City Research Online

City St George's, University of London

Citation: Amato, R., Lacasa, L., Díaz-Guilera, A. & Baronchelli, A. (2018). The dynamics of norm change in the cultural evolution of language. Proceedings of the National Academy of Sciences, 115(33), pp. 8260-8265. doi: 10.1073/pnas.1721059115

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/20390/

Link to published version: https://doi.org/10.1073/pnas.1721059115

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

The Dynamics of Norm Change in the Cultural Evolution of Language

Roberta Amato[a,b], Lucas Lacasa[c], Albert Díaz-Guilera[a,b], and Andrea Baronchelli[d,*]

[a]Departament de Física de la Matèria Condensada, Universitat de Barcelona, Barcelona (Spain)

[b]Universitat de Barcelona Institute of Complex Systems (UBICS), Barcelona (Spain)

[c]School of Mathematical Sciences, Queen Mary University of London, London E1 4NS (UK)

[d]Department of Mathematics, City, University of London, London EC1V 0HB (UK)

Abstract

What happens when a new social convention replaces an old one? While the possible forces favoring norm change - such as institutions or committed activists - have been identified since a long time, little is known about how a population adopts a new convention, due to the difficulties of finding representative data. Here we address this issue by looking at changes occurred to 2,541 orthographic and lexical norms in English and Spanish through the analysis of a large corpora of books published between the years 1800 and 2008. We detect three markedly distinct patterns in the data, depending on whether the behavioral change results from the action of a formal institution, an informal authority or a spontaneous process of unregulated evolution. We propose a simple evolutionary model able to capture all the observed behaviors and we show that it reproduces quantitatively the empirical data. This work identifies general mechanisms of norm change and we anticipate that it will be of interest to researchers investigating the cultural evolution of language and, more broadly, human collective behavior.

Introduction

Social conventions are the basis for social and economic relations [1–4]. Examples range from driving on the right side of the street, to language, rules of politeness or moral judgments. Broadly speaking, a convention is a pattern of behavior shared throughout a community, and can be defined as the outcome that everyone expects in interactions that allow two or more equivalent actions (e.g., shaking hands or bowing to greet someone) [5, 6]. Conventions emerge either thanks to the action of some formal or informal institution, or through a self-organized process in which group level consensus is the unintended consequence of individual efforts to coordinate locally with one another [2,6]. Crucially, since conforming to a convention is in everyone's best interest when everyone else is conforming too, social conventions are self-enforcing [5]. Yet behaviors change all the time and old conventions are constantly replaced by new ones: words acquire new meanings [7], orthography evolves [8], rules of politeness are updated [9], and so on. In isolated groups, shifts in conventions may be driven by the same forces that determine the emergence of a consensus from a disordered state, i.e., institutions or self-organization [6, 7]. However,

*Corresponding author: Andrea.Baronchelli.1@city.ac.uk

a quantitative understanding of the processes of norm change has remained elusive so far, probably hindered by the difficulty of accessing adequate empirical data [10].

Here, we address this issue by focusing on shifts in orthographic and linguistic norms through the lenses of about 5 million written texts covering the period from 1800 to 2008 from the digitized corpus of Google Ngram [9] dataset. Following the same approach that has allowed quantification of processes such as the regularization of English verbs [12] or the role of random drift in language evolution [13], we analyze the statistics of word occurrences for a set of specific linguistic forms that have been historically modified either by language authorities or spontaneously by language speakers in English or Spanish. These include words that have changed their spelling in time and competition between variants of the same word or expression. To explore the mechanisms of norm change we consider three separate cases:

1. **Regulation by a formal institution.** We analyze the effect of the deliberations of the Royal Spanish Academy, *Real Academia Española* (RAE), the official royal institution responsible for overseeing the Spanish language, on the spelling of 23 Spanish words (complete list in SI Sec. 3) [2–8].

2. **Intervention of informal institutions.** We investigate the effect of dictionary publishing in the US on the updating of American spelling for 900 words [21, 22] (complete list in SI Sec. 10.A).

3. **Unregulated (or 'spontaneous') evolution.** We consider the alternation between forms that are either unregulated or described as equivalent by an institution but have nonetheless exhibited a clear evolutionary trajectory in time (i.e., we do not consider the case of random drift as primary evolutionary force [13]). In particular, we examine (i) the evolution over time of the use of two equivalent forms for the construction of imperfect subjunctive verbal time in Spanish, for $1,571$ verbs (complete list in SI Sec. 10.C; verbs and declination for each form in [23, 24]), (ii) the alternation of two written forms of the Spanish adverb *solo/sólo* ('only') [25], and (iii) 46 cases of substitution of British forms (e.g., words) with American ones in the US [26] (complete list in SI Sec. 10.B).

We show that these mechanisms leave robust and markedly distinct stylized signatures in the data, and we propose a simple evolutionary model able to reproduce quantitatively all of the empirical observations. When a formal institution drives the norm change, the old convention is rapidly abandoned in favor of the new one [7, 27–30]. This determines a universal process of norm adoption which is independent of both word frequency and corpus size. A qualitatively similar pattern is also observed for norm adoption driven by an informal institution, although in this case the adoption of the new form is smoother and word dependent. In the case of unregulated norm change, the transition from the old to the new norm is slower, potentially occurring over the course of decades, and is often driven by some asymmetry between the two forms, such as the presence of a small fraction of individuals committed to one of the two alternatives [31–33].

Data and historical background

Spanish

Founded in 1713, the *Real Academia Española* (Royal Spanish Academy, RAE) is the official institution responsible for overseeing the Spanish language. Its mission is to plan language by applying linguistic prescription in order to promote linguistic unity within and across Spanish-speaking territories, to ensure

a common standard in accordance with Article 1 of its founding charter: "... to ensure the changes that the Spanish language undergoes [...] do not break the essential unity it enjoys throughout the Spanish-speaking world." [34–36]. Its main publications are the *Dictionary of Spanish Language* (23 editions between 1780 and today) and its *Grammar*, last edited in 2014. Particularly interesting for our study is the standardization process that the RAE carried out during the 19th century, which enforced the official spelling of a number of linguistic forms [3, 37].

Our data set contains 23 spelling changes that occurred in four different reforms, in 1815, 1884, 1911 and 1954 (additional details are in the SI Sec. 3) [2–8]. To illustrate this, Fig. 1*A* shows the temporal evolution of the spelling change of the word *quando* ('when') into *cuando* –regulated in the 1815 reform– in the Spanish corpus, showing a sharp transition (or "S-shaped" behavior [27–29]). Different is the case of the adverb *solo* ('only'), whose spelling variant *sólo* was added to the RAE dictionary in 1956 after a long unofficial existence supported by a number of academics [25,33,38]. We will consider the coexistence of these latter two forms as an example of unregulated evolution (since 2010, the RAE discontinued the *sólo* variant again [25], but our dataset does not include such recent data).

A major example of unregulated norm change is offered by the Spanish past subjunctive, which can be constructed in two - equivalent [39, 40] - ways by modifying the verbal root with the (conjugated) ending *-ra* or *-se* (additional details are in SI Sec. 3). For example, the first person of the past subjunctive of the verb *colgar* ('to hang') could be indistinctly *colga-ra* or *colga-se*. Fig. 1*B* shows the growth of the *-ra* variant, for all verbal persons, over two centuries. A similar behavior is found in most Spanish verbs, the form *-se* being the most used at the beginning of XIX century (preferred $\approx 80\%$ of the times) to the less used at the beginning of the XXI century (chosen $\approx 20\%$ of the times). This peculiar phenomenon has attracted the attention of researchers for the last 150 years and has not been entirely clarified [39].

Recent results suggest that, whereas individuals typically use only one of the two forms, the alternation between the two variants tends to be found only in speakers who prefer the *-se* form [40, 41], as also confirmed by a recent analysis of written texts [42]. Thus, the users of $-ra$ appear to be effectively committed to this unique form. As we will see below, the possibility of such asymmetries of behavior have been incorporated into our model.

British English vs. American English

The emergence of American English was encouraged by the initiative of academics, newspapers and politicians – e.g., US President Theodore Roosevelt [31] – who over time introduced and supported new reforms [43]. The process gained momentum in the 19th century, when a debate on how to simplify English spelling began in the United States [32, 44–46], which was also influenced by the development of phonetics as a science [47]. As a result, in 1828 Noah Webster published the first *American Dictionary of the English Language*, beginning the Merriam-Webster series of Dictionaries that is still in use nowadays [32, 48]. Some changes, such as *color* instead of *colour* or *center* for *centre*, would become the distinctive features of American English. Fig. 1*C* shows the transition from the British spelling *centre* to the American *center*. The complete list of the 900 words examined is reported in SI Sec. 10.A.

The phenomenon of 'Americanization' of English [26] is not limited to spelling but includes also the introduction of different words or expression which over time replaced the British ones. Recent works [26, 49] report how the globalization of American culture might be favoring the affirmation of their specific form of English. We will consider a list of 46 American-specific expressions in relation to their British counterpart ([26], complete list in SI Sec. 10.B), such as *garbage* vs *rubbish* reported in Fig. 1*D*, or *biscuit* vs *cookie*. In all cases, we will consider only books listed in the American English Corpus of

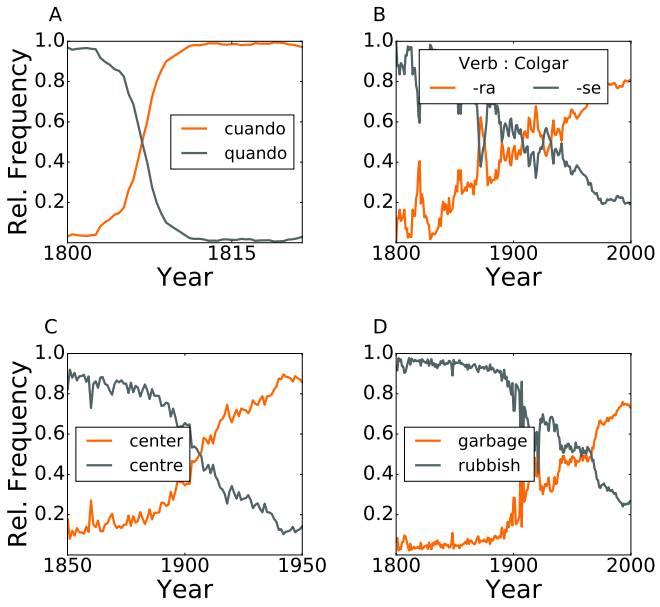

Figure 1: Illustrative examples of competing conventions in our dataset (relative frequencies). (A) Formal institution: the spelling of the Spanish word "quando" (when) was changed into "cuando" by a RAE reform in 1811. (B) Unregulated evolution of two equivalent forms for the past subjunctive, *-ra* and *-se*, for the verb "colgar" (to hang). (C) Informal institution: the American Spelling "center" versus the British spelling "centre". (D) Unregulated evolution of "garbage", the American variant of the British "rubbish".

Google Ngram.

Model

We introduce a simple model that describes the evolution in time of two alternative forms of a word (i.e., two alternative conventions). For example, the two norms may represent two spelling alternatives (*-or* vs *-our* as in *color/colour*), two ways to form a verbal tense (*-ra* vs *-se*) or two different words to refer to the same concept (*biscuit* vs *cookie*).

The model describes a system of books where instances of the two conventions are added by authors through the publication of books. Authors select which convention to use (i.e., which form to introduce in the system) either by following the indications of an institution or considering the current state of language. In the first case, authors simply adopt the recommended norm (or 'new norm', for simplicity, as we focus on cases of norm change). In the latter case, the convention to be used is selected with a probability proportional to its current frequency, as in the neutral model for evolution [50]. Additionally, some authors can be committed to one specific form, thus being indifferent to any external influence, as suggested by the literature on the study of orthographic norm change in both English [31, 32] and Spanish [33]. When an authority is present, the presence of commitment is revealed by the (empirically verified, see SI Sec. 4 and Fig S1) persistence in time of the old norm, and translates into the model such phenomena as, for example, the re-editions of past books whose orthography is not updated [15]. In the case of unregulated evolution, committed authors privilege the initially less popular new norm, contributing to its success.

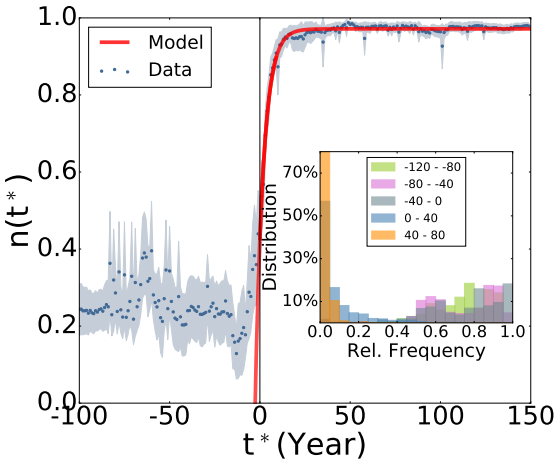

Figure 2: Regulation by a formal institution (Spanish, RAE). Main panel: Relative frequency of the new spelling form as a function of the rescaled time t^*. Blue points represent the average over all the considered pairs of words and the gray area the standard deviation of the data. The solid line is the prediction of the model outcome (eq. (4)) after parameter fit ($\chi^2 = 6 \cdot 10^{-5}$, $p = 0.99$). The black vertical line denotes the rescaled regulation year $t^* = 0$. Inset: Frequency histogram of the old spelling form for all pair of word forms, for different time periods (negative time refers to periods before the regulation).

The two different conventions are labeled as 'new' and 'old', and their number is $\mathcal{N}$ and $\mathcal{O}$ respectively and the total number of conventions at time t is given by $\mathcal{W}(t) = \mathcal{N}(t) + \mathcal{O}(t)$. For a more transparent comparison with the data, aggregated on a yearly basis, we adopt a discrete-time formulation of the model where one time step corresponds to one year. The evolution of the densities $n(t) = \mathcal{N}(t)/\mathcal{W}(t)$ and $o(t) = \mathcal{O}(t)/\mathcal{W}(t) = 1 - n(t)$ is described by the following equations

$$\begin{aligned} n(t+1) &= (1-c)\,(1-\gamma)\,n(t) + (1-c)\,\gamma E_{\mathcal{N}} + c\,, \\ o(t+1) &= (1-c)\,(1-\gamma)\,o(t) + (1-c)\,\gamma E_{\mathcal{O}}\,. \end{aligned} \tag{1}$$

New words are inserted by writers (authors). A writer is committed to the use of one specific convention, with probability c, or neutral, with probability $1-c$. Neutral writers follow the institutional enforcement, with probability γ, or sample the current distribution of norms, with probability $1-\gamma$. For simplicity, we assume that each writer inserts just one convention and that the probabilities c and γ are constant. When an institution promotes the norm $\mathcal{N}$, it makes an effort $E_{\mathcal{N}} = 1$ and $E_{\mathcal{O}} = 0$ otherwise. If the institution is impartial, both forms are a priori equivalent and $E_{\mathcal{N}} = E_{\mathcal{O}} = \frac{1}{2}$. Again for simplicity, in the equations all committed writers privilege the same convention [31–33].This is the new norm in the above equations, while expressions for the symmetric case of committed agents that support the old form are reported in SI Sec. 1. The general solution of the system of equations (1) is:

$$n(t) = \frac{B}{(1-A)}\left(1 - A^t\right) + n_0 A^t\,, \tag{2}$$

where $A = (1-c)\,(1-\gamma)$, $B = (1-c)\,\gamma E_{\mathcal{N}} + c$ and $n_0 = n(t=0)$ (see SI Sec. 2). It is worth noticing that, when $E_{\mathcal{N}} = 1$, for $\gamma = 1$ eq. (4) describes an instantaneous transition in which the new norm immediately saturates to $B/(1-A)$ (with $B/(1-A) = 1$ if the commitment supports the new norm, as

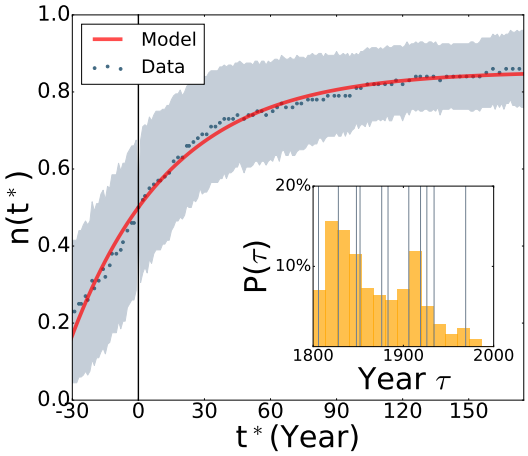

Figure 3: Regulation by informal institutions. Main panel: Relative frequency of the American spelling for 900 English words as a function of the rescaled time t^* ($t^* = 0$, denotes the surpassing year, for all pairs of words considered). Blue dots represent the average over all the pairs of words and the gray area the standard deviation of the data. The solid line is the model outcome, eq. (4), after parameter fit ($\chi^2 = 8 \cdot 10^{-4}$, $p = 0.97$). Inset: Distribution $P(\tau)$ of the years τ in which the American form overcame the British variant for each word. Vertical lines denote important moments of informal regulations of the US spelling such as dictionary editions or spelling updates (additional details are in SI Sec. 5).

here, or $B/(1-A) = 1-c$ if it supports the old norm, see SI). In this sense, values $\gamma < 1$ correspond to a situation in which the response of the system to an institutional intervention is not immediate. In the following sections we show that, by appropriately varying the parameter values, the analytic solution Eq. (4) reproduces all the empirical observations.

Results

Regulation by a formal institution

In the main panel of Fig. 2 we consider the relative frequency, $n(t)$, of appearance of the new spelling for the 23 words in our dataset affected by RAE reforms [2–8] (See SI Sec. 4). By a simple rescaling (translation) of the time axis as $t^* = t - t_r$ (where t_r is the regulation year for each specific pair of conventions), we find that all the experimental curves collapse. The regulatory intervention ($t^* = 0$) determines an abrupt transition towards the adoption of the new norm. This discontinuity is captured by the distribution of the old spelling among the words before and after the regulation in the inset panel of Fig. 2. Importantly, such rescaling indicates that the transition is size-independent. For example, for the 1815 regulation, our dataset consists of $B_{1815} = 59$ books and $S_{1815} = 4,149,151$ words, whereas for the regulation enforced in 1954 we have $B_{1954} = 2774$ books and $S_{1954} = 244,138,299$ words, but transition between the old and new form occurs over approximately the same amount of time in the two cases. Model parameters for the case of formal regulation are $E_{\mathcal{N}} = 1$ and commitment supporting the old convention for which $B = (1-c)\,\gamma$ (eq. (1) of SI Sec. 1). The main panel of Fig. 2 shows that the fit of eq. (4) matches the empirical data ($\gamma = 0.2$, $c = 0.006$ and $n_0 = 0.42$ from the data). As we will see below, different values of γ correspond to different roles played by institutions in the process of

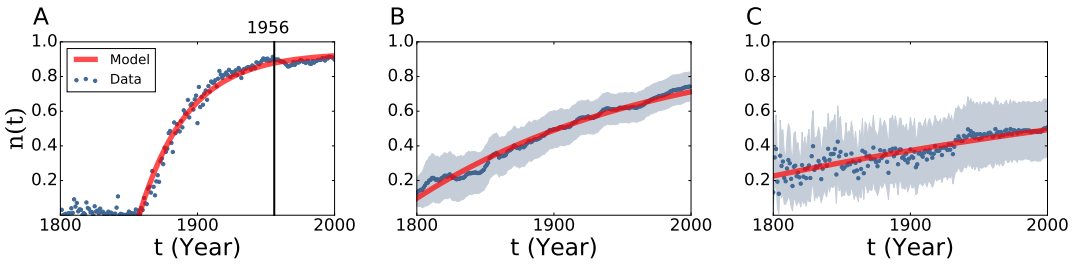

Figure 4: Unregulated norm change. (A) Case *sólo* versus *solo*. Blue dots represent the relative frequency of the Spanish adverb *sólo* (increasing in detriment of the alternative form *solo*). The solid line is the prediction of eq. (4) for this case, after parameter fit ($\chi^2 = 4 \cdot 10^{-4}$, $p = 0.98$). The vertical line signs the year 1956 when RAE intervened explicitly on the case [25, 33, 38]. The curve saturates to a value smaller than 1 probably due the presence of a percentage of adjectives, indistinguishable from the adverb in the data. (B) Case of '*-ra*' versus '*-se*' in Spanish subjunctive. Blue dots represent the relative frequency of the form *-ra* (increasing in detriment of the alternative but equivalent form *-se*) in Spanish past subjunctive conjugation of verbs, averaged over all verbs considered. Solid line is the specific prediction of eq. (4) for this case, after parameter fit ($\chi^2 = 2 \cdot 10^{-3}$, $p = 0.96$). (C) Case of Americanization of English in US. Blue dots represent the relative frequency of the American variant (with respect to the British variant) in US corpus, averaged over all the expressions examined. Solid line is the specific prediction of eq. (4) for this case, after parameter fit ($\chi^2 = 6 \cdot 10^{-3}$, $p = 0.93$). For all the cases the grey area identifies the standard deviation of the data.

norm change (see SI Sec. 7, Fig. S3*A* and Fig S4*A* for the behavior of individual curves, and SI Sec. 8 Fig.S4*B* for the corresponding distribution of γ).

Intervention of informal institutions

We now focus on the dynamics occurring between American and British spelling through the analysis of 900 words as they appear in our US corpus (complete list in SI Sec. 10.A). As in the case of formal institution, we have $E_{\mathcal{N}} = 1$ and the commitment supporting the old (i.e., the British, here) spelling. For each pair of conventions we identify the year τ in which the British form was surpassed in popularity by the American one (the inset panel of Fig. 3 shows the empirical frequency distribution of these surpassing times $P(\tau)$). The main panel of Fig. 3 shows that by rescaling time via simple translation $t^* = t - \tau$ all experimental curves collapse, similarly to the above case of formal institution. The model eq. (4) reproduces the data. The value of $\gamma = 0.02$ obtained in this case is much smaller than the one relative to the above case of formal institution ($\gamma = 0.2$), quantifying the weaker role played by informal institutions (other parameters $c = 0.003$ obtained by the fit, and $n_0 = n(t^* = 0) = 0.5$ by construction). This result is confirmed by analyzing each pair of competing conventions in isolation (see SI Sec. 7, Fig. S3*B* and Sec. 8, Fig.S4).

Unregulated evolution

As a third case we explore the process of unregulated norm change by considering the relative frequency of appearance of the form *sólo* (vs *solo*) (Spanish for 'only') [25] in the Spanish corpus, the relative frequency of appearance in the Spanish corpus of the past subjunctive form ending in $-ra$ and the one ending in $-se$ for 1,571 verbs (See SI Sec. 10.C), and the relative frequency of appearance in the US corpus of 46 cases, among words and expressions, of substitution of British forms for American ones (See

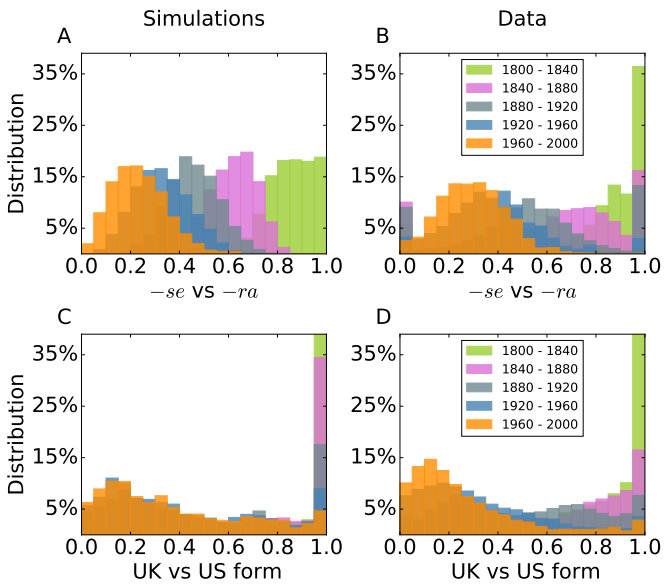

Figure 5: Empirical and simulated distributions for the relative frequencies of a given form. Top: Spanish subjunctive case, *-se* vs *-ra*. Simulation reported in (A) reproduced the empirical observation of the equivalent distributions (B). Bottom: Intervention of informal institution case, UK vs US variant. Simulations reported in (C) can be compared with the actual empirical distribution (D). Simulated distributions from 200 simulation runs with parameters informed by the fitting procedure and $\mathcal{W} = 2000$.

SI Sec. 10.B). Since the institution is impartial, we have $E_{\mathcal{N}} = E_{\mathcal{O}} = \frac{1}{2}$. Figs. 4*A*, *B* and *C* show that the solution Eq. (4) describes well the data relative to growth of the form *sólo* ($\gamma = 3.10^{-3}$, $c = 0.02$), the growing of the $-ra$ form for the subjunctive of Spanish verbs ($\gamma = 10^{-17}$, $c = 0.005$) and the growth of American forms ($\gamma = 10^{-14}$, $c = 0.002$), respectively (solid lines correspond to the model predictions after parameter fitting). The values of γ obtained here are significantly smaller than the ones observed for the cases of formal and informal institutions, and corroborate the fact that centralized authorities played essentially no role in this case (see also Fig. S4 in the SI Sec. 8 for the analysis of individual curves). It is worth noting that *solo* (without accent) can be also used as adjective and that, while the competition *solo/sólo* concerns only the adverb, the data do not allow us to distinguish between the adverb or adjective use. Our analysis shows that the adverb is dominant, as the adverb-specific *sólo* is nowadays the most used form, but the non-saturation of the curve in Fig. 4*A* can be interpreted as a signature of the presence of a percentage of adjectives in our dataset.

Microscopic dynamics

As a further assessment, we ran stochastic simulations of the model to reproduce the microscopic evolution of each pair of conventions for the case of spontaneous transition and for the case of the intervention of informal institution. In each numerical experiment we impose the parameters recovered through the fitting procedure described above. We initially consider the case of unregulated (spontaneous) norm adoption. In Fig. 5*A* and *B* we report probability distribution of observing a relative frequency $n(t)$ for the verbal form *-se*, estimated by simulating the evolution of all verbs for which we have empirical record. The simulation results suggest that our model captures well the ensemble evolution over time of the whole empirical distributions. Similarly, Fig. 5*C* and *D* show empirical and numerical results for

the class of norm adoption via informal authority, in the case of American spelling change. To account for multiple interventions of informal institutions, numerical experiments were run by 'switching on' the parameter γ at different, randomly chosen, times. Moreover, the American case consists of conventions that manifest themselves through specific set of words, i.e. the use of *or* instead of *our* in *behavio(u)r* or *colo(u)r* or *-ize* instead of *-ise* in verbs. Thus, for each simulation we extract γ from a Gauss distribution centered in $\gamma = 0.02$ (as informed by the data, and $\sigma = 0.005$) to reproduce the fact that, in this case, the transition from the old to the new convention is word-dependent.

By visually comparing empirical distributions of conventions over time for each norm adoption class (inset panel of Fig. 2, Fig. 5*B*, and Fig. 5*D* for formal authority, spontaneous and informal authority respectively) it is evident that, microscopically, the transition from the old to the new convention is governed by different dynamics. For enforcements by formal authorities (inset panel of Fig. 2), when the norm is regulated the system simply switches to the new convention. On the other hand, for unregulated (spontaneous) norm change (Fig. 5*B*) the distribution essentially remains unaltered but for a translation of its mean value which gradually shifts from 1 to 0. Finally, the word dependent transition of the informal institution case yields a broadening of the shape of the distribution over time (Fig. 5*D*).

Conclusion

In this work we have capitalized on a recently digitized corpus to analyze the process of norm change in the context of the cultural evolution of written English and Spanish. Through the analysis of $2,541$ cases of convention shifts occurring over the past two centuries, we identified three distinct mechanisms of norm change corresponding to the presence of an authority enforcing the adoption of a new norm, an informal institution recommending the normative update and a bottom-up process by which language speakers select a new norm. Each of these mechanisms displayed different stylized patterns in the data. We rationalized these findings by proposing a simple evolutionary model that describes the actions of the drivers of norm change previously identified in the literature, namely institutions and language users committed to the use of one of the two competing conventions. We showed that this single model captures the dynamics of norm change in each of the three cases described above, quantitatively matching the empirical data in all circumstances. In doing so, it differentiates the empirical curves in three classes according to the measured strength of the institutional intervention (fitted values of γ and single curve evaluation, see SI Sec. 8), thus confirming a posteriori the validity of our approach. Finally, through numerical simulations we were also able to reproduce the observed microscopic dynamics of norm adoption.

When a formal institution is present, the transition is sharp and does not depend either on the properties of the considered system (e.g., year or number of published books) or the relative importance of the linguistic convention subject to the norm change. The effect of informal institutions is weaker, resulting in a slower reaction of the system and a smoother transition. Finally, in the bottom-up process of spontaneous change the mechanisms of imitation and reproduction are key in bringing about the relatively slower onset of the new norm, catalyzed by the presence of 'committed activists' [31–33].

It is important to delimit the scope of our findings. First, we only considered cases for which historical records show that a norm change did occur and we did not attempt to predict whether a specific form is at risk of being substituted or not [12]. Second, we considered that the new convention had an advantage over the old one, represented either by the intervention of an institution or by the presence of committed users [31–33], and we did not consider examples where random drift is the dominant

evolutionary force [13]. In this respect, it is worth mentioning that the role of a committed minority has been investigated in the context of various multi agent models where it has been shown to play an important role on the final consensus provided its size exceeds a certain threshold [6, 52–54], as observed also in recent laboratory experiments [55]. Third, we focused on the case where the competition takes place between two alternative norms, but more complex cases where more conventions concur to the process of norm change could exist [26]. Fourth, the model we introduced describes the process of norm change for an isolated linguistic group and does not address the important case of language change resulting from the contact between two linguistically independent populations or conflict between different languages [7]. Finally, our analysis did not consider regional differences or any geographical factors. All these points represent directions for future work.

Taking a broader perspective, our results shed new light on the dynamics leading to the adoption of new linguistic conventions and have implications on the more general process of norm change. Today's technology, and in particular online social networks, are reportedly speeding up the process of collective behavioral change [56,57] through the adoption of new norms [55,58–60]. Understanding the microscopic mechanisms driving this process and the signature that it may leave in the data will lead to a better understanding of our society as well as to possible interventions aimed at contrasting undesired effects. In this perspective, we anticipate that our work will be of interest also to researchers investigating the emergence of new political, social, and economic behaviors [10, 61].

1 Methods

Google Ngram data set provides about 4% of the total number of books ever printed [9]. We analyzed the following data. Regulation by a formal institution: 23 Spanish words that change their spelling recovered in [2–8] (See SI Sec. 3). Intervention of informal institutions: 900 words with the double American and British spelling as reported in SI Sec. 10.A. The list is extracted from [22] and the double spelling verified with the Merriam-Webster dictionary [21]. Unregulated evolution: (i) Case of Spanish past subjunctive, 1571 Spanish verbs, 325 of which irregular. All the verbs, together with their declination, are listed in [23, 24]. (ii) Case of Americanization of English, 46 among words and expressions (the complete list is provided in SI Sec. 10.C).

For the microscopic dynamics, we performed numerical simulations of the model. At the beginning, we set $\mathcal{W}_0$ conventions in the state $\mathcal{O}$. At each time authors extract and replace the conventions of the previous time with the following rules. With probability c the author is committed. If the commitments support the new conventions, a convention in the state $\mathcal{N}$ is added, otherwise a convention in the state $\mathcal{O}$ is added; with probability $(1-c)(1-\gamma)$ the author reproduce the convention extracted; and with probability $(1-c)\gamma$ the author follows the institution effort: with probability $(1-c)\gamma E_{\mathcal{N}}$ a convention in the state $\mathcal{N}$ is added while with probability $(1-c)\gamma E_{\mathcal{O}}$ a convention in the state $\mathcal{O}$ is added. We impose the values recovered by the fitting procedure to set the parameters c and γ in the simulations.

2 Acknowledgements

We are very grateful to JMR Parrondo and J. Cuesta, who contributed with ideas and discussions to earlier stages of this work. RA and AD-G acknowledge support from Ministerio de Economia y Competitividad of Spain project no. FIS2012-38266-C02-02 and no. FIS2015-71582-C2-2-P (MINECO/FEDER) and Generalitat de Catalunya grant no. 2014SGR608. LL acknowledges funding from EPSRC Early

Career Fellowship EP/P01660X/1.

References

[1] Young HP (2015) The evolution of social norms. *Annual Reviews of Economics* 7(1):359–387.

[2] Ehrlich PR, Levin SA (2005) The evolution of norms. *PLoS Biol* 3(6):e194.

[3] Bicchieri C (2005) *The grammar of society: The nature and dynamics of social norms.* (Cambridge Univ Press).

[4] Marmor A (2009) *Social conventions: From language to law.* (Princeton Univ Press).

[5] Lewis D (1969) *Convention: A philosophical study.* (Blackwell).

[6] Baronchelli A (2018) The emergence of consensus: a primer. *Royal Society Open Science* 5:172189.

[7] Croft W (2000) *Explaining language change: An evolutionary approach.* (Pearson Education).

[8] Andersen H (1989) Understanding linguistic innovations. *Language Change: Contributions to the study of its causes* 43:5.

[9] Watts RJ (2005) *Politeness in language: Studies in its history, theory and practice.* (Walter de Gruyter).

[10] Nyborg K, et al. (2016) Social norms as solutions. *Science* 354(6308):42–43.

[11] Jean-Baptiste M, et al. (2011) Quantitative analysis of culture using millions of digitized books. *science* 331(6014):176–182.

[12] Lieberman E, Michel JB, Jackson J, Tang T, Nowak MA (2007) Quantifying the evolutionary dynamics of language. *Nature* 449(7163):713–716.

[13] Newberry MG, Ahern CA, Clark R, Plotkin JB (2017) Detecting evolutionary forces in language change. *Nature* 551(7679):223.

[14] Vaquera MLC, de Molina Redondo JA (1986) *Historia de la gramática española:(1847-1920).* (Gredos).

[15] B PO (1992) Rendimiento funcional de las nuevas normas de prosodia y ortografia de la real academia española de la lengua (1959). *Cauce, 1992,(14-15): 171-219.*

[16] Wagner C (2016) Andrés bello y la gramática castellana latinoamericana. *Revista Documentos Lingüísticos y Literarios UACh* (29).

[17] Merín MQ (2014) La academia literaria i zientífica de instruczion primaria: defensa razonada (y apasionada) de su ortografía filosófica en 1844.

[18] Alcoba S (2007) Ortografía y DRAE. algunos hitos en la fijación léxica y ortográfica de las palabras. *Español Actual* 88:11–42.

[19] RAE (2017) Nuevo tesoro lexicogrfico de la lengua espaola (NTLLE). *Española (NTLLE):¡ http://ntlle. rae. es/ntlle.*

[20] Casares J (1954) La academia y las nuevas normas. *BRAE XXXIV*.

[21] Merriam-Webster I (2017) Merriam-webster.com. *https://www.merriam-webster.com*.

[22] www.tysto.com/uk-us-spelling list.html (2017) Comprehensive* list of american and british spelling differences.

[23] List of Spanish irregular verbs. *Available at:*
https://gist.github.com/robertaamato/99f646670e9f8d0f261807a0185503be.

[24] List of Spanish regular verbs *Available at:*
https://gist.github.com/robertaamato/4cd3682cdab72bce0019c0e5b356f121.

[25] RAE (2017) El adverbio solo y los pronombres demostrativos, sin tilde. *Consultas lingsticas. Available at: http://www.rae.es*.

[26] Gonçalves B, Loureiro-Porto L, Ramasco JJ, Sánchez D (2018) Mapping the Americanization of English in space and time. *PLoS ONE 13(5): e0197741*.

[27] Osgood CE, et al. (1954) Psycholinguistics: a survey of theory and research problems. *The Journal of Abnormal and Social Psychology* 49(4p2):i.

[28] Blythe RA, Croft W (2012) S-curves and the mechanisms of propagation in language change. *Language* 88(2):269–304.

[29] Ghanbarnejad F, Gerlach M, Miotto JM, Altmann EG (2014) Extracting information from s-curves of language change. *Journal of The Royal Society Interface* 11(101):20141044.

[30] Lass R (1997) *Historical linguistics and language change*. (Cambridge Univ Press) Vol. 81.

[31] Vivian JH (1979) Spelling an end to orthographical reforms: Newspaper response to the 1906 Roosevelt simplifications. *American Speech* 54(3):163–174.

[32] Scragg DG (1974) *A history of English spelling*. (Manchester Univ Press) Vol. 3.

[33] Martin Rodrigo I, Moran D, De la Fuente M, Doria S (2014) Sólo/solo: la tilde que enfrenta a la RAE con los escritores. *ABC Cultura*.

[34] RAE (2017) Orgenes y fines de la rae. *Available at: http://www.rae.es/la-institucion/historia/origenes*.

[35] RAE (2017) Estatutos y reglamento de la real academia española. *Available at: http://www.rae.es/la-institucion/organizacion/estatutos*.

[36] Pochat MT (2001) Historia de la real academia española. *Olivar* 2(2):249–253.

[37] Española RRA (2010) *Ortografía de la lengua española*. (Espasa).

[38] Alvarez Mellado E (2017) Sólo y la tilde de la nostalgia. *eldiario*.

[39] Guzmán Naranjo M (2017) The se-ra alternation in spanish subjunctive. *Corpus Linguistics and Linguistic Theory* 13(1):97–134.

[40] Kempas I (2011) Sobre la variación en el marco de la libre elección entre cantara y cantase en el español peninsular. *Moenia* 17:243–264.

[41] Gómez AB, Co GVE (2002) *Corpus de conversaciones coloquiales*, Oralia (Madrid). (Arco).

[42] Rosemeyer M, Schwenter SA (2017) Entrenchment and persistence in language change: the spanish past subjunctive. *Corpus Linguistics and Linguistic Theory*.

[43] Weinstein B (1982) Noah webster and the diffusion of linguistic innovations for political purposes. *International Journal of the Sociology of Language* 1982(38):85–108.

[44] Venezky RL (1999) *The American way of spelling: The structure and origins of American English orthography*. (Guilford Press).

[45] Hodges RE (1964) A short history of spelling reform in the united states. *The Phi Delta Kappan* 45(7):330–332.

[46] Zachrisson RE (1931) Four hundred years of english spelling reform. *Studia neophilologica* 4(1):1–69.

[47] Wijk A (1961) Regularized english. *Educational Research* 3(2):157–159.

[48] Micklethwait D (2005) *Noah Webster and the American dictionary*. (McFarland).

[49] Leech G (2009) *Change in contemporary English: A grammatical study*. (Cambridge Uni.).

[50] Kimura M (1983) *The neutral theory of molecular evolution*. (Cambridge Univ Press).

[51] Pechenick EA, Danforth CM, Dodds PS (2015) Characterizing the google books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PloS one* 10(10):e0137041.

[52] Xie J, et al. (2011) Social consensus through the influence of committed minorities. *Physical Review E* 84(1):011130.

[53] Mistry D, Zhang Q, Perra N, Baronchelli A (2015) Committed activists and the reshaping of status-quo social consensus. *Physical Review E* 92(4):042805.

[54] Niu X, Doyle C, Korniss G, Szymanski BK (2017) The impact of variable commitment in the naming game on consensus formation. *Scientific Reports* 7.

[55] Centola S, Becker J, Brackbill D, Baronchelli A (2018) Experimental evidence for tipping points in social convention *Science* 360(6393):1116–1119.

[56] Kooti F, Yang H, Cha M, Gummadi PK, Mason WA (2012) The emergence of conventions in online social networks. in *ICWSM*.

[57] Centola D, Baronchelli A (2015) The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences* 112(7):1989.

[58] Becker B, Mark G (1999) Constructing social systems through computer-mediated communication. *Virtual reality* 4(1):60–73.

[59] Bicchieri C, Fukui Y (1999) The great illusion: Ignorance, informational cascades, and the persistence of unpopular norms. *Business Ethics Quarterly* 9(1):127–155.

[60] Del Vicario M, et al. (2016) The spreading of misinformation online. *Proceedings of the National Academy of Sciences* 113(3):554–559.

[61] Valenzuela S, Arriagada A, Scherman A (2014) Facebook, twitter, and youth engagement: A quasi-experimental study of social media use and protest behavior using propensity score matching. *International Journal of Communication* 8:25.

Supporting Information

Symmetric case for eq. (1)

When the commitment supports the old convention $\mathcal{O}$, eq. (1) takes the form:

$$\begin{aligned} n(t+1) &= (1-c)\,(1-\gamma)\,n(t) + (1-c)\,\gamma E_{\mathcal{N}} \\ o(t+1) &= (1-c)\,(1-\gamma)\,o(t) + (1-c)\,\gamma E_{\mathcal{O}} + c\,. \end{aligned} \tag{3}$$

The general solution of the system of equations 3 is:

$$\begin{aligned} n(t) &= \frac{B}{(1-A)}\left(1-A^t\right) + n_0 A^t \\ o(t) &= 1 - n(t)\,, \end{aligned} \tag{4}$$

where $A = (1-c)\,(1-\gamma)$, $B = (1-c)\,\gamma E_{\mathcal{N}}$ and $n_0 = n(t=0)$.

Spanish past subjunctive

In Spanish two equivalent forms exist to construct the past subjunctive : the one ending in $-ra$ and the one ending in $-se$ (as in *pensa-ra* and *pensa-se* had thought). The form $-se$ evolved from the Latin plusquamperfect subjunctive, while the form $-ra$ evolved from the Latin plusquamperfect indicative [1].

Spanish spelling reforms

We present the complete list of the 23 words examined in the Spanish spelling change case, grouped into their respective reforms [2–8].

- 1815 :
  - antiquario → anticuario (antiquarian)
  - quaderno → cuaderno (notebook)
  - quadro → cuadro (picture)
  - quando → cuando (when)
  - quanto → cuanto (how much)
  - quarto → cuarto (fourth)
  - quatro → cuatro (four)
  - quociente → cociente (quotient)
  - quota → cuota (quote)
  - quotidiano → cotidiano (daily)
  - Equador → Ecuador (Ecuador)
  - iniquo → inicuo (iniquitous)

  - obliquo → oblicuo (oblique)

- 1884
  - guion → guión (script)
  - truhan → truhán (rogue)
  - virey → virrey (viceroy)
  - vireina → virreina (viceroy's wife)
  - vireinato → virreinato (viceroyalty)

- 1911
  - ó → o (or)
  - á → a (to)

- 1954
  - dió → dio (it gave)
  - fué → fue (it was)
  - vió → vio (it saw)

Persistence in time of the old norm

Fig. 6 shows that typically the frequency of appearance of the old spelling in the case of formal regulation does not go to zero but rather, after a sudden drop, stabilizes on a plateau as if a small number of writers ignored the intervention of the institution (i.e., as if they were committed against the new norm). This is mainly due to the presence of historical books, compendiums of the language as well as re-editions of past books (see Supporting Online Material of [9]). For example the list of Spanish books still containing the word "quando" after regulation is composed of texts such as "Biblioteca histórica de la filología castellana" ("Historical library of Castillan filology"), published in 1893, "Documentos para la historia de la Revolución de 1809" ("Documents for the history of 1809 revolution") published in 1954 or "Manuscritos litúrgicos de las bibliotecas de España" ("Liturgical manuscripts of Spanish libraries") published in 1977, which clearly refer to its historical use.

American Spelling: Important Moments

Important moments in the history of American vs British spelling:

- 1806 - Noah Webster published *'A Compendious Dictionary of the English Language'*
- 1828 - First American Dictionary *'An American Dictionary of the English Language'*
- 1848 - Alexander John Ellis published *'A Plea for Phonetic Spelling'*
- 1876 - American Spelling Reform Association were founded and start to adopt the reforms
- 1883 - The Chicago Tribune newspaper start to adopt the reforms

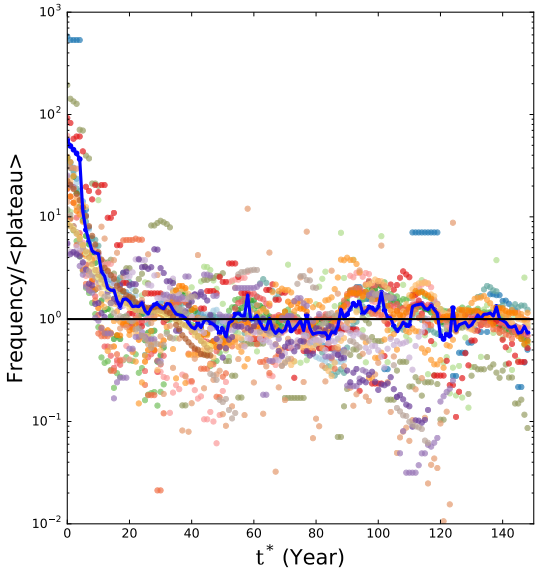

Figure 6: Frequency of appearance of the old spelling form as a function of the rescaled time $t*$ for the case of formal regulation. Colored dots represent all the considered words and the blue solid line is the average. The frequency of appearance of each word is divided by the average made on the respective plateau.

- 1906 - The Simplified Spelling Board was founded and President of the United States Theodore Roosevelt signed an executive order imposing the use of reformed spelling in the official communications of the Congress.
- 1919 - H.L. Mencken published the first edition of *The American Language*
- 1926 - Henry Fowler published the first edition of *Dictionary of Modern English Usage*
- 1969 - Harry Lindgren published *Spelling Reform: A New Approach*

British and American spelling conflict

In our analysis of the case of spelling conflicts, the various inflections of a term, such as singular or plural, or, for a verb, present, past, etc, were considered separately because they do behave differently. As en example, we report in Fig. 7 the evolution of the singular and the plural of the word "behavior/behavior" (American/British spelling). The American spelling of the singular exceeds that of the British almost a century before the plural one. As mentioned in the Main Text, investigating the reasons of this phenomenon is beyond the scope of our analysis, and an interesting starting point for future work.

Individual variations beyond the average

For the case of formal institution, time is rescaled according to the regulation year (i.e., 1815, 1884, 1911 or 1954, in our dataset, see Sec. 2 above), which in all cases mark the sudden adoption of new norm. Fig. 8*A*, shows how individual pair of norms react to the corresponding regulation event.

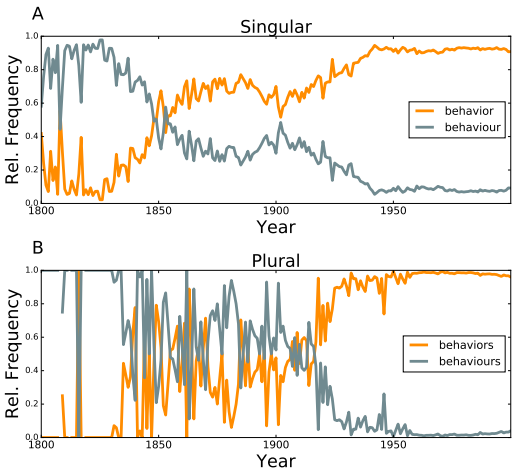

Figure 7: (A) Evolution of *behavior*, the American form and *behaviour*, the Bristish form. (B) Evolution of the plural *behaviors* and *behaviours*.

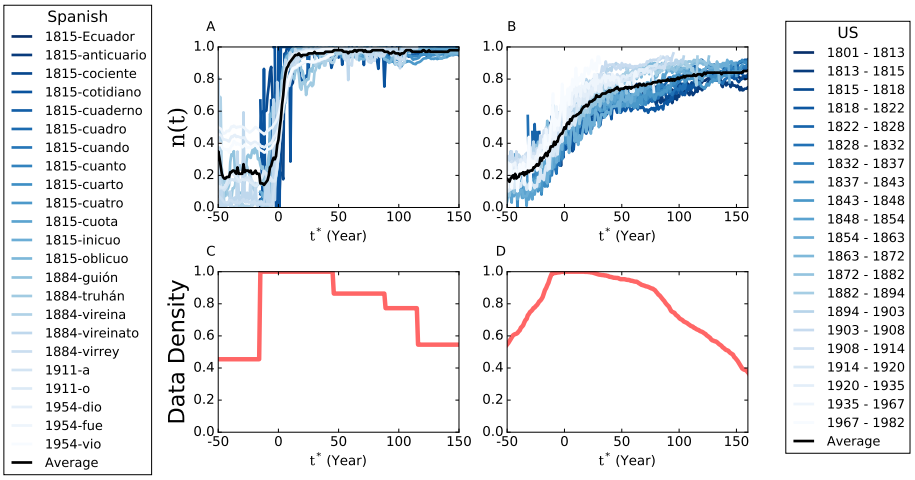

Figure 8: A, C Formal Regulation Case. A Relative frequency of the new spelling form as a function of the rescaled time $t*$. Colored lines represent all the considered pairs of words, while the black line is the average. C shows the density of data that contribute to the average for each year. B, D Informal Regulation Case. B Relative frequency of the new spelling form as a function of the rescaled time $t*$. The words are grouped so that each time interval contains 5% of the curves. D shows the density of data that contribute to the average for each year. Colored lines represent the average in these intervals and the black line is the total average.

For the case of informal regulation, a first dictionary proposing the considered spelling reforms was edited in the 1828 following two decades of discussions (Sec. 2). The moment in which each American form surpassed in popularity the corresponding British one is distributed over 1800 and 2000. Fig. 8 B shows how pairs of conventions corresponding to different surpassing times contribute to the global

average. For the sake of clarity, individual curves were grouped so that each time window contains 5% of the data. Thus, for example, the curve labeled as 1837-1843 represents the average over the words for which the American spelling surpassed the British one in that time interval.

Figs. 8 *C* and *D* report the density of data contributing to the total average for the corresponding year for the Formal and Informal Regulation cases, respectively.

Comparison of the three classes

To further assess the contribution of the individual pairs of conventions to the three classes we identified in the paper (formal institution, informal institution, and unregulated change), Fig. 9 *A* shows the temporal evolution of the relative frequency of the new forms,$n(t)$, confirming that the time scales involved in the process of norm change are different for the three cases.

We further investigated the variability between cases by using the model discussed in the Main Text as an inspecting tool. Fig. 9 *B* reports the cumulative distributions of the parameter γ, $C(\gamma)$, obtained by fitting individual curves for the three cases. The Kolmogorov-Smirnov returns $p = 10^{-12}$, $p = 10^{-20}$ and $p = 0.0$, for the first with the second, the first with the third and the second with the third curve respectively. We therefore reject the hypothesis that the three empirical distributions are instances of the same probability distribution.

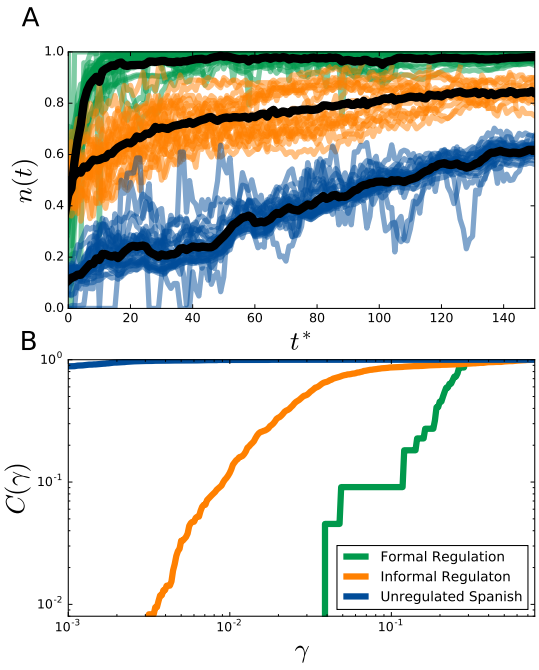

Figure 9: *A* Relative frequency of the new forms as a function of the rescaled time $t^* = t - \tau$. For the case of formal regulation, τ is the year of regulation, for the case of informal regulation is the year in which the new norm becomes more frequent than the old one, while for the unregulated case $\tau = 0$ as the curves are not rescaled. Black lines represent the average. *B* Cumulative distribution of γ, $C(\gamma)$, for each pair of curves for the Formal Regulation, Informal Regulation and Unregulated (Spanish) cases. By comparing the three distributions with the KS test we recover $p = 10^{-12}$, $p = 10^{-20}$ and $p = 0.0$, respectively for the first with the second, the first with the third and the second with the third distribution.

Google Books Corpus

Google Ngram Corpora offers an unprecedented opportunity to analyze linguistic and cultural change in a quantitative way [9–14]. The dataset is the product of a massive effort in text digitization, in collaboration with thousands of the world's libraries [9]. The resulting 5 million books contain over half a trillion words, 361 billion of which are in English [9, 15].

Although Google Ngram can serve as a useful barometer of lexical change [14] it is important to mention some of its limitations. To avoid breaking any copyright laws, the datasets are not accompanied by any metadata regarding the texts the corpora consist of [12] and it does not account for the popularity of a text, namely each text contributes with equal weight to token counts [14,15]. Furthermore, as recently pointed out, the inclusion of scientific texts, which have become an increasingly substantive portion of the corpus after the 1900s risks to artificially skew the statistical composition of the dataset [15].

As far as our analysis is concerned, the problem of increasing scientific publications would affect mainly the American English dataset after the 1900s. However, the fact that (i) 71% of the norm-change events we observe in the American English corpus date before 1900, and (ii) we find no significant change in behavior between the transitions observed before and after 1900s or (iii) between the transitions occurring in American English and Spanish (in the case of unregulated change, where we can compare them) are all indications in favor of the robustness of our results. Moreover, it is reasonable to assume that our American English dataset, thanks to its size (946 cases of norm change, in total), includes words whose frequency of use in the scientific literature vary considerably. Thus, the homogeneity of behavior revealed by the curve-by-curve analysis reported in Figs. 8 and 9 further confirms the validity of our findings.

References

[1] J M Wilson. The-ra and-se verb forms in Mexico: a diachronic examination from non-literary sources. 1983.

[2] M L C Vaquera and J A de Molina Redondo. *Historia de la gramática española:(1847-1920)*. Gredos, 1986.

[3] Palomo Olmos B. Rendimiento funcional de las nuevas normas de prosodia y ortografia de la real academia española de la lengua (1959). *Cauce, 1992,(14-15): 171-219*, 1992.

[4] C Wagner. Andrés bello y la gramática castellana latinoamericana. *Revista Documentos Lingüísticos y Literarios UACh*, (29), 2016.

[5] Mercedes Quilis Merín. La academia literaria i zientífica de instruczion primaria: defensa razonada (y apasionada) de su ortografía filosófica en 1844. 2014.

[6] Santiago Alcoba. Ortografía y DRAE. algunos hitos en la fijación léxica y ortográfica de las palabras. *Español Actual*, 88:11–42, 2007.

[7] RAE. Nuevo tesoro lexicogrfico de la lengua espaola (NTLLE). *Española (NTLLE):¡ http://ntlle. rae. es/ntlle*, 2017.

[8] Julio Casares. La academia y las nuevas normas. *BRAE XXXIV*, 1954.

[9] M Jean-Baptiste, Y K Shen, A P Aiden, A Veres, M K Gray, J P Pickett, D Hoiberg, D Clancy, P Norvig, J Orwant, et al. Quantitative analysis of culture using millions of digitized books. *science*, 331(6014):176–182, 2011.

[10] Peter Sheridan Dodds, Kameron Decker Harris, Isabel M Kloumann, Catherine A Bliss, and Christopher M Danforth. Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PloS one*, 6(12):e26752, 2011.

[11] Peter Sheridan Dodds, Eric M Clark, Suma Desu, Morgan R Frank, Andrew J Reagan, Jake Ryland Williams, Lewis Mitchell, Kameron Decker Harris, Isabel M Kloumann, James P Bagrow, et al. Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences*, 112(8):2389–2394, 2015.

[12] Alexander Koplenig. The impact of lacking metadata for the measurement of cultural and linguistic change using the google ngram data setsreconstructing the composition of the german corpus in times of wwii. *Digital Scholarship in the Humanities*, 32(1):169–188, 2017.

[13] Marcelo A Montemurro and Damián H Zanette. Coherent oscillations in word-use data from 1700 to 2008. *Palgrave Communications*, 2:16084, 2016.

[14] Tyler J Gray, Andrew J Reagan, Peter Sheridan Dodds, and Christopher M Danforth. English verb regularization in books and tweets. *arXiv preprint arXiv:1803.09745*, 2018.

[15] Eitan Adam Pechenick, Christopher M Danforth, and Peter Sheridan Dodds. Characterizing the google books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PloS one*, 10(10):e0137041, 2015.

Dataset composition

See complete SI on the PNAS website.