ROYAL SOCIETY OPEN SCIENCE

royalsocietypublishing.org/journal/rsos

Research

Cite this article: Nunes Amaral LA, Capozzi A, Helbing D. 2026 Breaking the code: Multi-level learning in the Eurovision Song Contest. *R. Soc. Open Sci.* **13**: 251727.
https://doi.org/10.1098/rsos.251727

Received: 9 September 2025
Accepted: 6 February 2026

Subject Category:
Computer science and artificial intelligence

Subject Areas:
behaviour

Keywords:
contests, learning, culture

Author for correspondence:
Luis A. Nunes Amaral
e-mail: amaral@northwestern.edu

Supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.8440854.

Breaking the code: Multi-level learning in the Eurovision Song Contest

Luis A. Nunes Amaral[1], Arthur Capozzi[2] and Dirk Helbing[2,3]

[1]Department of Engineering Sciences and Applied Math, Northwestern University, Evanston, IL, USA
[2]Computational Social Science, ETH Zürich, Zürich, Switzerland
[3]Complexity Science Hub, Vienna, Austria

LANA, 0000-0002-3762-789X; AC, 0000-0002-1996-1800; DH, 0000-0002-9898-0101

Organizations learn from market, political and societal responses to their actions. While in some cases both the actions and responses take place in an open manner, in many others, some aspects may be hidden from external observers. The Eurovision Song Contest offers a mostly open-data case in which to study organizational level learning at the levels of organizers and participants. We present here evidence for changes in the rules of the Contest in response to undesired outcomes such as runaway winners. We also find strong evidence of participant learning in the characteristics of competing songs over the 70 years of the Contest. English has been adopted as the *lingua franca* of the competing songs and pop has become the standard genre. The number of words of lyrics has also grown in response to this collective learning. Remarkably, we find evidence that France, Italy, Portugal and Spain have chosen to ignore the 'lesson' that English lyrics increase winning probability, consistent with utility functions that award greater value to featuring national culture than to winning the Contest. These countries—but not Germany—appear to be less susceptible to Anglo-Saxon cultural influence than their peers, a resistance that may extend beyond cultural matters.

1. Introduction

Contests occur in a multitude of contexts [1]. Animal contests determine access to food and mating opportunities [2]. Sport contests determine popularity, income and recognition of teams and individuals [3]. Product contests determine the access of business firms to markets [4]. Political contests determine access to power by individuals and parties [5–7]. Contest theory explores how

THE ROYAL SOCIETY PUBLISHING

experiences—i.e. learning opportunities—inform decision making and the role of evolutionary dynamics [8,9]. In many of these settings, the rules of the contest can be seen as constant, hiding the possible interplay between contestants' actions and rule setting. The Eurovision Song Contest—which from now on we will refer to simply as ESC or Contest—provides a novel and exciting context in which to study such interactions. Scholarly work has demonstrated that the ESC is a stage for more than just musical choices (see [10–13] and, especially, [14]). The ESC involves interactions at multiple organizational levels—organizers, national selection organizers, competing musicians and voters. We will focus here on two of those levels: organizers of the Contest and the participating countries.

The organizers' goal is to maximize television audiences. To achieve this goal, they must ensure that the rules of the Contest are clear and that the outcomes are not too predictable. If the rules are not clear, the audience may feel that the winner was selected in advance. If the outcome is predictable either before the Contest or even during the voting stage, then the audience will be bored and uninterested. Participating countries are de facto represented by a public broadcaster, but have wide latitude in determining how they select their competing songs. For example, the song representing Italy has typically been the winner of the Sanremo Music Festival. However, both the artist and song could be selected internally by a committee appointed by the broadcaster. Typically, there is a clear tension between a country's specific music traditions and the aggregate music taste of voting groups (electronic supplementary material, figures S1 and S2). This is evident in markets where domestic music competitions, such as Italy's Sanremo Festival, have a significantly higher audience than the ESC. In contrast, other countries show a much higher relative audience share for the ESC.

Because the sizes of global and national audiences, competing songs and voting records are public information, there is ample opportunity for individual and collective learning by all parties. Here, we investigate these learning processes. In this, our study differs from others that have usually focused on the characteristics of successful songs, not on interaction and learning effects.

The triumphant winning song of the 2024 ESC is well remembered to start with the words 'Welcome to the show. Let everybody know… I broke the code'. While the theme of this song is actually the breaking of gender stereotypes, one might wonder whether they were also suggesting that the band had figured out the formula for success of the Contest itself. Below, we show that some song characteristics can indeed increase the chances of success. However, those rules were not uncovered by a small group of ingenious songwriters, but figured out by all the participants collectively.

1.1. Background

Songs were the object of study in prior research concerning the dynamics of success. Salganik *et al.* [15] demonstrated that the ranking of songs leads to more unpredictable and unequal distributions of success, and that both the timing and order in which the songs are presented also matters. Rosati *et al.* [16] reported that the dynamics of song popularity are consistent with a contagious process. More recently, the wide availability of databases with information about songs, including metrics of their success, has allowed the development of models that predict song popularity with some reasonable success [17–20].

Importantly, however, most of these studies focus on the success of individual songs, ignoring the context in which those songs were created and presented. Here, we will investigate this overlooked aspect. Success of songs—and other creative products [21,22]—is better understood in the context of a specific competitive environment. This is motivated by theories of (co-)evolution [23], where 'fitness' is not an absolute property, but largely influenced by the other species. Similar approaches have also been applied to the study of innovation dynamics, interactive behavioural change [24] and cultural change [25]. An interesting example for this approach is the 'Art Exhibition Game', a behavioural experiment that studied the (co-)evolution of innovation and success in a decision science laboratory [26].

2. Data and methods

We retrieved data for all the Contests from its start in 1956 until 2024 from Wikipedia [27]. The dataset comprises 1763 songs from 51 countries, including historical nations that no longer exist, such as Yugoslavia. For each contest, these pages list the set of competing countries, the performer names, the song titles, the song languages and the points earned. The 'points earned' information is quite detailed, including not only total points in the finals and the semi-finals (from 2004 onwards) but also a breakdown of points by awarding country and by origin (jury or televoting).

We enriched these data by obtaining song lyrics from the site *Letras.com* [28]. We then used AI tools to obtain (i) the music genre for each song, (ii) the characteristic words representing the subject matter of a song, and (iii) the audio features for each track. We obtained the translations of the lyrics from the website *EurovisionWorld.com*. To quantify the evolution of the themes in song lyrics, we leveraged the 12 song themes taxonomy identified by Henard & Rossetti [29], which can be associated with distinct sets of keywords. We focus on the 11 most frequent themes: Aspiration, Breakup, Confusion, Desire, Desperation, Escapism, Inspiration, Loss, Nostalgia, Pain and Rebellion. We excluded 'Jaded' because it appears in fewer than 5% of the songs. Some of these 11 themes can be seen as relating to painful feelings (Breakup, Confusion), while others may relate to positive states (Aspiration, Desire, Escapism, Inspiration). Desperation, Pain and Rebellion may speak of dissatisfaction with the present, while both Loss and Nostalgia may reflect a longing for the past. Recent studies have demonstrated the effectiveness of large language models in topic modelling [30,31], lyric emotion classification [32,33] and lyric generation [34,35]; accordingly, we employed *GPT-4o* to annotate each song's lyrics with up to 3 of the 11 selected themes. More information on this task can be found in the electronic supplementary material, text section.

To characterize the musical characteristics of songs, we employed the API provided by the streaming service Spotify [36]. Spotify's API offers 13 audio features for each track. However, our study focused on six key features that we found to be the most relevant and informative: Acousticness, Danceability, Energy, Loudness, Tempo and Valence. It is worth noting that these features are only based on the audio track and do not take the song's lyrics into account. *Acousticness* is a confidence measure of whether the track is acoustic or not, with values between 0 and 1. Typically, high values correspond to recordings featuring non-amplified instruments, while low values indicate a high presence of electronic synthesis, amplification or digital production effects. *Energy* is a perceptual measure of intensity and activity that relies on dynamic range, perceived loudness, timbre, onset rate and general entropy. *Loudness* is the overall volume of a track in decibels (dB). *Danceability* is based on a combination of musical elements—including tempo, rhythm stability, beat strength and overall regularity—relating to whether one can dance easily to the song or not. *Tempo* measures the pace at which a section of music is played. *Valence* describes the musical positivity conveyed by a piece of music. Songs with high valence sound happy and cheerful, while pieces with low valence sound sad or angry.

While Spotify does not disclose how exactly its internal system assigns audio features to songs, its API is widely used in research and considered sufficiently reliable. For example, the API has been used for the identification of emotions [37], to explain song popularity [38,39], to understand why people listen to music [40] and to explore how music streaming behaviour relates to the personality traits of Spotify users [41].

Finally, we categorized the audio tracks according to musical genre. We employed a genre classification model available on Huggingface [42], fine-tuned from the *wav2vec2-base-960h* model [43], using the labelled GTZAN Dataset [44]. This dataset consists of 1000 audio samples, each 30 s long, evenly distributed across 10 music genres: 'Blues', 'Classical', 'Country', 'Disco', 'Hiphop', 'Jazz', 'Metal', 'Pop', 'Reggae' and 'Rock'. After 24 epochs, the model obtained in the training phase had a cross-entropy validation loss of 0.7037, an area under the ROC curve of 0.964 and a weighted accuracy of 0.815.

The model assigns each track a percentage of relevance for one or more of the 10 genres listed above. To make our analysis more robust and to better fit these genres to the songs competing, we aggregated the 10 genres into three macro categories: *Pop*, which also includes disco; *Rock*, which also includes metal; and *Other*, which aggregates the remaining six genres.

3. Results

We start by extracting aggregated metrics for each Contest (figure 1). We find that the number of countries participating in the Contest shows three periods, which we denote as *Formation*, *Consolidation* and *Expansion* (figure 1). The 'Formation' period includes the Contests taking place prior to 1974. During this period, the number of participating countries increased from fewer than 10 to over 15, with an average of 15.6 participants. The voting was not yet standardized and changed very much from year to year and from participant to participant (see electronic supplementary material, figure S3, for the impact of these changes). The second period, 'Consolidation', includes the Contests from 1974 to 2003. While there is a steady growth in the number of participating countries from 15.6 to 22.2, the voting rules remain stable. This also results in a stable partitioning across Contests in terms of the fraction of all votes earned by each of the top three songs (figure 1b). The third period, 'Expansion', includes the Contest from 2004

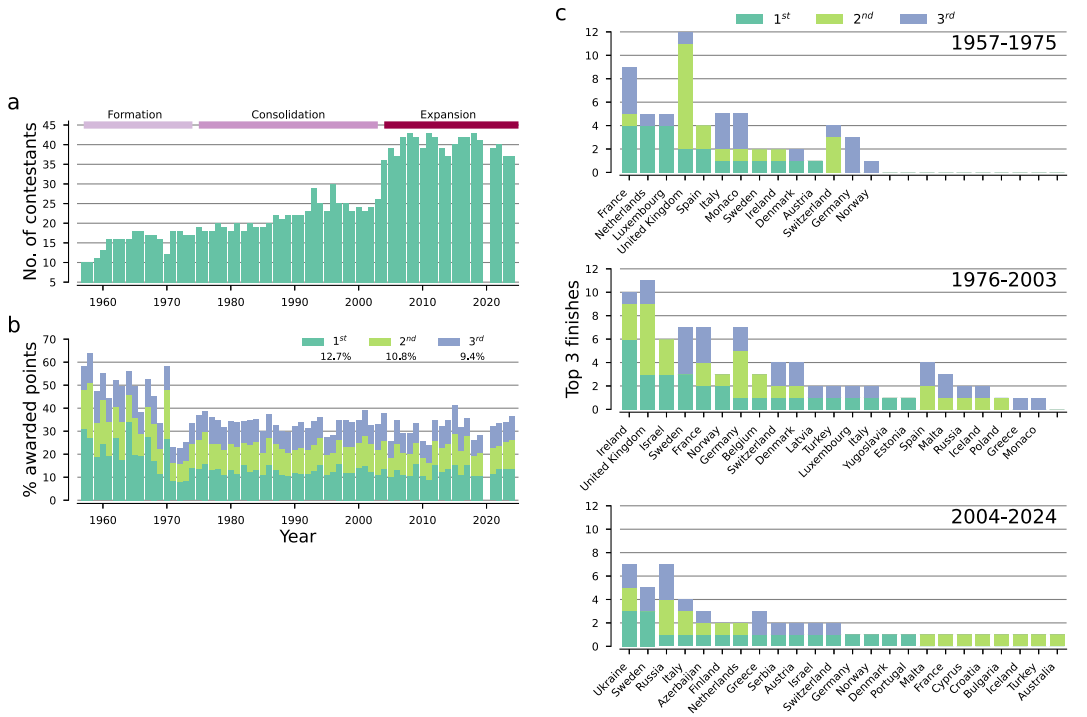

Figure 1. Over its history, the ESC has experienced dramatic changes in the range of countries participating and voting, in the rules underlying the contest and in its outcomes. These changes can be interpreted as a response of the organizers to an increased homogeneity of songs competing. (a) The number of countries participating in the Contest has increased steadily over time. A major change, however, occurred in 2004 when there was an increase in the number of participants and countries allowed to vote by nearly 50%. This increase required the use of semi-final contests. We split this 68 year history into three periods that we denote 'Formation', 'Consolidation' and 'Expansion'. (b) Changes in the voting process resulted in changes in the distribution of the votes received by participating countries (electronic supplementary material, figure S3). Importantly, it also affected the fraction of total votes received by the top three performing songs. It is visually apparent, nonetheless, that the system reached a stationary state after 1974. (c) While the voting share of the winners remained stationary after 1974, the voting patterns did not. The top panel shows the number of top three finalists for countries that reached the podium in the 1957 to 1975 Contests. It is visually apparent that some countries systematically do well while many others do not. Indeed, the top performing four countries took the majority of top three finishes. This concentration of winning decreases over time to almost parity in the Expansion period. This evolution suggests that the advantage held by some countries initially was lost over time.

onwards. Its start coincides with a very large increase in the number of participating countries (average 40.2) and in the number of voting countries. Furthermore, a semi-final stage was introduced for most of the participating countries. Finally, since 2004, there has been an increase in the impact of televoting. These latest changes resulted in a greater degree of variability in the vote fraction earned by the top three songs (figure 1b). The third period, 'Expansion', includes the Contest from 2004 onwards. Its start coincides with a very large increase in the number of participating countries (average 40.2) and in the number of voting countries. Furthermore, a semi-final stage was introduced for most of the participating countries. Finally, since 2004, there has been an increase in the impact of televoting. These latest changes resulted in a greater degree of variability in the vote fraction earned by the top three songs (figure 1b) and in a significant increase in the entropy of the distribution of votes (electronic supplementary material, figure S3).

The initial group of participating countries—Belgium, France, Germany, Italy, Luxembourg, The Netherlands and Switzerland—are located in close proximity, but differ in a number of cultural aspects, including language. By 1975, the set of participating countries also included the Scandinavian countries, the Iberian countries, Ireland, Turkey, the UK and the former Yugoslavia, thereby dramatically increasing the number of languages and cultures represented. One might surmise that such diversity could lead to the formation of voting blocs based on language, religion or geographic proximity. To test this hypothesis, we first investigate the structural outcomes of the voting process.

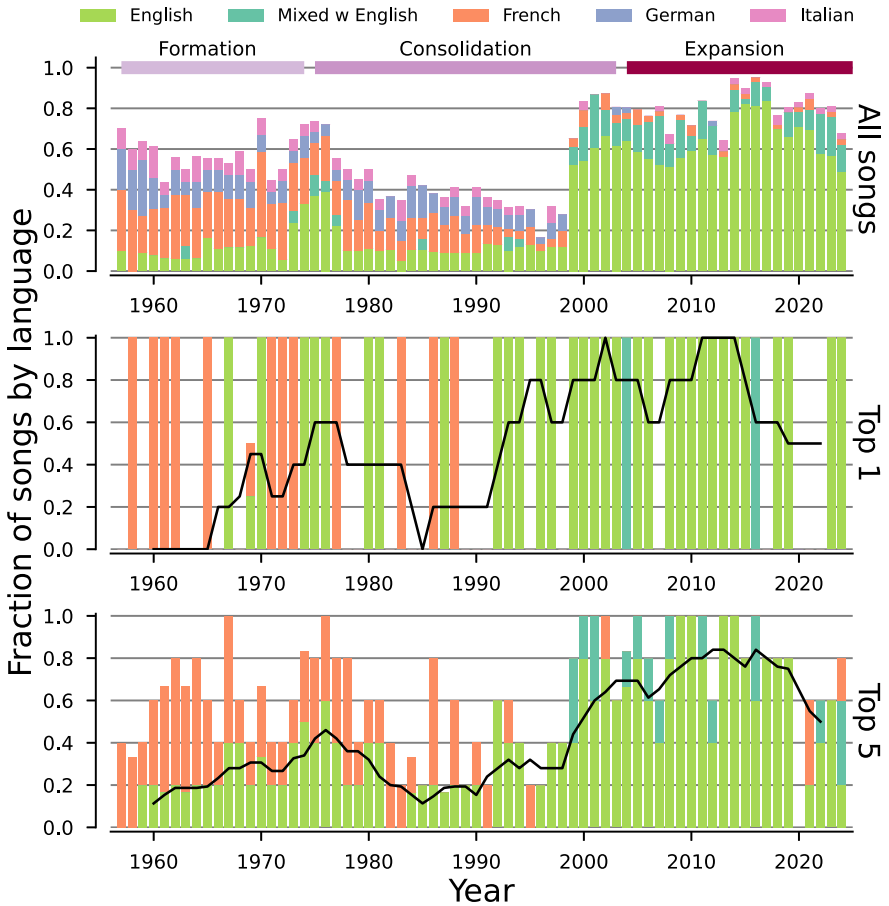

Figure 2. Despite the diversity of European languages, English becomes the de facto language of the ESC**.** Fraction of songs using the four most common languages of most initial participating countries. During the 'Formation' period, there is a great diversity in the language of the competing songs. This diversity increases through most of the 'Consolidation' period. In 1998, when there was a permanent lift of the restriction on using non-native languages of the contesting country, there is a transition to a new state, where English becomes the dominant language. In this new state of affairs, over 70% of songs are entirely in English or a mix of English and a native contestant language. When considering the languages of the top performing songs, the situation is quite different. Initially, French is the dominant language, but already by the 1970s, nearly 40–60% of the winning songs are in English (see full black line, which shows 5 year running average). By the early 1990s, nearly 80% of winning songs are in English. While the rate of increase is less pronounced for the top three (not shown) and top five, the data are consistent with the learning by participants that songs with English lyrics have an advantage. The increased dominance of English as the language of the top songs is accompanied by a decrease of French, which nearly disappears as the language of choice among the top songs after 1990.

Figure 1c shows that the rate at which different countries place in the top three positions changes over time. In the Formation period, top placement is concentrated within four countries: France, The Netherlands, Luxembourg and the UK. This concentration may be due to a number of factors, including voting rules. However, during the Consolidation and Expansion stages, the voting rules become set, making the interpretation of the results easier. Thus, it is still surprising that the top three positions are rather more concentrated during the Consolidation period than during the Expansion period. This suggests that the recent parity is not just a statistical artefact resulting from having more participants. Rather, it indicates that the systematic dominance previously held by specific countries has effectively vanished.

3.1. Language

If the number of individuals speaking a given language would explain these voting consensuses, then it would be expected that bigger countries such as France, Germany and the UK should perform particularly well (electronic supplementary material, table S1). In order to test this hypothesis, we calculate the fraction of songs with lyrics using a given language for each year (figure 2a). These data reveal two striking observations. First, there is a steady decline over time in the fraction of songs that have French

or German lyrics. Second, in 1999, there is a sudden, step-like increase in the use of English lyrics. Unlike the rise in the use of English around the mid-1970s, this rise shows no sign of reversing. So, we may ask, why did this step change happen?

These shifts can be understood in the context of the evolving Contest rules concerning language use. Language choice was free in the early years (1956–1965), but was then restricted to national languages, with the exception of a 3 year suspension between 1973 and 1976. Electronic supplementary material, table S2, provides a more detailed overview of how language restrictions have changed over time. However, these restrictions allowed for the use of multiple languages in the lyrics, with some countries bending the rules and including a majority of English verses. The language restrictions were lifted during 1973–1976 and again after 1998 [45]. These regulatory changes had a significant impact on the linguistic landscape of the contest (figure 2a). During the period from 1973 to 1976, when the national language rule was temporarily lifted, there was a noticeable increase in the number of songs performed in English. However, this increase was modest compared to the dramatic rise observed after 1999, when around 80% of songs were either entirely in English or mixed English with a national language.

To answer this question, we calculate how this percentage changes over time. Figure 2b,c displays the fraction of top ranked songs that have English lyrics. They make clear that—whether we consider the winning song, the top 3 songs or the top 5 songs—by 1970 the top-rated songs are primarily using English. Moreover, while the best-performing songs, prior to 1970, had frequently used French lyrics, French lyrics became a strong predictor of 'not top 5' after 1990. These results make it clear that, during the Consolidation period, participating countries learnt that submitting songs with lyrics in English dramatically increased the competitiveness of the song.

Next, we investigate whether there are other winning song features that are being learnt over time. Two features likely to be important in determining the popularity of a song are the characteristics of the lyrics, including complexity and theme, and the music attributes, including audio features and music genre.

3.2. Lyrics

First, we focus on the characteristics of song lyrics. A straightforward measure of lyrics complexity is the size of the lyrics as measured by number of words. Figure 3a shows the time evolution of the mean lyrics size. By 1975, once rules of the competition stabilize and some learning has had an opportunity to occur, lyrics stabilize around a size of about 230 words. This range is valid across countries (electronic supplementary material, figure S4). However, by 1999, there is a step-like 20% increase in lyrics size. As for song language, we find that songs placing in the top 3 already have larger lyrics size by the late 1980s. However, the lyrics size of the winning song shows much greater fluctuations.

The themes present in song lyrics also display interesting patterns. While each of these 11 themes appears in at least 5% of the songs, their co-occurrence is not random. Aspiration and Inspiration appear more frequently together than one would expect from chance alone, as do Pain, Breakup and Loss (electronic supplementary material, figure S5).

Four song themes—Nostalgia, Pain, Rebellion and Desperation—display significant temporal trends. Surprisingly, an initially strong focus on Nostalgia has been in steep decline. This may be related to the increasing distance from the horrors of World War II and improving economic conditions. In contrast to Nostalgia's steep decline, Pain, Rebellion and Desperation have been increasing significantly in representation. Intriguingly, Confusion and Escapism show nearly a step change, stretching from 1970 to 1980, to higher levels of representation, whereas Loss shows an opposite change over the same period to lower levels of representation (electronic supplementary material, figures S6 and S7). This could be a response to all the crises of the 1970s.

These findings suggest that the songs competing in the ESC are picking up on emotional trends permeating culture. However, a question is whether top performing songs merely ride these trends or whether they contribute to their emergence. The data are particularly interesting for Pain and Desperation (electronic supplementary material, figure S8). Between 1990 and 2000, top performing songs refer less to Pain than the average competing song. Since 2000, however, the representation of Pain among top performing songs has been increasing steadily and exceeds the average annual level since 2010. It may not be a coincidence that this is occurring after the Great Recession. In contrast to Pain, Desperation has had lower representation among top performing songs than average songs. This may be due to the emotional weight of a feeling such as desperation, which could turn voters off.

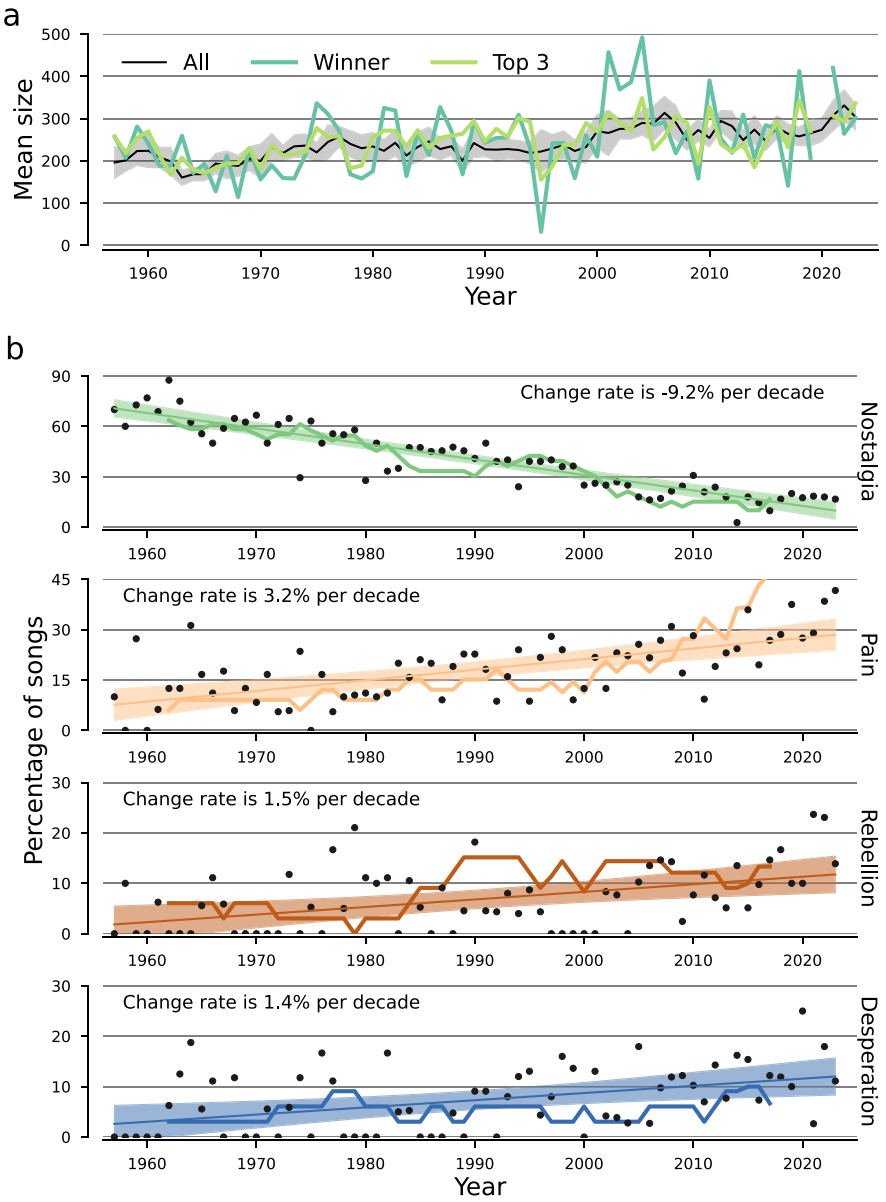

Figure 3. Song lyrics also show signs of change over time consistent with participant learning. (a) Time evolution of mean lyrics size (as measured by number of words) for all songs, winning song and top three songs. For all songs, the shaded region shows the 95% confidence interval for the estimate of the mean. As for song language, we find a step-like increase in lyrics size in 1999. Demonstrating the presence of a signal to be learnt, we find that the lyrics size of top three songs was already increasing during the period 1975–1998. Interestingly, the winning songs are characterized by extreme lyrics sizes, either very small or very large. As shown in electronic supplementary material, figure S4, there is no significant dependence of lyrics size on country. (b) Time evolution of the percentage representation of the four most informative song themes for all songs and top three songs. The black circles show the average over all competing songs for a given year, the line and shaded regions show 99% confidence intervals for the linear fit, and the coloured lines show 10 year running average of theme representation among top three songs. All four themes display strong and statistically significant temporal trends. While top performing songs follow the common trend for Nostalgia, they focused less on Pain between 1990 and 2000, but since 2010 are focusing more on Pain than the average of all competing songs.

3.3. Music

Now, we address the songs' audio features and genres. We calculate the distribution and cross-correlations for six audio features identified by the Spotify API [36] for all songs (electronic supplementary material, figure S9) and for the top performing songs (electronic supplementary material, figure S10). It is visually apparent that several are strongly correlated. Thus, we focus on just three of them

that we believe provide the greatest insight: Acousticness, Danceability and Valence (electronic supplementary material, figure S10). While displaying no significant temporal auto-correlations (electronic supplementary material, figure S11), the time evolution of these features displays clear patterns (figure 4a). Acousticness, which was dominating earlier on, has been decaying for the entire existence of the Contest. At first, the decline was smooth and exponential in nature, but then, it happened in a step-like way. The early primacy of Acousticness has since been taken over by Danceability. Valence and Danceability follow a similar oscillating pattern. They both increase until they reach a peak in the late 1970s. Then, they both decrease until the mid-1990s and increase towards an early 2000s peak. That peak is again followed by a decrease until the mid-2010s and an increase since.

We also examined the specific audio characteristics of the winning songs. While their long-term evolution mirrors the aggregate trends shown in figure 4a, their short-term dynamics are distinct. The lack of significant temporal auto-correlations after 1984 (electronic supplementary material, figure S11) suggests that a winning song's audio profile is not simply a replica of the previous year's winner. However, the songs are far from random. They constitute the target signal for collective learning. As we show in the following section (figure 5b,c), the danceability of top performing songs is a statistically significant attractor. Participants actively adjust their entries to match the standards set by recent winners.

The time evolution of audio features in competing songs is mirrored in the change in music genre, as the two types of classification are correlated (electronic supplementary material, figure S12). Figure 4b shows the mean genre score for Pop and Rock for all songs and for the winning song. It is visually apparent that Pop songs have come to dominate the Contest, although not as much as the English language, and not in such an abrupt manner. As for language and lyric characteristics, there is a clear signal regarding the characteristics of the top placed songs that allow all participants to learn. This is visible by the fact that the fraction of pop songs among the top performing echelon occurs earlier (electronic supplementary material, figure S13), even though there are no significant temporal auto-correlations (electronic supplementary material, figure S14). Interestingly, the situation is not as unequal as when considering language with regard to the genre of the winning song. Nowadays, pop songs tend to make up only about half of the top performing songs, and there are no signs of complete homogenization, as occurred with language. This suggests that the contest retains a degree of musical diversity.

3.4. Country-specific learning

Up to now, we have focused on system-wide learning. We found that competing countries have learnt from each other about lyrics language, size and topic, and about audio characteristics and music genre. What we have not investigated is whether this learning has occurred equally across countries or whether some countries have eschewed some of those lessons. Figure 5a compares the usages of Acousticness and Danceability by country for the pre- and post-1980 periods. As expected, we find adoption of Danceability across the board. We find a similar adoption for lyric sizes (electronic supplementary material, figure S4) and for music genre (electronic supplementary material, figure S15).

To determine whether these changes regarding music characteristics are due to learning, we fit a linear model to the danceability score of a country's song in year y to the average danceability scores of different subsets of competing songs—Top 1, Top 3, Top 5 or all—with a time delay of 1, 3 or 5 years and averaged over time windows with width of 1, 3 or 5 years (figure 5b). We find that the best model occurs when averaging the danceability scores over Top 5 songs in the previous 5 years. Figure 5c shows that we find similar models for eight countries for which the fit is statistically significant. We find similar results for acousticness (electronic supplementary material, figure S16). Note that the fit does not converge for all countries.

Figure 6a shows language utilization by country for the period pre- and post-system-wide transition to English. Prior to the 1999 relaxation of the rules concerning the choice of language, countries submitted songs in one of their national languages, some mixing in other languages, with a particular focus on English. Post-1999, the lyrics of the songs submitted by most countries switched to English. To determine whether these changes regarding song language are due to learning, we fit a logistic model to the probability that a country's song's lyrics include English in year y to the average fraction of songs with English lyrics for different subsets of competing songs—Top 1, Top 3, Top 5 or all—with a time delay of 1, 3 or 5 years and averaged over time windows with width of 1, 3 or 5 years (figure 6b). We find that the best model occurs when averaging the English usage of Top 3 songs in the previous 5 years. Figure 6c shows that we find similar models for 10 countries for which the fit is statistically significant. Note that the fit does not converge for all countries.

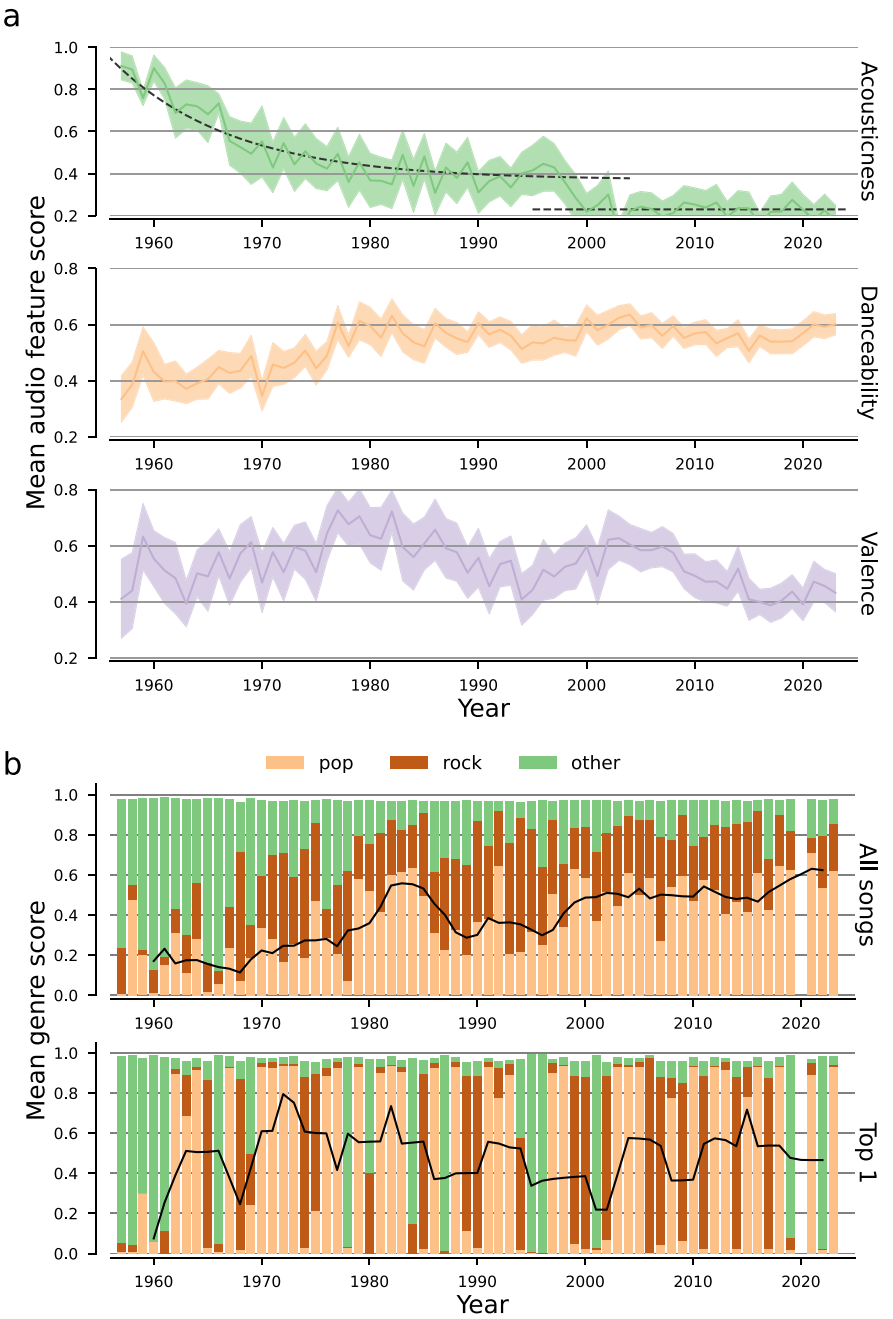

Figure 4. Despite the diversity of music styles and traditions in Europe, pop has become the dominant music genre for participating songs, but not for winning songs. (a) Annual mean audio feature scores over time of competing songs. The solid lines show the mean value of the attribute for the songs competing in a given year, while the red shaded areas display the 95% confidence interval for the mean. The grey lines in the rightmost plot show fits to the functional forms discussed in the text. (b) Mean annual genre score of all competing songs and winning songs. The mean pop score of competing songs has been increasing steadily over the existence of the competition. Intriguingly, the genre of the winning songs shows much different trends (see black line, which shows the running average for 5 year time windows). Since about 1970, half of the winning songs have been pop songs. This pattern is also present when considering the top three or top five songs (electronic supplementary material, figure S13) but occurs at a slower rate. These results suggest that, unlike language, where English is the standard, there is more opportunity for 'bucking the trend' in song genre.

The exceptions to English hegemony are France, Italy, Portugal and Spain. It is unlikely that these four countries did not learn about the advantage provided by the use of English, prompting the question of why they eschewed that advantage. The likely explanation for why France, Portugal or Spain stick with their own languages is that they are spoken by large numbers of people around the world. Thus,

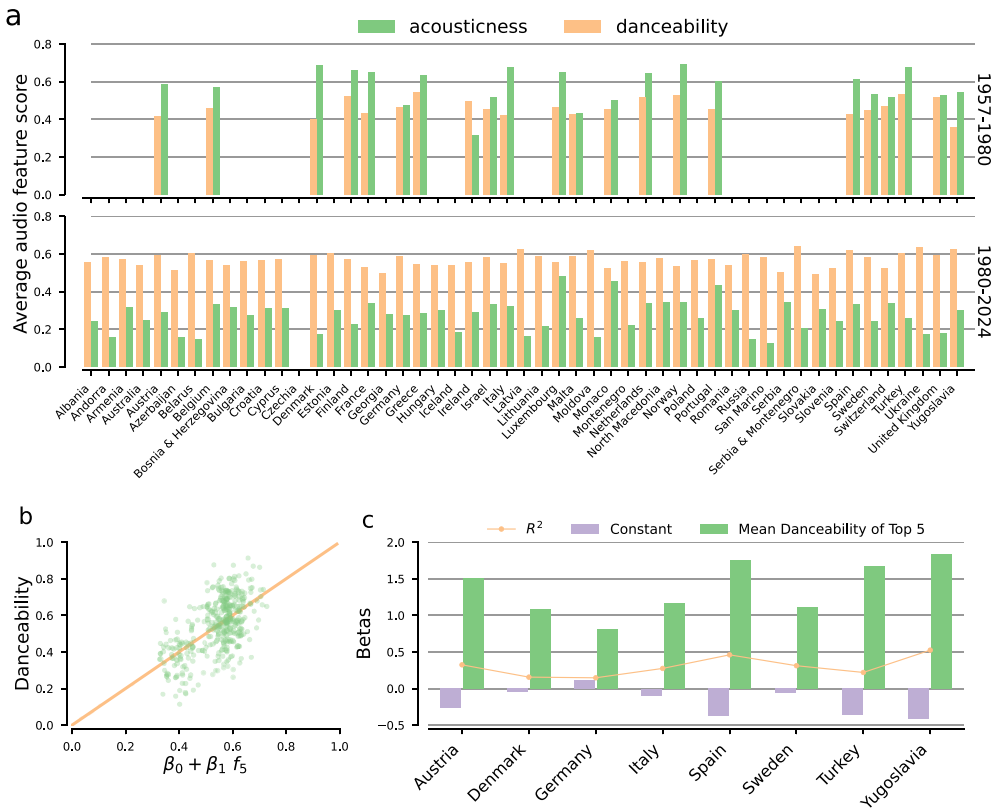

Figure 5. Participant learning of winning strategies—music features. (a) Average Acousticness and Danceability scores of songs submitted to the ESC by different countries across two time periods. It is visually apparent that, prior to 1980, competing songs had high Acousticness, but after 1980, competing songs—across all countries—had high Danceability. We find similar results for lyric sizes (electronic supplementary material, figure S4) and music genre (electronic supplementary material, figure S15). (b) Linear regression of danceability of a country's song in a given year against the average danceability of top five songs in the previous 5 years. (c) Regression coefficients for the eight countries for which the linear model is statistically significant and R^2 of the linear fit.

paying a cost for promoting their own languages is a rational choice in a geopolitical context that extends beyond winning at the ESC [14].

Finally, it is interesting to consider the extent to which learning, and the consequent cultural homogenization, is embraced by different countries. If one considers population size, language utilization in Europe or economic might, then Germany would appear to be a prime case for setting trends instead of following them. Instead, we find that France, Spain and Italy—but especially Portugal—are the countries shunning 'peer pressure'. It is perhaps not surprising that some of these countries have shown a more independent streak from Anglo-Saxon influence on some of the most controversial political matters now facing Europe.

4. Discussion

Our study provides insights into how learning and adaptation occur at multiple levels in a cultural market. From the organizers, who aim to maximize television audiences by making the Contest's outcome unpredictable and exciting, to the participating countries, who want to maximize the chances that their representatives' songs win the Contest while attending to other geopolitical necessities, to the voters, who may want to be fair in the ratings but are influenced by cultural and linguistic baggage. All these participants have left traces in the Contests' records that enable us to determine what they have collectively learnt and how they have decided to use those lessons.

Organizers tweaked the voting rules during the Formation period of the ESC, fostering a more predictable process, but less predictable outcomes. These changes led to a stable Consolidation period, which has eventually pushed organizers to expand voting options and voting participants. The recency of those changes does not yet allow us to provide final insights. In particular, the impact of televoting and the

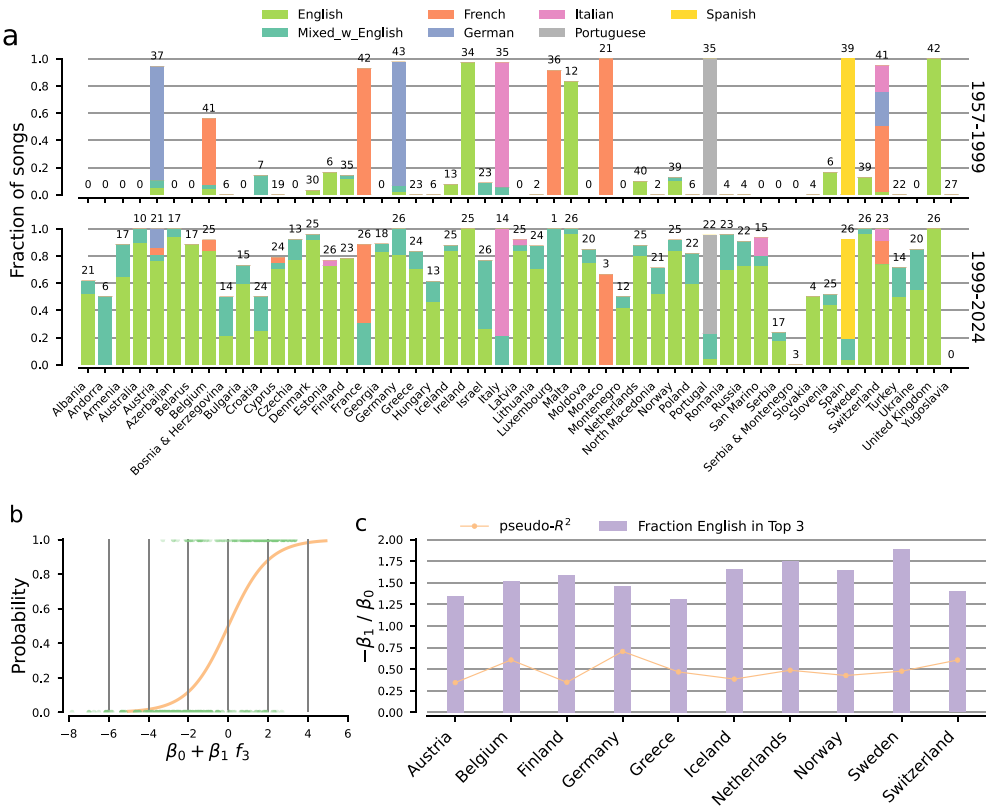

Figure 6. Participant learning of winning strategies—language. (a) Fraction of songs with lyrics in a given language submitted to Eurovision countries across two time periods. Prior to 1999, most countries submitted songs in one of their national languages (see Switzerland pre-1999). Post-1999, most countries submitted songs with English lyrics. The exceptions were France, Italy, Portugal and Spain, which were mostly submitting songs in their national languages. (b) Logistic regression of a country's song having English lyrics in a given year to the average fraction of songs placing in the Top 3 with English in their lyrics in the previous 5 years. (c) Ratio of the regression coefficients for the 10 countries for which the logistic model is statistically significant and pseudo-R^2 of the fit. We find similar results for acousticness (electronic supplementary material, figure S16).

possibility of its manipulation with 'bot farms' is not without precedent. Such manipulation has been attempted in the context of political elections [46]. While European Union's General Data Protection Regulation provides some level of protection against such manipulation, the desire for changing hearts and minds is likely to foster such manipulation, particularly in times of information wars and propaganda.

On the participant side, we have revealed the generalized adoption of strategies that increase song competitiveness: use of English language, increased danceability, increased use of Pop styles, more diverse lyrics and the focus on themes that better capture the cultural *zeitgeist*. However, the widespread adoption of these optimal strategies raises an interesting question. Why do some countries adopt these features but still fail to succeed? This phenomenon can be understood as a 'Red Queen' effect in cultural evolution [47,48]. As the adoption of English, Pop and high danceability become prevalent (in the sense of increased homogenization), these features shift from being a competitive advantage to a baseline requirement. In the current 'Expansion' phase, compliance with the learnt code is necessary to avoid failure, but not sufficient to guarantee success. Furthermore, the statistical distribution of voting outcomes could be interpreted through the lens of geopolitical 'friendships'. The transition to equilibrium observed in figure 1c and the increasing entropy of vote distributions (electronic supplementary material, figure S3) suggest that historical voting blocs have lost their decisive power. In a field of over 40 participants driven by televoting, the 'signal' of optimized song features appears to have largely overridden the pattern of traditional alignments, creating a more meritocratic, albeit homogenized, competitive landscape. Specifically, the 'Red Queen' race manifests in two distinct modes of cultural adaptation. The first is hegemonic convergence, in which a strategy—specifically, the shift to English—becomes a rigid constraint that establishes the baseline necessary to avoid failure. The second is attractor-based learning, which is observed in musical features such as danceability. In this case, top performing songs not only meet the baseline but also deviate from the average to set new standards that others subsequently chase.

Top performing songs drive the trend, but they do not simply replicate the average competing song; rather, they deviate from the norm to set new standards that others subsequently imitate. Furthermore, our study quantifies the limits of this pattern. While countries like Germany or Sweden have fully optimized their entries for the international market, France, Italy, Portugal and Spain have demonstrated a measurable resistance to this cultural pressure. By consistently ignoring the 'English advantage', these nations reveal a preference for cultural distinctiveness over competitive optimization. This shows that cultural identity can effectively constrain the drive for market optimization.

These results prompt the question of whether one could combine data analytics with generative AI to create songs that would ideally resonate with the diverse audience of the Contest. Whether that will be possible or not, one has to ask whether it would be desirable.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. Code and data for replicating this study can be found at https://github.com/amarallab/RS-Open_Science-Breaking_the_Code/tree/main.

Electronic supplementary material is available online [49].

Declaration of AI use. Large language models were used to classify themes of song lyrics.

Authors' contributions. L.A.N.A.: conceptualization, data curation, formal analysis, methodology, software, supervision, visualization, writing—original draft, writing—review and editing; A.C.: data curation, formal analysis, methodology, software, visualization, writing—review and editing; D.H.: conceptualization, funding acquisition, methodology, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. L.A.N.A. thanks the project 'CoCi: Co-Evolving City Life', which received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 833168, for partial financial support of his research stay at ETH Zurich. A.C. acknowledges partial financial support by the European Union—Horizon 2020 Program under the scheme 'INFRAIA-01-2018-2019—Integrating Activities for Advanced Communities', grant agreement no. 871042, 'SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics' (http://www.sobigdata.eu). D.H. has enjoyed coordinating and contributing to research in the CoCi and SoBigData++ projects.

References

1. Van Long N. 2015 The theory of contests: a unified model and review of the literature. In *Companion to the political economy of rent seeking* (eds RD Congleton, AL Hillman), pp. 19–52. Cheltenham, UK: Edward Elgar Publishing. (doi:10.4337/9781782544944.00010)
2. Reichert MS, Quinn JL. 2017 Cognition in contests: mechanisms, ecology, and evolution. *Trends Ecol. Evol.* **32**, 773–785. (doi:10.1016/j.tree.2017.07.003)
3. Fernandez-Navarro J, Fradua L, Zubillaga A, Ford PR, McRobert AP. 2016 Attacking and defensive styles of play in soccer: analysis of Spanish and English elite teams. *J. Sports Sci.* **34**, 2195–2204. (doi:10.1080/02640414.2016.1169309)
4. Mohr JJ, Sengupta S. 2002 Managing the paradox of inter-firm learning: the role of governance mechanisms. *J. Bus. Ind. Mark.* **17**, 282–301. (doi:10.1108/08858620210431688)
5. De Dreu CKW, Pliskin R, Rojek-Giffin M, Méder Z, Gross J. 2021 Political games of attack and defence. *Phil. Trans. R. Soc. B* **376**, 20200135. (doi:10.1098/rstb.2020.0135)
6. Iyengar S, Simon AF. 2000 New perspectives and evidence on political communication and campaign effects. *Annu. Rev. Psychol.* **51**, 149–169. (doi:10.1146/annurev.psych.51.1.149)
7. Jungherr A. 2016 Twitter use in election campaigns: a systematic literature review. *J. Inf. Technol. Polit.* **13**, 72–91. (doi:10.1080/19331681.2015.1132401)
8. Ribeiro HV, Mendes RS, Lenzi EK, del Castillo-Mussot M, Amaral LAN. 2013 Move-by-move dynamics of the advantage in chess matches reveals population-level learning of the game. *PLoS One* **8**, e54165. (doi:10.1371/journal.pone.0054165)
9. Pastor L, Veronesi P. 2009 Learning in financial markets. *Annu. Rev. Financ. Econ.* **1**, 361–381. (doi:10.3386/w14646)
10. García D, Tanase D. 2013 Measuring cultural dynamics through the Eurovision Song Contest. *Adv. Complex Syst.* **16**, 1350037. (doi:10.1142/S0219525913500379)
11. Dekker A. 2007 The Eurovision Song Contest as a 'friendship' network. *Connections* **27**, 53–58.
12. Yair G. 1995 'Unite Unite Europe' the political and cultural structures of Europe as reflected in the Eurovision Song Contest. *Soc. Networks* **17**, 147–161. (doi:10.1016/0378-8733(95)00253-K)
13. Yair G. 2019 Douze point: Eurovisions and Euro-divisions in the Eurovision Song Contest—review of two decades of research. *Eur. J. Cult. Stud.* **22**, 1013–1029. (doi:10.1177/1367549418776562)
14. Yair G. 2022 A march for power: the variety of political programmes on the Eurovision Song Contest stage. In *The Eurovision Song Contest as a cultural phenomenon* (eds A Dubin, D Vuletic, A Obregón), pp. 83–95. London, UK: Routledge. (doi:10.4324/9781003188933-8)

15. Salganik MJ, Dodds PS, Watts DJ. 2006 Experimental study of inequality and unpredictability in an artificial cultural market. *Science* **311**, 854–856. (doi:10.1126/science.1121066)
16. Rosati DP, Woolhouse MH, Bolker BM, Earn DJD. 2021 Modelling song popularity as a contagious process. *Proc. R. Soc. A* **477**, 20210457. (doi:10.1098/rspa.2021.0457)
17. Votter M, Mayerl M, Specht G, Zangerle E. 2021 Novel datasets for evaluating song popularity prediction tasks. In *2021 IEEE Int. Symp. on Multimedia (ISM)*, *Naples, Italy, 29 November–1 December 2021*, pp. 166–173. (doi:10.1109/ISM52913.2021.00034)
18. Kamal J, Priya P, Anala MR, Smitha GR. 2021 A classification based approach to the prediction of song popularity. In *2021 Int. Conf. on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, *Chennai, India, 24–25 September 2021*. (doi:10.1109/ICSES52305.2021.9633884)
19. Lee J, Lee JS. 2018 Music popularity: metrics, characteristics, and audio-based prediction. *IEEE Trans. Multimedia* **20**, 3173–3182. (doi:10.1109/TMM.2018.2820903)
20. Kang I, Mandulak M, Szymanski BK. 2022 Analyzing and predicting success of professional musicians. *Sci. Rep.* **12**, 21838. (doi:10.1038/s41598-022-25430-9)
21. Wasserman M, Zeng XHT, Amaral LAN. 2015 Cross-evaluation of metrics to estimate the significance of creative works. *Proc. Natl Acad. Sci. USA* **112**, 1281–1286. (doi:10.1073/pnas.1412198112)
22. Wasserman M, Mukherjee S, Scott K, Zeng XHT, Radicchi F, Amaral LAN. 2015 Correlations between user voting data, budget, and box office for films in the internet movie database. *J. Assoc. Inf. Sci. Technol.* **66**, 858–868. (doi:10.1002/asi.23213)
23. Helbing D. 1996 A stochastic behavioral model and a 'Microscopic' foundation of evolutionary game theory. *Theory Decis.* **40**, 149–179. (doi:10.1007/BF00133171)
24. Centola D. 2010 The spread of behavior in an online social network experiment. *Science* **329**, 1194–1197. (doi:10.1126/science.1185231)
25. Creanza N, Kolodny O, Feldman MW. 2017 Cultural evolutionary theory: how culture evolves and why it matters. *Proc. Natl Acad. Sci. USA* **114**, 7782–7789. (doi:10.1073/pnas.1620732114)
26. Balietti S, Goldstone RL, Helbing D. 2016 Peer review and competition in the art exhibition game. *Proc. Natl Acad. Sci. USA* **113**, 8414–8419. (doi:10.1073/pnas.1603723113)
27. Wikipedia contributors. Eurovision Song Contest. See https://en.wikipedia.org/wiki/Eurovision_Song_Contest (accessed 5 March 2025).
28. Academy Platform contributors. Letras de canciones. See https://www.letras.com (accessed 5 March 2025).
29. Henard DH, Rossetti CL. 2014 All you need is love? Communication insights from pop music's number-one hits. *J. Advert. Res.* **54**, 178–191. (doi:10.2501/JAR-54-2-178-191)
30. Pham C, Hoyle A, Sun S, Resnik P, Iyyer M. 2024 TopicGPT: a prompt-based topic modeling framework (eds K Duh, H Gomez, S Bethard, K Duh, H Gomez, S Bethard). In *Proc. 2024 Conf. North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 2956–2984. Stroudsburg, PA: Association for Computational Linguistics. (doi:10.18653/v1/2024.naacl-long.164)
31. Wang H, Prakash N, Hoang NK, Hee MS, Naseem U, Lee RKW. 2023 Prompting large language models for topic modeling. In *2023 IEEE Int. Conf. on Big Data*, *Sorrento, Italy, 15–18 December 2023*, pp. 1236–1241. (doi:10.1109/BigData59044.2023.10386113)
32. Miyakawa S, Utsuro T. 2024 Emotion classification of lyrics through summarization by large language models. In *2024 IEEE Int. Conf. on Big Data*, *Washington, DC, USA, 15–18 December 2024*, pp. 2999–3006. (doi:10.1109/BigData62323.2024.10825406)
33. Revathy VR, Pillai AS, Daneshfar F. 2023 LyEmoBERT: classification of lyrics' emotion and recommendation using a pre-trained model. *Procedia Comput. Sci.* **218**, 1196–1208. (doi:10.1016/j.procs.2023.01.098)
34. Lei S *et al*. 2024 Songcreator: lyrics-based universal song generation. *Adv. Neural Inf. Process Syst.* **37**, 80107–80140.
35. Ma X, Wang Y, Kan MY, Lee WS. 2021 Generating music and vocabulary constrained lyrics. In *Proc. 29th ACM Int. Conf. on Multimedia MM '21*, pp. 1002–1011. New York, NY: ACM. (doi:10.1145/3474085.3475502)
36. Spotify. Spotify API. See https://developer.spotify.com/documentation/web-api (accessed 2 July 2024).
37. Panda G, Redinho H, Gonçalves C, Malheiro R, Paiva RP. 2021 How does the Spotify API compare to the music emotion recognition state-of-the-art? In *18th Sound and Music Computing Conf. (SMC 2021)*, pp. 238–245. Axea sas/SMC Network.
38. Nijkamp R. 2018 Prediction of product success: explaining song popularity by audiofeatures from Spotify data. BS thesis, University of Twente, The Netherlands.
39. Gulmatico JS, Susa JAB, Malbog MAF, Acoba A, Nipas MD, Mindoro JN. 2022 SpotiPred: a machine learning approach prediction of Spotify music popularity by audio features. In *2022 2nd Int. Conf. on Power, Control and Computing Technologies (ICPC2T)*, *Raipur, India, 1–3 March 2022*. (doi:10.1109/ICPC2T53885.2022.9776765)
40. Duman D, Neto P, Mavrolampados A, Toiviainen P, Luck G. 2022 Music we move to: Spotify audio features and reasons for listening. *PLoS One* **17**, e0275228. (doi:10.1371/journal.pone.0275228)
41. Hongpanarak T, Mongkolnavin J. 2021 A study of relationship between music streaming behavior and big five personality traits of Spotify users. In *Proc. 12th Int. Conf. on Advances in Information Technology IAIT '21*. New York, NY: ACM. (doi:10.1145/3468784.3469854)
42. Dmytro I. 2023 Music genres classification. See https://huggingface.co/dima806/music_genres_classification (accessed 26 July 2024).
43. Baevski A, Zhou Y, Mohamed A, Auli M. 2020 wav2vec 2.0: a framework for self-supervised learning of speech representations. *Adv. Neural Inf. Process Syst.* **33**, 12449–12460.
44. Sturm BL. 2013 The GTZAN dataset: its contents, its faults, their effects on evaluation, and its future use. (https://arxiv.org/abs/1306.1461)
45. Motschenbacher H. 2016 *Language, normativity and Europeanisation: discursive evidence from the Eurovision Song Contest*. London, UK: Palgrave Macmillan. (doi:10.1057/978-1-137-56301-9)

46. Berghel H. 2018 Malice domestic: the Cambridge analytica dystopia. *Computer* **51**, 84–89. (doi:10.1109/MC.2018.2381135)
47. Van Valen L. 1973 A new evolutionary law. *Evol. Theory* **1**, 1–30.
48. Barnett WP, Hansen MT. 2007 The red queen in organizational evolution. *Strateg. Manag. J.* **17**, 139–157. (doi:10.1002/smj.4250171010)
49. Amaral L, Capozzi A, Helbing D. 2026 Supplementary material from: Breaking the Code: Multi-level Learning in the Eurovision Song Contest. Figshare. (doi:10.6084/m9.figshare.c.8440854)