A Nature Portfolio journal

0

https://doi.org/10.1038/s44271-025-00219-4

Temporal patterns in the complexity of child-directed song lyrics reflect their functions

Check for updates

Pierre Labendzki 1, Louise Goupil 12 & Sam Wass 1

Content produced for young audiences is structured to present opportunities for learning and social interactions. This research examines multi-scale temporal changes in predictability in Child-directed songs. We developed a technique based on Kolmogorov complexity to quantify the rate of change of textual information content over time. This method was applied to a corpus of 922 English, Spanish, and French publicly available child and adult-directed texts. Child-directed song lyrics (CDSongs) showed overall lower complexity compared to Adult-directed songs (ADsongs), and lower complexity was associated with a higher number of YouTube views. CDSongs showed a relatively higher information rate at the beginning and end compared to ADSongs. CDSongs and ADSongs showed a non-uniform information rate, but these periodic oscillatory patterns were more predictable in CDSongs compared to ADSongs. These findings suggest that the optimal balance between predictability and expressivity in information content differs between child- and adult-directed content, but also changes over timescales to potentially support multiple children's needs.

Infants are powerful learners, sensitive to statistical variabilities in their environment, and capable of selectively orienting their attention towards stimuli with the maximal potential for learning^{1,2}. It has been suggested that adults structure the information that they present to young learners, whether implicitly or explicitly, to maximise learnability³. In other words, the form of child-directed content typically reflects their functions⁴. This has been shown, for example, for child-directed speech (CDS)⁵, child-directed gestures⁶, child-directed television⁷, and child-directed singing⁸.

The first and best-known way in which we alter the information content of stimuli that we present to young listeners for learnability is by structuring it to allow for the signal to be easily differentiable from the noise⁹⁻¹⁴. For example, in many languages, CDS contains exaggerated pitch, elongated words, and expanded vowel space with stretched formant frequencies¹⁵. Previous research suggests that these changes entail exaggerating the critical features that distinguish one phoneme from another¹⁶⁻¹⁸. Children learn better from CDS than they do from adult-directed speech¹⁹, and CDS evokes greater attention-related electrophysiological responses than ADS¹⁷. This suggests that the specific form of CDS is tailored to its transmissive function, namely, to allow infants to learn the regularities of their native language.

A second way in which we alter the content of information that we present to young listeners is by balancing the trade-off between predictability and complexity (see Fig. 1). An upcoming stimulus is highly predictable when a particular learner can fully anticipate and encode it based on their prior knowledge without too much cognitive effort (i.e., when the agent's mental model of the environment does not need to be expanded to process the stimulus fully). Contrarily, a stimulus is complex when it is relevant for the individual by yielding substantial information gain (reduction of uncertainty) for a particular learner in a particular environment. These two needs (predictability and complexity), which are fundamental and antagonist, are essential to enable learners to reduce their uncertainty by creating models of their environment in order to make future predictions²⁰⁻²³, but also aid effective communication and cooperation²⁴⁻²⁷. This idea has been extensively discussed in the context of adult-directed content, such as music²⁸⁻³¹. The balance between these two opposite forces, often referred to as the Goldilocks zone, has also been shown to govern infants' allocation of their attention to visual and auditory stimuli^{32,33}. For instance, 8-month-old infants already orient their attention to maximise learning², preferring stimuli that cause a model to update (the complexity force) while still being able to encode novel information (the predictability force)²⁶. Importantly, the predictability-complexity trade-off is not an absolute constant universal value but is modulated depending on the intention, context, and agents' prior representations and past experiences.

This balance between predictability and complexity is also particularly important when considering the acoustic properties of child-directed songs³⁴⁻³⁷. Previous research has investigated these differences from the perspective of fine-grained acoustics: for example, we know that similar acoustic differences exist between child- and adult-directed music as

¹Department of Psychology, University of East London, London, UK. ²Université Grenoble Alpes, CNRS, LPNC, Grenoble, France. 🖂 e-mail: u2176677@uel.ac.uk

Fig. 1 | The preferred information rate is the result of a balance between predictability and complexity. The mean (preferred information rate) and standard deviation (tolerance for a different information rate) can change over development, intention, and socio-environmental context. For instance, speakers spontaneously increase their speech redundancy when talking in a noisy environment to maintain the same signal-to-noise ratio¹¹. CR (Compression Ratio) indicates complexity measured as the length of the shortest description of the message divided by the length of the original message (see Methods—Text informativeness metric). A small CR (~0) indicates low information density, high signal-to-noise ratio, and high predictability; a large CR (\simeq 1) indicates high information density, low signal-to-noise ratio, and low predictability.



between child- and adult-directed speech⁵; and that, in music, phonetic and acoustic information are mapped more clearly than in speech alone³⁸, which increases intersensory redundancy and thus guides perceptual differentiation³⁹. Even 10-month-old infants can parse a continuous auditory stream into discrete word segments from a song⁴⁰, and speech sung along with music helps language acquisition³⁸. Furthermore, caregivers spontaneously alter the information-theoretic characteristic of their prosodic contours, with pitch trajectories being less predictable in infant-directed speech than in adult-directed speech, possibly to attract infant attention by inducing uncertainty-driven curiosity⁴¹.

However, thus far the literature has focused on static measures of predictability in information content. Studies have looked at how a single time-invariant information content measure of a specific stimulus impacts infants' attention or learning. Here, we concentrate instead on how information content is structured over time in child-directed content. To do so, we focused on a different aspect of child-directed songs, namely, the form of the information input carried over time by the song lyrics, in terms of textual predictability. Extensive evidence suggests that semantic and orthographic complexity tend to be aligned in language^{26,42-45} (for example, simpler contexts tend to have similar forms). Even though child-directed songs like nursery rhymes contain more than textual information, for instance, melody, harmony, rhythm, and movement⁴⁶, or in this case, even video, we focus here on textual predictability as a relevant feature to investigate because of the role of nursery rhymes in word learning. We acknowledge that the overall complexity carried by the song is not solely dependent on the textual content. One of the limitations of studying only the form of lyrics, in a medium that also contains musical features, is that we are missing some of the cues that infants use to help them encode the audiostream⁴⁷. However, as the vocal part of the nursery rhymes attract infants' attention more than an instrumental version⁴⁸, and as vocals are present in more than half the musical experience of infants⁴⁹, the lyrical content alone should still be constrained by the infant's needs as they are a salient part of the audiostream. Specifically, we hypothesised that textual information is introduced differently in child-directed songs (CDSongs) compared to adult-directed songs (ADSongs) and child-directed stories (CDStories). We think that those characteristics emerged from a long evolutionary process (through copy, variation, competition, and selection) carried by children⁵⁰ along with caregivers⁵¹. By looking at the shape of textual content as cultural artefacts⁵², we can try to unveil the selective pressures that shaped them in the first place.

We hypothesise that the form of CDSongs lyrics should be determined by the functions they have and the multiple caregiving roles that they support. In other words, the introduction of semantic information through lyrics, like any other musical feature, is potentially shaped to fulfill specific needs for each agent^{51,53}. Accordingly, we posited that the differences in the patterns of textual information introduction, between CDSongs and ADSongs, and between CDSongs and CDStories, stem from the different functions of each register. CDSongs are used by caregivers across cultures^{54,55} to assist them on a variety of tasks⁵⁶; from early social bonding^{57,58}, and soothing^{59,60}, to word learning⁶¹⁻⁶³ and collective creative endeavours^{64,65}. In addition to acoustic differences between CDSongs, ADSongs, and CDStories, the form of information introduction patterns in CDSongs could also partly explain infants' preference for them compared to ADSongs⁶⁶ and CDSpeech⁶⁷. Accordingly, we compared the form of CDSongs to ADSongs and the form of CDSongs to CDStories. We expected the CDSongs form to differ from the ADSongs form by showing a relative predictability level adapted to a younger audience, in line with the idea that CDSongs are used by caregivers to teach new words, soothe, and enhance social bonding with their child. We also expected CDSongs to differ from CDStories because of the more salient role of music for social bonding and joint creation compared to storytelling65,68,69, but also because of the differences in their performance by caregivers: namely that CDSongs are usually performed without written support contrary to CDStories.

We examined differences in information content between CDSongs, ADSongs, and CDStories in three analyses spanning nested timescales. In the first analysis, we compared the overall complexity of the three text types (CDSongs, ADSongs, and CDStories). Complexity was measured by calculating the Compression Ratio (CR), which is the length of the shortest description of the message divided by the length of the original message. This approach, based on minimum length description⁷⁰, gives us insight into the relative quantity of information contained within the text (see Methods). Our first analysis assessed whether there was a relationship between overall complexity and the song's popularity, indexed as the number of views of that content on YouTube, normalised by time since the content had been posted (Analyses 1a). Based on the Goldilocks principle^{32,33}, we predicted that popularity should follow an inverted U-shaped as a function of complexity, meaning that the most popular songs should be of intermediate complexity. Our second hypothesis was that CDSongs have greater redundancy in the information content presented, compared to ADSongs or CDStories (Analyses 1b), manifesting as lower complexity. This lower complexity compared to ADSongs should result from infants' improved word encoding for higher redundancy language^{71,72}. CDSongs lower complexity compared to CDStories should result from the caregiver's constraint to have easily remembered lyrics to sing without written support^{68,69,73}.

In the second analysis (Analysis 2), we examined the rate of change of information, in order to discover: while presenting a set quantity of information, when should the information rate be higher? According to the Serial-Position Effect, information presented at the beginning (primacy effect) and at the end (recency effect) have greater recall chances than information presented in the middle of a sequence⁷⁴, and this effect is present from 6 months⁷⁵. This phenomenon was also measured for long-term

Table 1 | Number of text per condition and languages

	English	Spanish	French	Total
CDSongs	207	67	39	313
ADSongs	100	96	71	267
CDStories	146	100	96	342
Total	453	263	206	922

recollection of real-world lyrics^{76,77}. We thus predicted that CDSongs will take advantage of the Serial-Position Effect, i.e., show a higher relative information rate at the beginning and end compared to ADSongs.

Our final set of analyses (Analysis 3) examined the rate of change of new information over time. For this analysis, we examined temporal regularities in the timing with which new information was presented. Temporal regularities are present at numerous levels in early child-caregiver interactions, from sub-second level variabilities in amplitude fluctuations in CDS78 to periodic fluctuations at the time-scale of seconds, minutes, and hours in socio-communicative behaviours and stress physiology^{79,80}. These temporal regularities may serve important functions-first, they allow for the development of predictions^{22,81}; and second, they allow for oscillatory entrainment to develop between periodic endogenous oscillations and external information content^{80,82}. In line with the Smooth Signal Redundancy Hypothesis⁸³ and the Uniform Information Density Principle^{84,85}, we predicted that, compared to CDSongs and ADSongs, CDStories should distribute the information load equally throughout the transmission, in order to be more robust to random noise^{86,87}. Additionally, consistent with the micro-level regularities already noted at the sub-second level in CDS, we expected to observe macro-level oscillatory structure at the second to minute scale in how new information content was presented in CDSongs and ADSongs, and that oscillations in information predictability should be less variable and thus more easily trackable in the CDSongs compared to ADSongs. Endogenous oscillatory activity is present at multiple levels from early life, and evidence from a variety of sources suggests that oscillatory patterns in external information content facilitate learning⁸⁰.

Methods

Analyses were not pre-registered.

Corpora

We identified three contemporary databases of child-directed songs (CDSongs), adult-directed songs (ADSongs), and child-directed stories (CDStories) in three languages (English, Spanish, and French). Our analyses in the main text are based on the entire corpus; in addition, in the Supplementary Materials, we also include a breakdown of analyses into individual language corpora.

Table 1 shows a breakdown of the number of individual texts included per language. For the English language content, the CDSongs is composed of 313 song lyrics from traditional and contemporary nursery rhymes taken from the YouTube channel *Cocomelon—Nursery Rhymes* (e.g., Incy Wincy Spider, Wind the Bobbin Up, etc—see SM Section 4 for a full list). The Spanish corpora is taken from the Youtube channel *Little Angel Español— Canciones Infantiles* (e.g., Cinco Patitos Salieron a Nadar, El Viejo Mac-Donald tenía una Granja, etc) and French language CDSongs were taken from the Youtube channel *Titounis—Comptine bébé—Chanson enfant* (e.g., Frère Jacques, Il était un petit navire, etc).

The ADSongs corpus is composed of 267 song lyrics (100 song lyrics from the UK top 100 Charts from 07/11/2022; Spanish content from Youtube playlist "Pop Latino 2022–2023 Videos Oficiales" (96 song lyrics); French content from the Youtube playlist "French top songs (most viewed 2022)") (71 song lyrics). The child-directed story corpus (CDStories) is composed of 342 Æsop fables, (146 English Æsop fables, 100 Spanish Æsop fables, and 96 French Æsop fables). The full list of titles included for all three corpora is given in the SM (Section 4). For CDSongs, and ADSongs, the lyrics were extracted either from the video description, as provided by the YouTube channel, or from an online lyrics catalogue, and the accuracy of the lyrics was manually checked for all of them. For CDStories, the corpus was directly extracted as a text file. In order to assess whether the textual content of these transcriptions, which is what we are analysing in this paper, is identical between the versions that we used and other transcriptions of the same nursery rhymes, songs, and stories extracted from different online sources, we also included some additional analyses in the Supplementary Materials (see SM Section 2), in which we compare the textual similarity of the same content taken from different sources online. These analyses confirm that the written text transcriptions used by Cocomelon, for example, are highly similar to the texts used by other sources.

The mean (std) length in characters (individual letters and punctuations) of the texts extracted was 1023.31 (382.83) for CDSongs, 1603.55 (898.70) for ADSongs, and 955.41 (411.02) for CDStories. Because of these differences in mean lengths, for analyses concerning overall complexity (Analyses 1.a and 1.b), we present results based on normalised data, i.e., divided by the length of the original text.

YouTube views as a popularity metric

For Analysis 1a, following other previously published papers⁸⁸⁻⁹⁰, we extracted the number of YouTube views as a measure of popularity. Viewer recommendations for YouTube are made based on the viewer's personal history and what others with similar viewing history are choosing, as described in the following references. In these pages^{91,92}, YouTube describes how the aim of their algorithms is to mimic and continue the selection process initiated by children and caregivers, and describe in some detail the exact metrics that YouTube uses to measure this (e.g., watchtime and repeat views).

Text informativeness metric

Kolmogorov complexity is the quantitative aspect of the principle of Minimum Length Description. The Kolmogorov complexity K(s) of a finite discrete sequence *s* is the length of the shortest programme that generates *s*. Such a programme that can losslessly describe s in K(s) symbols can be thought of as the representation of the original message s but reduced to its core information without any redundancy⁹³. Exact Kolmogorov complexity is not computable, but lossless compression algorithms offer a reasonable approximation⁹⁴. PyLZMA, a Python implementation of the Lempel-Ziv-Markov chain algorithm, was used for its compression power and simplicity of use⁹⁵ (see Fig. 2). LZMA algorithms use dynamic dictionaries to count the occurrences of every n-sequence and then use Huffman coding to allocate shorter symbols to the most frequent string sequences. This method allows tracking of statistical regularities not only in the zero-order, character frequency scale, but also in hierarchically increasing n-sequences of characters, to match human language comprehension⁹⁶. The size of the compressed version represents the actual number of characters needed to losslessly encode the sequences. As the cumulative window of analysis increases, the compressed size increases when the model needs updating to encode the sequence (Fig. 3A). This increasing cumulative window analysis (left-conditioned) is supposedly closer to how we perceive an incoming continuous stream of information: not as a sliding window but as an integrative one⁹⁷.

We examined how new information was introduced over time by calculating the derivative of the cumulative window. As the cumulative window of analysis increases, when the model needs updating to encode the sequence, the compressed size increases and thus the derivative becomes positive (Fig. 3B). This can also be seen as the difference in minimum description length between consecutive timepoints.

This LZMA lossless cumulative compression process can be seen as a model for predictive processing⁹⁸ and for cognitive constructivism^{20,99,100} where data are sequentially unveiled to a patternseeking agent that keeps track of past statistical regularities to fit the incoming data into a condensed representation (see Fig. 2). This



low-level statistical technique, already in use for image analysis¹⁰¹, can also be adapted for textual analysis^{102,103}.

A CR was calculated, defined as the compressed size of the lyrics divided by the total length of the uncompressed lyrics to control for the fact that each story/song had a different length (see Fig. 2). A CR close to 1 indicates that no compression was done because of maximal information content, whereas a CR close to 0 indicates minimal information content in the lyrics. The CR was calculated iteratively using a cumulative window that increased in increments of one character between iterations. Analyses 1a and 1b are based on the average CR of the entire song, which is effectively the complexity of the song, standardised across different song lengths.

For Analyses 2 and 3, we examined how complexity changed over time during each song or story in two ways. Analysis 2 examines coarse-grained changes over the course of the entire song. To this end, we calculated the point within the progression of the song when given percentages of information content (25, 50, 75, 85, 95%) were reached (Fig. 4C for total result and see Fig. 5C for illustrative examples), and compared this between corpora using an analysis of variance (ANOVA). We also summed the derivative of the compressed size into 8 equal bins, divided those sums by the total amount of new information, and compared the amount of information in the first and last segments (segment 1 vs 8) using an ANOVA (see Fig. 4D).

Analysis 3 examines fine-grained moment-by-moment changes. We conducted three analyses to examine fine-grained changes in the derivative of the compressed size during the song. First, we plotted the compressed size against the progression of the song (see Fig. 3A) and calculated the coefficient of determination R-squared¹⁰⁴ (R^2). A high R^2 indicates that the information profile closely matches a straight line, indicating that the information is introduced uniformly over the entire course of the story/ song, which matches what would be expected under the Uniform Information Principle. Second, after computing the derivative of the compressed size (see Fig. 3B) we calculated its standard deviation. A high std derivative indicates that new information is introduced at variable rates along the text, with periods of high informativeness and periods of high redundancy, capturing the staircase aspect of the information profile.

Third, we assessed oscillatory structures in the introduction of information content. We could not do this using standard frequency domain analyses, as the period and phase of the oscillations differed between songs of different lengths. Instead, we computed the average span (the average character length) and variability (std) of redundancy (i.e., low complexity) periods within one text. Periods of redundancy were computed as sequences where the derivative of the compressed size was lower than a fixed threshold, chosen to differentiate high from low information input moments (Fig. 3C). This threshold was set arbitrarily at a value of 3; but to check that the results were not dependent on this specific threshold, the same set of analyses was repeated with each possible threshold and results were found to be consistently significant regardless of the threshold (see SM section 1). Less within-song variability in the span of the redundancy periods indicates more periodic fluctuations in the rate of change of information content over time.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Results

Our results section is in four parts. First, in Analyses 1a and 1b, we examine differences in overall complexity (standardised for length between songs). In Analysis 1a we examine the relationship between overall complexity and number of YouTube views. In Analysis 1b we compare differences in overall complexity between the different corpora (CDSongs, ADSongs, and CDStories).

Next, in Analyses 2 and 3, we examine differences in the rate of change of information content. In Analysis 2 we examine coarse-grained changes in how the rate of change of new information content changed over the course of the song, using two separate analyses. In Analysis 3 we examine finegrained changes in how the rate of new information content changed within the song, using three separate analyses.

All analyses were conducted using ANOVAs as the texts being examined came from different authors, and so the assumption that observations are independent was verified. In the Supplementary Materials, we have also included the full breakdown for English, Spanish and French corpora considered separately (Section 3). Overall, individual languages almost always show similar results and directions, although not always in a significant way, so unless specified, the results were similarly significant for all languages taken separately.

Analysis 1a-overall complexity and number of views

First, we examined whether there was a relationship between stories' and songs' overall complexity (total amount of information, as indexed by the compressed size, standardised by song length) and the number of views (standardised by number of days since the video was uploaded). Our prediction was that the relationship would be an inverted U, in line with the Goldilocks principle. To test this, we fitted both a linear and quadratic model to the CDSongs corpus, and found that the linear model $F_{(1, 311)} = 17.99, p < 0.001, AIC = 752.7, BIC = 760.2$ was a better fit than the quadratic model $F_{(2, 310)} = 10.28$, p < 0.001, AIC = 752.2, BIC = 763.4. The linear relationship observed was a negative one, indicating that increased complexity is associated with fewer views (Fig. 4A). However, we did find the expected inverted U shape for the ADSongs, as the quadratic model was significant $(F_{(2, 264)} = 12.19, p < 0.001,$ AIC = 588.7, BIC = 599.5), but the linear one was not significant (p = 0.33). No popularity index was available for the CDStories corpus. Virtually identical patterns were observed when we examined the English, Spanish, and French corpora independently (Fig. SM3-5). Overall, these results show that only in the CDSongs corpus was decreased complexity associated with increased views and that this relationship is a linear one rather than the predicted inverted U shape. By contrast, the adult corpus reflected the expected inverted U shape relationship between complexity and popularity.



Fig. 3 | **Illustration of the cumulative lossless compression technique.** Example of a compressed size analysis with a cumulative window (**A**) and its derivative (**B**) at the beginning of the song "Baby Shark". **C** shows onset and offset of periods of high

(plateau = 1) and low (plateau = 0) derivative of compressed size, identified with a threshold of 3. All time series are plotted over character progression.

Analysis 1b—comparison of overall complexity between corpora We first ran a one-way ANOVA to test whether there were any differences in overall complexity (indexed by the CR) between the CDSongs, ADSongs, and CDStories corpora. This showed a significant main effect of condition $(F_{(2, 914)} = 978.5, p < 0.001)$. We thus ran a series of pairwise comparisons between all three conditions (Fig. 4B). There was a significant difference in CR between CDSongs and ADSongs, $t_{(578)} = -4.03$, p < 0.001 $(M_{\text{CDSongs}} = 0.322, std_{\text{CDSongs}} = 0.137; M_{\text{ADSongs}} = 0.362, std_{\text{ADSongs}} = 0.091,$ Cohen's d = -0.33, $CI_{CDSongs} = [0.30; 0.33]$, $CI_{ADSongs} = [0.35; 0.37]$), with CDSongs having lower complexity, although there was no statistically significant differences in the Spanish and French corpora taken individually. We also compared the CR between CDSongs and CDStories, and found it to be higher for the CDStories corpus compared to the CDSongs corpus, $t_{(653)} = -39.1, p < 0.0001$ ($M_{CDStories} = 0.647, std_{CDStories} = 0.065,$ Cohen's d = -3.05, CI_{CDStories} = [0.64; 0.65]), this was found for all languages taken separately (see SM3, SM4, SM5). Another pairwise comparison revealed that ADSongs had significantly lower average CR than CDStories $(t_{(607)} = -44.80, p < 0.0001,$ Cohen's d = -3.65), this was found for all languages taken separately (see SM3, SM4, SM5). Overall, these results suggest that CDSongs tend to be less complex than ADSongs, and that both CDSongs and ADSongs are less complex than CDStories.

Analysis 2—rate of change of new information content—changes across song

In order to quantify changes in information input over time we computed the derivative of the cumulative compressed size. A positive and high derivative indicates new information occurring, and a derivative close to zero indicates that a segment is redundant—i.e., that it adds no novel information compared to the cumulative amount of information already introduced thus far (see Fig. 3B). Samples of raw data from this analysis are shown in Fig. 5A, C.

In order to quantify whether the way information is introduced over time differed between corpora we performed two analyses. First, we calculated the point within the progression of the song when given percentages of information content (25, 50, 75, 85, 95%) were reached (see Fig. 4C for average results and Fig. 5B for raw data illustration). We then calculated a two-way ANOVA to assess differences between corpora (CDSongs, ADSongs, and CDStories). The ANOVA revealed an interaction between the condition and the percentage of total information content $(F_{(8, 914)} = 2150, p < 0.0001)$. We thus ran a series of pairwise comparisons between conditions. Pairwise comparison showed that CDSongs reached 25% of their total amount of information earlier than the ADSongs, $t_{(578)} = -4.58$, p < 0.001 ($M_{CDSongs} = 0.113$, $std_{CDSongs} = 0.051$; $std_{ADSongs} = 0.051$, Cohen's d = -0.38, $M_{\rm ADSongs} = 0.131$, CI_{CDSongs} = [0.108; 0.119], CI_{ADSongs} = [0.126; 0.135]) although no statistical difference was observed in the Spanish corpus taken individually; and that CDSongs songs reached 95% of their total amount of information later than ADSongs songs, $t_{(578)} = 3.32$, p < 0.0001 ($M_{CDSongs} = 0.847$, $std_{CDSongs} = 0.171; M_{ADSongs} = 0.804, std_{ADSongs} = 0.133, Cohen's d = 0.27,$ CI_{CDSongs} = [0.828; 0.866], CI_{ADSongs} = [0.788; 0.820]) although there was no significant statistical difference in the French corpus taken individually. This can be interpreted as an earlier fade in and a later fade out (Fig. 4C) of information input in CDSongs compared to the ADSongs. CDSongs also reached 25% of their total amount of information sooner than the CDStories, $t_{(653)} = -27.5$, p < 0.0001 (M_{CDStories} = 0.193, std_{CDStories} = 0.011, Cohen's d = -2.15, CI_{CDStories} = [0.151; 0.159]), but they reached 95% of their total amount of information sooner than the CDStories, $t_{(653)} = -10.0$, p < 0.0001 (M_{CDStories}=0.941, std_{CDStories} = 0.008, Cohen's d = -0.78,



Fig. 4 | CDSongs are shown blue (n = 313), ADSongs in red (n = 267), and CDStories in green (n = 342). A shows the relation between the log normalised number of YouTube views as a function of complexity (i.e., the Compression Ratio). B shows the Compression Ratio (total amount of information divided by length) across conditions. C shows the relative progression at which texts reach 25, 50, 75, 85 95% of their total information content. For instance, CDSongs reached 95% of their

CI_{CDStories} = [0.940; 0.942]), similar results were found for all languages ho taken individually (Fig. SM3–5). This can be interpreted as an earlier fade in and fade out of information input in CDSongs compared to the CDStories.

Second, as a complementary analysis, we summed the derivative of the compressed size into 8 bins and then divided those sums by the total amount of information (final compression size) (Fig. 4D for results and Fig. 5D for raw data illustration). This allows us to compare the dynamics of information input across songs with different overall information content. In order to test whether the beginning and end had higher probability of new information, a two-way ANOVA was performed to analyse the effect of corpora (CDSongs, ADSongs, and CDStories) and progression (segment from 1 to 8) on the probability of information input. The ANOVA revealed an interaction between condition and progression segments $(F_{(8, 914)} = 59.91, p < 0.0001)$. We thus ran a series of pairwise comparisons between conditions. CDSongs showed higher values compared to ADSongs in the first $t_{(578)} = 3.33$, p < 0.0001; and eighth segment $t_{(578)} = 7.69$, p < 0.0001, although the higher value in the first segment was not significantly different in Spanish and French considered separately. CDSongs also showed higher values compared to CDStories in the first segment, $t_{(653)} = 17.9$, p < 0.0001; however, CDSongs showed lower information input in the eighth segment compared to CDStories, $t_{(653)} = -9.42$, p < 0.0001. Similar patterns were observed when the English, Spanish, and French corpora were considered independently (see Fig. SM3-5). This is fully consistent with the prior analysis in showing differential patterns in

total information content later than ADSongs. **D** shows the probability of information input in segments. **E** shows the R^2 values of the compressed size across conditions. **F** shows the standard deviation of the derivative of the compressed size. **G** shows the mean redundancy period duration in each song. **H** shows within-song redundancy period variability (std). Fig. SM3–5 shows the same analyses broken down into English, French, and Spanish corpora separately.

how information is introduced towards the beginning and end of CDSongs as compared to both ADSongs and CDStories. Importantly, though, this probability of information input over progression comes from a normalisation by overall information content, and thus we can only make claims about the relative repartition of information in each text type, not their absolute information content.

Overall, these results suggest that: i) Relatively to the overall information content of each song, more novel information is presented at the beginnings and ends of songs in CDSongs compared with ADSongs; ii) When looking at the beginning, middle, and end, the introduction of novel information is more evenly distributed along the text in CDStories compared to CDSongs and ADSongs.

Analysis 3—rate of change of new information content—finegrained changes

In order to compare the micro differences in information input between CDSongs, ADSongs, and CDStories, we first plotted the derivative of the compressed size against the progression of the song and calculated the R^2 (see Fig. 4E). R^2 was found to be different across conditions $F_{(2, 914)} = 143.2$, p < 0.0001, with the R^2 in the CDSongs condition being lower compared to the ADSongs, $t_{(578)} = -4.90$, p < 0.001 ($M_{\text{CDSongs}} = 0.882$, $std_{\text{CDSongs}} = 0.132$; $M_{\text{ADSongs}} = 0.926$, $std_{\text{ADSongs}} = 0.068$, Cohen's d = -0.408, $\text{CI}_{\text{CDSongs}} = [0.868; 0.897]$, $\text{CI}_{\text{ADSongs}} = [0.918; 0.935]$) (although there was not statistically significant difference in Spanish); and



Fig. 5 | CDSongs (CoComelon—Animal Dance) is shown in blue, ADSongs (Taylor Swift—Anti-Hero) in red, and CDStories (Aesop—The Fox and the Crow) in green. A Raw compressed size over song progression (in characters, i.e., individual letters and punctuation). B Standardised compressed size adjusted to have similar length (1000) and similar final compressed size (100), over song

progression (in characters). **C** Derivative of standardised compressed size over song progression (in characters). **D** Probability of new information over segment progression computed from the derivative divided by the total amount of information and binned into eight segments.

compared to CDStories, $t_{(653)} = -15.6$, p < 0.0001 ($M_{\text{CDStories}} = 0.995$, $std_{CDStories} = 0.002$, Cohen's d = -1.22, CI_{CDStories} = [0.995; 0.996]). We also found that the R^2 were significantly lower in the ADSongs compared to the CDStories, $t_{(607)} = -18.4$, p < 0.0001. This indicates that the rate at which new information is presented during each item is more evenly distributed in CDStories compared with songs, and in ADSongs compared with CDSongs. Second, we calculated the standard deviation of the derivative of the compressed size (Fig. 4F). In line with the previous finding, we also found differences in the variability of information introduction, $F_{(2, 914)} = 988.2$, p < 0.0001, with variability being significantly higher in the CDSongs condition compared to the CDStories condition, $t_{(653)} = 38.2.6$, p < 0.001 $(M_{\text{CDSongs}} = 3.10, std_{CDSongs} = 0.47; M_{\text{CDStories}} = 2.00, std_{\text{CDStories}} = 0.22,$ Cohen's d = 2.98, $CI_{CDSongs} = [3.050; 3.155]$, $CI_{CDStories} = [1.981; 2.030]$). Variability of information introduction was also higher in the ADSongs condition compared to the CDStories condition, $t_{(607)} = 46.3$, p < 0.0001 $(M_{ADSongs} = 3.15, std_{ADSongs} = 0.37, Cohen's d = 3.78, CI_{ADSongs} = [3.110;$ 3.201]), however no statistically significant difference was found between CDSongs and ADSongs (p = 0.15). Overall we found more variability in the introduction of information in the CDSongs and ADSongs compared to CDStories, and this was found in all languages considered separately (Fig. SM3–5). Taken together, the lower R^2 values and higher variability of information introduction in CDSongs and ADSongs compared to CDStories, indicates that in CDSongs and ADSongs, information is more likely to be allocated in bursts with periods of high informativeness and periods of high redundancy.

The CDStories showed an information profile that closely matched a straight line (Fig. 4E, i.e., R^2 values close to 1), so we excluded them from the next analysis. To further examine the oscillating patterns in the two remaining conditions, we computed the mean and variability of the redundancy periods, when less (or no) new information is presented (Fig. 3B; here, we present results using a threshold of 3, see Methods and SM

section 1 for different threshold values). There was a main effect of conditions on variability (F_(2, 562) = 98.16, p < 0.001, see Fig. 4G, H), where variability was measured as the standard deviation of the redundancy spans within one song. The variability was lower for CDSongs compared to ADSongs, $t_{(562)} = -9.90$, p < 0.0001 ($M_{\text{CDSongs}} = 43.21$, $std_{\text{CDSongs}} = 45.37$; $M_{\text{ADSongs}} = 86.22$, $std_{\text{ADSongs}} = 57.50$, Cohen's d = -0.83, $\text{CI}_{\text{CDSongs}} = [38.04; 48.39]$, $\text{CI}_{\text{ADSongs}} = [79.27; 93.16]$), this was found even when considering language separately. We also found that the average redundancy spans were shorter in the CDSongs compared to the ADSongs condition, $t_{(652)} = -2.29$, p < 0.025 ($M_{\text{CDSongs}} = 59.33$, $std_{\text{CDSongs}} = [51.20; 67.46]$, $\text{CI}_{\text{ADSongs}} = [65.17; 78.02]$), although there was no statistically significant differences in English and Spanish considered separately.

Taken together, our findings show that within-song variability and duration in the spans of the redundancy periods is lower in CDSongs than in ADSongs. This indicates that fluctuations in the rate of change of information content over time are more periodic and less spaced out in CDSongs compared to ADSongs.

Discussion

We compared the rate of change of information content between three open-source corpora: child-directed song (CDSongs), adult-directed song (ADSongs) and child-directed stories (CDStories), in English, Spanish and French.

In Analysis 1a, we examined the relationship between complexity and popularity in CDSongs (number of YouTube views) (Fig. 4A). The inverted U shape relationship predicted based on the literature^{31–33} was observed for the ADSongs corpus. For CDSongs, contrary to our predictions, we did not find the same inverted U shape. Instead, we observed that decreased complexity was associated with increased popularity. Three reasons might



Fig. 6 | On the left—preference vs complexity: different hypothetical distributions of information rate processing capacities. Each curve representing the relationship between preference and complexity is a balance between opposing needs: predictability and complexity (see Fig. 1). Those needs can change according to contexts, and those changes are represented by different means and spread, with their mean value extended to the right-hand plot. From top to bottom: a preference for high complexity with a low tolerance (small range), a preference for medium complexity with a wide tolerance, and a preference for low complexity with medium tolerance. On the right- complexity vs time: CDSongs, ADSongs, and CDStories archetypal complexity temporal profiles (in blue, red, and green, respectively) and different information rate processing capacities (dashed lines). A variable complexity profile can satisfy a wider range of preferences for complexity, as illustrated by the archetypal CDSongs curve having a rewarding complexity level for more than one preference distribution.

explain this. First, it remains possible, despite the evidence in its favour, that the Godlilocks effect is more restricted than previously thought, and cannot be observed using real textual information contents, and that young children do not prefer stimuli of intermediate predictability. Second, it is possible that the conditions to apply the Goldilocks principle might not be met here, e.g., YouTube users might select musical extracts that are more predictable for other reasons than knowledge transmission (e.g., self-soothing¹⁰⁵). Third, it is possible that the expected U shape was not observed because extremely predictable text sequences, for instance, a song in which just one word is repeated, are simply not present in the corpus, as lyricists and music producers would themselves consider these unlikely to be popular, and therefore not produce them in the first place.

In Analyses 1b, we found that overall CDSongs showed lower complexity than either ADSongs or CDStories (Fig. 4B), although this was not significant for the Spanish and French corpora considered individually. This is in line with the literature on CDSongs showing a reduced lexicon; but also "infra-word" structure and other predictable sequences, like counting or verse/chorus oscillations68,106,107. This increased redundancy could allow for more overall infant word encoding^{99,100,108}; for easier context recognition, in turn enhancing speech neural tracking¹⁰⁹ and word learning¹¹⁰. As CDSongs are usually sung without written support and are transmitted through oral tradition^{68,69}, they require sequence repetition to be more easily remembered, declared, and transmitted^{73,111,112}. Potentially, then, one could also posit that this reduced complexity is the result of generational transmission constraints¹¹³, where simpler messages are more replicated over time as they require fewer cognitive resources to be remembered and transmitted^{114,115}. The simpler shared songs would then reinforce social bonding (but also social outing) and intergenerational links, as people sing and share the same musical experience more easily, in turn increasing their chances of being replicated. Indeed, the more redundant form of CDSongs compared to CDStories can also be seen as a result of their different "storage of information" technics¹¹⁶. For CDStories, however, the redundancy constraint is reduced, as the textual trace largely assists the storytelling. The lower absolute complexity found in CDSongs compared to ADSongs and CDStories in Analyses 1b seems to point in that direction.

Second, in Analysis 2, we found that, even if the total amount of information was higher in ADSongs compared to CDSongs, CDSongs had more information at the beginning and end compared to ADSongs relative to their total amount of information (Fig. 4C, D), and individual language analysis reveal similar patterns. For example, in the song "Baby Shark" (Fig. 3), information is introduced at the beginning of the song ("Baby shark doo doo..."), followed by high levels of repetition/redundancy (i.e., introducing new family members following the same pattern and repeating "doo doo doo doo doo doo"), followed by new content at the end of the song ("Let's go dance doo doo..."). These patterns are in contrast to ADSongs, which seems to follow the general pattern of constant entropy rate instead, over the song progression, found in naturally occurring languages, consistent with the Smooth Signal Redundancy hypothesis and the Uniform Information Density^{117–119}.

In Analysis 3, consistent with this, we found that information was more evenly allocated, at an even finer time-scale in CDStories compared to CDSongs and ADSongs (Fig. 4F), and this was found for all languages even when considered individually. According to the Smooth Signal Redundancy Hypothesis⁸³, in order to convey information efficiently while still being robust to noise occurring at random times, one should use a code that distributes information and redundancy equally throughout the transmission^{86,87}. This is what we observed in the CDStories corpus but not the CDSongs and ADSongs corpora. CDStories are thus more optimal for efficient communication from an information theory perspective, as they use the information density principle to their advantage by having a near-constant information rate. This could potentially indicate that efficient information transmission is not the sole function of CDSongs and ADSongs.

The U-shaped input of information that we observed in Analysis 2 could potentially be helpful for learning by virtue of the serial-position effect. The serial-position effect states that information at the start and end of a sequence is better recalled than the one introduced in the middle of the sequence⁷⁵. This strategic adaptation could enhance word learning, as new words are introduced at the preferred serial position for word recall¹²⁰. Indeed, caregivers speaking English infant-directed speech spontaneously place unfamiliar, less frequent words, (words carrying more information by our metric) at the final position of a sequence¹²¹. Although not a universal feature of infant-directed speech, the serial position effect could also be a driving force behind the different temporal structure of information at the beginning-middle-end scale in CDSongs compared to ADSongs, possibly supporting both attention elicitation and word learning.

In addition, the more variable information rate that we observed in CDSongs could be beneficial due to being less predictable, as we know that infants tend to pay more attention to variable stimuli^{6,41}. One reason why the variable information rate could be more popular among infant audiences might be due to the potential higher range of information processing rates that they can accommodate (Fig. 6). If we assume that infants' preferred information rate is lower and more variable than that of adults, CDSongs create more opportunities for infants to be on both sides of their preferred information rate. As caregivers might not be aware of their infant's current processing capacity, having a signal that carries a broad range of information rates over time might increase the chances that at least a part of the message will be processed.

In Analysis 3 we also observed oscillating patterns in the information profile of CDSongs and ADSongs, and that those oscillating patterns were more predictable (less variable in span) in the CDSongs compared to the ADSongs (Fig. 4H), and this results was found even when considering languages individually. These observations are consistent with the finding that, during early caregiver-child interaction, information tends to be packaged in oscillatory temporal structures across multiple timescales, from sub-second-scale amplitude fluctuations (which are stronger in CDS) through to daily routines and rhythms^{79,80,122}. These periodicities may help to generate predictability, which, in turn, may facilitate information transfer¹²³. Periodicities may also drive oscillatory entrainment between exogenous rhythms that are present in the external environment and endogenous neural and physiological rhythms that operate across multiple timescales^{123,124}. Importantly, the findings from Analyses 1a (i.e., overall

reduction of information content in CDSongs) and those from Analysis 3 (i.e., information appearing in bursts in CDSongs) are interrelated. Indeed, it is during the redundant segments that no new information is added, contributing to the overall reduction of information in relation to the increasing length of the analysis. These oscillations may even help to separate the new information from the redundancy, creating obvious sub-event segmentation guiding infant attention¹²⁵ and training them for event segmentation, an important part of comprehension and anticipation¹²⁶.

More speculatively, our findings—and, in particular, the differences we observed between CDSongs and CDStories—may also be compatible with the idea that one of the functions of CDSongs is not simply to maximise information transfer, but also to create a common ground for synchrony and shared interpretation, developing early social bonding. As Gratier et al. pointed out: "The repetition-variation dialectic of all musical idioms is intrinsically a culture-producing process. It provides a basic architecture for intersubjective engagement"⁶⁵. As we discussed above, CDSongs offers a platform for co-creation, as high redundancy lyrics are easier to remember and in turn to sing along^{127–129} potentially leading to blur the adult/child division compared to the top-down didacticism of CDStories⁶⁹.

The increased periodicities in new information content that we documented in CDSongs may also permit the temporal alignment of behaviours, as the next information peak can be anticipated⁸⁰. Isochronous oscillations are more easily trackable as they show less variability¹³⁰, which may enhance caregiver-infant synchronisation^{81,131}. It has been suggested that interpersonal synchrony may help infant word segmentation¹³² and foster early social bonding^{57,133-135}, for instance by blurring the self-other dichotomy during shared actions and/or by aligning multiple actions towards a shared goal¹³⁴.

Figure 6 is a visualisation of archetypal information profiles from each corpus. The lines were drawn with respect to the different analysis being made across multiple timescales. First, CDSongs has an overall lower complexity (area under the curve) compared to ADSongs and CDStories (Analysis 1b). Second, CDSongs have relatively more information at the beginning and end compared to ADSongs (Analysis 2). Third, CDSongs and ADSongs have oscillating unpredictability, with CDSongs having more isochronous oscillations than ADSongs (Analysis 3).

Limitations

One limitation of this study is that, while we measured the difference in temporal patterns of complexity across text type, we can only posit theories that would explain these differences in terms of functions, and constraints. Indeed, it remains possible that the forms we describe here either happened for other reasons (i.e., under the selective pressures of other-non-measures -factors) and/or randomly, and that after being repeatedly exposed to these forms, infants learned how to learn from them optimally. Because we only measure forms here, but not how they impact learning or attention, we cannot strictly conclude concerning the form-function relationships. Another limitation concerning the YouTube view as a popularity metric, is that, to our knowledge, it is unknown how much caregivers interfere in the song selection process. Another limitation of our study is that-as is the case for most of the available literature on this topic unfortunately-the texts we analysed come from a small subset of languages, all coming from WEIRD cultures¹³⁶, and not necessarily the one listened to everyday by children (e.g., the Aesop stories). Moreover, some effects did not hold when considering each language separately. Therefore, our findings have limited generalisation power¹³⁷, and even if we tried to mitigate this by choosing three languages that are spoken worldwide, we acknowledge that similar methods applied to a more diverse and naturalistic stimuli are needed to provide a more complete picture. Establishing a stronger relationship between the form of child-directed content and the functions that they support could be implemented in intervention or serve as guidelines for early years educators.

Conclusion

In line with the literature suggesting that the characteristics of nursery rhymes are constrained by form-function relationships, we examined the multiscale information dynamics of child-directed songs (CDSongs), adultdirected songs (ADSongs), and child-directed stories (CDStories). We developed a technique to quantify the rate of change of multilingual textual information content over several timescales. We found that ADSongs showed an inverted U relationship between complexity and popularity, but CDSongs only showed a negative relationship. Overall textual complexity was, however, lower in CDSongs. We also found that information was more unevenly allocated in CDSongs, with more information contained at the starts and ends of songs, and more periodic oscillatory patterns in the rate of change of information content. Taken together, these findings may suggest that nursery rhymes act as a form of pattern recognition training for infants and children to learn structure and novelty in a targeted way, while also enabling them to connect with others. These findings also highlight the need to treat infant preference for middle complexity as something to extend and analyse in the time dimension.

Data availability

The texts (CDSongs and ADSongs lyrics, and CDStories) used in this study are available to the public under a Creative Commons licence at: https://osf. io/fa7rz/?view_only=5294b351738749cd8e9d2408ff306091.

Code availability

The code used in this study are available to the public under a Creative Commons licence at: https://osf.io/fa7rz/?view_only=5294b351738749cd 8e9d2408ff306091.

Received: 11 October 2023; Accepted: 24 February 2025; Published online: 24 March 2025

References

- 1. Baer, C. & Kidd, C. Learning with certainty in childhood. *Trends Cogn. Sci.* **26**, 887–896 (2022).
- 2. Poli, F., Serino, G., Mars, R. & Hunnius, S. Infants tailor their attention to maximize learning. *Sci. Adv.* **6**, eabb5053 (2020).
- Eaves, B. S., Feldman, N. H., Griffiths, T. L. & Shafto, P. Infant-directed speech is consistent with teaching. *Psychol. Rev.* 123, 758–771 (2016).
- 4. Schick, J. et al. The function and evolution of child-directed communication. *PLoS Biol.* **20**, e3001630 (2022).
- Hilton, C. B. et al. Acoustic regularities in infant-directed speech and song across cultures. *Nat. Hum. Behav.* 6, 1545–1556 (2022).
- Meyer, M., van Schaik, J. E., Poli, F. & Hunnius, S. How infantdirected actions enhance infants' attention, learning, and exploration: evidence from EEG and computational modeling. *Dev. Sci.* 26, e13259 (2023).
- Wass, S. V. & Smith, T. J. Individual differences in infant oculomotor behavior during the viewing of complex naturalistic scenes. *Infancy* 19, 352–384 (2014).
- Cirelli, L. K., Jurewicz, Z. B. & Trehub, S. E. Effects of maternal singing style on mother–infant arousal and behavior. *J. Cogn. Neurosci.* 32, 1213–1220 (2019).
- Rost, G. C. & McMurray, B. Finding the signal by adding noise: the role of noncontrastive phonetic variability in early word learning. *Infancy*, 15, https://doi.org/10.1111/j.1532-7078.2010.00033.x (2010).
- Swingley, D. & Aslin, R. N. Lexical competition in young children's word learning. *Cogn. Psychol.* 54, 99–132 (2007).
- Cooke, M., King, S., Garnier, M. & Aubanel, V. The listening talker: a review of human and algorithmic context-induced modifications of speech. *Comput. Speech Lang.* 28, 543–571 (2014).
- Thiessen, E. D. Domain general constraints on statistical learning. *Child Dev.* 82, 462–470 (2011).
- Yu, C. & Smith, L. B. Embodied attention and word learning by toddlers. *Cognition* **125**, 244–262 (2012).
- 14. Bhandari, P., Demberg, V. & Kray, J. Semantic predictability facilitates comprehension of degraded speech in a graded manner. *Front. Psychol.* **12**, 714485 (2021).

- 15. Ferguson, C. Baby talk in six languages. *Am. Anthropol.* **66**, 103–114 (1964).
- 16. Pierrehumbert, B. Amour et attachement. *Spirale*, 28. https://doi. org/10.3917/spi.028.0031 (2003).
- 17. Zhang, Y. et al. Neural coding of formant-exaggerated speech in the infant brain. *Dev. Sci.* **14**, 566–581 (2011).
- Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H. & Mazuka, R. Vowels in infant-directed speech: more breathy and more variable, but not clearer. *Cognition* 166, 84–93 (2017).
- Swingley, D. Contributions of infant word learning to language development. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* (364, 3617–3632). Royal Society (2009).
- Chater, N. & Vitányi, P. Simplicity: a unifying principle in cognitive science? *Trends Cogn. Sci.* 7, 19–22 (2003).
- 21. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).
- Köster, M., Kayhan, E., Langeloh, M. & Hoehl, S. Making sense of the world: infant learning from a predictive processing perspective. *Perspect. Psychol. Sci.* **15**, 174569161989507 (2020).
- Peters, A., Mcewen, B. & Friston, K. Uncertainty and stress: why it causes diseases and how it is mastered by the brain. *Prog. Neurobiol.* **156**, 164–188 (2017).
- Wilson, D., & Sperber, D. Outline of relevance theory. *HERMES* Journal of Language and Communication in Business, 3, https:// tidsskrift.dk/her/article/view/21436 (2015).
- Woźniak, M. & Knoblich, G. Communication and action predictability: two complementary strategies for successful cooperation. *R. Soc. Open Sci.* 9, 220577 (2022).
- Carr, J. W., Smith, K., Culbertson, J. & Kirby, S. Simplicity and informativeness in semantic category systems. *Cognition* 202, 104289 (2020).
- Ravreby, I., Shilat, Y. & Yeshurun, Y. Liking as a balance between synchronization, complexity and novelty. *Sci. Rep.* **12**, 3181 (2022).
- Berlyne, D. E. Novelty, complexity, and hedonic value. https://link. springer.com/content/pdf/10.3758/BF03212593.pdf (1970).
- 29. Birkhoff, G. D. Aesthetic Measure. Harvard University Press (1933).
- Gold, B. P. et al. Predictability and uncertainty in the pleasure of music: a reward for learning? *J. Neurosci.* **39**, 9397–9409 (2019).
- Moles, A. A. Théorie informationnelle de la musique. *In:* Bulletin N°5 Du Groupe d'Acoustique Musicale (1964).
- Kidd, C., Piantadosi, S. & Aslin, R. The goldilocks effect: human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS One* 7, e36399 (2012).
- Kidd, C., Piantadosi, S. & Aslin, R. The goldilocks effect in infant auditory attention. *Child Dev.* 85, 1795–1804 (2014).
- Hannon, E. E. & Trehub, S. E. Tuning in to musical rhythms: Infants learn more readily than adults. *Proc. Natl Acad. Sci. USA* **102**, 12639–12643 (2005).
- 35. Meyer, L. Emotion and meaning in music. Chicago University Press (1956).
- Shimony, A. D. & Cohen, V. W. Training babies for "sweet anticipation": emphasizing expectation arousal in infant-directed mother singing. *Psychol. Music* 50, 1527–1543 (2022).
- Trehub, S. E. The developmental origins of musicality. *Nat. Neurosci.* 6, 669–673 (2003).
- Schön, D. et al. Songs as an aid for language acquisition. *Cognition* 106, 975–983 (2008).
- Bahrick, L. E. & Lickliter, R. Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Dev. Psychol.* 36, 190–201 (2000).
- Snijders, T. M., Benders, T. & Fikkert, P. Infants segment words from songs— an EEG study. *Brain Sci.* 10, 39 (2020).
- Räsänen, O., Kakouros, S. & Soderstrom, M. Is infant-directed speech interesting because it is surprising?—Linking properties of

IDS to statistical learning and attention at the prosodic level. *Cognition* **178**, 193–206 (2018).

- 42. Zipf, G. K. Human behavior and the principle of least effort. In: *Human behavior and the principle of least effort*. Addison-Wesley Press (1949).
- 43. Fenk-Oczlon, G. Familiarity, information flow, and linguistic form. In: *Frequency and the Emergence of Linguistic Structure* **45**, 431–448 (2001).
- 44. Kirby, S., Tamariz, M., Cornish, H. & Smith, K. Compression and communication in the cultural evolution of linguistic structure. *Cognition* **141**, 87–102 (2015).
- 45. Futrell, R. & Hahn, M. Information theory as a bridge between language function and language form. *Front. Commun.* **7**, (2022).
- Español, S. & Shifres, F. The artistic infant directed performance: a mycroanalysis of the adult's movements and sounds. *Integr. Psychol. Behav. Sci.* 49, 371–397 (2015).
- Keitel, A., Prinz, W., Friederici, A. D., von Hofsten, C. & Daum, M. M. Perception of conversations: the importance of semantics and intonation in children's development. *J. Exp. Child Psychol.* **116**, 264–277 (2013).
- Franco, F., Suttora, C., Spinelli, M., Kozar, I. & FADSolo, M. Singing to infants matters: early singing interactions affect musical preferences and facilitate vocabulary building. *J. Child Lang.* 49, 552–577 (2022).
- Mendoza, J. K. & Fausey, C. M. Everyday music in infancy. *Dev. Sci.* 24, e13122 (2021).
- Barrett, M. Meme engineers: children as producers of musical culture. *Int. J. Early Years Educ.* **11**, 195–212 (2003).
- Mehr, S. A., & Krasnow, M. M. Parent-offspring conflict and the evolution of infant-directed song. *Evol. Hum. Behav.* 38, 674–684 (2017).
- Eubank, E. E. The concepts of sociology, a treatise presenting a suggested organization of sociological theory in terms of its major concepts. D. C. Heath and Co. https://catalog.hathitrust.org/ Record/000972943 (1932).
- 53. Mehr, S. A., Singh, M., York, H., Glowacki, L. & Krasnow, M. M. Form and function in human song. *Curr. Biol.* **28**, 356–368.e5 (2018).
- Custodero, L. A. & Johnson-Green, E. A. Passing the cultural torch: musical experience and musical parenting of infants. *J. Res. Music Educ.* 51 (2003).
- 55. Trehub, S. E. Nurturing infants with music. *Int. J. Music Early Child.* **14**, 9–15 (2019).
- Nguyen, T., Flaten, E., Trainor, L. & Novembre, G. Early social communication through music: state of the art and future perspectives. *Dev. Cogn. Neurosci.* 63, 101279 (2023).
- 57. Trainor, L. J. & Cirelli, L. Rhythm and interpersonal synchrony in early social development. *Ann. N. Y. Acad. Sci.* **1337**, 45–52 (2015).
- 58. Savage, P. E. et al. Music as a coevolved system for social bonding. *Behav. Brain Sci.* 44, e59 (2021).
- Cirelli, L. K. & Trehub, S. E. Familiar songs reduce infant distress. Dev. Psychol. 56, 861–868 (2020).
- 60. Corbeil, M., Trehub, S. E. & Peretz, I. Singing delays the onset of infant distress. *Infancy* **21**, 373–391 (2016).
- 61. Phillips, J. R. Syntax and vocabulary of mothers' speech to young children: age and sex comparisons. *Child Dev.* **44**, 182–185 (1973).
- Horn, D., Jennings, P., Mcgraw, T., Zhou, P. & Ma, W. Singing facilitates word learning and memory. In: *Proceedings of the Annual Boston University Conference on Language Development.* 44, (2020).
- Ma, W., Bowers, L., Behrend, D., Hellmuth Margulis, E. & Forde Thompson, W. Child word learning in song and speech. *Q. J. Exp. Psychol.* **77**, 343–362 (2023).
- Gratier, M. & Trevarthen, C. Musical narrative and motives for culture in mother-infant vocal interaction. *J. Conscious. Stud.* **15**, 122–158 (2008).

- Gratier, M. & Magnier, J. Sense and synchrony: infant communication and musical improvisation. *Intermédialités* 19, 45–64 (2012).
- Trainor, L. J. Infant preferences for infant-directed versus noninfantdirected playsongs and lullabies. *Infant Behav. Dev.* 19, 83–92 (1996).
- Tsang, C. D., Falk, S. & Hessel, A. Infants prefer infant-directed song over speech. *Child Dev.* 88, 1207–1215 (2017).
- Resmond-Wenz, É. Rimes et comptines. 1001 bébés. https://doi. org/10.3917/eres.resmo.2008.01(2008).
- Scheiding, C. M. From nursery rhymes to childlore: orality and ideology. In: New Perspectives on Children's Literature, 22 (2019).
- Grunwald, P. Introducing the minimum description length principle. In: Advances in minimum description lenght: theory and application (2005).
- 71. Tal, S. & Arnon, I. Redundancy can benefit learning: evidence from word order and case marking. *Cognition* **224**, 1–111 (2022).
- Tal, S., Grossman, E. & Arnon, I. Infant-directed speech becomes less redundant as infants grow: implications for language learning. *Cognition*. 249, 105817 (2021).
- 73. Notopoulos, J. A. Mnemosyne in oral literature **69**, 469–493 (Lancaster Press, 1983).
- Cornell, E. H. & Bergstrom, L. I. Serial-position effects in infants' recognition memory. *Mem. Cogn.* 11, 494–499 (1983).
- Gulya, M., Galluccio, L., Wilk, A. & Rovee-Collier, C. Infants' longterm memory for a serial list: recognition and reactivation. *Dev. Psychobiol.* 38, 174–185 (2001).
- Kelley, M. R., Neath, I. & Surprenant, A. M. Three more semantic serial position functions and a SIMPLE explanation. *Mem. Cogn.* 41, 600–610 (2013).
- Maylor, E. A. Serial position effects in semantic memory: reconstructing the order of verses of hymns. *Psychonomic Bull. Rev.* 9, 816–820 (2002).
- Leong, V. & Goswami, U. Acoustic-emergent phonology in the amplitude envelope of child-directed speech. *PLoS One* **10**, e0144411 (2015).
- Feldman, R. From biological rhythms to social rhythms: physiological precursors of mother-infant synchrony. *Dev. Psychol.* 42, 175 (2006).
- Wass, S. v., Perapoch Amadó, M. & Ives, J. Oscillatory entrainment to our early social or physical environment and the emergence of volitional control. *Dev. Cogn. Neurosci.* 54, 101102 (2022).
- Lense, M. D., Shultz, S., Ast Esano, C. & Jones, W. Music of infantdirected singing entrains infants' social visual behavior. *Proc. Natl. Acad. Sci. USA*. **119**, e2116967119 (2022).
- Lakatos, P., Gross, J. & Thut, G. A new unifying account of the roles of neuronal entrainment. *Curr. Biol.* 29, R890–R905 (2019).
- Aylett, M. & Turk, A. The smooth signal redundancy hypothesis: a functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Lang. Speech* 47, 31–56 (2004).
- Frank, A. F. & Jaeger, T. F. Speaking rationally: uniform information density as an optimal strategy for language production. *Proceedings* of the Annual Meeting of the Cognitive Science Society, 30 (2008).
- 85. Jaeger, T. F. Redundancy and reduction: speakers manage syntactic information density. *Cogn. Psychol.* **61**, 23–62 (2010).
- Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 623–656 (1948).
- Levy, R. & Jaeger, T. F. Speakers optimize information density through syntactic reduction. *Adv. Neural Inf. Process. Syst.* 19, 849–856 (2006).
- Chatzopoulou, G., Sheng, C. & Faloutsos, M. A first step towards understanding popularity in YouTube. In: 2010 INFOCOM IEEE Conference on Computer Communications Workshops, 1–6. https://api.semanticscholar.org/CorpusID:3149082 (2010).
- 89. Schoeffler, M. & Herre, J. The influence of audio quality on the popularity of music videos: a youtube case study. In: *Proceedings of the First International Workshop on Internet-Scale Multimedia*

Management, 35–38, https://doi.org/10.1145/2661714.2661725 (2014).

- Mekouar, S., Zrira, N. & Bouyakhf, E.-H. Popularity prediction of videos in YouTube as case study: a regression analysis study. In: Proceedings of the 2nd International Conference on Big Data, Cloud and Applications. https://doi.org/10.1145/3090354.3090406 (2017).
- Goodrow, C. On YouTube's recommendation system: A deeper look into how YouTube's recommendation system works. https://blog. youtube/inside-youtube/on-youtubes-recommendation-system/ (2021).
- 92. Youtube for families help, https://support.google.com/youtubekids/ answer/6130531?hl=en (2025).
- Kolmogorov, A. N. Three approaches to the quantitative definition of information. *Int. J. Comput. Math.* 2, https://www.tandfonline.com/ doi/abs/10.1080/00207166808803030 (1968).
- 94. Delahaye, J.-P. *Complexité aléatoire et complexité organisée*. Editions Quae (2009).
- 95. Bauch, J., Theofilis, G., Rachum, R., Bovee, R. & Muñoz Pelayo González. *PyLZMA*. https://github.com/fancycode/pylzma (2015).
- Heilbron, M., Armeni, K., Schoffelen, J.-M., Hagoort, P. & de Lange, F. P. A hierarchy of linguistic predictions during natural language comprehension. *Proc. Natl Acad. Sci.* **119**, e2201968119 (2022).
- Hasson, U., Chen, J., & Honey, C. J. Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn. Sci.* **19**, 304–313 (2015).
- Schmidhuber, J. Driven by compression progress: a simple principle explains essential aspects of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes. https://arxiv.org/abs/0812.4360 (2009).
- Wolff, J. G. The SP theory of intelligence: benefits and applications. Information (Switzerland), 5, https://arxiv.org/abs/1307.0845 (2014).
- Wolff, J. G. Information compression as a unifying principle in human learning, perception, and cognition. In: *Complexity*. Hindawi Limited. 1–38 (2019)
- van Geert, E. & Wagemans, J. Order, complexity, and aesthetic appreciation. *Psychol. Aesthet. Creat. Arts* 14, 135–154 (2020).
- 102. Morris, C. Are pop lyrics getting more repetitive? An exercise in l language compression. In: *The Pudding* (2017).
- 103. Varnum, M. E. W., Krems, J. A., Morris, C., Wormley, A., & Grossmann, I. Why are song lyrics becoming simpler? a time series analysis of lyrical complexity in six decades of American popular music. *PLoS ONE*, 16. https://doi.org/10.1371/journal.pone. 0244576 (2021).
- 104. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat. Methods* **17**, 261–272 (2020).
- Elias, N. & Sulkin, I. YouTube viewers in diapers: an exploration of factors associated with amount of toddlers' online viewing. *Cyberpsychology* **11**, Article 2 (2017).
- Trehub, S. & Trainor, L. Singing to infants: lullables and play songs development of auditory scene analysis view project ManyBables View project. https://www.researchgate.net/publication/285504507 (1998).
- 107. Margulis, E. H. *On repeat: How music plays the mind*. Oxford University Press (2014).
- Lester, N. A. et al. Detecting structured repetition in childsurrounding speech: evidence from maximally diverse languages. *Cognition* 221, 104986 (2022)
- Broderick, M. P., Anderson, A. J. & Lalor, E. C. Semantic context enhances the early auditory encoding of natural speech. *J. Neurosci.* 39, 7564–7575 (2019).
- 110. Lavi-Rotbain, O. & Arnon, I. Children learn words better in low entropy. In: *Annual Meeting of the Cognitive Science Society*. https://escholarship.org/uc/item/37p381r7 (2019).
- Parry, M. Studies in the epic technique of oral verse-making. I. Homer and Homeric Style. In: *Classical Philology* **41** (1930).

- 112. Parry, M. Studies in the epic technique of oral verse-making: II. The Homeric Language as the Language of an Oral Poetry. In: *Classical Philology* **43** (1932).
- 113. Fay, N., Walker, B., Swoboda, N. & Garrod, S. How to create shared symbols. *Cogn. Sci.* **42**, 241–269 (2018).
- 114. Smith, K. How culture and biology interact to shape language and the language faculty. *Top. Cogn. Sci.* **12**, 690–712 (2020).
- 115. Tamariz, M. & Kirby, S. Culture: copying, compression, and conventionality. *Cogn. Sci.* **39**, 171–183 (2015).
- Zlotnik, G. & Vansintjan, A. Storage of information and its implications for human development: a dialectic approach. *Front. Psychol.* **11**, 1715 (2020).
- 117. Jaeger, T. F. Redundancy and syntactic reduction in spontaneous speech. http://www.bcs.rochester.edu/people/fjaeger/ (2006).
- 118. Coupé, C., Oh, Y., Dediu, D. & Pellegrino, F. Different languages, similar encoding efficiency: comparable information rates across the human communicative niche. *Sci. Adv.* **5**, eaaw2594 (2019).
- Gibson, E. et al. How efficiency shapes human language. In: *Trends in Cognitive Sciences* 23, 389–407. Elsevier Ltd. https://doi.org/10. 1016/j.tics.2019.02.003 (2019).
- 120. Jansen, J., Liu, Z. & Simon, Z. The effect of ad rank on the performance of keyword advertising campaigns. *Journal of the American Society for Information Science and Technology*, 64. https://doi.org/10.1002/asi.22910 (2013).
- 121. Fernald, A., Mcroberts, G. W. & Swingley, D. Infants' developing competence in recognizing and understanding words in fluent speech. https://benjamins.com/catalog/lald.23.08fer?srsltid= AfmBOopay8sCL9MZ5AVgMu3KDYaH7zYbrM9A_ DN3A3eyAXkoPYvvthkE (2001).
- 122. Bobin-Bègue, A. Rhythms in early development. *In*: Early interaction and developmental psychopathology: Volume I: Infancy (pp. 55–86). Springer International Publishing (2019).
- Haegens, S. & Golumbic, E. Z. Rhythmic facilitation of sensory processing: a critical review. *Neurosci. Biobehav. Rev.* 86, 150–165 (2018).
- 124. Buzsáki, G. *Rhythms of the brain*. In: Oxford University Press. https:// doi.org/10.1093/acprof:oso/9780195301069.001.0001 (2006).
- Kliesch, C., Parise, E., Reid, V. & Hoehl, S. The role of social signals in segmenting observed actions in 18-month-old children. *Dev. Sci.* 25, e13198 (2022).
- Zacks, J. M. Event perception and memory. Annu. Rev. Psychol. 71, 165–191 (2020).
- Calvert, S. L. & Tart, M. Song versus verbal forms for very-long-term, long-term, and short-term verbatim recall. *J. Appl. Dev. Psychol.* 14, 245–260 (1993).
- de l'Etoile, S. K. Infant behavioral responses to infant-directed singing and other maternal interactions. *Infant Behav. Dev.* 29, 456–470 (2006).
- Gratier, M. & Trevarthen, C. Voice, vitality and meaning: on the shaping of the infant's utterances in willing engagement with culture. Comment on Bertau's "On the notion of voice". *Int. J. Dialogical Sci.* 2, 169–181 (2007).
- 130. Vesper, C., der Wel, R., Knoblich, G. & Sebanz, N. Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Exp. Brain Res.* **211**, 517–530 (2011).
- Markova, G., Nguyen, T., Schätz, C. & de Eccher, M. Singing in tune being in tune: relationship between maternal playful singing and interpersonal synchrony. *Enfance* 5, 89–107 (2020).
- Vanoncini, M., Boll-Avetisyan, N., Elsner, B., Hoehl, S. & Kayhan, E. The role of mother-infant emotional synchrony in speech processing in 9-month-old infants. *Infant Behav. Dev.* 69, 101772 (2022).
- Launay, J., Tarr, B. & Dunbar, R. I. M. Synchrony as an adaptive mechanism for large-scale human social bonding. *Ethology* **122**, 779–789 (2016).

- Mogan, R., Fischer, R. & Bulbulia, J. A. To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. J. Exp. Soc. Psychol. 72, 13–20 (2017).
- Cirelli, L. K., Einarson, K. M. & Trainor, L. J. Interpersonal synchrony increases prosocial behavior in infants. *Dev. Sci.* 17, 1003–1011 (2014).
- 136. Henrich, J., Heine, S. J. & Norenzayan, A. The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–83 (2010).
- 137. Yarkoni, T. The generalizability crisis. Behav. Brain Sci. 45, e1 (2022).
- 138. Nelson, M., & Gailly, J. The data compression Book, 2nd edition. In: *History*. M&T Books (1995).

Acknowledgements

This research was funded by European Research Council grant number ONACSA 853251 (S.W.), and by European Research Council grant number JDIL 845859 (L.G.). The funders had no role in study design, data collection, and analysis, the decision to publish, or preparation of the manuscript. We thank Susanne Reisner for reading and commenting on earlier drafts of the manuscript, and members of the UEL BabyDevLab for useful discussions.

Author contributions

P.L., L.G., and S.W. conceived and planned the experiments. P.L. carried out the experiments. P.L., L.G., and S.W. contributed to the interpretation of the results. P.L. took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44271-025-00219-4.

Correspondence and requests for materials should be addressed to Pierre Labendzki.

Peer review information *Communications Psychology* thanks Florian Jaeger, Theano Kokkinaki and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Marike Schiffer. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/bync-nd/4.0/.

© The Author(s) 2025