GRADO EN FÍSICA

Treball de Grau

Models dinàmics de competició entre llengües

Juliol de 2014

Alumna: Clara Miralles Vila

Tutor (1): Manel Perucho Pla

Índex

Resum

En aquest treball es presenta l'estudi de la dinàmica de l'evolució de dues llengües en interacció, el valencià i el castellà, a partir de models matemàtics d'equacions diferencials. La investigació es centra en el model d'Abrams-Strogatz i el model de Mira-Paredes, els quals apliquem a dades empíriques obtingudes d'enquestes realitzades pel Servici d'Investigació i Estudis Sociolingüístics (SIES) de la Generalitat Valenciana i per l'Acadèmia Valenciana de la Llengua (AVL). Partint dels ajusts dels paràmetres que controlen els models, es presenta un estudi de la dependència de l'evolució de l'ús del valencià amb aquests. Finalment, ha estat possible establir uns certs rangs de valors possibles dels paràmetres que descriuen la interacció entre les dues llengües.

Abstract

This work presents the study of the dynamics of the evolution of two interacting languages, Valencian and Spanish, using mathematical models of differential equations. The research focuses on the Abrams-Strogatz model and the Mira-Paredes model, which are applied to empirical data obtained from surveys conducted by the Service of Sociolinguistic Research and Studies (SIES) of the Generalitat Valenciana and by the Valencian Academy of the Language (AVL). A study of the dependence of the evolution of the usage of the Valencian language on the parameters that control the models is presented, based on the fitting of these parameters. Finally, it has been possible to establish certain ranges of possible values of the parameters that describe the interaction between the two languages.

1 Introducció

> "El mite diu que al Paradís hom parlava una llengua i que després de Babel tot de llengües es van escampar pel nostre món. El ventall esplendorós de les llengües és el mirall de la humanitat, és la consciència que es manifesta en milers de veus diverses que canten una cançó única: la que parla de les arrels més fondes de la dignitat humana. És per això que l'amor a la pròpia llengua (sense xenofòbies, ni xovinismes) és l'amor al llenguatge i és l'estimació pregona de tot allò que ens fa humans i que ens permet, definitivament, viure en la coherència. No volem tornar, mai, a cap Paradís."
>
> Jesús Tuson. [1]

En l'actualitat, prop de 6000 llengües coexisteixen al món, però el patrimoni lingüístic de la humanitat es troba en una situació crítica [2]. La distribució de parlants és altament desigual: el 4% de les llengües són parlades pel 96% de la població, mentre que el 25% tenen menys de 1000 parlants [3]. Segons les previsions dels especialistes més optimistes, la meitat de les llengües desapareixerà aquest segle. Altres en prediuen la desaparició del 90% [2]. És curiós que, com senyala J. Tuson en algunes de les seues obres [4, 5], ningú posa en dubte la necessitat de protegir plantes o animals en perill d'extinció, mentre que quan es tracta de llengües l'actitud sol ser totalment la contrària: sembla que s'advoca per un reduccionisme lingüístic. Cal no oblidar, però, que les llengües són un dels patrimonis culturals més valuosos que tenim. Com afirma Crystal [3] són un reflex de la diversitat cultural i un element essencial de la identitat d'un poble; són també magatzems de la història i contribueixen a la suma del coneixement humà, a més de ser interessants en sí mateixes.

Fins fa poc més d'una dècada, l'estudi de l'evolució de les llengües i les relacions entre elles era tractat pràcticament sols per sociolingüistes. Tanmateix, la publicació al 2003 a la revista *Nature* de l'article *Modelling the dynamics of language death* [6] pel físic D. M. Abrams i el matemàtic S. H. Strogatz sembla haver estat el detonant de l'aplicació de models matemàtics en l'àmbit lingüístic. És important destacar la importància d'unir els esforços dels sociolingüistes i els físics en aquest tipus d'estudis. Ens sembla necessària la col·laboració entre ambdós grups per poder descriure coherentment la realitat sociolingüística i extraure'n conclusions pertinents dels models, com bé apunta J. Kabatek a l'article *Modelos matemáticos e substitución lingüística* [7].

La competició lingüística estudia la dinàmica de l'ús de les llengües degut a les interaccions socials en una societat multilingüe [8]. Podem classificar els estudis realitzats sobre la competició entre llengües en dos grans tipus [9]:

- Els models basats en equacions diferencials ordinàries no lineals que descriuen la dinàmica de la població de parlants de diferents llengües (*mean-field models*). Aquest és l'enfocament que es pren a l'article d'Abrams-Strogatz [6], així com el d'altres que segueixen la mateixa línia [10, 11, 12, 13].
- Els models basats en una descripció estadística en la qual cada individu es tractat de forma separada. Aquests models tracten la dinàmica de les llengües emprant sistemes de múltiples partícules (parlants) en termes d'una descripció microscòpica basada en regles d'interacció entre aquests (*agent-based models*). Estudis que implementen aquest tipus de tractament són, per exemple, [8, 14, 15].

En aquest treball ens centrarem en el primer grup: models basats en equacions diferencials per a la descripció de la competició entre llengües. Emprarem dos d'aquests models per estudiar l'evolució de l'ús del valencià a partir de dades obtingudes d'enquestes realitzades pel Servici d'Investigació i Estudis Sociolingüístics (SIES) de la Generalitat Valenciana i per l'Acadèmia Valenciana de la Llengua (AVL). En primer lloc, treballarem amb un dels models més simples, el model d'Abrams-Strogatz [6]. Aquest model simplifica la dinàmica de dues llengües a una equació diferencial per descriure com una llengua desapareix sota la preponderància d'una altra. En segon lloc, estudiarem la versió millorada d'aquest model proposada per J. Mira i A. Paredes [12] , que introdueix un nou grup d'individus bilingües al model i basa l'estabilitat d'aquest en la semblança entre les llengües en conflicte.

Altres autors han proposat també diverses millores del model d'Abrams-Strogatz, les quals no tractarem en aquest treball. Per exemple, M. Patriarca i T. Leppänen intenten generalitzar el model introduint una dependència espacial (per territoris) en termes d'una equació de reacció-difusió [10], mentre que J. P. Pinasco i L. Domanelli plantegen un model alternatiu del tipus Lotka-Volterra [11]. En aquest últim, Pinasco i Domanelli mesclen un model de dues espècies en competència en què la presa es converteix en depredador una vegada capturada (depredador de tipus vampir), i un model d'epidèmies en què els individus infectats no es recuperen ni moren immediatament, per tractar la dinàmica de competició lingüística.

Existeixen també altres tipus de problemes relacionats amb la llengua, que inclouen aquells relacionats amb l'evolució de l'estructura d'una llengua i el seu procés d'aprenentatge, així com l'estudi de la dinàmica no únicament de dues llengües en contacte, sinó de múltiples llengües en conjunt [16]. Aquests, però, excedeixen el propòsit del nostre treball.

2 Fonament teòric

2.1 Model d'Abrams-Strogatz

El model d'Abrams-Strogatz considera un sistema format per dues llengües en competició, X i Y. El nombre de parlants d'aquestes està determinat per:

- el nombre total de parlants de la llengua.
- el seu *estatus*, paràmetre que reflexa l'atractiu d'aquesta, és a dir, les oportunitats socials o econòmiques que s'ofereixen als seus parlants.

La dinàmica d'aquest sistema estarà regida per l'equació:

$$\frac{dx}{dt} = yP_{yx} - xP_{xy} \tag{1}$$

on x és la fracció de població que parla la llengua X, y la fracció de població que parla la llengua Y, P_{yx} la probabilitat de canviar de Y a X i P_{xy} la probabilitat de realitzar el canvi invers. El model considera un sistema tancat de parlants monolingües, és a dir, el nombre de persones que no parlen X, parlaran Y, i per tant $y = 1 - x$. És senzill vore que la variació en x serà igual a la fracció de parlants de la llengua Y multiplicada per la probabilitat de guanyar parlants (P_{yx}) menys la fracció de parlants de X multiplicada per la probabilitat de perdre'n (P_{xy}).

Definim el paràmetre s per tal de caracteritzar l'estatus percebut de la llengua X, amb $0 < s \leq 1$. Per simetria, si intercanviem les llengües hem d'obtindre la mateixa probabilitat de transició, de forma que: $P_{xy}(x, s) = P_{yx}(1 - x, 1 - s)$. Així mateix, el model també assumeix que ningú adoptarà una llengua sense parlants, $P_{xy}(0, s) = 0$, o sense prestigi, $P_{yx}(x, 0) = 0$, i que $P_{yx}(x, s)$ i $P_{xy}(x, s)$ són funcions suaus i monòtonament creixents en ambdós arguments. Abrams i Strogatz proposen que les funcions de probabilitat de transició tinguen la forma:

$$P_{yx}(x,s) = c\ s\ x^a \tag{2a}$$
$$P_{xy}(x,s) = c\ (1-s)(1-x)^a \tag{2b}$$

on a és un paràmetre a determinar en l'ajust i que caracteritza la dependència amb el nombre de parlants de la llengua, i c és un factor de normalització relacionat amb l'escala temporal.

Ajuntant tots aquests elements obtenim l'equació:

$$\frac{dx}{dt} = c\ (1-x)\ x\ \left[x^{a-1}s - (1-x)^{a-1}(1-s)\right] \tag{3}$$

En el cas $s < 0.5$, l'equació modelitza la competència entre una llengua de menor prestigi X contra una de més prestigiosa Y.

L'equació (3) té tres punts fixos[1] per a $a \neq 0$: $x_0 = 0$, $x_0 = 1$ i un tercer $0 < x_0^* < 1$ que depén d'a i s i té la forma:

$$x_0^* = \frac{1}{1 + \left(\frac{s}{s+1}\right)^{\left(\frac{1}{a-1}\right)}} \tag{4}$$

Per a $a > 1$, $x = 0$ i $x = 1$ són estables, mentre que x_0^* és inestable. Per a $a < 1$ els punts estables es fan inestables i viceversa [8], com vorem en l'estudi d'estabilitat que realitzarem en el següent subapartat.

El paràmetre a en les expressions de les probabilitats de transició d'una llengua a una altra representa la velocitat de la disminució de parlants. Al seu article sobre la mort de les llengües [6], Abrams i Strogatz ajusten el seu model a 42 conjunts de dades corresponents a parelles de llengües en vies d'extinció. Sorprenentment troben que aquest paràmetre és constant: $a = 1.31 \pm 0.25$.

En el cas $a = 1$, que implica una dependència lineal de la probabilitat de transició entre llengües amb la fracció de parlants, l'equació (3) es simplifica i dóna lloc a una equació similar a l'equació logística:

$$\frac{dx}{dt} = c\ (2s-1)(1-x)x \tag{5}$$

Aquesta equació té sols dos punts fixos, $x = 1$ i $x = 0$ i té solució analítica:

[1]Un punt fix o punt crític d'una equació diferencial és aquell que satisfà $dx/dt = 0$. Es tracta d'una solució d'equilibri, ja que si el sistema es troba en un punt crític x_0, romandrà en aquest punt indefinidament si el sistema no es veu pertorbat.

$$x = \frac{1}{e^{-c\ (2s-1)(t-C)} + 1} \tag{6}$$

on C és la constant d'integració. A més, per al cas de dues llengües equivalents, $s = 0.5$, la fracció de parlants roman constant en el temps.

2.1.1 Estabilitat lineal dels punts crítics

Estudiarem ara l'estabilitat dels punts crítics de l'equació d'Abrams-Strogatz a partir de l'anàlisi lineal de l'expressió (3) al voltant d'aquests.

Com ja s'ha intruduït prèviament, el model d'Abrams-Strogatz presenta tres punts fixos: $x_0 = 0$, $x_0 = 1$ i x_0^* donat per (4). Aquests es troben senzillament en imposar que x no varie amb el temps, és a dir, obtenint les solucions que satisfan $f(x) = 0$, on

$$f(x) = \frac{dx}{dt}$$

que ve donada per l'equació (3).

Una vegada coneguts els punts fixos x_0, realitzem una expansió de Taylor al voltant de cadascun d'aquests punts:

$$f(x) = f(x_0) + \left.\frac{df}{dx}\right|_{x=x_0} (x - x_0) + \mathcal{O}((x - x_0)^2) \tag{7}$$

En qualsevol dels punts fixos, $f(x_0) = 0$. A més, prop d'aquests els termes d'ordre superior seran menyspreables front a $(x - x_0)$, de forma que podem aproximar $f(x)$ com:

$$f(x) = m(x - x_0) \tag{8}$$

on hem definit $m = \left.\frac{df}{dx}\right|_{x=x_0}$. Si definim $w = x - x_0$ i tenim en compte que $f(x) = \frac{dx}{dt} = \frac{dw}{dt}$ obtenim:

$$\frac{dw}{dt} = mw \tag{9}$$

que correspon a l'equació diferencial de variació exponencial de w. Per tant, arribem trivialment a:

$$(x - x_0) = Ae^{mt} \tag{10}$$

on A és una constant. Si $m > 0$ i partim d'un entorn de x_0, $x(t)$ s'allunyarà exponencialment del punt fixe. Per contra, si $m < 0$, $x(t)$ s'aproximarà de forma exponencial a x_0. En conseqüència, el punt fixe serà estable si $m < 0$ i inestable si $m > 0$.

Si tornem a l'equació d'Abrams-Strogatz, obtenim que la derivada de $f(x)$ ve donada per:

$$\begin{aligned}\frac{df}{dx} =& c(1-x)\left[x^{a-1}s-(1-x)^{a-1}(1-s)\right] \\ & - cx\left[x^{a-1}s-(1-x)^{a-1}(1-s)\right] \\ & + cx(1-x)(a-1)\left[sx^{a-2}+(1-s)(1-x)^{a-2}\right]\end{aligned} \tag{11}$$

Si avaluem aquesta expressió en els punts crítics, obtindrem l'estabilitat dels nostres punts. Cal fer notar que per a $a<1$, l'expressió no està ben definida per als punts $x_0=0$ i $x_0=1$, ja que els termes x^{a-1} i $(1-x)^{a-1}$ donen infinits. Per tant sols podrem estudiar analíticament l'estabilitat d'aquests dos punts per a valors de $a>1$.

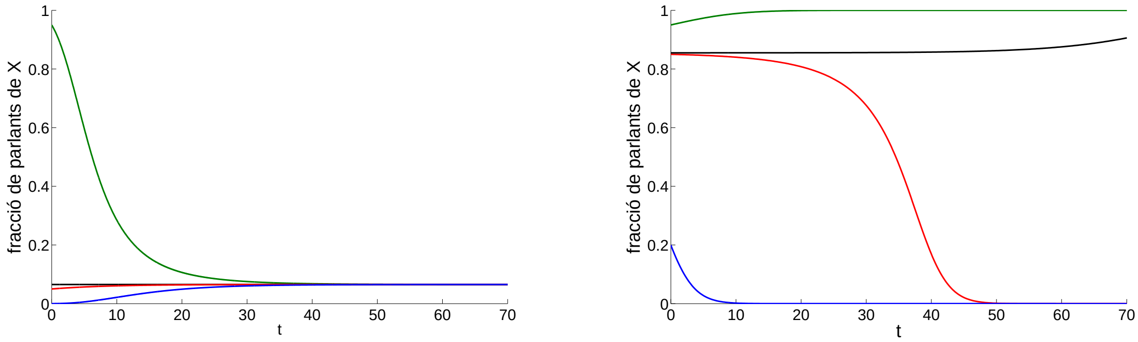
Substituint a (11):

$$\begin{aligned} x_0 = 0 \quad &\rightarrow \quad m = -c\,(1-s) \qquad \text{per a } a>1 \\ x_0 = 1 \quad &\rightarrow \quad m = -c\,s \qquad\qquad \text{per a } a>1 \\ x_0 = x_0^* \quad &\rightarrow \quad m = c\,(a-1)\left[s\,(x_0^*)^{a-1}(1-x_0^*)+(1-s)\,x_0^*(1-x_0^*)^{a-1}\right]\end{aligned}$$

Donat que $0<s<1$, tant $x_0=0$ com $x_0=1$ són estables per a $a>1$. En el cas de x_0^*, observem que la seua estabilitat dependrà del signe de $(a-1)$. Si $(a-1)>0$, el que és equivalent a imposar $a>1$, x_0^* serà inestable, mentre que si $a<1$, serà estable.

Per a $a<1$ l'estudi de l'estabilitat de $x_0=0$ i $x_0=1$ s'ha de fer numèricament. La Figura 1 mostra l'evolució del sistema segons el model d'Abrams-Strogatz per a diferents condicions inicials. S'observa clarament l'estabilitat dels punts crítics per a $a<1$ i $a>1$. Així, veiem que per a $a<1$ tant $x_0=0$ com $x_0=1$ són inestables. Comprovem també en aquesta Figura l'estabilitat predita en la nostre anàlisi algebraica en la resta dels casos.

En el model de competició entre llengües s'espera que les solucions $x=0$ (desaparició de la llengua X) i $x=1$ (tots els individus són parlants d'X) siguen estables. Per tant, és raonable considerar sols valors d'a majors a la unitat.



(a) $a = 0.8$, $s = 0.37$ i $c = 1$.
En aquest cas, $x_0^* = 0.0653$ és estable i $x = 0$ i $x = 1$ són inestables.
C.I.: $x_i = 0.95$ (verd), $x_i = x_0^*$ (negre), $x_i = 0.06$ (roig) i $x_i = 0$ (blau).

(b) $a = 1.3$, $s = 0.37$ i $c = 1$.
En aquest cas, $x_0^* = 0.8550$ és inestable i $x = 0$ i $x = 1$ són estables.
C.I.: $x_i = 0.95$ (verd), $x_i = x_0^*$ (negre), $x_i = 0.85$ (roig) i $x_i = 0.2$ (blau).

Figura 1: Fracció de parlants d'X front al temps. Solució numèrica del l'equació diferencial d'Abrams-Strogatz. Notem que per a $a < 1$, $x_0 = 0$ i $x_0 = 1$ són inestables i x_0^* és estable, mentre que per a $a > 1$, x_0^* es fa inestable i $x_0 = 0$ i $x_0 = 1$ passen a ser estables.

2.2 Model de Mira-Paredes

Com a millora del model d'Abrams-Strogatz, J. Mira i A. Paredes incorporen la possibilitat de parlants bilingües al seu article *Interlinguistic similarity and language death dynamics* [12]. En el seu treball suggereixen que el bilingüisme estable és possible i el fet que es done o no depén del grau de semblança entre les dues llengües que competeixen.

El model de Mira-Paredes generalitza el d'Abrams-Strogatz incorporant un tercer grup de parlants bilingües, B. Denotarem per b la fracció de població bilingüe, de forma que en aquest cas $x + y + b = 1$, sent x la fracció de parlants de la llengua X i y la de la llengua Y.

La dinàmica dels parlants de les llengües ve donada per:

$$\frac{dx}{dt} = yP_{yx} + bP_{bx} - x(P_{xy} + P_{xb}) \tag{12a}$$

$$\frac{dy}{dt} = xP_{xy} + bP_{by} - y(P_{yx} + P_{yb}) \tag{12b}$$

$$\frac{db}{dt} = bP_{xb} + yP_{yb} - b(P_{by} + P_{bx}) \tag{12c}$$

on P_{ij} denota la probabilitat que un individu pertanyent al grup i passe al grup j. Anàlogament al model d'Abrams-Strogatz, els autors assumeixen que la probabilitat que un monolingüe d'una llengua siga reemplaçat per un monolingüe de l'altra o un bilingüe és proporcional a l'estatus de la nova llengua i a una potència de la fracció de la població que parla aquesta nova llengua. Tanmateix, inclouen un nou paràmetre k, $0 \leq k \leq 1$, que reflecteix la facilitat que la desaparició d'un monolingüe de la llengua X siga compensada per l'aparició d'un parlant del grup bilingüe en compte d'un monolingüe de la llengua Y. Aquest paràmetre s'identifica amb la *semblança* entre les llengües. Així, les probabilitats de transició s'expressen com:

$$P_{xb} = c\ k\ (1-s)(1-x)^a \tag{13a}$$
$$P_{yb} = c\ k\ s\ (1-y)^a \tag{13b}$$
$$P_{bx} = P_{yx} = c\ (1-k)\ s\ (1-y)^a \tag{13c}$$
$$P_{by} = P_{xy} = c\ (1-k)\ (1-s)(1-x)^a \tag{13d}$$

on s és l'estatus de la llengua X.

S'ha pres $P_{bx} = P_{yx}$ ja que el pas de B a X i d'Y a X involucra la pèrdua de parlants d'Y, que ocorre principalment després de la mort dels parlants d'Y. De forma similar, es pren $P_{by} = P_{xy}$.

Cal notar que habitualment els parlants d'una llengua no passen a parlar-ne una de nova oblidant la seua. És per això que aquest model s'ha d'entendre més com un model de renovació de població que no pas com un model en què els individus canvien d'una llengua a una altra diferent. Tanmateix, el model permet, a diferència del d'Abrams-Strogatz, que els parlants monolingües es convertisquen en bilingües [13].

Quan $k = 0$, és a dir, quan la comunicació entre parlants de les llengües X i Y és impossible, el model es redueix al d'Abrams-Strogatz, o acaba decaient a aquest si inicialment la població b és diferent de zero. El cas contrari, $k = 1$, implica que X=Y.

Finalment, substituint a (12) les expressions de les probabilitats de transició donades per (13), s'obté el parell d'equacions diferencials acoblades per a x i y:

$$\frac{dx}{dt} = c\ \left[(1-x)(1-k)s(1-y)^a - x(1-s)(1-x)^a\right] \tag{14a}$$
$$\frac{dy}{dt} = c\ \left[(1-y)(1-k)(1-s)(1-x)^a - ys(1-y)^a\right] \tag{14b}$$

que es redueixen a la d'Abrams-Strogatz, equació (3), quan $k = b = 0$.

2.2.1 Estabilitat del model de Mira-Paredes

Per tractar l'estabilitat del sistema d'equacions acoblades del model de Mira i Paredes (14) ens referirem al treball realitzat per Mira, Seoane i Nieto [13]. En el seu article, presenten el resultat de l'estudi numèric d'aquestes equacions acoblades per a $a = 1.3$ (corresponent al valor obtingut per Abrams i Strogatz [6]) i diferents valors de s i k, amb el propòsit de trobar els resultats possibles a llarg termini de la competició entre dues llengües.

Depenent de s i k, s'obtenen cinc situacions possibles:

I Un únic punt estable en $x = 1$ ó $y = 1$, i.e. una de les llengües s'extingeix fins i tot en el cas que inicialment fóra dominant. Aquest comportament es troba il·lustrat a la Figura 2(a).

II Dos estats estables en $x = 1$ i $y = 1$: una de les llengües desapareix, però quina de les dues depén de la distribució inicial de parlants (x_0, y_0) (Figura 2(b)).

III Dos estats estables en $x = 1$ i $y = 1$ i un tercer sota la línia $x + y = 1$ i que, per tant, correspon a un grup bilingüe (Figura 2(c)).

IV Un estat estable sota la línia $x + y = 1$, de nou corresponent a un grup bilingüe, juntament amb un estat monolingüe en $x = 1$ ó $y = 1$ (Figura 2(d)).

V Un únic estat estable corresponent a un grup bilingüe (Figura 2(e)).

La Figura 2(f) mostra a quina de les cinc situacions dóna lloc cada parella de valors (s, k).

Els trets més rellevants que extrauen de la seua anàlisi Mira, Seoane i Nieto són:

– que l'existència estable d'un grup bilingüe en el model de Mira-Paredes (situacions III, IV i V) requereix un valor de semblança k superior a un valor mínim que es troba al voltant de 0.35.

– que si $k \geq 0.6$, com major siga la diferència entre l'estatus de cada llengua, major ha de ser la semblança per poder tindre una situació en què el bilingüisme siga estable.

– que si les dues llengües tenen un estatus semblant, aquella que desapareix depén de la grandària inicial dels grups lingüístics si aquests són poc semblants ($k \leq 0.4$), però no si sí ho són ($0.4 < k < 0.6$).

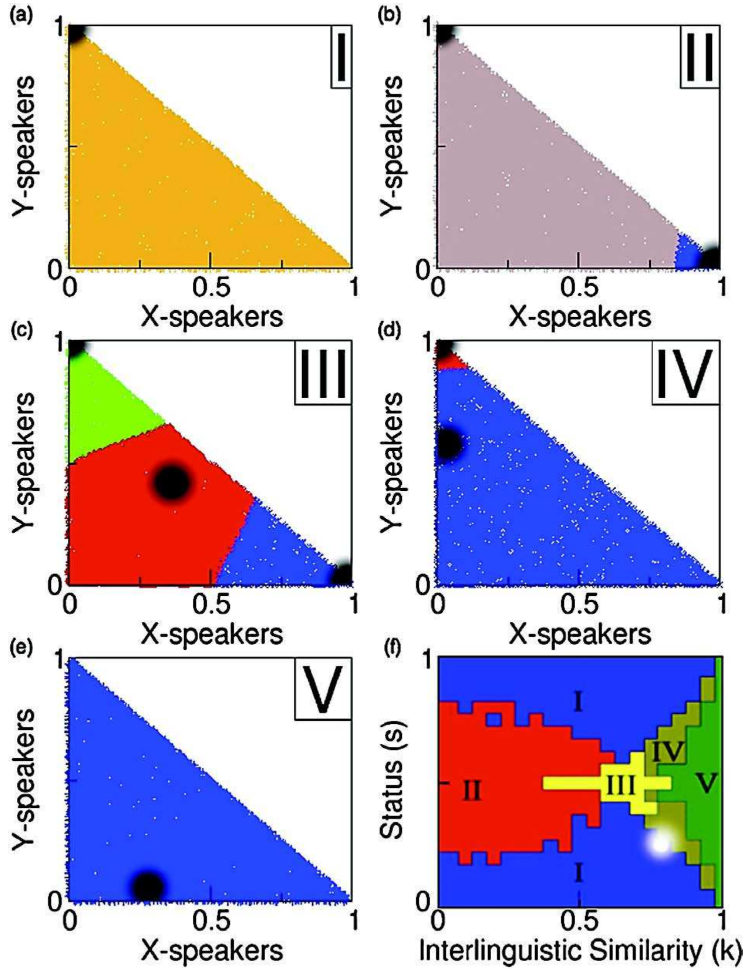


Figura 2: Figura extreta de [13]. Les cinc situacions estables possibles per al model de Mira-Paredes determinades per s i k. A les Figures (a-e) els punts representen les condicions inicials de les simulacions classificades per colors segons l'atractor cap al qual convergeixen. Les condicions inicials estan restringides a $x + y = 1$. Els atractors es representen per grans punts negres. Els valors de k i s determinen el nombre i la posició dels atractors. Les Figures mostren un exemple de cada situació: (a) tipus I, $s = 0.80$ i $k = 0.65$, (b) tipus II, $s = 0.40$ i $k = 0.20$, (c) tipus III, $s = 0.50$ i $k = 0.65$, (d) tipus IV, $s = 0.35$ i $k = 0.75$, (e) tipus V. En (f) es mostra el conjunt de (s, k) que dóna lloc a cada situació. Les parelles de paràmetres que donen lloc a la mateixa distribució topològica es troben pintades del mateix color. El punt blanc correspon als paràmetres k i s que els autors de la Figura han trobat per al sistema gallec-castellà ($s = 0.26, k = 0.80$).

3 Resultats

3.1 Recollida i tractament de dades

Per tal d'estudiar l'evolució de parlants del valencià a través dels models matemàtics prèviament introduïts hem fet ús de dues fonts de dades:

– les dades proporcionades pel Llibre Blanc de l'Ús del Valencià, corresponents a l'Enquesta sobre la Situació Social del Valencià promoguda per l'Acadèmia Valenciana de la Llengua durant l'any 2004 [17]. Més concretament, les dades tractades pel sociolingüista Ernest Querol al seu article *A cada bugada es perden molts llençols: el procés de substitució lingüística en l'àmbit familiar des de l'inici del segle XX: el rei va nu* publicat al Volum II del Llibre Blanc [17]. Les dades mostren l'ús del valencià l'any 2004 classificades per grups d'edat dins de l'àmbit familiar (llengua que l'enquestat parla amb sa mare, son pare, la seua parella i els seus fills).

– les dades proporcionades directament per l'arxiu del Servici d'Investigació i Estudis Sociolingüístics (SIES) a través de la pàgina web de la Conselleria d'Educació, Cultura i Esport de la Generalitat Valenciana [18], corresponents a les enquestes realitzades els anys 1989, 1992, 1995, 2005 i 2010 sota la direcció de Rafael L. Ninyoles. En aquestes, els percentatges d'ús es classifiquen en diferents àmbits: a casa, al carrer, en botigues tradicionals i en grans superfícies.

En aquest treball hem estudiat l'evolució de l'ús del valencià, pel que ens hem centrat en les preguntes *"en quina llengua parla vosté...?"* de les enquestes esmentades. Les possibles respostes que s'oferien són:

– Sempre en valencià

– Generalment en valencià

– Més en valencià que en castellà

– Indistintament

– Més en castellà que en valencià

– Generalment en castellà

– Sempre en castellà

– Altres/no sap, no respon

Depenent del model que vulguem aplicar, classificarem aquestes respostes en dos grans grups: valencianoparlants i castellanoparlants; o en tres, afegint el grup dels bilingües.

Com s'indica a la fitxa tècnica de cada enquesta, els percentatges d'ús es refereixen només a les zones valencianoparlants, que varien lleugerament d'una enquesta a una altra.

L'univers de les mostres correspon a individus de més de 15 o 16 anys, residents a la C.Valenciana l'any de realització de l'enquesta. El marge d'error estadístic de cada resposta depén de la grandària de la mostra (nombre d'enquestes realitzades) i varia entre $\pm 1.2\%$ i $\pm 2.9\%$, com queda indicat a cada fitxa tècnica. Els errors associats als percentatges d'ús finals, una vegada classificats, fetes les mitjanes i normalitzats, s'han calculat seguint la forma habitual de propagació d'errors.

Un dels problemes més rellevants d'aquest treball és la falta de punts i les limitacions que açò implica. Efectivament, sols disposem de cinc o sis punts que tractarem d'ajustar a models matemàtics que depenen de tres o quatre paràmetres. Aquesta situació comporta incerteses grans en els paràmetres a determinar, que tractarem de forma raonada per a cada model en les següents seccions.

3.2 Model d'Abrams-Strogatz

El model d'Abrams-Strogatz considera un conjunt d'individus que parlen o bé una llengua X o bé una llengua Y. Per poder aplicar aquest model a les nostres dades, ha sigut necessari separar els parlants en valencianoparlants i castellanoparlants i considerar-los com el total del grup a estudiar. Hem considerat valencianoparlants aquells inclosos en les tres primeres possibles respostes (parla "sempre valencià", "generalment valencià" o "més valencià que castellà"), i castellanoparlants el cas oposat. Finalment hem renormalitzat els percentatges per excloure els corresponents a "indiferent" o "altres".

El model d'Abrams-Strogatz té tres paràmetres lliures que cal determinar a partir de l'ajust: l'estatus s, el paràmetre de potència a i la constants de normalització del temps c. Atès que sols disposem de cinc o sis punts experimentals, és possible trobar més d'una combinació d'aquests paràmetres que s'adeqüe a les dades de les enquestes. Determinar quina d'aquestes combinacions representa millor la realitat de l'evolució de l'ús del valencià és la part més delicada i requeriria la crítica d'experts en sociolingüística, a més d'un nombre major de punts que corroboraren l'elecció. Nosaltres intentarem trobar les combinacions més raonables i establir uns rangs de variació acceptables dels paràmetres.

3.2.1 Dades del Llibre Blanc (2004)

En primer lloc, hem emprat les dades proporcionades pel sociolingüista Ernest Querol, extretes al seu torn del Llibre Blanc de l'Ús del Valencià [17]. Per poder aplicar el model d'Abrams-Strogatz ha sigut necessari transformar els punts *(edat, fracció de parlants)* per a un temps fixe (any 2004), a punts *(any, fracció de parlants)* per a un grup d'edat determinat. Açò ho hem fet considerant que cada grup d'edat representa l'any en el qual els individus tenien de 20 a 30 anys. Cal fer notar que l'aspecte important d'aquest canvi és que estem observant l'evolució de l'ús del valencià en una franja d'edat concreta. A quina franja d'edat ens referim no és rellevant. Per exemple, podríem haver triat que cada grup d'edat representara l'any en què els individus tenien entre 40 i 50 anys. Açò simplement provoca un desplaçament global dels punts al llarg de l'eix temporal, però els paràmetres que descriuen l'evolució romanen invariants. Les dades resultants es troben a la Taula 1.

Fanja d'edat	15-24	25-34	35-44	45-54	55-64	>64
% sobre el total	15.88	18.53	16.96	13.26	13.25	22.11
Nascuts a	1980/90	1970/80	1960/70	1950/60	1940/50	1930/40
Representen l'any	2010	2000	1990	1980	1970	1960
Valencià (%)	26.1±1.3	30.0±1.3	36.4±1.3	39.9±1.3	48.1±1.3	53.6±1.4
Castellà (%)	73.9±1.7	70.0±1.6	63.6±1.5	60.1±1.4	51.9±1.3	46.4±1.3

Taula 1: Fracció de valencianoparlants i castellanoparlants per edats l'any 2004. Dades extretes del Llibre Blanc de l'Ús del Valencià [17] i recollides pel sociolingüista Ernest Querol. Es mostra el percentatge d'enquestats (% sobre el total) corresponent a cada grup d'edat, l'any en què van nàixer i l'any que representen (quan tenien 20/30 anys), així com el percentatge de valencianoparlants i castellanoparlants.

Comencem fixant $a = 1$, valor per al qual el model d'Abrams-Strogatz té solució analítica (equació (6)). Realitzant un ajust dels punts hem trobat que els valors de l'estatus i la constant de normalitazació que millor descriuen els punts experimentals són, respectivament, $s \simeq 0.38$ i $c \simeq 0.1$. El resultat d'aquest ajust es mostra a la Figura 3.

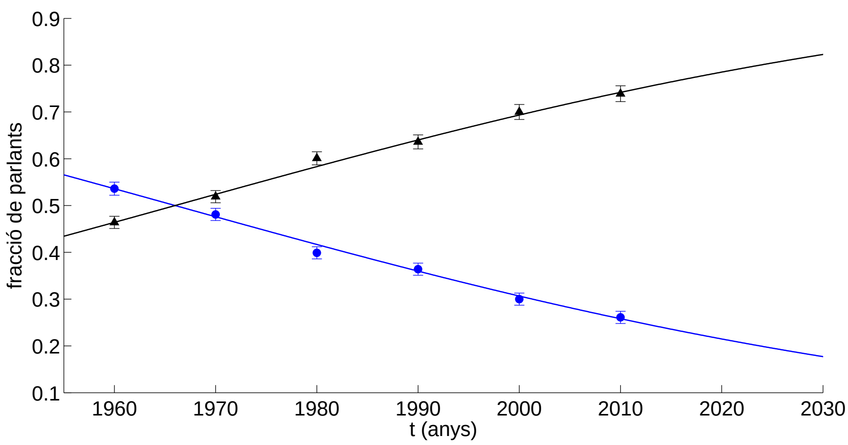


Figura 3: Fracció de parlants front al temps. Dades extretes del Llibre Blanc de l'Ús del Valencià [17] pel sociolingüista Ernest Querol i classificades en dos grups: valencianoparlants (punts blaus) i castellanoparlants (triangles negres). Les corbes corresponen a la solució analítica del model de d'Abrams-Strogatz per a $a = 1$, $s = 0.38$, i $c = 0.1$.

Tanmateix, també podem trobar de forma analítica (cas $a = 1$) altres parelles (s, c) que s'ajusten als punts si fixem diferents valors de c i obtenim els valors de l'estatus s que millor s'adeqüen. El resultat es troba resumit a la Taula 2, i representat a la Figura 4.

c	0.04	0.06	0.08	0.10	0.12	0.2	0.3
s	0.20	0.30	0.35	0.38	0.40	0.44	0.46

c	0.4	0.5	0.6	0.7	0.8	0.9	1
s	0.470	0.476	0.480	0.483	0.485	0.486	0.488

Taula 2: Valors dels paràmetres (c, s) que millor ajusten les fraccions d'ús del valencià de la Taula 1 a la solució analítica del model d'Abrams-Strogatz per a $a = 1$

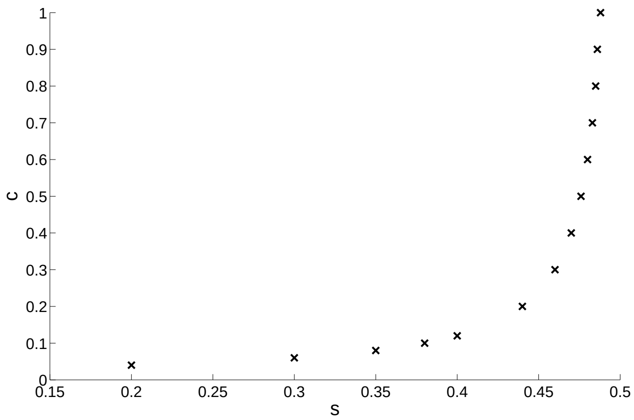


Figura 4: Paràmetre de l'estatus s front al paràmetre de normalització c que ajusten als punts experimentals de la Taula 1 per a la solució analítica $a = 1$. Els valors es troben recollits a la Taula 2

Observem que existeix una asímptota en $s = 0.5$. Açò és degut al fet que, per a aquest valor, x roman constant, com hem explicat a l'apartat Fonament Teòric. Per tant, és impossible ajustar punts de tendència decreixent, siga quin siga el valor de c. Aquesta asímptota implica que, segons anem augmentant c, es necessita una precisió major a l'hora de determinar s per obtindre un ajust adequat. Tanmateix, exigir un valor de l'estatus de la llengua amb tanta precisió no és coherent. Per una part, pel fet que les incerteses en els paràmetres han de ser necessàriament grans com a resultat de l'escassetat de punts i, per tant, no podem exigir una precisió alta en cap d'aquests. Per altra part, pel significat intrínsec del paràmetre de l'estatus: el model hauria de ser capaç de permetre que aquest es trobara en un cert rang sense que la resposta del sistema canviara de forma substancial, ja que es tracta d'un paràmetre difícil de quantificar. Aquest raonament ens permet reduir el rang de c a $c \in [0.04, 0.2]$. A més a més, existeix també una altra raó per excloure valors de c majors a 0.3: per a $c > 0.3$, els valors de l'estatus satisfan $s > 0.45$, una situació molt pròxima a $s = 0.5$ en la qual les dues llengües en competència serien socialment equivalents. Segons el nostre punt de vista, aquesta no és la situació real de la interacció valencià-castellà, pel que el model no estaria descrivint correctament la realitat.

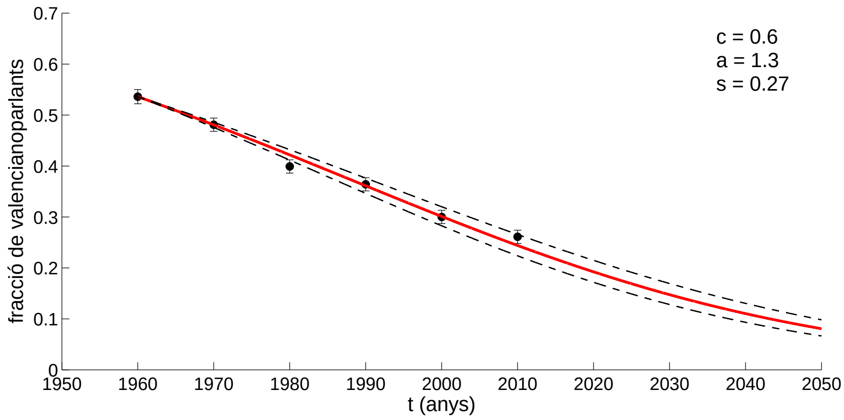
Per a valors de $a \neq 1$ és necessari l'estudi numèric de l'equació d'Abrams-Strogatz (3). Hem obtingut la solució numèrica d'aquesta emprant el mètode Ruge-Kutta 4 (RK4) partint de la condició inicial corresponent a l'any 1960 (grup d'edat >64 anys). Així ha sigut possible no sols trobar altres conjunts dels paràmetres que s'adeqüen als punts sinó també estudiar la dependència de la resposta del sistema amb aquests. En l'Apèndix s'inclou l'explicació del mètode RK4 per a la resolució numèrica d'equacions diferencials.

En primer lloc, hem fixat $a = 1.3$, corresponent el valor obtingut per Abrams i Strogatz ($a = 1.3 \pm 0.25$) [6]. Variant de forma sistemàtica c per a $c < 0.3$ hem buscat els valors de s que millor s'ajustaven. Per a valors de c fora del rang $c \in [0.04, 0.12]$ no s'han trobat ajusts adequats (valors més grans de c prediuen disminucions massa ràpides, mentre que valors de c més menuts prediuen disminucions massa lentes), el que corrobora el raonament realitzat anteriorment sobre el rang de c. Una vegada trobades les parelles de (c, s) més apropiades, hem variat a en el rang $a \in [1, 1.5]$ per contemplar la desviació estàndard del valor de a. Així, hem seleccionat els valors de a que millor s'ajustaven als punts i hem obtingut finalment els diferents conjunts (a, c, s) que descriuen les dades experimentals dins dels rangs que hem comentat. Aquestes ternes es mostren a la Taula 3.

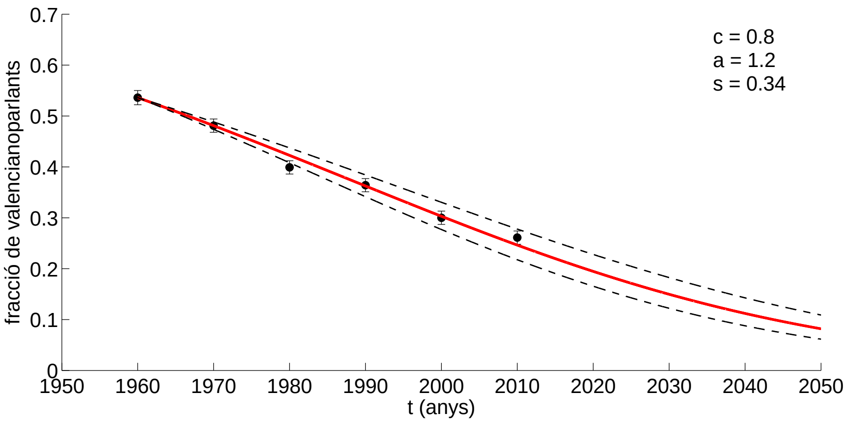
c	0.04	0.06	0.08	0.1	0.12
a	1.3	1.3	1.2	1.1	1
s	0.15	0.27	0.34	0.38	0.40

Taula 3: Conjunt de paràmetres (c, a, s) que millor ajusten el model d'Abrams-Strogatz als punts de l'ús del valencià extrets del Llibre Blanc (2004) de la Taula 1.

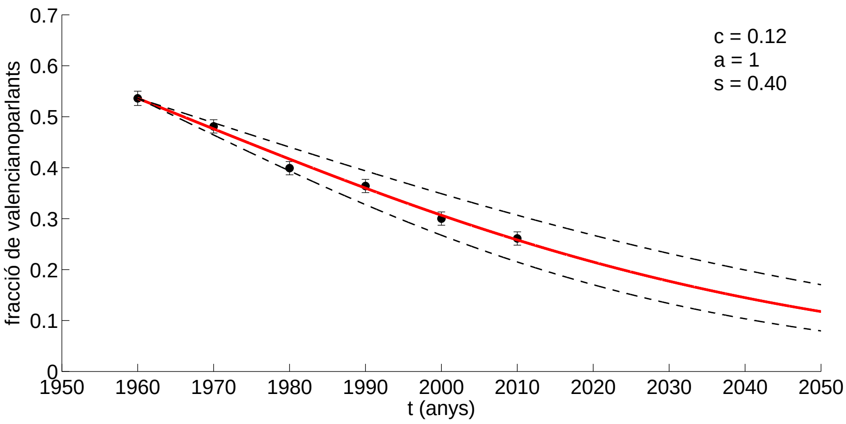
Finalment, per estudiar la dependència del sistema amb canvis de l'estatus, hem fet variar el valor que hem trobat de s en ± 0.02. Hem observat que aquestes variacions són majors quan major és c, el que indica que valors de c alts requereixen major exactitud en la determinació de l'estatus. La Figura 5 mostra tres exemples d'aquests ajusts.



(a) $c = 0.06,\ a = 1.3,\ s = 0.27$



(b) $c = 0.08,\ a = 1.2,\ s = 0.34$



(c) $c = 0.12,\ a = 1,\ s = 0.40$

Figura 5: Fracció de valencianoparlants front al temps. Ajusts dels punts de les enquestes de la Taula 1 al model d'Abrams-Strogatz per a diferents ternes dels paràmetres (c, a, s). La corba roja sòlida representa el valor de l'estatus s que millor s'hi ajusta. Les corbes discontínues negres corresponen a ± 0.02 aquest valor.

3.2.2 Dades del SIES

Passem ara a considerar les dades obtingudes de la base d'enquestes del SIES [18]. La classificació en dos grups, valencianoparlants i castellanoparlants, s'ha realitzat de forma anàloga a la realitzada sobre les dades de Querol. En aquest cas, a més, s'ha fet una mitjana per als àmbits en què es classificaven les respostes: llengua que l'enquestat parla a casa, al carrer, en botigues tradicionals i en grans superfícies. Els valors resultant es troben recopilats a la Taula 4 i representats a la Figura 6.

S'observa clarament una tendència desigual entre els tres primers punts, que semblen romandre constants o augmentar lleugerament, i els punts del 2005 i 2010, que presenten una reducció notable en la fracció de valencianoparlants. Aquest canvi de tendència i la falta de dades entre 1995 i 2005, sumat al fet de tindre sols cinc punts experimentals, fan molt difícil poder modelitzar l'evolució de l'ús de la llengua. No hem trobat cap conjunt de paràmetres del model d'Abrams-Strogatz que s'adeqüe als punts.

Any	Valencianoparlants (%)	Castellanoparlants (%)
1989	44.0 ± 1.4	56.0 ± 1.5
1992	45.2 ± 1.3	54.8 ± 1.4
1995	45.2 ± 3.0	54.8 ± 3.2
2005	31.8 ± 1.2	68.2 ± 1.5
2010	28.8 ± 1.1	71.2 ± 1.5

Taula 4: Percentatge de valencianoparlants i castellanoparlants a les regions valencianoparlants de la C. Valenciana. Resultats obtinguts a partir de les dades de les enquestes del SIES realitzades els anys indicats [18].

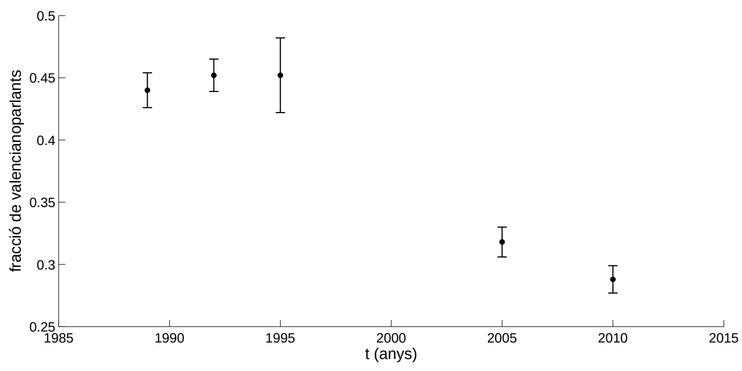


Figura 6: Fracció de valencianoparlants front al temps. Dades extretes de les enquestes realitzades pel SIES [18] i recopilades a la Taula 4.

Podem obtindre un nou punt si considerem els percentatges de l'enquesta de 2004 (Taula 1) i fem la mitjana de les fraccions de parlants de cada grup d'edat pesada pel percentatge d'enquestats que pertanyen a cada grup. Com que aquestes dades es refereixen sols a l'àmbit familiar, hem considerat sols l'àmbit de l'ús del valencià "a casa" de la resta de les enquestes del SIES. En aquest àmbit, els percentatges d'ús del valencià són lleugerament majors i l'error dels punts augmenta. A la Taula 5 es mostren les dades amb els corresponents errors.

Any	Valencianoparlants (%)	Castellanoparlants (%)
1989	48 ± 3	52 ± 3
1992	53 ± 3	47 ± 3
1995	51 ± 7	49 ± 7
2004	39 ± 5	61 ± 6
2005	40 ± 3	60 ± 3
2010	35 ± 3	65 ± 3

Taula 5: Percentatge d'ús del valencià/castellà "a casa" a les regions valencianoparlants de la C. Valenciana, obtinguts a partir de les dades de les enquestes del SIES [18], exceptuant els percentatges de l'any 2004, que s'han obtingut del Llibre Blanc de l'Ús del Valencià [17].

Per a aquests nous valors, restringits a l'àmbit familiar, observem que la tendència decreixent comença a partir de 1992. Prendrem aquest punt com a condició inicial, ja que el model d'Abrams-Strogatz no és capaç de descriure l'augment que es dóna entre 1989 i 1992. Ens trobem novament front a un ajust de solament cinc punts amb errors considerables i tres paràmetres a determinar, per la qual cosa és possible obtindre diferents conjunts de valors que s'adeqüen a la tendència correctament.

De nou, hem fet un estudi dels valors (c, s) per al cas amb solució analítica, $a = 1$. La Figura 7 mostra la mateixa tendència que en l'apartat anterior. Tanmateix, donat que ara la disminució de valencianoparlants és més ràpida, per als mateixos valors de l'estatus es requereix un valor de c major (recordem que un augment de c accelera el canvi entre llengües ja que aquest paràmetre normalitza la probabilitat de transició entre aquestes). Tornem a restingir el valor de c a $c \leq 0.3$ ja que, per a valors superiors l'estatus és poc realista ($s > 0.45$) i requereix una precisió massa alta.

Procedint altra vegada de forma sistemàtica hem estudiat el valor de l'estatus per a $c \in [0.06, 0.3]$ en els casos en què la probabilitat de transició entre llengües és lineal ($a = 1$), quadràtica ($a = 2$) i els casos intermedis $a = 1.3$ (valor d'Abrams-Strogatz) i $a = 1.5$. Els valors obtinguts es trobem recollits a la Taula 6.

A partir de la Taula 6 observem que per als valors més grans de c, l'estatus no varia pràcticament quan variem el paràmetre de potència a. Per altra part, numèricament hem

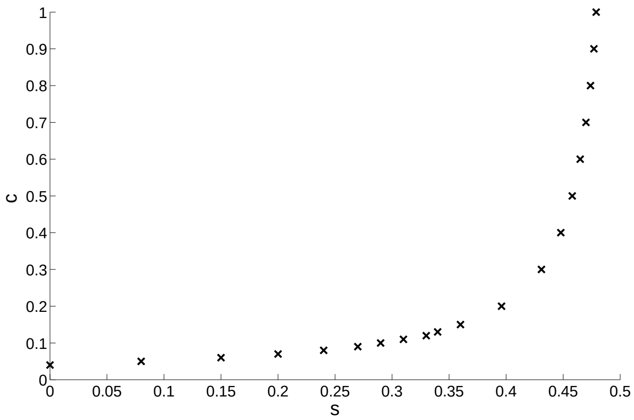


Figura 7: Paràmetre de l'estatus s front al paràmetre de normalització c que ajusten als punts experimentals de la Taula 5 per a la solució analítica $a = 1$.

	a=1	a=1.3	a=1.5	a=2
c=0.06	0.15	0.10	0.05	-
c=0.08	0.24	0.20	0.16	0.04
c=0.1	0.29	0.26	0.23	0.13
c=0.12	0.33	0.30	0.28	0.20
c=0.15	0.36	0.34	0.33	0.33
c=0.2	0.40	0.39	0.38	0.34
c=0.3	0.43	0.43	0.42	0.40

Taula 6: Conjunts de paràmetres (c, a, s) que ajusten els punts de la Taula 5 al model d'Abrams-Strogatz. Els valors centrals corresponen a l'estatus, s.

observat que com major és c, major és la precisió necessària de l'estatus per descriure correctament els punts, com ja hem vist que ocorria en l'apartat anterior. També hem notat que per a valors alts de c, $a = 2$ comença a no ajustar correctament les dades, pel que intuïm que $a < 2$. La Figura 8 mostra tres exemples per a diferents ternes de paràmetres. A la figura s'observa que la disminució de parlants és més acusada quan més gran és c i que la tendència decreixent es suavitza per a valors de a baixos.

Donat l'excés de graus de llibertat, la mancança de dades i els errors dels punts de què disposem, no podem concloure quin d'entre tots els conjunts de la Taula 6 és el més adequat per al nostre sistema. Per una part, caldria fer una anàlisi de caire sociolingüístic per tal de reduir el rang de variació del paràmetre de l'estatus del valencià, de forma que s'adequara a la realitat de la nostra llengua. Per altra part, necessitaríem més dades experimentals per saber quina és la rapidesa de disminució de l'ús del valencià i afinar els valors de a i c. Tanmateix, queda patent que si la tendència es manté durant els pròxims anys, en

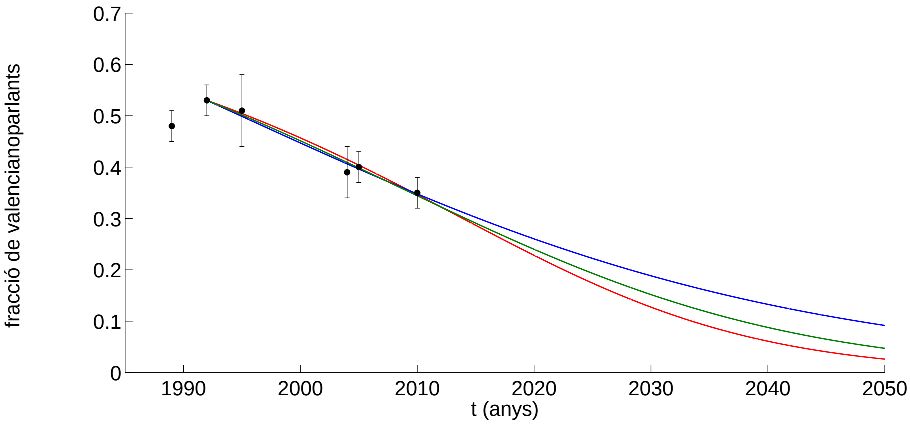


Figura 8: Fracció de valencianoparlants front al temps, dades extretes de la Taula 5. Les corbes corresponen a la solució numèrica de l'equació d'Abrams-Strogatz prenent com a condició inicial la fracció de parlants a $t_0 = 1992$ i els paràmetres: $c = 0.08$, $a = 1$, $s = 0.24$ (blau); $c = 0.12$, $a = 1.3$, $s = 0.30$ (verd); $c = 0.15$, $a = 1.5$, $s = 0.33$ (roig).

l'any 2050 el percentatge de valencianoparlants serà inferior al 10% segons la predicció del model d'Abrams-Strogatz. Aquesta tendència podria ser modificada si es porten a terme polítiques lingüístiques que augmenten l'estatus del valencià i permeten l'estabilització del seu ús.

3.2.3 Comparació de dades

Per finalitzar aquesta secció volem fer notar que, tot i que en les dues seccions anteriors hem aplicat el model d'Abrams-Strogatz per descriure l'evolució de l'ús del valencià, les dades de cada apartat no són comparables. És a dir, no podem emprar les dades de la Taula 1 conjuntament amb les de la Taula 4. El motiu principal és que les primeres (extretes del Llibre Blanc de l'Ús del Valencià) descriuen com varia l'ús del valencià en el temps dins d'un determinat grup d'edat que prenem com a representatiu de l'evolució del conjunt (que nosaltres hem agafat de 20-30 anys), mentre que les segones (extretes del SIES) consideren tots els grups d'edat a l'hora d'estudiar l'evolució del nombre de valencianoparlants. Tanmateix, esperaríem obtindre uns paràmetres semblants o, al menys, compatibles. Efectivament, si comparem els resultats podem dir que, segons els nostres ajusts, els paràmetre es troben en uns rangs aproximats de:

$$c \in [0.06, 0.2]$$

$$a \in [1, 1.5]$$

$$s \in [0.15, 0.4]$$

3.3 Model de Paredes-Mira

El model de Paredes i Mira introdueix un nou grup possible, les persones bilingües. Per fer ús d'aquest model hem dividit de forma diferent els percentatges d'ús: hem considerat valencianoparlants sols aquells que han respost que parlen "sempre valencià" o "generalment valencià", bilingües els que parlen "indistintament", "més valencià que castellà" i viceversa, i castellanoparlants els que parlen "sempre castellà" o "generalment castellà". La Taula 7 mostra els percentatges obtinguts una vegada tractades les dades de les enquestes del SIES. Les dades mostrades són la mitjana de tots els àmbits amb els errors corresponents. Per a aquest model no podem emprar les dades de 2004 del Llibre Blanc, perquè els percentatges de l'enquesta corresponen a la llengua que es parla amb una persona concreta: la mare, el pare, la parella o els fills. Com que la situació habitual és parlar o bé el valencià o bé el castellà amb una determinada persona i no ambdues llengües alhora, no té sentit considerar un grup de bilingües per a aquestes respostes.

Any	Valencianoparlants (%)	Bilingües (%)	Castellanoparlants (%)
1989	37.2±1.1	14.9±1.1	48.0±1.2
1992	34.6±0.9	21.4±1.0	44.0±1.0
1995	36.8±1.9	18.4±2.3	44.7±1.9
2005	25.9±0.9	16.8±1.0	57.3±1.2
2010	23.0±0.8	16.6±1.0	60.4±1.2

Taula 7: Precentatge de valencianoparlants, castellanoparlants i bilingües. Dades tractades a partir de les enquestes del SIES realitzades els anys indicats [18].

Hem aplicat el model de Mira-Paredes (sistema d'equacions (14)), a les fraccions de parlants de cada grup (Taula 7), prenent x com la fracció de parlants del valencià, y com la fracció de parlants del castellà, i b com la fracció de bilingües obtinguda a partir de $b = 1 - x - y$. Com que es tracta d'un sistema de dues d'equacions diferencials ordinàries acoblades sense solució analítica, hem calculat la solució de forma numèrica utilitzant el mètode RK4 per a sistemes d'equacions (vegeu l'Apèndix) i prenent com a condició inicial els valors per a l'enquesta més antiga, la de 1989.

El model de Mira i Paredes té quatre paràmetres a determinar, els corresponents al model d'Abrams-Strogatz (c, a, s) més el paràmetre de semblança k. El càlcul de la semblança entre llengües és extremadament difícil en casos pràctics, com bé senyalen Mira i Paredes al seu article [12], i el seu valor s'ha d'estimar a partir dels ajusts a les dades.

Donada la dificultat tècnica de realitzar un ajust amb quatre paràmetres lliures als cinc valors empírics de què disposem, hem partit dels valors dels paràmetres del model d'Abrams-Strogatz que havíem trobat en les seccions anteriors i hem tractat de trobar el paràmetre de semblança que millor s'adeqüe. A més, hem analitzat la dependència entre els paràmetres i les corbes que resulten de l'estudi numèric.

Fixant $a = 1.3$ i $c = 0.1$, hem obtingut un valor de l'estatus de $s = 0.35$ i de semblança $k = 0.35$. Amb aquests valors de l'estatus i la semblança, ens trobaríem en el règim II d'estabilitat en la Figura 2 de Mira, Seoane i Nieto, que prediu que una de les dues llengües acaba desapareixent, en el nostre cas el valencià, i que el grup de bilingües no és estable i també desapareix. La Figura 9 mostra els punts i l'ajust per a aquests valors. Destaquem que el valor de l'estatus s es troba dins del rang que havíem obtingut en aplicar el model d'Abrams-Strogatz, pel que podria ser una estimació raonable. Tanmateix, notem que el valor del paràmetre de semblança k és més baix del que cabria esperar (Mira i Paredes obtenen $k = 0.8$ en el cas del gallec-castellà, per exemple), i que l'ajust no acaba de caracteritzar correctament els punts de castellanoparlants i bilingües per a l'any 1992 i 1995. Conseqüentment, no podem concloure amb certesa que el model o els valors dels paràmetres siguen els correctes.

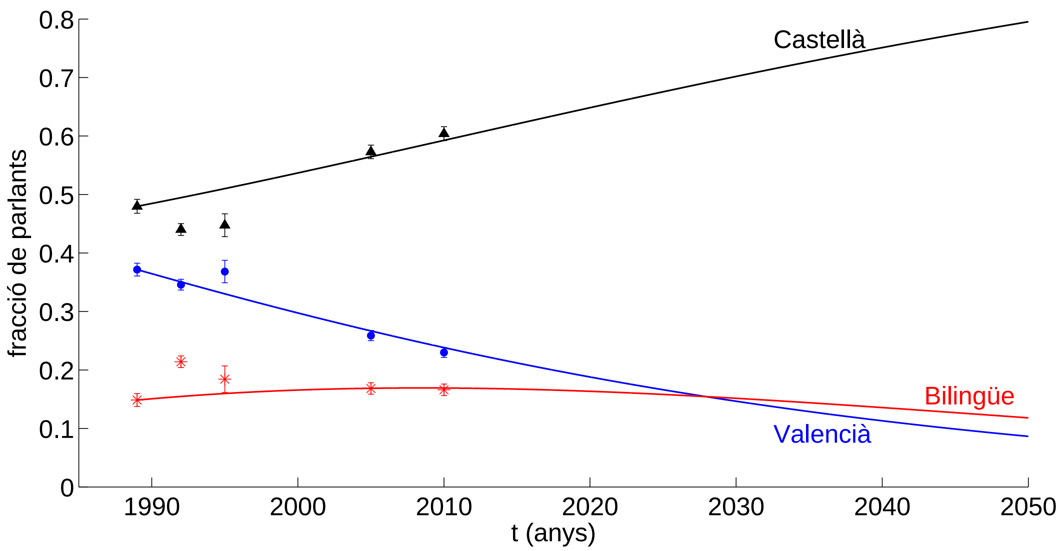


Figura 9: Fracció de parlants front al temps. Dades extretes de les enquestes del SIES [18] i classificades en tres grups: valencianoparlants (punts blaus), castellanoparlants (triangles negres) i bilingües (estreles roges). Les corbes corresponen a l'ajust del model de Mira i Paredes per als paràmetres $s = 0.35$, $k = 0.35$, $a = 1.3$ i $c = 0.1$.

Per altra part, per tal d'estudiar com varien les corbes d'evolució amb els paràmetres, hem realitzat modificacions sobre aquests i hem observat que:

- Es requereix un valor de c baix, de l'ordre de $c \sim 0.1$, per suavitzar la disminució de l'ús del valencià. Per a valors majors, la fracció de valencianoparlants cau massa ràpidament.
- La modificació del paràmetre de semblança ens mostra que: valors de k alts fan augmentar la fracció de bilingües, en detriment principalment de la llengua dominant (castellà); valors baixos de k fan augmentar la fracció de castellanoparlants i, en

menor mesura, també de valencianoparlants.

– Per a estatus majors, la caiguda del valencià és més lenta, mentre que per a estatus menors, la caiguda és més ràpida.

– Si $c \sim 0.1$, canvis del paràmetre de potència a entre 1 i 2 no afecten pràcticament a la corba.

4 Conclusions

La diversitat lingüística és un fenomen complex que depén de múltiples factors. El seu equilibri és sempre precari donat que es tendeix a l'economia del llenguatge. Tanmateix, defensar la riquesa lingüística és defensar diferents formes de veure el món i preservar el patrimoni cultural que representen les llengües. Des del camp de la física, s'han desenvolupat recentment models matemàtics que descriuen aquests tipus de situacions sociolingüístiques i de competència entre llengües. Aquests models permeten predir situacions de futur a què es pot arribar si no es canvien els patrons vigents i es prenen mesures polítiques i socials adients per tal de reconduir la situació.

En aquest treball hem estudiat dos d'aquests models: el model d'Abrams-Strogatz i el model de Mira-Paredes. Ambdós s'engloben dins del grup de models que enfoquen el problema de la competició lingüística fent ús d'equacions diferencials que descriuen el comportament global del sistema. En primer lloc, hem analitzat els paràmetres i les suposicions que es fan en cada model, així com les seues solucions estables. Posteriorment hem tractat d'aplicar-los al cas de la interacció entre el valencià i el castellà fent ús de les dades de les enquestes realitzades pel SIES i l'AVL als territoris valencianoparlants de la C.Valenciana.

El model d'Abrams-Strogatz descriu la competició entre dues llengües X i Y, en el qual l'*atractiu* d'una llengua augmenta amb el nombre de parlants i l'*estatus* percebut. Tanmateix, el model calcula la transició d'individus d'un grup a un altre sense fase intermèdia i sense la possibilitat de grups bilingües. Evidentment, les societats bilingües existeixen i la situació valencià-castellà n'és un exemple, per la qual cosa aquest model no pot descriure de forma apropiada la realitat de la nostra situació sociolingüística. Ens ha servit, però, per donar-nos una idea del rang de variació dels paràmetres que hem emprat posteriorment en el model de Mira-Paredes. Mira i Paredes plantegen un model més general per tal de refinar el d'Abrams-Strogatz. Aquests investigadors incorporen al seu model un nou grup d'individus, els parlants bilingües, i introdueixen un nou paràmetre que representa la probabilitat que un parlant monolingüe no passe a l'altra llengua sinó que adopte una pràctica bilingüe. Els autors identifiquen aquest paràmetre amb la *semblança* entre les dues llengües.

El problema fonamental amb què ens hem trobat durant l'estudi ha sigut la mancança de dades. Com que sols disposàvem de cinc o sis punts experimentals, no hem pogut trobar valors numèrics concrets dels paràmetres sinó més bé rangs dins dels quals aquests poden variar. Així, l'aplicació dels models ens ha permés quantificar, tot i que amb incerteses considerables, les variables que modelen la interrelació entre el valencià i el castellà i trobar conjunts que descriuen apropiadament les dades de les enquestes. Aplicant el model d'Abrams-Strogatz, hem observat que la transició d'una llengua a una altra està frenada pel paràmetre de normalització temporal, c, que es troba entre 0.04 i 0.3 aproximadament. Així mateix, hem comprovat que el paràmetre de potència a ha de ser inferior a 2 i hem trobat que els millors ajusts corresponen a $a \in [1, 1.5]$ aproximadament, cosa que coincideix amb el resultat presentat per Abrams-Strogatz [6]. Finalment, pel que fa al valor de l'estatus de la llengua hem considerat adequat reduir el seu rang de variació a $s \in [0.15, 0.4]$, ja que, per a valors superiors es requeria una precisió poc versemblant en la determinació d'aquest paràmetre. Així mateix, tot i tractar-se d'una afirmació subjectiva, considerem que valors fora d'aquest rang no descriurien de forma realista la situació del valencià.

Partint dels rangs establerts amb el model d'Abrams-Strogatz, hem aplicat el model de Mira-Paredes a les dades de què disposem. Hem obtingut un valor de l'estatus $s = 0.35$, coherent amb els resultats extrets del model d'Abrams-Strogatz, i un valor de semblança $k = 0.35$ que sospitem és massa baix. Donat els pocs punts i el fet que les dades no acaben d'ajustar-se bé al model, creiem necessari un estudi més profund i amb més dades per poder arribar a afirmacions concloents sobre la validesa del model i dels valors dels paràmetres.

És obvi que els models presenten una sèrie de limitacions. En primer lloc, no hem d'oblidar que els models són sempre una simplificació a sols uns pocs factors d'una situació complexa, com és la interacció entre parlants. Per altra part, els models estudiats tampoc tenen en compte que els paràmetres, com l'estatus de la llengua, poden variar al llarg del temps. Així mateix, tampoc consideren possibles variacions de la proporció relativa de parlants degut a immigració, emigració, etc.

Com a futur treball, proposem la millora dels models per tractar de superar aquestes limitacions i remarquem la necessitat de disposar de dades fiables i comparables i, a ser possible, extretes d'enquestes dissenyades per a ser aplicades als models. Així mateix, també trobem interessant promoure el desenvolupament d'altres tipus de models, com els basats en xarxes complexes i en les interaccions individuals dels parlants (*agent-based models*). Per últim, volem fem notar de nou la necessitat d'una col·laboració oberta i constructiva amb sociolingüistes per tal d'afinar els models a la realitat social i reflexionar sobre el significat dels seus paràmetres (*estatus*, *semblança*, etc.).

Conclusions

Linguistic diversity is a complex phenomenon that depends on many factors. His equilibrium is always precarious as it is common to tend to the economy of language. However, defending the linguistic diversity is to defend different ways of seeing the world and to preserve the cultural heritage that languages represent. From the field of Physics, mathematical methods have been recently developed to describe this type of sociolinguistic situations of competition between languages. These models allow predicting future situations to which we can arrive if current patterns are not changed and appropriate measures and social policies are not taken in order to redirect the situation.

In this work we have studied two of these models: the model of Abrams-Strogatz and the model of Mira-Paredes. Both are included within the group of models that focus on the problem of linguistic competition using differential equations that describe the behavior of the overall system. First, we have analyzed the parameters and assumptions made in each model, as well as its stable solutions. Then we have tried to apply them to the case of the interaction between the Catalan and Spanish using data from surveys conducted by the SIES and the AVL over Valencian-speaking territories of C.Valenciana.

The Abrams-Strogatz model describes the competition between two languages, X and Y, in which the *attractiveness* of a language increases with the number of speakers and the *perceived status*. However, the model calculates the transition of individuals from one group to another without intermediate phase and without the possibility of bilingual groups. Obviously, bilingual societies exist and the situation of Spanish-Valencian is an example, so this model cannot appropriately describe the reality of our sociolinguistic situation. However, it has been useful to give us an idea of the range of variation of the parameters that we have used later in the Mira-Paredes model. Mira and Paredes propose a more general model in order to refine the Abrams-Strogatz one. These researchers include in their model a new group of individuals, the bilingual speakers, and they introduce a new parameter that represents the probability that a monolingual speaker do not change to the other language but adopts a bilingual practice. The authors identify this parameter with the *similarity* between the two languages.

The fundamental problem that we have found during the study has been the lack of data. Since we had only five or six experimental points of data, we have not been able to find specific numerical values of the parameters but ranges within which they can vary. Thus, the application of the models has allowed us to quantify the variables that shape the relationship between Valencian and Spanish, although with considerable uncertainties, and find sets that describe properly the survey data. Applying the Abrams-Strogatz model, we have found that the transition from one language to another is constrained by the time-scale normalization parameter, c, which lies between 0.04 and 0.3 approximately. Furthermore, we have found that the power parameter a must be smaller than 2 and that the best fittings corresponds to $a \in [1, 1.5]$ approximately, which agrees with the results presented by Abrams -Strogatz [6]. Finally, with regard to the value of the status of the language we have considered appropriate to reduce its range of variation to $s \in [0.15, 0.4]$, because for higher values an unlikely precision is required in the determination of this parameter. In addition, despite being a subjective statement, we believe that values outside this range do not describe realistically the situation of Valencian.

Based on the ranges established using the Abrams-Strogatz model, we have applied the Mira-Paredes model to the available data. We have obtained a value of the status $s = 0.35$, which is consistent with the results deduced from the Abrams-Strogatz model, and a value of similarity $k = 0.35$ which we suspect that is too low. Given the lack of data and that it does not fit completely the model, we think that a deeper study and more data are needed to reach conclusive statements about the validity of the model and the values found for the parameters.

It is obvious that the models have several limitations. First, we should not forget that the models are always a simplification to only a few factors of a complex situation, such as the interaction between speakers. Moreover, the models do not take into account that parameters such as the status of the language may vary over time. Likewise, they do not consider possible changes in the relative proportion of speakers due to immigration, emigration, etc.

As future work, we propose the improvement of the models to try to overcome these limitations and we emphasize the need for reliable and comparable data, drawn from surveys designed to be applied to the models if possible. Besides, we also find interesting to promote the development of other types of models, such as those based on complex networks and on the interaction of individual speakers (*agent-based models*). Finally, we want to reiterate the need for an open and constructive collaboration with sociologists to adjust the models to the social reality and reflect on the meaning of its parameters (*status*, *similarity*, etc.).

5 Bibliografia

[1] J. Tuson, *El luxe del llenguatge*, 12a edició. Empúries, Barcelona (1997)

[2] E. Amorrortu, A. Barreña, I. Idiazabal i E. Izaguirre. *Informe sobre les llengües del món. Síntesi*, UNESCO Etxea/ Centre UNESCO de Catalunya/Angle Editorial, Barcelona, (2005)

[3] D. Crystal, *Language Death.* Cambridge University Press, Cambridge (2000)

[4] J. Tuson, *Una imatge no val més que mil paraules*, 3a edició. Empúries, Barcelona (2004)

[5] J. Tuson, *Patrimoni natural*, 10a edició. Empúries, Barcelona (2005)

[6] D. M. Abrams and S. H. Strogatz, *Nature* **424** (2003) 900

[7] J. Kabatek, *Estudos de Lingüística Galega* **4** (2012) 27-43

[8] D. Stauffer, X. Castelló, V. M. Enguíluz and M. San Miguel, *Physica A* **374** (2007) 835-842

[9] D. Stauffer, S. Moss de Oliveira, P. M. C. de Oliveira and J. S. Sa Martins, *Biology, Sociology, Geology by Computational Physicists.* Elsevier, Amsterdam, (2006) pp. 151-177

[10] M. Patriarca and T. Leppänen, *Physica A* **338** (2004) 296-299

[11] J. P. Pinasco and L. Romanelli, *Physica A* **361** (2006) 355-360

[12] J. Mira and A. Paredes, *Europhys. Lett* **69** (2005) 1031-4

[13] J. Mira, L. F. Seoane, and J. J. Nieto *New Journal of Physics* **13** (2011) 033007

[14] X. Castelló, V. M. Enguíluz and M. San Miguel, *New Journal of Physics* **8** (2006) 308

[15] X. Castelló, L. Loureuro-Porto, V. M. Enguíluz and M. San Miguel, *The fate of bilingualism in a model of language competition.* Advancing Social Simulation: The First World Congress. Springer (2007) pp. 83-94

[16] C. Schulze and D. Stauffer, *Computing in Science & Engineering* **8** (2006) 60-67

[17] *Llibre Blanc de l'Ús del Valencià. Enquesta sobre la situació social del valencià. 2004.* Publicacions de l'Acadèmia Valenciana de la Llengua, València (2005).

[18] Servici d'Investigació i Estudis Sociolingüístics (SIES). Conselleria d'Educació, Cultura i Esports. Generalitat Valenciana. Fons estadístic i documental. http://www.cece.gva.es/polin/val/sies/sies_fonum.htm

6 Apèndix

6.1 Mètode Runge-Kutta 4 (RK4)

Existeixen diferents mètodes per integrar de forma numèrica sistemes d'equacions diferencials ordinàries donat un valor inicial. El mètode Runge-Kutta 4 (RK4) n'és un d'ells. Considerem l'equació diferencial:

$$\frac{dx}{dt} = f(t, x(t)) \tag{15}$$

amb condició inicial $x(t_0) = x_0$.

L'estratègia més simple d'integrar aquest tipus d'equació diferencial és emprar l'expansió de Taylor:

$$x(t+h) = x(t) + f(t,x)\ h + \mathcal{O}(h^2) \tag{16}$$

Així, si coneixem el valor de x en t podem utilitzar un pas finit h per calcular el següent valor de x en $t+h$ amb una precisió de $\mathcal{O}(h^2)$. En termes del límit discretitzat de la funció, podem escriure:

$$x_{n+1} = x_n + f(t_n, x_n)\ h \tag{17}$$

on hem utilitzat el subíndex n per fer referència a la funció x en el pas $t_n = t_0 + nh$. Aquesta primera aproximació s'anomena **mètode d'Euler**.

Els mètodes de Runge-Kutta milloren el mètode d'Euler utilitzant més d'un terme de la sèrie de Taylor per calcular la variació en la variable x. Ens centrarem en el mètode RK4, que utilitza quatre gradients per calcular-ne un valor mitjà i té una precisió d'ordre $\mathcal{O}(h^4)$. L'algoritme d'aquest mètode ve donat per:

$$x_{n+1} = x_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k4) \tag{18}$$

on

$$\begin{aligned}
k_1 &= h\ f(t_n, x_n)\\
k_2 &= h\ f(t_n + \frac{1}{2}h, x_n + \frac{1}{2}k_1)\\
k_3 &= h\ f(t_n + \frac{1}{2}h, x_n + \frac{1}{2}k_2)\\
k_4 &= h\ f(t_n + h, x_n + k_3)
\end{aligned}$$

Considerem ara un sistema de dues equacions diferencials:

$$\frac{dx}{dt} = f(t, x(t), y(t)) \tag{19a}$$

$$\frac{dy}{dt} = g(t, x(t), y(t)) \tag{19b}$$

amb la condició inicial $(x(t_0), y(t_0)) = (x_0, y_0)$. L'algoritme RK4 es generalitza fàcilment:

$$x_{n+1} = x_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{20a}$$

$$y_{n+1} = y_n + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4) \tag{20b}$$

on

$$\begin{aligned}
k_1 &= h\ f(t_n, x_n, y_n) & l_1 &= h\ g(t_n, x_n, y_n) \\
k_2 &= h\ f(t_n + \tfrac{1}{2}h, x_n + \tfrac{1}{2}k_1, y_n + \tfrac{1}{2}l_1) & l_2 &= h\ g(t_n + \tfrac{1}{2}h, x_n + \tfrac{1}{2}k_1, y_n + \tfrac{1}{2}l_1) \\
k_3 &= h\ f(t_n + \tfrac{1}{2}h, x_n + \tfrac{1}{2}k_2, y_n + \tfrac{1}{2}l_2) & l_3 &= h\ f(t_n + \tfrac{1}{2}h, x_n + \tfrac{1}{2}k_2, y_n + \tfrac{1}{2}l_2) \\
k_4 &= h\ f(t_n + h, x_n + k_3, y_n + l_3) & l_4 &= h\ g(t_n + h, x_n + k_3, y_n + l_3)
\end{aligned}$$