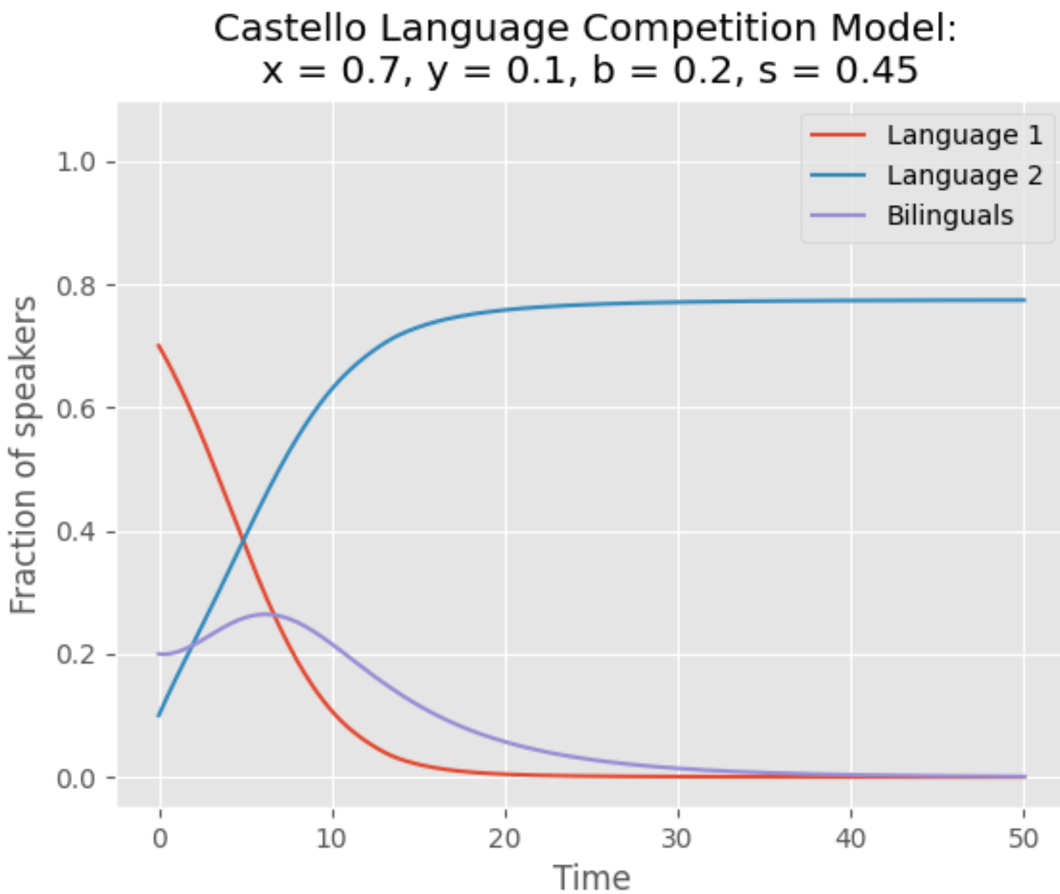

Figure 6 - Castello 3

The Castello model still suffers from the same fundamental issue as the Abrams-Strogatz Model. While it does account for bilingual speakers, the *s* parameter is still incredibly subjective with even the slightest difference yielding different results. Also, both Castello and Strogatz fail to account for similar languages. Essentially, if languages are more related to each other, the likelihood that a speaker of one language changes to the alternative will be greater than if the languages are very different.

Mira Model

Mira, *et al's* model is an extension of the Castello model. While it keeps the same set of differential equations, it adds a new constant, k, which indicates the similarity between two languages. It also gives each languages its own prestige constant, s_x and s_y. The k value is quite difficult to determine and, like the other parameters mentioned, it needs to be determined experimentally. Essentially, a k value close to 1 indicates a greater similarity between the two languages whereas a k value smaller than 1 indicates less of a similarity between the two languages. Sutantawibul *et al* found values of k as low as 0.18 and as high as 0.92 for various scenarios. More information about the k parameter is given in the Appendix. Sutantawibul *et al* also used different values for the a exponent. To acount for the new s, s and k parameters, the equations governing the transition probabilities need to be slightly modified. It is important to note that certain probabilities share the same value. Mira and Peredes explain that $P_{bx} = P_{yx}$ since both involve the loss of language y. Likewise, $P_{by} = P_{xy}$ since both involve the loss of language x. Below are the modified probabilities.

Equation 5 - Mira Probabilities

$$P_{xb} = ks_y(1-x)^a$$

$$P_{by} = P_{xy} = (1-k)s_y(1-x)^a$$

$$P_{yb} = ks_x(1-y)^a$$

$$P_{bx} = P_{yx} = (1-k)s_x(1-y)^a$$

The following graphs will represent three of the scenarios discussed by Sutantawibul *et al*. The first one is the case of Catalan and Spanish. The second one is the case of French in English in Montreal. The final one Welsh and English. All three of these cases have different values for each of the parameters.

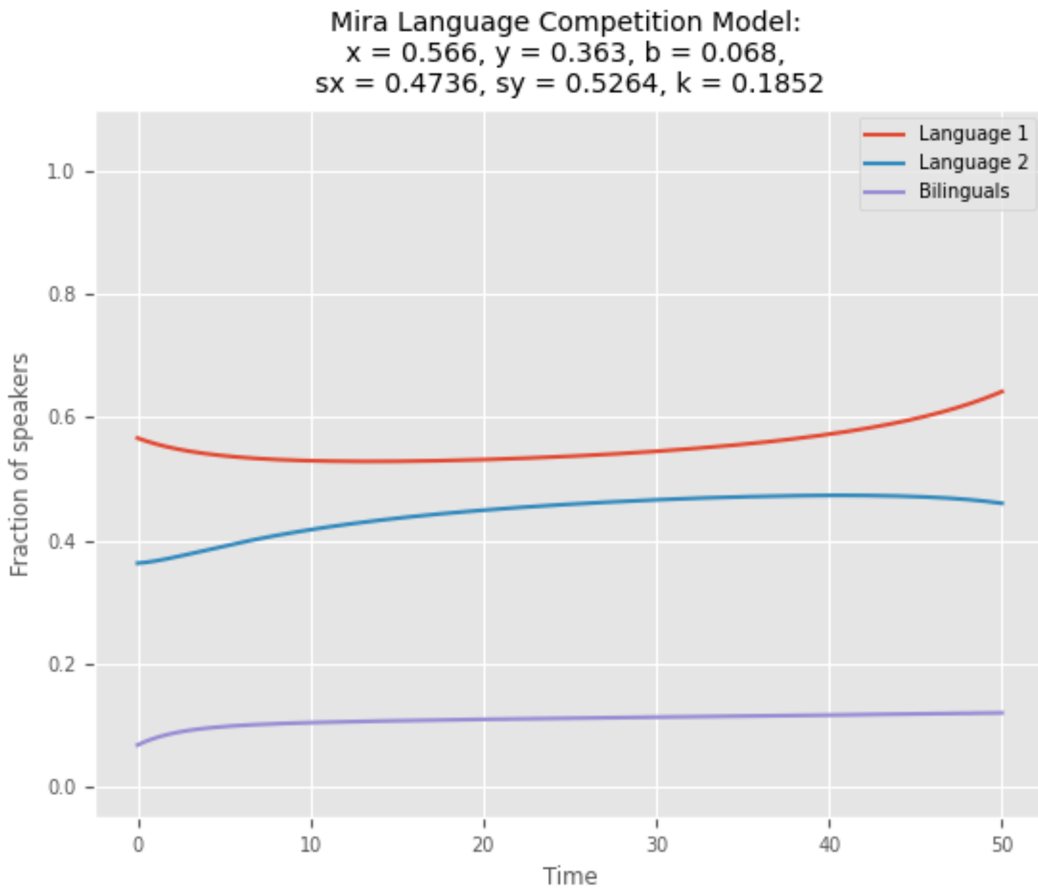

Figure 7 - Catalan/Spanish 2013

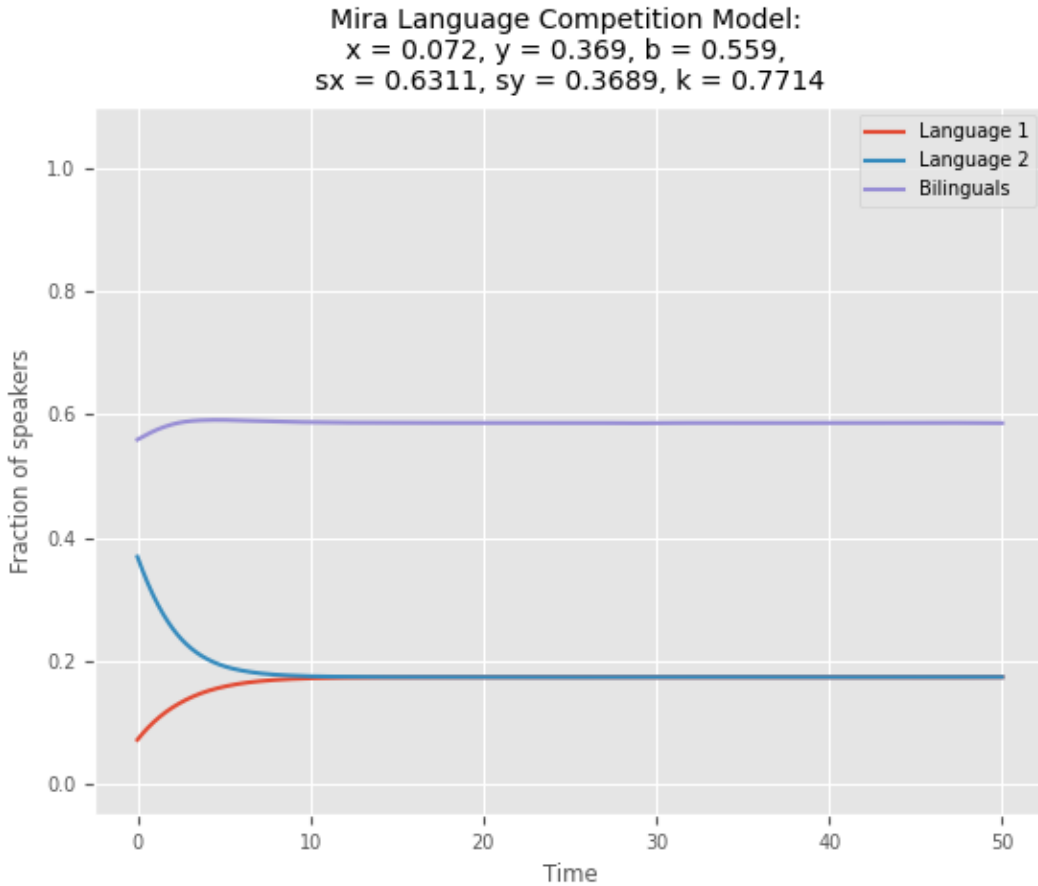

Figure 8 - French/English Montreal 2017

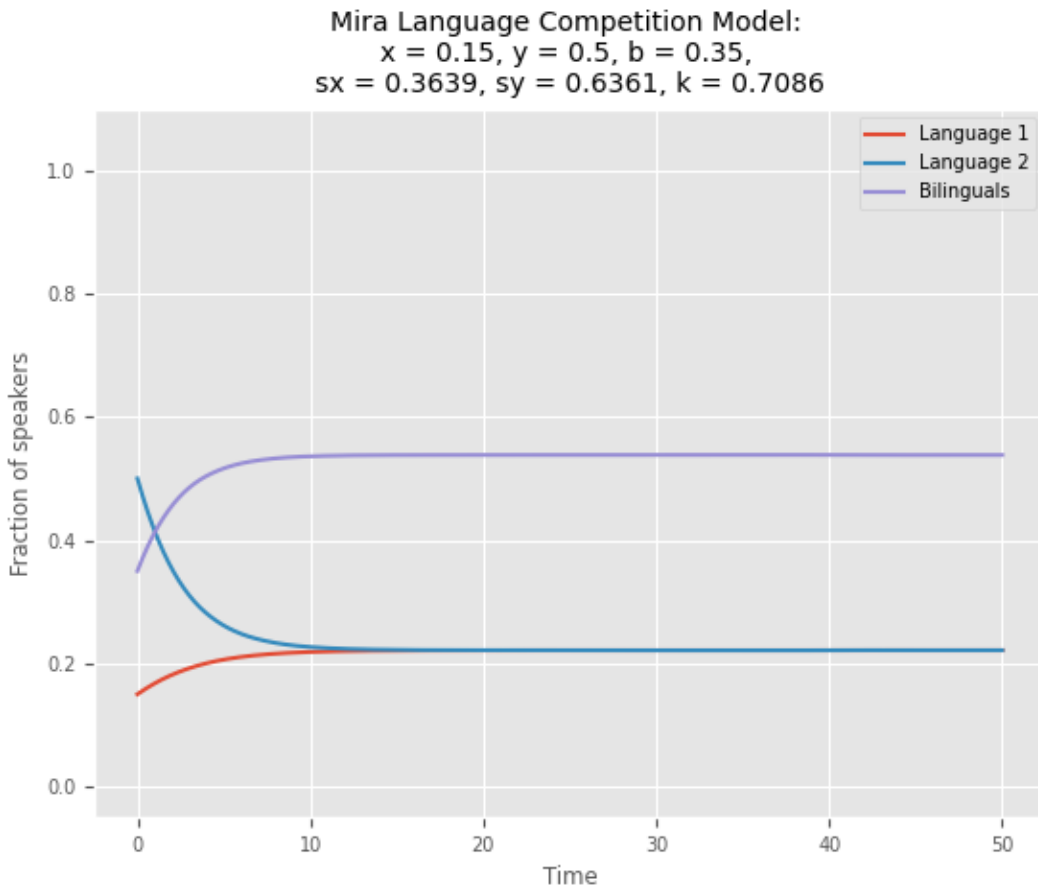

Figure 9 - Welsh/English 1901

While the Mira model is the superior alternative out of the three models, there remain some flaws. Again, all the parameters need to be determined experimentally and can sometimes produce erroneous results. For example, applying the Mira model on the Welsh/English values from 1901 shows that Welsh will not die out. This is obviously false as monolingual Welsh speakers no longer exist. Overall, the Mira model can be used to model real-world examples of linguistic decline, albeit with some caution. Compared to Abrams-Strogatz and Castello, however, it is the best choice.

Conclusion

In summary, this paper explored linguistic competition and decline using three different models. The Abrams-Strogatz model was used to explore linguistic competition amongst monolingual speakers of two languages. The model predicts that one language will always die out unless it is already dead or is spoken by 100% of the population. The parameter used to quantify the popularity was the s parameter. The higher the value for s, the more prestigious the language is. The exponent a was found to be 1.31 ± 0.25 for all the cultures investigated by Abrams and Strogatz. Both s and a are used to determine the probability that a speaker of the first language switches to the second language and vice-versa. As mentioned previously, the central flaw of the model is its lack of bilingualism and its drastic prediction of language death.

The Castello model extended the previous model by accounting for bilingual speakers. In essence, it adds new probabilities that represent the chance that a monolingual speaker of either language becomes bilingual and vice-versa. The model still retains the s and a parameters from the Abrams-Strogatz model. However, it is not as drastic as the Abrams-Strogatz. For instance, whereas the Abrams-Strogatz model predicts that two languages cannot coexist without one going extinct, Castello fixes this issues by including bilingualism. While scenarios exist where a language goes extinct, it is possible to have cohabitation among monolinguals and bilinguals.

The final model discussed was the Mira model. The Mira model uses the same equations as the Castello model, yet adds a new parameter, k. This new parameter, meant to be determined experimentally, shows the degree of similarity between two languages. More similarity, of course, will make it easier for speakers of one language to move to the competitor and vice-versa. While the Mira model can be used for real-world examples, as shown in this paper, flaws remain. The value of the k parameter determined experimentally can be confusing to interpret.

For example, Sutantawibul's tean found a k value of 0.1852 for Catalan and Spanish. This would seem to suggest that the two languages are barely similar. This is not the case, however, as Catalan and Spanish are quite similar to each other.

Further models are available in the domain of lingustic death. Models from Nowak, Stauffer and Schulze, and Patriarca and Leppanen, either extend the models in this paper or derive completely new ones. For example, Nowak uses the replicator equation to show how j different languages evolve over time. Stauffer and Schulze, on the other hand, use a probabilistic model. All in all, the field of language death modeling remains as complex as it does important.

Acknowledgments

I would like to thank Prof. Ivan T. Ivanov of the Vanier College Mathematics Department for his help throughout the project.

Bibliography

[1] D. Stauffer and C. Schulze, "Microscopic and macroscopic simulation of competition between languages," *Physics of Life Reviews,* vol. 2, pp. 89-116, 2005.

[2] D. M. Abrams and S. S. Strogatz, "Modelling the dynamics of language death," *Nature,* vol. 424, 21 August 2003.

[3] C. Sutantawibul, P. Xiao, S. Richie and D. Fuentes-Rivero, "Revisit Language Modelling Competition and Extinction: A Data-Driven Validation," *Jounal of Applied Mathematics and Physics,* vol. 6, pp. 1558-1570, 2018.

[4] J. Mira and A. Paredes, "Interlinguistic similarity and language death dynamics," *Europhysics Letters,* vol. 69, no. 6, pp. 1031-1034, 2005.

Appendix

When varied, the k parameter of the Mira model produces rather interesting results regarding language competition and decline. Note that many of k yield incorrect results since they were not determined experimentally and were instead generated via the model. Also, the incorrect results are due to the initial conditions being entirely identical for all of the graphs.

Below are 10 different graphs with varying parameters of k from 0 to 0.9 in increments of 0.1. As k increases, the languages become more similar according to Mira. This translates to the bilinguals becoming the dominant group as switching between the 2 languages becomes easier. A k value of 0.5 produces a very interesting result. The model essentially predicts an equal share of monolinguals and bilinguals.

In essence, the Mira model is trying to tell us that 2 similarly related languages will result in bilingualism becoming mainstream. Perhaps, another way to view it would be the emergence of a new language or dialect that combines many of the 2 older languages' previous rules since they are similar.

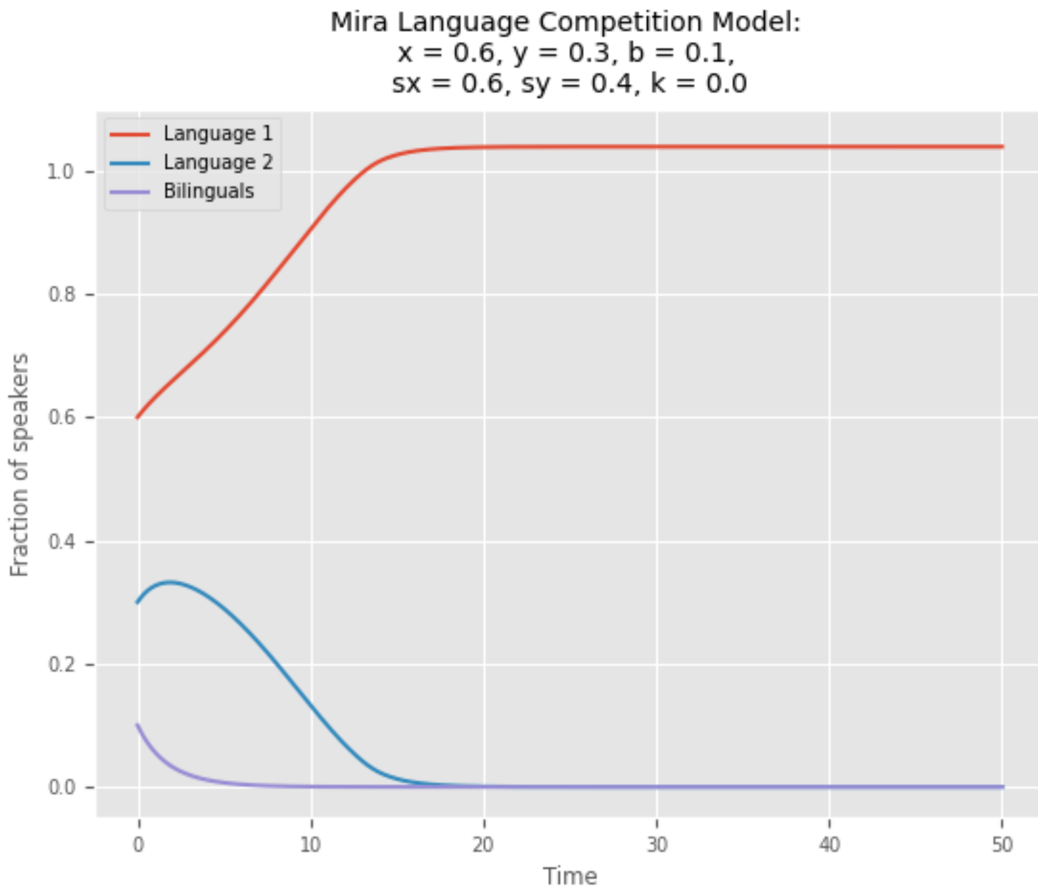

Figure 10 - Mira model with k = 0

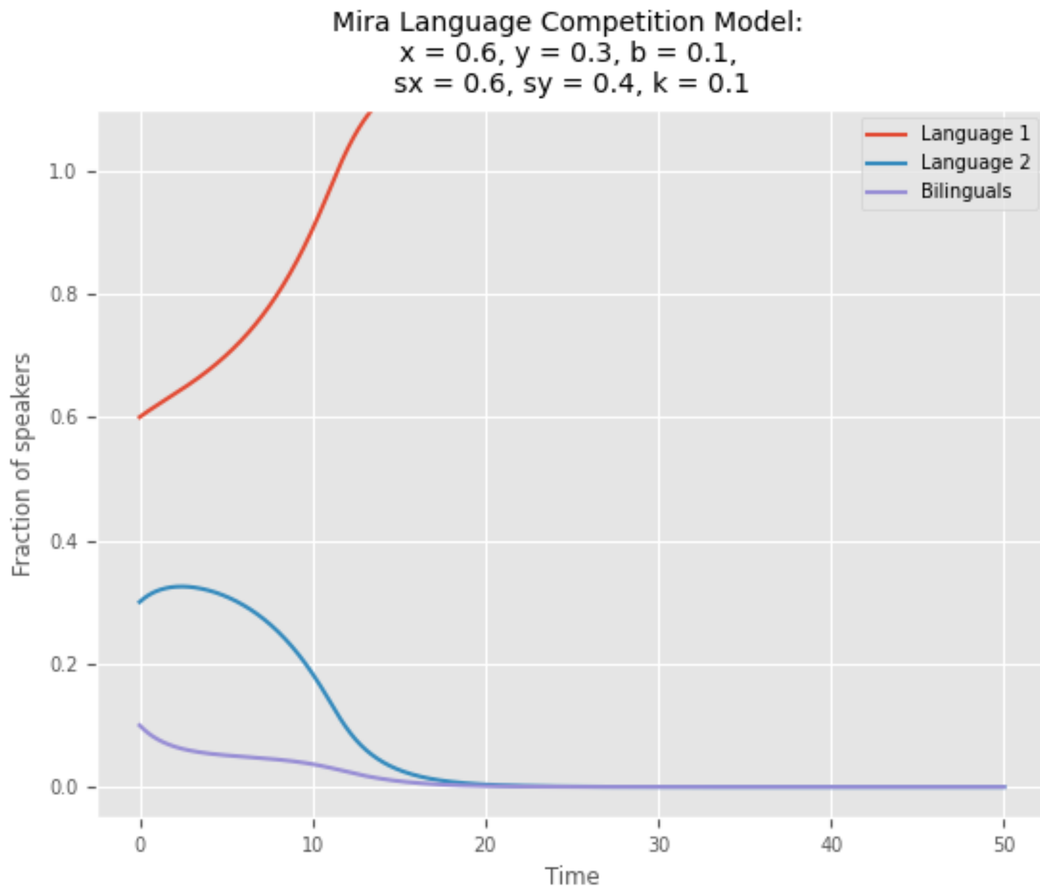

Figure 11 - Mira model with k = 0.1

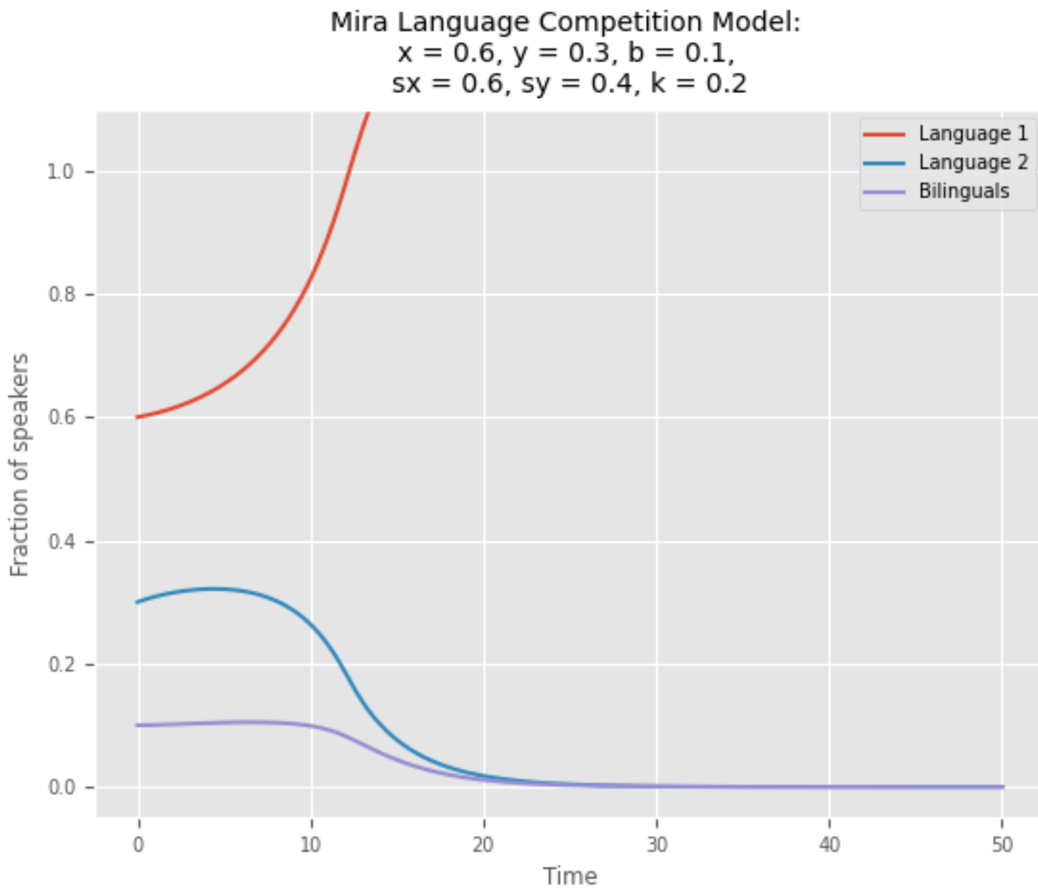

Figure 12 - Mira model with k = 0.2

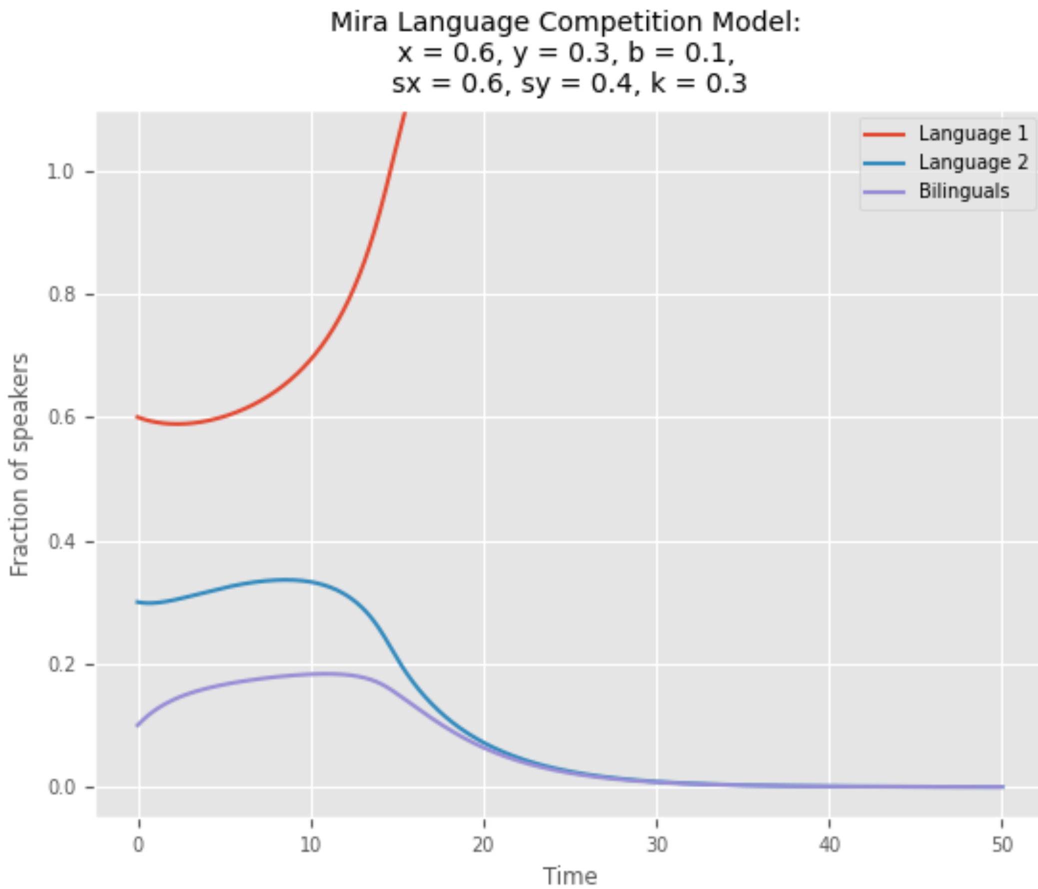

Figure 13 - Mira model with k = 0.3

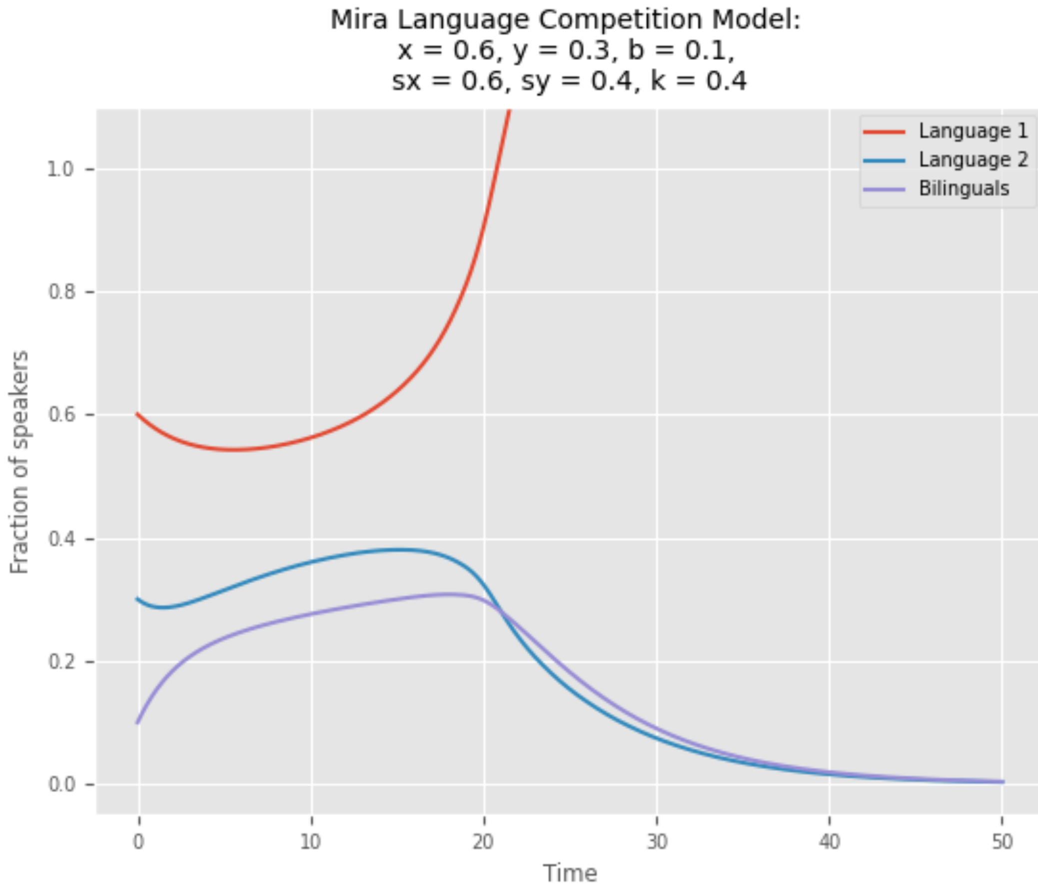

Figure 14 - Mira model with k = 0.4

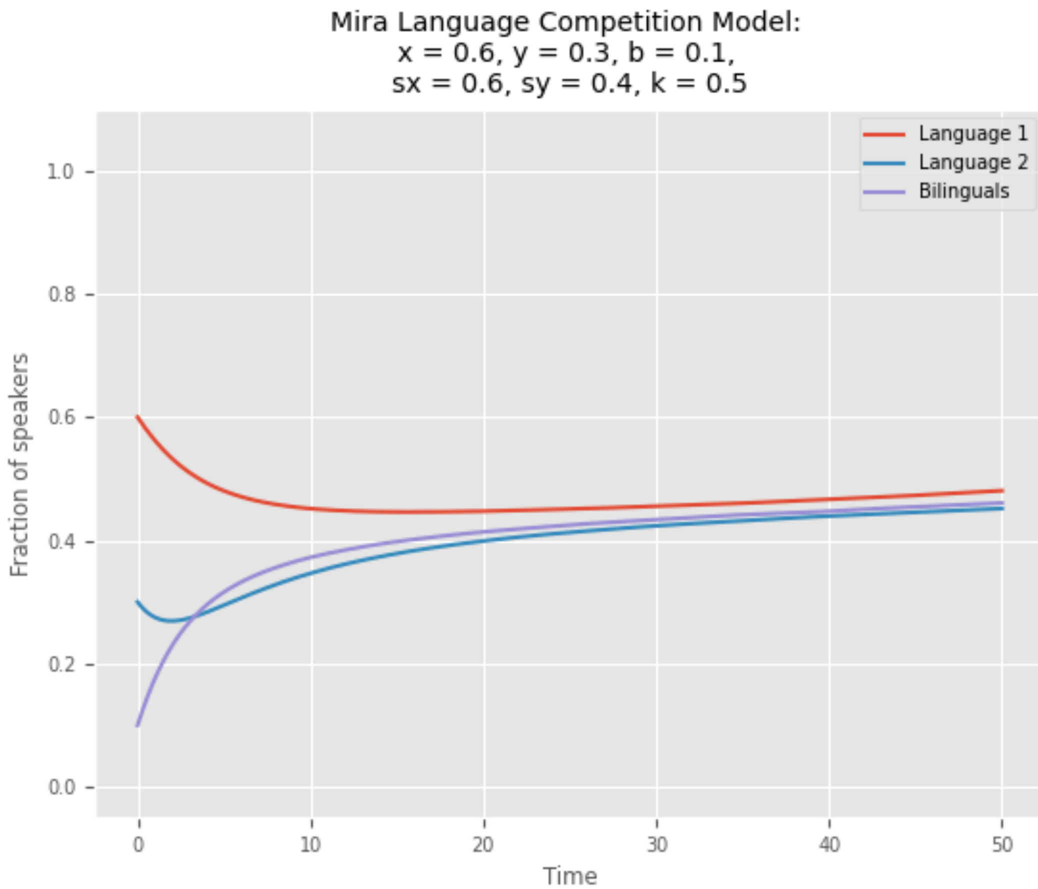

Figure 15 - Mira model with k = 0.5

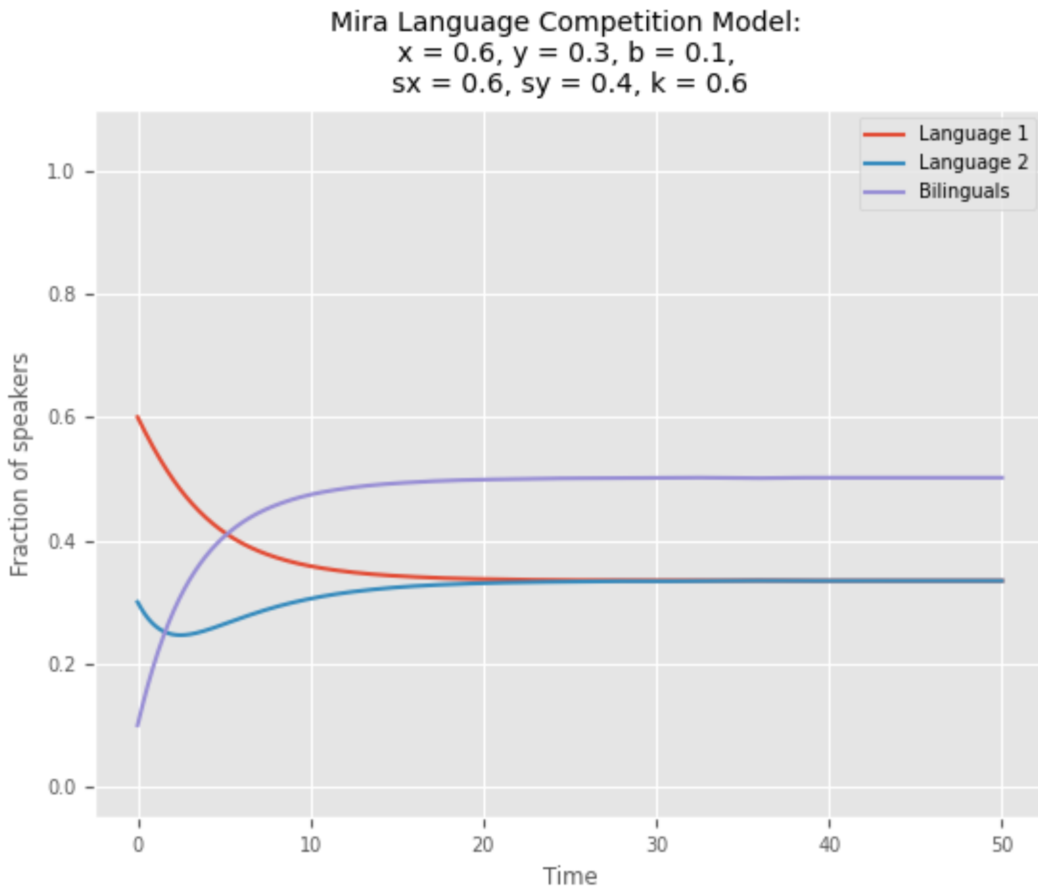

Figure 16 - Mira model with k = 0.6

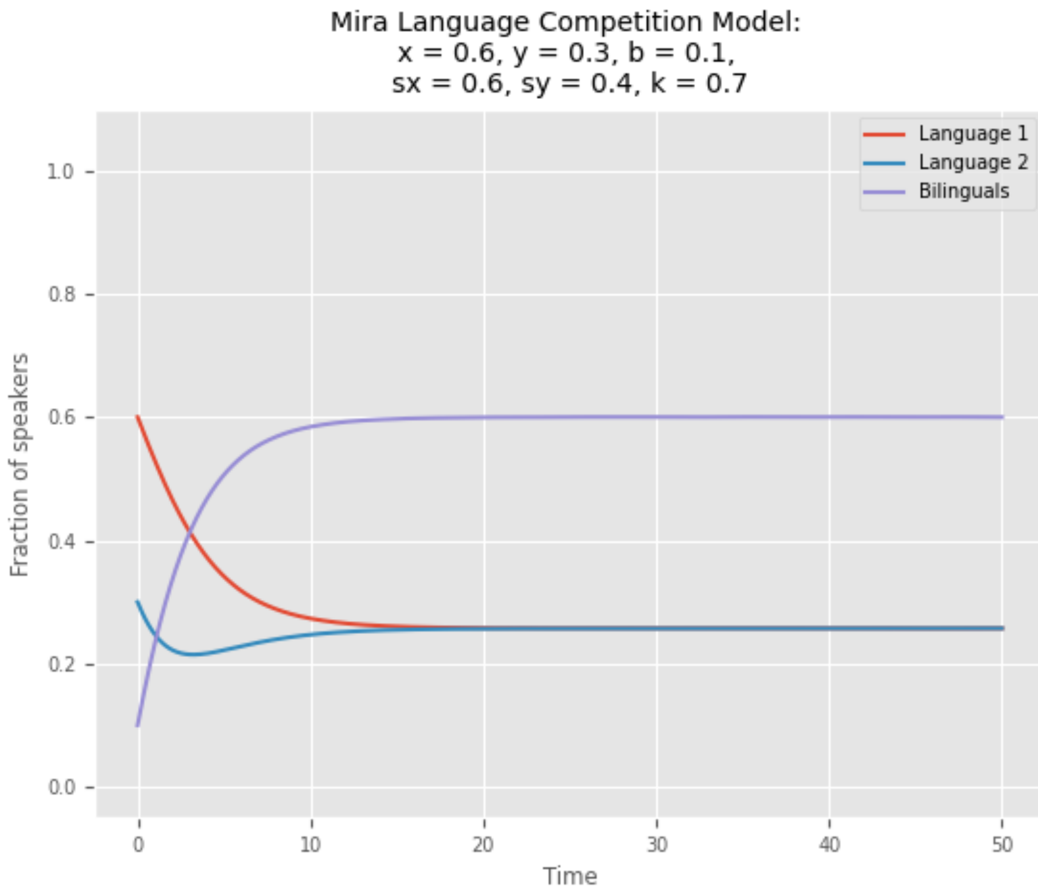

Figure 17 - Mira model with k = 0.7

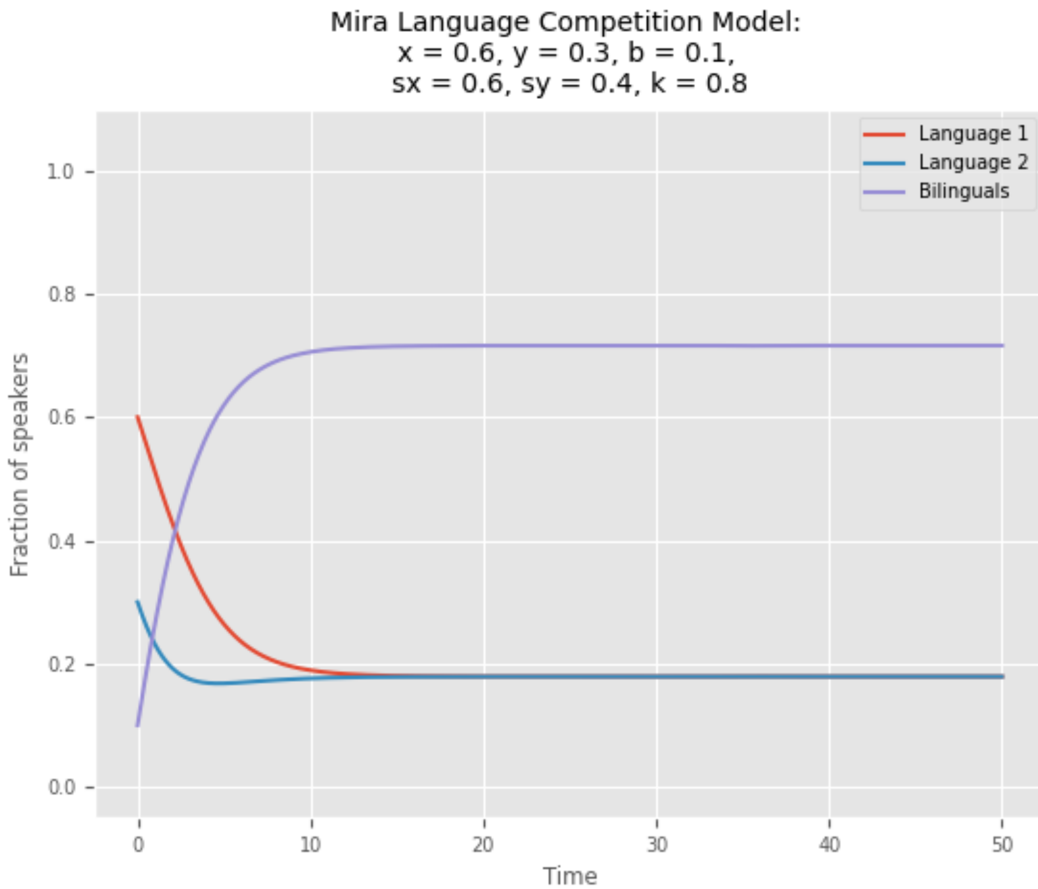

Figure 18 - Mira model with k = 0.8

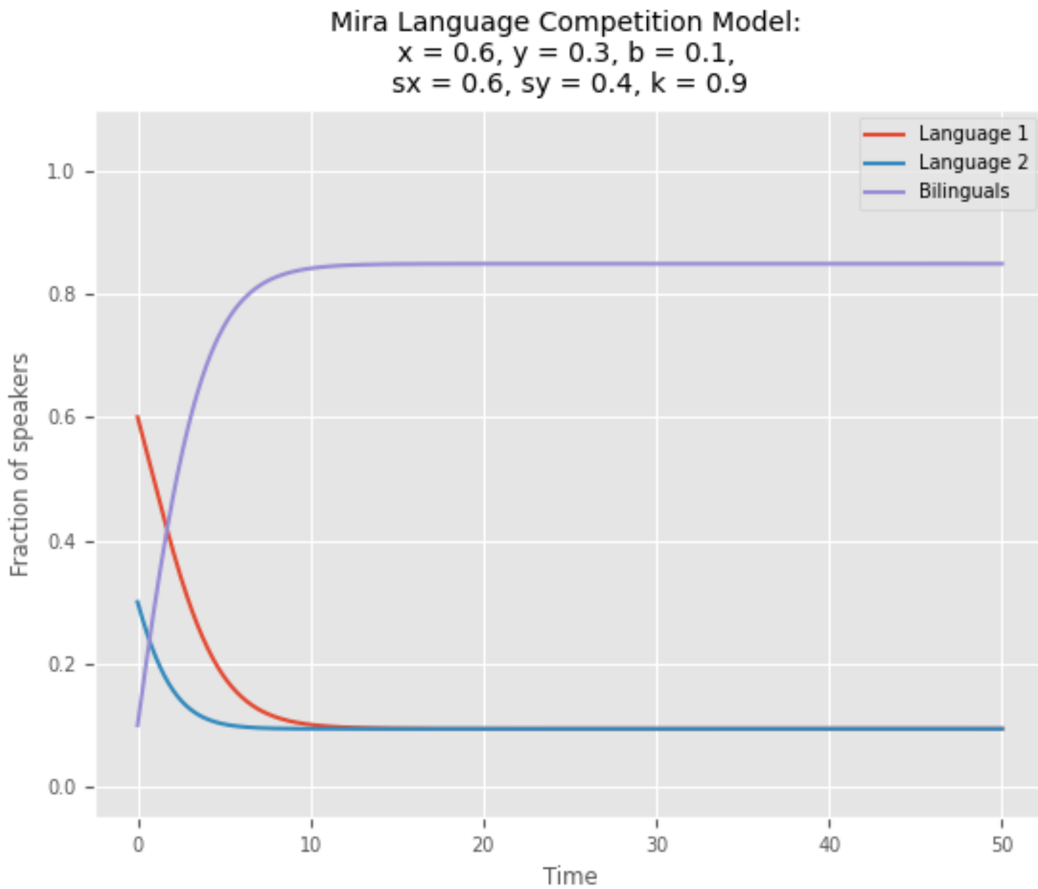

Figure 19 - Mira model with k = 0.9