

# Can two languages coexist within the same community of speakers?

Jorge Mira,<sup>1</sup> Luís F. Seoane,<sup>1</sup> Ángel Paredes<sup>2</sup>

1. Department of Applied Physics, University of Santiago de Compostela, Santiago de Compostela, Spain

2. Department of Fundamental Physics, Faculty of Physics, University of Barcelona, Barcelona, Spain

**Resum.** S'analitza l'evolució històrica de la situació lingüística de Galícia en el marc d'un model de competència entre el gallec i el castellà. El model, malgrat que és la simplificació matemàtica d'una complexa situació sociolingüística, n'ajusta amb notable precisió l'evolució des de final del segle XIX. A més, com a generalització del resultat, s'observa la possibilitat de formar una situació bilingüe estable, sense que això impliqui la segregació de dues comunitats monolingües coexistents sobre un mateix territori, en contra d'algunes prediccions teòriques prèvies. Proposem que la clau determinant per a l'estabilitat del bilingüisme és un factor que defineix la distància interlingüística entre les llengües en competició, i que aquí es calcula per al binomi gallec-castellà. A causa dels bons resultats del model, a l'hora d'ajustar una situació real com la convivència entre el gallec i el castellà a Galícia, es du a terme un estudi exhaustiu de l'estabilitat de les equacions involucrades. Aquest estudi dona resultats esclaridors sobre els possibles resultats de la competició entre dues llengües i destaca el fet que, a partir de certa semblança entre els dos idiomes, la convivència entre tots dos es torna estable. Amb referència al cas pràctic que va inspirar el treball, aquest estudi de l'estabilitat permet localitzar el sistema gallec-castellà en un context més ampli que, al seu torn, permetria proposar estratègies millors per a conservar ambdues llengües.

**Paraules clau:** bilingüisme · competència lingüística · distància interlingüística · estudi d'estabilitat

**Summary.** This article analyzes the historic evolution of the linguistic situation in Galicia within the framework of a competition model between Galician and Spanish. The model, despite being a mathematical simplification of a complex socio-linguistic situation, calibrates its evolution with remarkable precision from the end of the 19th century onward. Indeed, in a generalization of the outcome, it shows that establishing a stable bilingual situation is possible and, contrary to prior theoretical predictions, that it does not imply the segregation within one territory of the two coexisting monolingual communities. We propose that the key determinant of the stability of bilingualism is a factor that defines the *interlinguistic distance* between the competing languages and which here is calculated for the binomial Galician-Spanish. Given the model's positive results in describing a real situation such as the coexistence of Galician and Spanish in Galicia, the stability of the involved equations was exhaustively studied. The results led to enlightening conclusions on the possible outcomes of competition between the two languages, highlighting the fact that, based on a certain likeness between them, their coexistence becomes stable. In reference to the practical case that inspired this work, our study of stability allows the Galician-Spanish system to be situated within a wider context and, at the same time, suggests improved strategies for the safeguarding of both languages.

**Keywords:** bilingualism · linguistic competition · interlinguistic distance · stability studies

## Introduction

The study of complex self-adaptive systems is an area of research in which classical domains of knowledge such as economics, physics, mathematics, biology or sociology converge. Languages are an example of such systems [4].

Based on the Distinguished Lecture given by Jorge Mira at the Institute for Catalan Studies, Barcelona, on 8 April 2010.

Correspondence: Jorge Mira, Departamento de Física Aplicada, Facultade de Física, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain. Tel. +34-981563100. Fax +34-981520676. E-mail: jorge.mira@usc.es

In a world that tends towards globalization there is increasing competition between them that—just like the competition between local economic structures—forces the modification of the competing languages and might even lead to the extinction of the minority language against the hegemonic one. In fact, most of the world's languages are disappearing at an alarming speed which is parallel to the increase in interconnectivity between different parts of the world. It is estimated that of the approximately 6000 languages spoken in the world 90% might disappear in this generation [6], a situation that also concerns European languages [3].

This fact has generated a substantial increase of activity for this line of research. Simulation models typically used in socio-

physics are based on statistical learning theories [8]. In more recent studies, simulations are based on bit-string models [5], which bare some analogy to those used in economics and biology. The approach is therefore somewhat local in nature.

The present work tackles the problem of the linguistic situation in Galicia from a unique perspective. It adopts an extensive focus, describing the states of the system and the competition between its parts. The averages established give way to a system of differential equations whose resolution is only possible using IT tools.

This perspective considers languages as entities in competition by their speakers and has enjoyed recent success in the study of specific cases, as undertaken by Abrams and Strogatz (2003) [1]. In their work, now considered an essential reference, the authors did a remarkable job in tracking the historical evolution of the number of speakers from 42 of the planet's territories where two languages prevail, places such as in Peru, Scotland, Wales, Bolivia, Ireland, and Alsace-Lorraine (France). The well executed descriptions of Welsh's competition with English in Wales, Quechua's with Spanish in Peru, and Scottish Gaelic's with English in Scotland are particularly impressive.

Nevertheless, one of the main conclusions is that existing bilingual societies are unstable as a result of the recent mixing of communities with different languages that before were separated. Our work maintains that bilingualism is possible and that the determining factor for its existence is a parameter that we refer to as the *interlinguistic distance*. We also have an advantage in our familiarity with the situation in Galicia, a region that accordingly affords us a preeminent sociological laboratory in which results can be compared, extrapolated into the future, and the consequences of the respective findings analyzed.

## Model

In Abrams' and Strogatz's view, bilingualism is strictly societal: two monolingual groups coexist, not by the presence of bilingual individuals, but simply as two segregated communities occupying the same territory. In this way, of the two first languages, a speaker will clearly have only one of them. While this approach may at first seem odd, the situation is actually possible, as shown by the fact that in some areas of Peru in which Quechua and Spanish compete, cases can be found in which children are incapable of communication with their grandparents. The speaker is unable to make both languages his own and ends up prioritizing one over the other, with the consequent *theft* of speakers between them.

However, this model fails in its description of situations like those of Galicia and Catalonia, where the outcome of the competition between two Romance languages (one of them being Spanish, which has historically enjoyed high status) has been the establishment of large bilingual groups.

So, where is the problem? In their work, Abrams and Strogatz only considered competition between very different languages [1]. Communication between monolingual speakers of Quechua and Spanish is impossible, as it is between Welsh and English or Gaelic and English. An ideal example would cite

the hypothetical situation of a Galician-speaking community mixing with a Chinese-speaking one (with communication between them being impossible). With this as the starting point, it is not surprising that bilingualism has no chance at being established within this framework.

From our view, this situation limits the model of Abrams and Strogatz, as it is not directly applicable to cases comparable to Galicia's. Galician and Spanish's Latin roots make conversation between its monolingual speakers perfectly possible. Furthermore, the similitude of their vocabulary and grammar makes learning one of the languages easy if a person knows the other, as every Galician does.

This circumstance (odd at first for a monolingual person and absolutely obvious for people who, like Galicians, live within a bilingual society) has led us to consider the possibility that the emergence and survival of a socially significant bilingual group depends upon the degree of similitude between the conflicting languages. We have therefore generalized the Abrams-Strogatz model as described below, in order to incorporate this notion:

- X, Y, B being the population sets that are monolingual in the languages X, Y, and bilingual, respectively. X, y, b being the population fractions that belong to these groups (as such,  $x + y + b = 1$ )
- $P_{xy}$  being the probability, per unit time, of a speaker of language X converting to language Y, the notation of which analogously extends to the migrations between the remaining groups.

In this way, the change of x per unit of time is:

$$\frac{dx}{dt} = yP_{yx} + bP_{bx} - x(P_{xy} + P_{xb}) \quad (1)$$

with analogous equations for  $dy/dt$  and  $db/dt$ .

We should now define these probabilities. As in the Abrams-Strogatz model, we consider the probability of a speaker's change from language X to language Y as increasing the number of the latter's speakers to  $b+y$ , but also account for the change in perception of status associated with the switch, i.e., the social and economic opportunities that the individual perceives as being enjoyed by the speakers of a determined language. Referring to this parameter as  $S_Y$  (language Y's status, with the analogous notation for language X), this probability—in the case of the strict existence of monolingual speakers—should theoretically take the form

$$P_{xy} = cS_Y(1 - x)^a \quad (2)$$

where  $c$  and  $a$  are constants and the status parameter is defined between 0 and 1 (maximum status) and is complementary to the other language's status, such that

$$S_X = 1 - S_Y \quad (3)$$

Now, once we assume bilingualism as a possible situation, this probability for change should split into the probability of becoming bilingual and the probability of using only language

Y; that is, the departure of speakers of language X to language Y or to the bilingual situation is:

$$P_{XB} = ckS_Y(1 - x)^a \tag{4}$$

and

$$P_{XY} = c(1 - k)S_Y(1 - x)^a \tag{5}$$

where we introduce for the first time a parameter, *k*, that reflects the tendency toward bilingualism and as such, in accordance with our hypothesis, the similitude between the two languages. As one can see, *k* = 0 represents situations in which monolingual speakers exist, that is, conversation is impossible between monolingual speakers of each language (as in the cases studied by Abrams and Strogatz [1]), and *k* = 1 simply implies that languages X and Y are the same.

It is worth mentioning that while, until now, the similitude or distance between languages has been described from a theoretical perspective [8], the calculation of either one in practical cases has been nearly impossible.

Similarly, for the departure of speakers from language Y

$$P_{YB} = ckS_X(1 - y)^a \tag{6}$$

and

$$P_{YX} = c(1 - k)S_X(1 - y)^a \tag{7}$$

For the transfers of speakers from a bilingual situation to X, we consider  $P_{BX} = P_{YX}$  (given that the changes from B to X and from Y to X imply losing language Y). Equally,  $P_{BY} = P_{YX}$ .

Thus, we obtain two coupled differential equations matched for X and Y:

$$\frac{dx}{dt} = c[(1 - x)(1 - k)S_X(1 - y)^a - x(1 - S_X)(1 - x)^a] \tag{8}$$

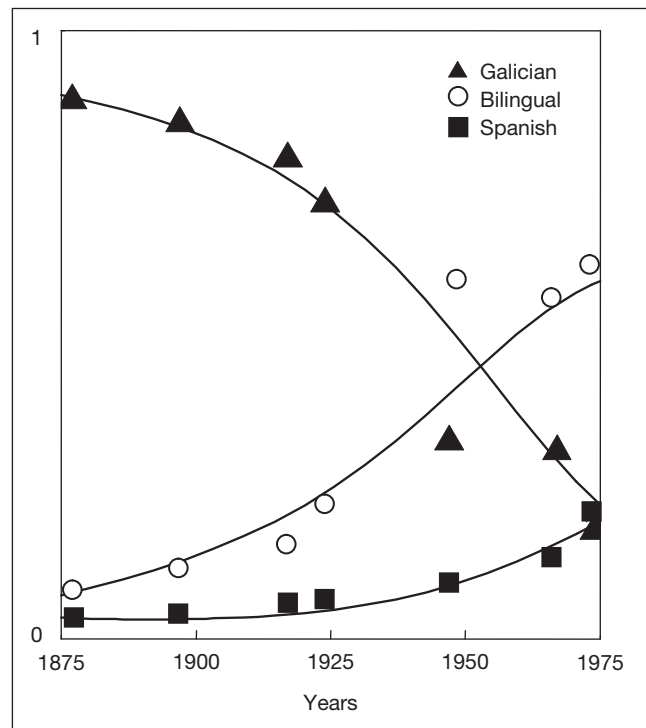
$$\frac{dy}{dt} = c[(1 - y)(1 - k)(1 - S_X)(1 - x)^a - yS_X(1 - y)^a] \tag{9}$$

### Outcome

The model successfully tracks the historical evolution of the number of bilingual and monolingual speakers in Galician and Spanish in Galicia since the year 1877, the date of the first register contained in the information from the Real Academia Galega [9], and, for the first time also permits measurement of the inter-linguistic distance between Galician and Spanish [7], which, with *k* = 0.80 (that is, an 80% similitude) seems to reasonably quantify the extensive common stem of the two languages (Fig. 1).

The outcome of the status parameter deserves separate mention: a value of 0.26 for Galician (thus, 0.74 for Spanish), which is very low, underscores the historical reality of this language in the period under consideration.

At this point, the meaning of the system's two parameters *c* and *a* should be considered. As their interpretation has poten-



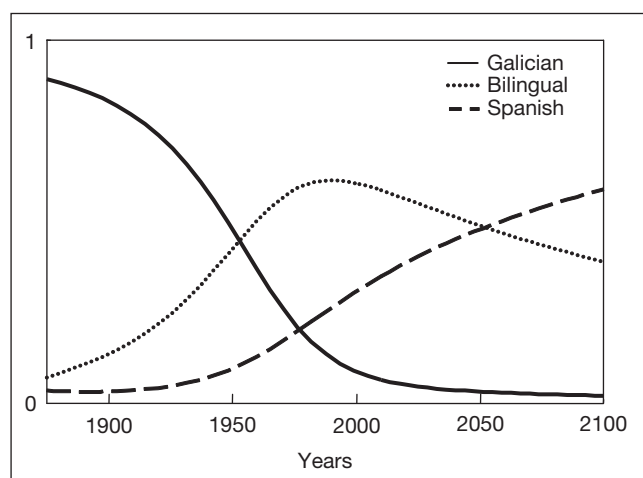
**Fig. 1.** Fraction of speakers according to year during the period 1877–1975 in Galicia (data from the Real Academia Galega). In this era, Galician suffers a low status, whereas the status of the language improves after democracy is restored. The data corresponding to 1945 are clearly an anomaly, possibly reflecting the political situation after the WWII. The lines demonstrate the successful tracking by our model, reproducing the historical evolution with the parameters *a* = 1.50, *S*<sub>Galician</sub> = 0.26, *c* = 0.1 and *k* = 0.80.

tially interesting consequences, an initial approach is described in the following.

Parameter *c*—obtained from empirical data, as was the case for *k* and *s*—is related to the speed at which the described phenomenon happens. In the event of one language's extinction when replaced by another, this constant would give us the speed at which said extinction takes place. As such, different values for *c* would have no qualitative effects on the stability of the system, nor on any other outcome.

Parameter *a* is even more interesting; in studies of different cultures, it was found to be relatively constant [1]. The outcome for the Galician-Spanish case is within the range of experimental error. This parameter escalates the influence that the number of speakers of one language has in terms of attracting new speakers. Thus, societies with different *a* values would be societies that place distinct importance on the "status" or on the number of speakers of a language in deciding which language to adopt. The fact that this parameter is constant in different civilizations allows us to reason that the importance conceded to potential speakers at the moment of deciding on a language is a universal attribute of humankind.

The work is also clearly of practical interest because it opens up the possibility of carrying out long-term, projectional studies on behavior. In this article, we present our analysis up to the year 2100, using the same tracking parameters obtained from the historical data (Fig. 2). First, the Spanish monolingual fraction can be seen to constantly rise while the fraction of Galician



**Fig. 2.** Projection of the fraction of Galician speakers, Spanish speakers, and bilinguals up to the year 2100, based on the tracking shown in Fig. 1. According to the parameters of this tracking system, the group of monolinguals in Spanish steadily grows while that of bilinguals, which had experienced a surge throughout the 20th century, reaches a maximum around the year 2000, at which point it steadily declines.

monolinguals notably drops until their presence becomes residual by around the year 2050. By the year 2100, Galician speakers can be considered as extinct in this simulation. It should be noted that the disappearance is of the monolingual group and *not of the Galician language itself*, which would continue to survive in the bilingual group.

It is worth separately pointing out the bilingual group's behavior: according to the model, speakers of both languages formed the majority group in the last part of the 20th century and into the start of the 21st, an easily corroborated situation because it responds to the current reality in Galicia (currently, the majority of its population claims to be bilingual).

Nevertheless, and despite the bilingual group's initial rise, which at first seemed to indicate its progressive hegemony, according to the model the system's dynamics takes an unexpected turn at the start of the 21st century. After reaching a maximum, the bilingual group starts to decrease, slowly, but surely. With Galician's status having a value of 0.26, the bilingual situation is, therefore, not stable.

One clear affirmation of the simulations is the progress of the group of Spanish monolinguals, an issue that tends to generate intense controversy but which is readily confirmed by an impartial analysis. Interpretations aside, the scientific fact is that studies undertaken in the past few years have shown a progressive and alarming loss of the use of Galician and a corresponding increase in the Spanish monolingual option in younger generations. The inverted population pyramid of monolinguals in Galician reflects, in fact, the social situation of Galicia, in which it can be clearly seen that the use of the Galician monolingual option has decreased in direct proportion to the decrease in the person's age, and along a line that coincides with the analysis derived from our system of equations.

For example, according to the latest data available to us [2], the percent of monolinguals in Galician has fallen from 65.7% in people over 65 years of age to 27% in those 16–25 years of age. At the same time, for monolinguals in Spanish, the ten-

dency is the inverse, rising from 10.4% in those over 65 to 33.1% in those under 16. Since the youth of today are the speakers of tomorrow, the sociolinguistic situation described by our model seems correct.

### Intensive study of the stability of the system of equations

In searching for definitive proof that the model could lead to situations in which the two languages stably coexist along with a bilingual group, we carried out an intensive study of the equations, numerically solving them for multiple arbitrary initial conditions and with 400 different pairs of values for the constants  $k$  and  $s$ . This also allowed us to verify the theory's consistency with other, earlier ones since the Abrams-Strogatz model is invoked for  $k = 0$  (i.e., a total lack of similarity between the two languages). Thus, it was quickly verified that for a great number of the pairs  $k$ - $s$ , the instability outcomes obtained by the authors remain valid and only one of the languages survives. This happens in situations of great asymmetry in status and for a large range of values in interlinguistic distance.

When the statuses of the two languages are similar but the interlinguistic similarity remains low, the extinction of one of the languages is once again inevitable, but which of them will disappear depends upon the initial distribution of speakers. At this point in the analysis, there is no trace of stability in bilingualism or in the coexistence of the two languages.

The situation changes drastically upon sufficiently increasing the interlinguistic similitude. For each status value, there is a critical similitude factor,  $k_c$ , such that if  $k > k_c$ , then the situation can revert to being stable. This stability presumably depends heavily upon the initial distribution of the speakers, with it being equally likely that the minor-status language disappears in its monolingual speakers, but survives in a share of bilingual speakers. If the similarity between the languages continues to increase, this dependence of the stability on the initial conditions disappears, replaced by situations in which both languages always survive amongst their corresponding monolingual speakers, with a measure of bilingual speakers always remaining.

Upon studying the Galician-Spanish system from this new point of view, we can see that the extinction of Galician is inevitable, with the initial distribution of speakers of each of the languages making no difference on the outcome. Effective modifications of the parameters—which could be a variation of *status* (a possibility already pointed out by Abrams and Strogatz) or the increment of the *interlinguistic similitude*—could bring the system to a point that would allow for the coexistence of the two languages, as well as the preservation of the set of bilingual speakers. This eventual outcome would depend in great part upon the distribution of speakers reaching the desired parameters.

### Conclusions

A model was presented that reproduces the historical evolution of the linguistic situation in Galicia, including a description of

the group that is bilingual. The model is a simplification of reality in that it reduces a complicated sociolinguistic situation to a few parameters in the equations. Nonetheless, the data available to us are described with remarkable precision and we believe that, as a whole, the results are in qualitative agreement with the observed linguistic evolution in Galicia.

A mathematical study of the stability of the model's equations revealed disparate behaviors that represent the possible distinct evolutions of a system's two languages in terms of competition or coexistence. Two of these evolutionary pathways were already described in the Abrams and Strogatz model [1] and both necessarily imply the extinction of one of the two languages. The other possibilities allow for the stable coexistence of the two languages, contrary to the original model. This coexistence does not entail the segregation of the two communities but is linked to the survival of a group of bilingual individuals.

The key to permitting the stability of bilingualism seems to be the introduction into our model of a factor such as interlinguistic distance, which softens the effect of the differences in status between the languages in competition. To this factor—which has been included in different models—we can now assign a quantitative value based on our tracking: 0.80 for the Galician-Spanish binomial (an 80% coincidence between both languages).

Projections on the future of the sociolinguistic situation in Galicia indicate that a failure of the Galician language to reach a status of 0.335 would result in the sustained growth of the group of Spanish monolinguals—to the detriment of the group of Galician monolinguals or even of the bilinguals such that neither would continue to be a stable group. Thus, the survival of Galician would be endangered (Fig. 3).

Which of the predicted behaviors prevails depends upon the parameters  $k$  and  $s$ , and the “winner” will place the Galician-Spanish system within a more global framework. This, in turn, will allow the pursuit of strategies to achieve a stable, desirable coexistence of the two languages.

## References

- [1] Abrams DM, Strogatz SH (2003) Modeling the dynamics of language and death. *Nature* 424:900
- [2] Consello da Cultura Galega (2005) A Sociedade galega e o idioma. A evolución sociolingüística de Galicia (1992-2003)
- [3] Euromosaic (1996) Production and reproduction of the minority linguistic groups of the European Union. European Commission
- [4] Hawkins JA, Gell-Mann M (1972) *The Evolution of Human Languages*. Addison-Wesley, Reading, Massachusetts
- [5] Kosmidis K, Halley JM, Argyrakis P (2007) Language evolution and population dynamics in a system of two interacting species. *Physica A* 353:595-612
- [6] Krauss M (1992) The world's languages in crisis. *Language* 68:4-17
- [7] Mira J, Paredes A (2005) Interlinguistic similarity and language death dynamics. *Europhysics Letters* 69:1031-1034
- [8] Nowak MA, Komarova NL, Niyogi P (2002) Computational and evolutionary aspects of language. *Nature* 417:611-617
- [9] Real Academia Galega (1994) Usos lingüísticos en Galicia. Real Academia Galega (Santiago de Compostela)

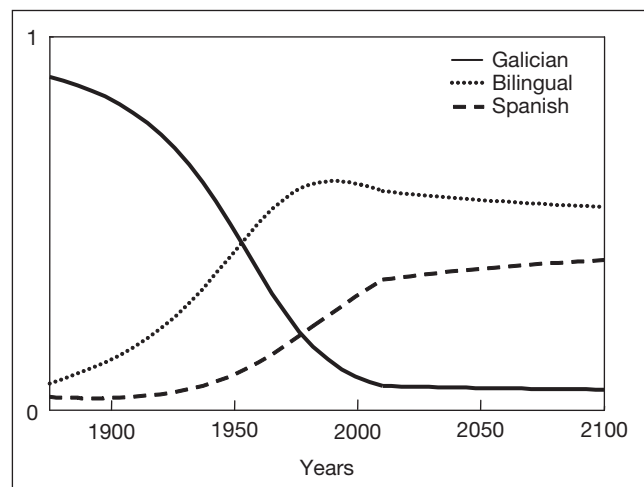
## About the authors

**Jorge Mira** is a physicist and scientific communicator. He is currently professor in the Area of Electromagnetism at the University of Santiago de Compostela (USC). His research is mainly focused in the electric and magnetic properties of solids, a field in which he has authored over 100 articles. He has been a finalist in the Award of the Royal Spanish Soci-

ety of Physics to Novel Researchers (1999), and the recipient of the Award of the Royal Galician Academy of Sciences (2002). He is the director of the Department of Applied Physics of the USC and of the ConCiencia program, a series of activities for the popularization of science with the participation of Nobel Prizes (or their equivalents in Mathematics and Computer Science). Besides this, he is a collaborator in newspapers and radio

and television programs. For such activities of science communication he has received several national awards.

**Luis F. Seoane** earned his degree in Physics (with outstanding award) at the University of Santiago de Compostela in 2009, with one-year stays at the Universities of Heidelberg (Erasmus program) and Granada (Séneca program). During his studies, he also worked as assistant



**Fig. 3.** Future projection of the sociolinguistic situation in Galicia up to the year 2100 in a scenario with the interlinguistic distance  $k = 0.80$  and a Galician status  $S_{\text{Galician}} = 0.36 > S_{\text{min}}$ . It can be seen that, once the minimum status is overcome, the survival of all three groups of speakers is guaranteed in time.



at the Max Plank Institute for Astronomy (Heidelberg) and as young researcher at GSI (Darmstadt). In 2009, he was awarded the Special Prize of the Arquímedes Program, the Spanish contest of scientific research for undergraduate students. He is currently studying a Master's degree in Computational Neurosciences at the Bernstein Center for Computational Neuroscience (Berlin), granted by the Barrié de la Maza Foundation.

**Ángel Paredes** is a physicist, and has mainly worked in the area of theoretical high energy physics. He obtained his degree (with outstanding award) and Ph.D. (with outstanding Ph.D. award) in Physics at the University of Santiago de Compostela. He carried out postdoctoral research stays in three leading European Institutions: École Polytechnique (France), University of Utrecht (the Netherlands) and University of Barcelona. He has pub-

lished 25 articles in leading international journals and has been invited to present his work at specialized seminars in 15 different countries. He has been awarded several prizes such as the first position of the Spanish Prize of Secondary Education, in 1995, and Gold Medals at both the Physics and the Mathematics Spanish Olympiad. He also won the Bronze Medal in the International Mathematics Olympiad.