

Identifying the Periodicity of Information in Natural Language

Yulin Ou¹, Yu Wang², Yang Xu¹, Hendrik Buschmeier²

¹ CLCS Lab, Department of Computer Science and Engineering,
Southern University of Science and Technology

² Digital Linguistics Lab, Department of Linguistics, Bielefeld University

Correspondence: xuyang@sustech.edu.cn

Abstract

Recent theoretical advancement of information density in natural language have raised the following question: To what degree does natural language exhibit *periodicity* pattern in its encoded information? We address this question by introducing a new method called AutoPeriod of Surprisal (APS). APS adopts a canonical periodicity detection algorithm and is able to identify any significant periods that exist in the surprisal sequence of a single document. By applying the algorithm to a set of corpora, we have obtained the following empirical results: Firstly, a considerable proportion of human language demonstrates a strong pattern of periodicity in information. Secondly, new periods that are outside the distributions of typical structural units in text (e.g., sentence boundaries, elementary discourse units, etc.) are found and further confirmed via harmonic regression modeling. We conclude that the periodicity of information in language is a joint outcome from both structured factors and other driving factors that take effect at longer distances. The advantages of our periodicity detection method and its potentials in LLM-generation detection are further discussed. Our code is available as public repositories.¹

1 Introduction

Understanding how information transmits in language is a long-standing research topic in linguistics (Shannon, 1948, 1949, 1951; Bell et al., 2003). One of the most influential assumptions is uniform information density (UID; Aylett and Turk, 2004; Jaeger, 2010; Jaeger and Levy, 2006; Meister et al., 2021), which states that information in language, measured by surprisal, distributes uniformly out of a need for efficient communication, as evidenced in its effect on, e.g., word duration in conversation and word abbreviation (Demberg et al., 2012;

Mahowald et al., 2013). Recent work further supplements this assumption with different empirical findings, such as the fluctuation of information with discourse boundaries (Xu and Reitter, 2016, 2018); the thematic structures in dialogue that influence the change of information from speakers (Maës et al., 2022).

These empirical findings give rise to an emerging research topic: the *periodicity* of information in natural language. Periodicity is a bold hypothesis that goes a step further than the simple “fluctuation” of information, which is supported by recent evidence: the spectral features of text surprisal are distinguishable from white noise (Xu and Reitter, 2017; Yang et al., 2023); spectra of surprisal are useful in machine-generation detection (Xu et al., 2024; Liu et al., 2025); surprisal variation can be modeled by harmonic regression, which confirms periodic patterns across various linguistic structures (Tsipidi et al., 2025).

However, some key questions about this topic remain unanswered. First of all, how can we find **direct evidence of the periodicity** of information in language? The Fourier-analysis-based spectral methods proposed in Yang et al. (2023) and Liu et al. (2025) provide useful insights, but can only serve as indirect evidence of periodicity, because the spectrum in frequency-domain was not transformed back to explicit periods in time-domain. On the other hand, the harmonic regression method proposed by Tsipidi et al. (2025), provides a way to validate periodicity in information, given “candidate” periods selected from certain linguistic structures (e.g., sentences, paragraphs, etc.). Therefore, it potentially filters out periods that cannot be attributed to any pre-defined linguistic units. And its reliance on regression analysis prevents it from determining if a text segment is truly periodic. This leads to our second question: can we **detect periodicity at the document level** more precisely and with higher quality? That is, given

¹Github: <https://github.com/CLCS-SUSTech/APS>
Zenodo: <https://doi.org/10.5281/zenodo.19567886>

a piece of text, we want to find all periods that can pass a certain detection threshold with some clearly defined confidence levels, or otherwise to determine its non-existence. As detection occurs at the document level, such a method would allow us to explore whether periodicity correlates with other theoretically interesting factors, such as text genre, authorship, and the distinction between human-written texts and those generated by large language models (LLMs).

In this study we address the above-mentioned questions by proposing a periodicity detection algorithm named AutoPeriod of Surprisal (APS), which is able to directly identify any existing periods in an input sequence of surprisal at certain confidence levels (see Section 3). We apply this algorithm to common linguistic corpora and have obtained positive evidence for periodicity of information (Section 5). The validity of our results is verified with an alternative method (Section 6). The potential merits of periodicity detection in distinguishing LLM-generations (Section 7).

2 Background

2.1 UID and the Natural Fluctuation of Information

The UID theory describes that a general trend of how the information of linguistic units distributes within the course of communication is to stay relatively uniform or constant. However, token-level uniformity/constancy is an asymptotic property that would only manifest when observation is conducted over a sufficiently long sequence, during which the natural fluctuation of information would inevitably occur. For example, Genzel and Charniak (2002, 2003) observed the increasing surprisal near the beginning of paragraphs and the sudden drop near the end. It can be imagined that if multiple paragraphs of similar lengths are organized together as a chapter, then an overall regular fluctuation (or even *periodicity*) will obviously appear.

Building on the coupling between uniformity and fluctuation illustrated above, recent works have attempted to model the latter more formally. A recurring convergence pattern of surprisal is found within topic boundaries in dialogue (Xu and Reitter, 2016, 2018). It is also found that the uniformity trend is only salient in certain contextual units (Giulianelli et al., 2021) and not always observed (Giulianelli and Fernández, 2021). More recently, surprisal fluctuation over discourse has been modeled with

hierarchically-structured Bayesian regression (Tsipidi et al., 2024) and harmonic regression (Tsipidi et al., 2025).

2.2 Why Study the Periodicity of Surprisal

Given that surprisal can describe how linguistic information changes dynamically in human language, its potential uses have been explored, such as (i) measuring creative use of language such as metaphor and humor (Bunescu and Uduehi, 2022; Xie et al., 2021); (ii) evaluating the quality of generated text (Kharkwal and Muresan, 2014; Yang et al., 2023; Liu et al., 2025); (iii) and, recently, detecting LLM-generated text (Xu et al., 2024).

A further study on the periodicity of surprisal will largely facilitate the above-listed research directions. While UID theory and its variants can make asymptotic predictions about the general trends of information in language, using periodicity as a lens would provide more fine-grained insights into the instantaneous dynamics of information, which better meets the purpose of studying the token-by-token generation process. What’s missing in the picture is a method and toolkit that can effectively extract periodicity as a clear feature from text – current regression-based methods (Tsipidi et al., 2025) provide ways to validate but not to comprehensively identify periodicity; Fourier-based methods (Xu et al., 2024) lack direct interpretation. Therefore, the goal of this study is to fill this gap.

3 Periodicity Detection

The method proposed in this study for detecting periodicity in information is based on the AutoPeriod algorithm (Vlachos et al., 2005). Originally designed for detecting periodicity in economics and social time series data, this algorithm demonstrates reliable performance in identifying periodic patterns. We adaptively modify and re-implement this algorithm – tailoring its parameters to the characteristics of surprisal – and release it as a standalone toolkit called AutoPeriod of Surprisal (APS).

The algorithm runs two sub-procedures:

- 1) **GETPERIODHINTS**: Gets the periodogram spectrum of an input sequence and extracts the candidate periods called *period hints*, whose power values are above a certain threshold obtained from random permutations.
- 2) **ACFFILTERING**: Filters the period hints using the auto-correlation function (ACF), and only

keep those that lie on the hills of the ACF curves. These are considered as the valid periods.

We first introduce the mathematical prerequisites for the Fourier transform and the periodogram in Section 3.1, as these foundational concepts are essential for understanding the AutoPeriod procedure.

3.1 Prerequisites: Fourier Transform and Periodogram

Given a piece of text that contains N tokens, we first estimate the surprisal of each token by computing the negative log probability $x_n = -\log p(t_n | t_{<n})$, where p is the probability of token t_n predicted by a language model based on its preceding context $t_{<n}$. The resulting sequence of surprisal values $\mathbf{x} := x_0, \dots, x_{N-1}$ will be the input to the discrete Fourier transform (DFT), defined as:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi \frac{k}{N}n}$$

where X_k is the cross correlation of \mathbf{x}_n and a complex sinusoid, $e^{-i2\pi \frac{k}{N}n} = \cos(2\pi \frac{k}{N}n) - i \sin(2\pi \frac{k}{N}n)$ ($n = 0, \dots, N-1$).

The norm $\|X_k\| = \sqrt{\text{re}(X_k)^2 + \text{im}(X_k)^2}$ measures to what degree the original input can be reconstructed from the periodic signal at frequency $2\pi \frac{k}{N}$. The raw DFT output $\mathbf{X} := X_0, \dots, X_{N-1}$ has the property of being *even symmetric*, that is, $\|X_k\| = \|X_{N-k}\|$, $\forall k \in 0, \dots, N-1$. Therefore, it is sufficient to use the first half of \mathbf{X}_k , $k \in 0, \dots, \lceil \frac{N-1}{2} \rceil$ to encode the full spectral information. This truncated magnitude sequence is also known as **periodogram**,² a widely-used technique in signal processing (Schuster, 1898; Oppenheim et al., 1999). We denote the periodogram as $\mathcal{P} = \{\|X_k\|\}$ ($k \in 0, \dots, \lceil \frac{N-1}{2} \rceil$).

3.2 Step 1: Extract Period Hints From the Periodogram

The GETPERIODHINTS procedure is described in Algorithm 1. Within this step, the surprisal sequence is first permuted for m times, and the maximum spectrum powers from these permuted sequences are recorded in *maxPower*. The *percentile*-th largest

²The periodogram algorithm we introduced here follows the design in Vlachos et al. (2005). In our experiments, we tried variants of periodogram, and selected the Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982) in order to enhance the performance of APS in low-frequency range (i.e., long periods).

value in *maxPower* is used as the power threshold $\mathcal{P}_{\text{threshold}}$ to find those frequencies/periods whose powers are above this threshold. Therefore, the period hints returned are the most significant ones with a *percentile*/100 confidence level.

Algorithm 1 First step of AUTOPERIOD: Extract period hints.

```

1: function GETPERIODHINTS( $\mathbf{x}, m, \text{percentile}$ )
2:    $N \leftarrow \mathbf{x}.\text{length}$   $\triangleright$  length of input sequence
3:    $m \leftarrow 100$   $\triangleright$  number of permutations
4:    $\text{percentile} \leftarrow 99$ 
5:    $\text{maxPower} \leftarrow \{\}$   $\triangleright$  empty set
6:    $\text{periodHints} \leftarrow \{\}$ 
7:   for  $i = 1$  to  $m$  do
8:      $\mathbf{x}^{\text{rand}} \leftarrow \text{PERMUTE}(\mathbf{x})$ 
9:      $\mathcal{P}^{\text{rand}} \leftarrow \text{GETPERIODOGRAM}(\mathbf{x})$ 
10:     $\text{maxPower}.\text{ADD}(\text{max}(\mathcal{P}^{\text{rand}}))$ 
11:  end for
12:   $\text{maxPower}.\text{SORT}()$   $\triangleright$  ascending
13:   $P_{\text{threshold}} \leftarrow \text{maxPower}[m \cdot (\frac{\text{percentile}}{100})]$ 
14:   $\mathcal{P} \leftarrow \text{GETPERIODOGRAM}(\mathbf{x})$ 
15:  for  $k = 1$  to  $\mathcal{P}.\text{length}$  do
16:    if  $\mathcal{P}[k] > P_{\text{threshold}}$  then
17:       $\text{periodHints}.\text{ADD}(\frac{N}{k})$ 
18:    end if
19:  end for
20:  return  $\text{periodHints}$ 
21: end function

```

Here we use document #0976 from PTB-WSJ as a complete example of analysis. The texts are snippets of TV listings with highly structured format of content – with weekdays followed by channel names and short summaries (see Figure 1, left). The result of running GETPERIODHINTS on this document is shown in Figure 1 (right). It can be seen that when the confidence level (CL) decreases from .99 to .50, more period hints are extracted. The most confident (CL = .99) period hint occurs at the frequency peak $f \approx .019$, and the corresponding period $T = \frac{1}{f} \approx 53.1$ (tokens), suggesting that a period of surprisal occurs roughly every 53 tokens. CL = .90 is used throughout this study. Before a period hint is validated in the next step, we deem it as quasi-periodic.

3.3 Step 2: Period Validation via ACF Filtering

The brief version of ACFFILTERING procedure at Step 2 is described in Algorithm 2. This involves applying a fine-grained validation to the period *hints*

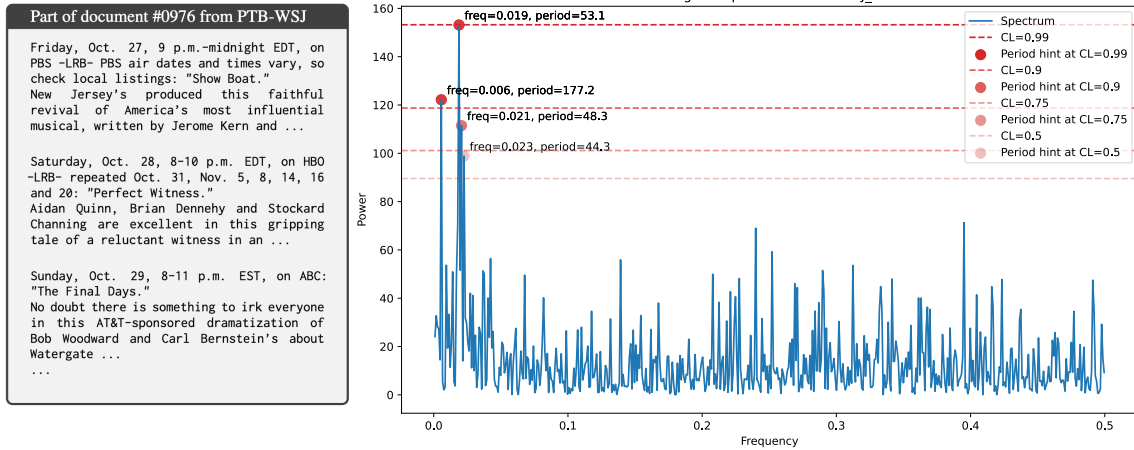


Figure 1: Periodogram spectrum obtained from the surprisal sequence of document #0976 in PTB-WSJ corpus. Right: Red dots indicate the period hints returned by Algorithm 1. Dashed lines are the power thresholds used at different confidence levels.

found in Step 1, which allows spurious periods and false alarms to be filtered out and the *valid* periods to be identified following further refinement. This filtering step is initially designed to increase the detection precision for periodicity in generic time series, and briefly, the motivation is that a *dominant* period τ of a time series x (of length N) should have a relatively large value of the autocorrelation function (ACF):

$$\text{ACF}(\tau) = \frac{1}{N} \sum_{n=0}^{N-\tau} x(n) \cdot x(n + \tau),$$

which measures how similar the sequence is to its previous values for different τ lags. If a period hint τ from the periodogram lies on a hill of the ACF curve, then we are more confident to consider it as a valid period, otherwise it is likely to be a false alarm. This validation is done to each period hint $\tau = \frac{N}{k}$, by examining the nearby points on the ACF curve within a window of size $W_{N/k} = [a, b]$. If τ does not lie on a hill, this period hint is discarded, otherwise a further *refinement* is done by searching within the nearby region and finding the closest local maximum of the ACF scores. As can be seen in Figure 2 (bottom), the raw/unrefined period hints $\tau'_1 = 53.2$ and $\tau'_2 = 177.2$ are refined to the final integer periods $\tau_1 = 52$ and $\tau_2 = 163$, respectively.

4 Corpora and Language Models

This study utilizes several text corpora. Penn Treebank (PTB; Marcus et al., 1993) and Chinese Treebank (CTB; Xue et al., 2013) are adopted as representatives of human-written texts in English and

Algorithm 2 ACF Filtering (brief version)

```

1: function ACFFILTERING( $\mathbf{x}$ , periodHints)
2:    $N \leftarrow \mathbf{x}.length$ 
3:    $ACF \leftarrow \text{GETACF}(\mathbf{x})$ 
4:    $M \leftarrow \text{periodHints}.length$ 
5:    $refinedPeriods \leftarrow \{\}$  ▷ empty set
6:   for  $i = 1$  to  $M$  do
7:      $\tau \leftarrow N/k \leftarrow \text{periodHints}[i]$ 
8:      $\tau' \leftarrow$  local maximum within a range
9:     if  $\tau'$  is on hill of ACF then
10:        $refinedPeriods.ADD(\tau')$ 
11:     end if
12:   end for
13:   return  $refinedPeriods$ 
14: end function

```

Chinese, respectively. Specifically, the Wall Street Journal (WSJ) and Brown portions are used from PTB. Moreover, to explore the relationship between the global periodicity detected by APS and the local linguistic structures focused on by Tsipidi et al. (2025), we incorporate the Rhetorical Structure Theory (RST) based Discourse Treebank (Carlson et al., 2001) for English and the GCDT corpus (Peng et al., 2022) for Chinese. These corpora provide discourse-level annotations, enabling a comparative analysis. The detection results of the corpora mentioned above are mainly discussed in Section 5 and Section 6.

To investigate the periodicity differences between human-written and LLM-generated texts, we use two datasets: FACE collected by Liu et al. (2025)

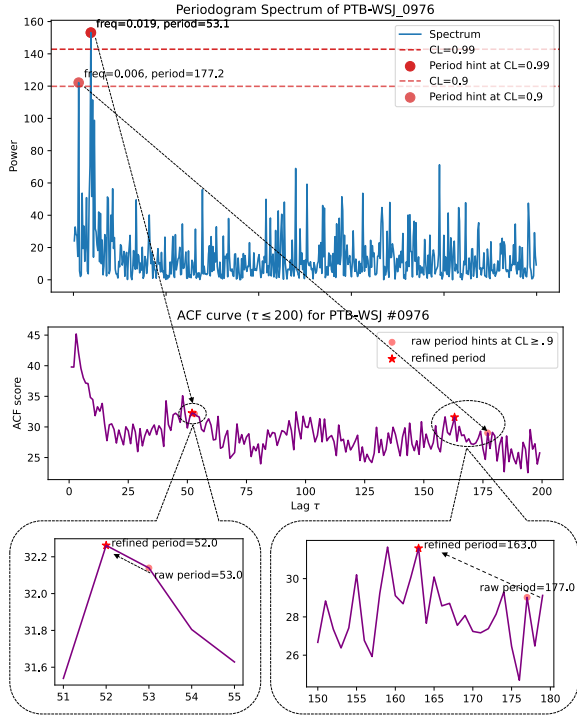


Figure 2: ACF filtering for surprisal sequence of document #0976 from the PTB-WSJ corpus. Period Hints (red dots; top) are mapped to the raw periods in the ACF curve (red dots; middle). ACF filtering refines and returns the final periods (red stars; bottom).

and EvoBench by Yu et al. (2025). Our analysis focuses on the English news and Chinese news portions of FACE, and the XSum portion of EvoBench. For each portion, we select a subset generated by the largest models (LLaMA3-70B for English FACE, Qwen2-72B for Chinese FACE, GPT-4o for EvoBench). The specific subsets used are detailed in Table 8, with results presented in Section 7.

To extract surprisal sequences from the corpora, LLMs are used to encode the input texts. For English texts, LLaMA3-8B (Touvron et al., 2023) is adopted. For Chinese texts, Qwen2-7B (Yang et al., 2024) is adopted. Specifically, to ensure comparability with the experiments of Tsipidi et al. (2025), we use Yarn-LLaMA2-7B (Peng et al., 2024) for RST Discourse Treebank.

5 Result: Identified Periods of Information

5.1 Proportions of Periodic Documents

Applying APS on a text corpus, indicated by Σ , would partition it into three subsets: P_1 , the set of documents that contain at least one period hint (not necessarily a valid period); P_2 , the set of documents

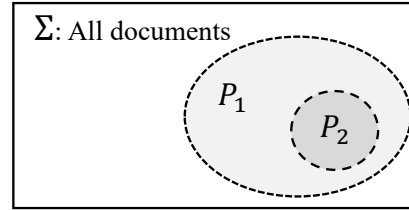


Figure 3: Partitioning a corpus based on the periodicity detection outcomes from APS.

that contain at least one valid period; and $\Sigma - P_1$, the remaining set containing no period hints or valid periods (see Figure 3).

Across the corpora examined, we consistently find that **a non-negligible proportion of texts exhibit strong periodicity of information**. Table 1 shows the counts of documents $|\Sigma|$, $|P_1|$ and $|P_2|$, and the proportions $|P_1|/|\Sigma|$, $|P_2|/|P_1|$ and $|P_2|/|\Sigma|$. P_2 are the most strictly periodic documents, whose proportion $|P_2|/|\Sigma|$ reaches above 5% for all corpora. P_1 is the quasi-periodic set, and its odds $|P_1|/|\Sigma|$ is significantly higher, reaching about 15% on average. The relatively high ratio of $|P_2|/|P_1|$ (about 66%) indicates that once a hint is detected by GETPERIODHINTS (Step 1), it has a large chance to pass ACFFILTERING (Step 2).

All three ratios are higher in Chinese corpora than in English. The exceptionally high values of GCDT against the rest also suggests that text genre is an important factor affecting periodicity – GCDT contains five genres of formal texts (academic articles, biographies (bio), interview conversations, news, and how-to guides (whow)), while WSJ, Brown, and CTB are solely collected from news articles. We will leave the comprehensive examination on how text genre (such as the dimensions of verse vs. prose, oral vs. written, etc.) influences the periodicity of information as future work.

5.2 Comparing APS Periods With Linguistic Structural Units

To assess the relationship between the global periodicity detected by APS and the local linguistic structures focused on by Tsipidi et al. (2025), we conducted a quantitative comparison between length distribution of the valid periods detected by APS and those of three structural units: elementary discourse units (EDUs) annotated in the RST Discourse Treebank, sentences, and paragraphs. The distribution of the detected periods is presented in Figure 4, along with the length distributions of

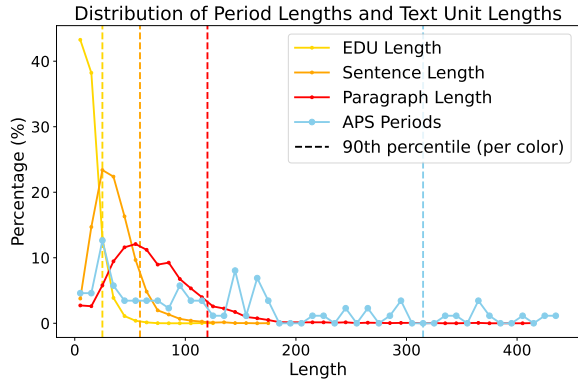


Figure 4: Distribution of valid periods identified by APS on Discourse Treebank, and the length distributions of structural units (EDUs, sentences, paragraphs). The dashed lines indicate the 90th percentiles of the corresponding distributions.

the structural units. The most prominent peak in the period distribution is observed around 25 tokens, which aligns closely with the peaks of length distributions from the three structural units. This alignment suggests that a subset of the detected periodic patterns can indeed be attributed to these fundamental structural units.

However, a substantial proportion of the detected periods fall outside the typical length distributions of the three considered structural units. Notably, the *paragraph* (the longest among the three) has 3.63% of its instances exceeding 150 tokens, and only 1.11% exceeding 200 tokens. In contrast, there are 33.33% of the valid periods detected by APS that exceed 150 tokens, and 21.84% that exceed 200 tokens. The significant portion of long-period patterns identified by APS cannot be accounted for by the structural units. This clear mismatch indicates that **the global periodicity of information identified by APS is not fully explicable by the succession of local discourse structures alone**. Instead, it suggests the presence of additional, yet-to-be-identified, factors contributing to the long-periodicity of information in text.

6 Result: Validity-Check with Harmonic Regression Modeling

In this section, we use an alternative method to carry out a validity-check on the detection results from APS. Tsipidi et al. (2025) formulates periodicity of information as the effect of harmonic components in predicting the surprisal of a token w_t using harmonic regression (HR) models. In their models, periods are defined as the lengths of structural units

(EDUs, sentences, and paragraphs). The prediction task is formulated as:

$$s(w_t) \sim \text{baseline} + \text{HR}(U_t) \quad (1)$$

in which “baseline” indicates the same basic predictors (e.g., relative positions of w_t , etc.) used in Tsipidi et al. (2025), and the $\text{HR}(U_t)$ term is the linear combination of sine and cosine components corresponding to a certain unit length U_t :

$$\begin{aligned} \text{HR}(U_t) = \beta_0 + \sum_{k=1}^K \left(\beta_{1,k} \cdot \sin\left(\frac{k2\pi t}{U_t}\right) \right. \\ \left. + \beta_{2,k} \cdot \cos\left(\frac{k2\pi t}{U_t}\right) \right) \end{aligned} \quad (2)$$

Significant coefficients in the form of amplitude $A_k = \sqrt{\beta_{1,k}^2 + \beta_{2,k}^2}$ from the $\text{HR}(U_t)$ terms would indicate that certain degrees of periodicity exist within the scope of U_t . Based on this paradigm, we proceed with two inquiries:

1. Does using the hints and periods from APS (our method) as the scaling factor U_t result in significant HR modeling outcomes?
2. Do the documents that return positive in APS detection (P_2) result in better overall performance in HR modeling?

6.1 Hints/Valid Periods Directly Used as Scaler in HR

We replace U_t in Equation (1) with the *hints* and valid *periods* returned by APS, and then examine the corresponding HR modeling outcomes. If there are multiple hints or periods in one document, then the longest one is selected. The modeling formulas then become:

$$\begin{aligned} m_{\text{hints}} : s(w_t) &\sim \text{baseline} + \text{HR}(\text{hints}) \\ m_{\text{periods}} : s(w_t) &\sim \text{baseline} + \text{HR}(\text{periods}) \end{aligned}$$

The models m_{hints} and $m_{\text{valid period}}$ are fit against the P_1 and P_2 portions of WSJ, respectively (to be directly comparable to Tsipidi et al.’s (2025) results, only WSJ is used). The coefficients of the harmonic terms $\text{HR}(\text{hints})$ and $\text{HR}(\text{periods})$ with the top-3 largest amplitude values are reported in Tables 2 and 3. The results show that both terms are indeed significant predictors in HR modeling, which confirms the periodicity of hints and valid periods that APS returns.

Table 1: Counts of document with non-empty period hints ($|P_1|$) and valid periods ($|P_2|$) identified by APS, along with corresponding ratios relative to the total document counts ($|\Sigma|$) across four corpora. The last two rows are the mean and median lengths of the hints/valid periods identified in P_1 and P_2 , respectively.

	WSJ _{en}	Brown _{en}	GCDT _{zh}	CTB _{zh}	Avg.
$ \Sigma $	2499	500	50	2773	—
$ P_1 $	221	52	15	304	—
$ P_2 $	131	25	13	212	—
$ P_1 / \Sigma $	8.84%	10.40%	30.00%	10.96%	15.05%
$ P_2 / P_1 $	59.28%	48.08%	86.67%	69.74%	65.94%
$ P_2 / \Sigma $	5.24%	5.00%	26.00%	7.65%	10.97%
P_1 : mean (median)	127.68 (100)	111.37 (28)	113.76 (81)	97.53 (51)	112.36 (74)
P_2 : mean (median)	135.65 (108)	144.57 (84)	110.53 (56)	89.69 (34)	113.14 (70)

Table 2: Significance test of the top-3 harmonic components with the largest amplitudes in HR models using **APS periods** as scaling factors on WSJ (P_2 portion).

k	A_k	β	coef	$P > t $
1	0.2022	$\beta_{1,1}$	0.1213	0.000
		$\beta_{2,1}$	0.1571	0.000
3	0.0652	$\beta_{1,3}$	0.0524	0.000
		$\beta_{2,3}$	0.0405	0.000
4	0.0467	$\beta_{1,4}$	0.0367	0.002
		$\beta_{2,4}$	0.0289	0.010

There is no significant differences between the A_k values of hints and valid periods. This suggests that, although the latter is stronger than the former by the standard of passing the periodicity threshold, their modeling effects on predicting local surprisal are almost equivalent. We also find that the A_k values from hints and valid periods are smaller than those from EDU, the best predictor reported in Tsipidi et al. (2025). This is to be expected, as hints/valid periods are fixed values within a document and therefore less dependent on the position of tokens. This naturally makes them less informative predictors. However, this does not affect their periodicity. Rather, it strengthens our conclusion that APS can detect global periodic patterns that differ from local units.

6.2 Filtering Data Using Hints/Valid Periods Improves HR Performance

To test whether APS can filter out data with better overall periodicity, we fit HR models on four different portions, $\{P_2, P_1, \Sigma, \Sigma - P_1\}$ of RST and GCDT corpora, using two structural units (EDUs,

Table 3: Significance test of the top-3 harmonic components with the largest amplitudes in HR models using **APS hints** as scaling factors on WSJ (P_1 portion).

k	A_k	β	coef	$P > t $
1	0.1772	$\beta_{1,1}$	0.0736	0.000
		$\beta_{2,1}$	0.1612	0.000
5	0.0731	$\beta_{1,5}$	0.0503	0.000
		$\beta_{2,5}$	-0.0531	0.000
7	0.0575	$\beta_{1,7}$	0.0036	0.723
		$\beta_{2,7}$	-0.0574	0.000

sentences) along with whole texts as different scaling factors. We assume that the resulting mean squared errors (MSE) should rank in a descending order, i.e., $P_2 < P_1 < \Sigma < \Sigma - P_1$, as P_2 is the most periodic portion according to APS, while $\Sigma - P_1$ is the least.

The results in Tables 4 and 5 confirm our assumption. The rankings remain consistent across different scaling factors and even so in baseline models (without HR terms). This indicates that the documents identified by APS as containing strong periodicity of information are more predictable in terms of token surprisal. Scaling by structural units further improves the performance on the filtered portions, and EDU scaling is the most effective one, which is consistent with Tsipidi et al.’s (2025) results. From this perspective, the global periodicity identified by APS and the local periodicity modeled by HR using structural units are compatible. It is possible that the periodic patterns of discourse structures contribute to the overall periodicity of information in text.

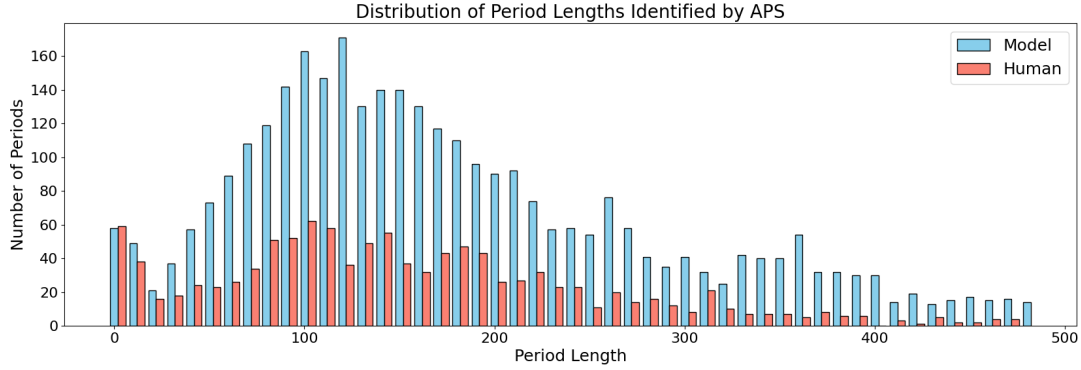


Figure 5: Distributions of valid periods from human (FACE BBC News) and LLM-generated (LLaMA-3-70B) texts.

Table 4: MSE of HR modeling on **WSJ** using different scaling factors and different data portions.

	Baseline	EDU	Sentence	Whole Text
P_2	16.3273	13.7788	15.5958	16.2915
P_1	16.8654	14.1434	16.1400	16.8353
Σ	18.2635	15.1843	17.5619	18.2470
$\Sigma - P_1$	18.5859	15.4160	17.8875	18.5692

Table 5: MSE of HR modeling on **GCDT** using different scaling factors and different data portions.

	Baseline	EDU	Sentence	Whole Text
P_2	6.8699	6.3902	6.8070	6.8489
P_1	6.9967	6.5097	6.9303	6.9789
Σ	7.2485	6.5961	7.1511	7.2279
$\Sigma - P_1$	7.2832	6.5644	7.1673	7.2554

7 Result: Human vs. Generated Texts

By applying APS on the corpora of human-written and LLM-generated text pairs (FACE and EvoBench; see Section 4), we can observe periodicity differences between them at the document level. These datasets contain generated texts obtained from LLMs by feeding the beginning of a human-written text as a prompt. In general, we find that LLM-generated texts contain stronger periodicity of information than human-written ones in both English and Chinese. For instance, the proportion of documents with valid periods in LLaMA3-70B-generated fake BBC news is 30.06%, which is more than twice the proportion in the real news (14.80%). Detailed results are listed in Table 8.

Figure 5 shows the distributions of the detected periods in human-written and LLM-generated texts in FACE BBC News data. It can be observed that the LLM-generated texts exhibit more periodic patterns than human-written texts, especially in the long-period range (> 50 tokens). One of the pos-

sible reasons for the stronger periodicity in LLM-generated texts is that LLMs may generate repetitive content. Sometimes, the generated text may contain repeated phrases or sentences, which can create a strong periodic pattern. We randomly sample 10% of the documents with valid periods in LLaMA3-70B-generated BBC news (150 documents in total), and find that 7.33% of them (11 documents) contain repeated phrases or sentences. Thus, repetition could be one of the factors contributing to the difference, especially in the short-period range. However, the gap between the ratios of documents with valid periods in human-written and LLM-generated texts is still significant (14.80% vs. 27.86%) after excluding the repeated portion. Another possible reason is that LLMs are more likely to generate similar structures across paragraphs, maintaining structural coherence under large spans. Human writers, however, tend to be more flexible in their narrative structures, avoiding orderly patterns. This could lead to stronger periodicity in LLM-generated texts in the long-period range.

The above finding lends interpretation to some existing studies that utilize periodicity of information as a feature for the task of detecting model-generated texts (Xu et al., 2024) and for automatic NLG evaluation (Liu et al., 2025). Through APS, we can explicitly tell *how* the periodicity of information in LLM-generated texts is different from that in human-written ones. Therefore, information periodicity holds promise as a distinctive feature for generation-related tasks.

8 Conclusions

In this study, we propose AutoPeriod of Surprisal (APS), an algorithm for revealing periods of surprisal in natural language with a high resolution,

which contributes to the emerging research topic of periodicity of information in human language. Our main observations and conclusions are as follows:

Periodicity can be detected. Our method confirms that, although it is a subtle phenomenon, the periodicity of information can be detected at the single document level. It is validated that there is a considerable proportion of human-written language that exhibits significant periodicity of information.

Sources of periodicity. We find evidence for periods that are not purely due to the linguistic structures, such as sentences, paragraphs, and EDUs. This indicates that there are other factors than structural information processing pressure that cause periodicity, which needs to be studied in future work. We speculate that the longer periods detected could be due to some latent structures (such as topic) that span across multiple paragraphs.

Usability in human-machine distinction. APS is not a regression-based technique that focuses on analysis, but rather a detection method that attempts to extract useful cues about how information distributes temporally given any input text. Therefore, it has potential in distinguishing human- vs. model-generated texts, under the basic assumption that human language production is a cognitively-constrained process.

Limitations

Within the scope of this study, we consider the following limitations, which we believe can be further addressed in future work. Firstly, this study only examines English and Chinese texts. As languages differ in structure, we believe that our proposed APS method might vary in performance across different texts from different language.

Secondly, our detection results in Section 5.1 suggest that the periodicity of information is influenced by text genre. We also find that parts of the detected periods cannot be explained by local structural units, as discussed in Section 5.2. However, we do not further analyze the sources of these periods or potential factors that influence the periodicity of information, leaving these inquiries to future work.

Thirdly, the performance of AutoPeriod might not be as accurate in detecting signals with short periods as those with long periods. The corresponding frequencies of the component signals returned by DFT distribute unevenly on the axis of period length. Components with short periods are more densely distributed than those with long periods.

When searching for period hints by random perturbation, short periods are thus more likely to be false positives. From the distribution of the detected periods, we observe that the prominent peak in the region of short periods tends to shrink when the confidence level is set higher. The observation indicates that the detection of short periods is less robust.

Ethics Statements

A potential ethical issue of this research results from the discovery of significant differences in periodicity of information between human- and model-generated texts (Section 7). While a promising application of this is the detection of LLM-generated text, such applications have the potential to be misused. Direct conclusions should not be drawn when a text is identified as having unusually high periodicity, as this could be caused by various factors, such as the use of an LLM-based writing assistant, the author's individual style, or the author writing in a second language (which can result in more regular, formulaic expression). Detection results should thus be approached with caution.

The text examples shown in Figures 1 and 6 are for illustrative purposes only, which is in compliance with the fair use license of the [Linguistic Data Consortium](#).

Acknowledgments

We sincerely thank Eleftheria Tshipidi and Mario Giulianelli for providing the data for harmonic regression analysis and for the insightful discussions. We also thank all the reviewers for their feedback on the paper. This study is funded by Shenzhen Science and Technology Program (No. JCYJ20240813094612017) and Guangdong Province ZJRC Program (No. 2024QN11X145). This study is also funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): TRR 318/3 2026 – 438445824, B07.

References

- Matthew Aylett and Alice Turk. 2004. [The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech](#). *Language and Speech*, 47(1):31–56.
- Alan Bell, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, Michelle Gregory, and Daniel Gildea.

2003. Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America*, 113(2):1001–1024.
- Razvan C. Bunescu and Oseremen O. Uduehi. 2022. Distribution-based measures of surprise for creative language: Experiments with humor and metaphor. In *Proceedings of the 3rd Workshop on Figurative Language Processing (FLP)*, pages 68–78, Abu Dhabi, UAE (Hybrid). ACL.
- Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski. 2001. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of the 2nd SIGdial Workshop on Discourse and Dialogue*, page 1–10, Aalborg, Denmark. ACL.
- Vera Demberg, Asad Sayeed, Philip Gorinski, and Nikolaos Engonopoulos. 2012. Syntactic surprisal affects spoken word duration in conversational contexts. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 356–367, Jeju Island, Korea. ACL.
- Dmitriy Genzel and Eugene Charniak. 2002. Entropy rate constancy in text. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 199–206, Philadelphia, PA, USA. ACL.
- Dmitriy Genzel and Eugene Charniak. 2003. Variation of entropy and parse trees of sentences as a function of the sentence number. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, pages 65–72, Sapporo, Japan. ACL.
- Mario Giulianelli and Raquel Fernández. 2021. Analysing human strategies of information transmission as a function of discourse context. In *Proceedings of the 25th Conference on Computational Natural Language Learning*, pages 647–660, Online. ACL.
- Mario Giulianelli, Arabella Sinclair, and Raquel Fernández. 2021. Is information density uniform in task-oriented dialogues? In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 8271–8283, Punta Cana, Dominican Republic. ACL.
- T. Florian Jaeger. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61(1):23–62.
- T. Florian Jaeger and Roger Levy. 2006. Speakers optimize information density through syntactic reduction. *Advances in Neural Information Processing Systems*, 19.
- Gaurav Kharkwal and Smaranda Muresan. 2014. Surprisal as a predictor of essay quality. In *Proceedings of the 9th Workshop on Innovative Use of NLP for Building Educational Applications*, pages 54–60, Baltimore, MA, USA. ACL.
- Zhichen Liu, Yongyuan Li, Yang Xu, Yu Wang, Yingfang Yuan, and Zuhao Yang. 2025. Evaluating text generation quality using spectral distances of surprisal. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 2444–2463, Suzhou, China. ACL.
- Nick R. Lomb. 1976. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39(2):447–462.
- Eliot Maës, Philippe Blache, and Leonor Becerra. 2022. Shared knowledge in natural conversations: Can entropy metrics shed light on information transfers? In *Proceedings of the 26th Conference on Computational Natural Language Learning (CoNLL)*, pages 213–227, Abu Dhabi, UAE (Hybrid). ACL.
- Kyle Mahowald, Evelina Fedorenko, Steven T. Piantadosi, and Edward Gibson. 2013. Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*, 126(2):313–318.
- Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2):313–330.
- Clara Meister, Tiago Pimentel, Patrick Haller, Lena Jäger, Ryan Cotterell, and Roger Levy. 2021. Revisiting the Uniform Information Density hypothesis. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 963–980, Online and Punta Cana, Dominican Republic. ACL.
- Alan V. Oppenheim, Ronald W. Schaffer, and John R. Buck. 1999. *Discrete-Time Signal Processing*. Prentice Hall, Upper Saddle River, NJ, USA.
- Bowen Peng, Jeffrey Quesnelle, Honglu Fan, and Enrico Shippole. 2024. YaRN: Efficient context window extension of large language models. In *Proceedings of the 12th International Conference on Learning Representations*, Vienna, Austria.
- Siyao Peng, Yang Janet Liu, and Amir Zeldes. 2022. GCDT: A Chinese RST treebank for multigenre and multilingual discourse parsing. In *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing*, pages 382–391, Online. ACL.
- Jeffrey D. Scargle. 1982. Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *Astrophysical Journal*, 263(2):835–835.
- Arthur Schuster. 1898. On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Terrestrial Magnetism*, 3(1):13–41.
- Claude E. Shannon. 1948. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423.

- Claude E. Shannon. 1949. [Communication theory of secrecy systems](#). *The Bell System Technical Journal*, 28(4):656–715.
- Claude E. Shannon. 1951. The redundancy of English. In *Transactions of the 7th Conference on Cybernetics*, pages 248–272, New York, NY, USA.
- Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023. [LLaMA: Open and efficient foundation language models](#). *Preprint*, arXiv:2302.13971.
- Eleftheria Tsipidi, Samuel Kiegeand, Franz Nowak, Tianyang Xu, Ethan Wilcox, Alex Warstadt, Ryan Cotterell, and Mario Giulianelli. 2025. [The harmonic structure of information contours](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 31636–31659, Vienna, Austria. ACL.
- Eleftheria Tsipidi, Franz Nowak, Ryan Cotterell, Ethan Wilcox, Mario Giulianelli, and Alex Warstadt. 2024. [Surprise! Uniform Information Density isn't the whole story: Predicting surprisal contours in long-form discourse](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 18820–18836, Miami, Florida, USA. ACL.
- Michail Vlachos, Philip S. Yu, and Vittorio Castelli. 2005. [On periodicity detection and structural periodic similarity](#). In *Proceedings of the 2005 SIAM International Conference on Data Mining*, pages 449–460, Newport Beach, CA, USA. SIAM.
- Yubo Xie, Junze Li, and Pearl Pu. 2021. [Uncertainty and surprisal jointly deliver the punchline: Exploiting incongruity-based features for humor recognition](#). In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 33–39, Online. ACL.
- Yang Xu and David Reitter. 2016. [Entropy converges between dialogue participants: Explanations from an information-theoretic perspective](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 537–546, Berlin, Germany. ACL.
- Yang Xu and David Reitter. 2017. [Spectral analysis of information density in dialogue predicts collaborative task performance](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 623–633, Vancouver, Canada. ACL.
- Yang Xu and David Reitter. 2018. [Information density converges in dialogue: Towards an information-theoretic model](#). *Cognition*, 170:147–163.
- Yang Xu, Yu Wang, Hao An, Zhichen Liu, and Yongyuan Li. 2024. [Detecting subtle differences between human and model languages using spectrum of relative likelihood](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 10108–10121, Miami, Florida, USA. ACL.
- Nianwen Xue, Xiuhong Zhang, Zixin Jiang, Martha Palmer, Fei Xia, Fu-Dong Chiou, and Meiyu Chang. 2013. [Chinese treebank 8.0 ldc2013t21](#). Linguistic Data Consortium, Philadelphia, PA, USA.
- An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, Guanting Dong, Haoran Wei, Huan Lin, Jialong Tang, Jialin Wang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Ma, and 43 others. 2024. [Qwen2 technical report](#). *Preprint*, arXiv:2407.10671.
- Zuhao Yang, Yingfang Yuan, Yang Xu, Shuo Zhan, Hua-jun Bai, and Kefan Chen. 2023. [FACE: Evaluating natural language generation with Fourier analysis of cross-entropy](#). In *Advances in Neural Information Processing Systems*, volume 36, pages 17038–17056. Curran.
- Xiao Yu, Yi Yu, Dongrui Liu, Kejiang Chen, Weiming Zhang, Nenghai Yu, and Jing Shao. 2025. [EvoBench: Towards real-world LLM-generated text detection benchmarking for evolving large language models](#). In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 14605–14620, Vienna, Austria. ACL.

Table 6: Significance test of all the amplitudes in HR models using **APS hints** as scaling factors on WSJ (P_1 portion).

k	A_k	β	coef	$P > t $
1	0.1772	$\beta_{1,1}$	0.0736	0.000
		$\beta_{2,1}$	0.1612	0.000
2	0.0255	$\beta_{1,2}$	0.0228	0.023
		$\beta_{2,2}$	0.0115	0.227
3	0.0539	$\beta_{1,3}$	0.0335	0.001
		$\beta_{2,3}$	0.0423	0.000
4	0.0462	$\beta_{1,4}$	0.0246	0.016
		$\beta_{2,4}$	0.0391	0.000
5	0.0731	$\beta_{1,5}$	0.0503	0.000
		$\beta_{2,5}$	-0.0531	0.000
6	0.0373	$\beta_{1,6}$	0.0034	0.749
		$\beta_{2,6}$	0.0371	0.000
7	0.0575	$\beta_{1,7}$	0.0036	0.723
		$\beta_{2,7}$	-0.0574	0.000
8	0.0318	$\beta_{1,8}$	0.0304	0.003
		$\beta_{2,8}$	0.0093	0.326
9	0.0191	$\beta_{1,9}$	0.0163	0.113
		$\beta_{2,9}$	0.0099	0.290
10	0.0406	$\beta_{1,10}$	0.0343	0.001
		$\beta_{2,10}$	-0.0217	0.021

Table 7: Significance test of all the amplitudes in HR models using **APS periods** as scaling factors on WSJ (P_2 portion).

k	A_k	β	coef	$P > t $
1	0.2022	$\beta_{1,1}$	0.1273	0.000
		$\beta_{2,1}$	0.1571	0.000
2	0.0449	$\beta_{1,2}$	0.0316	0.007
		$\beta_{2,2}$	0.0319	0.005
3	0.0662	$\beta_{1,3}$	0.0524	0.000
		$\beta_{2,3}$	0.0405	0.000
4	0.0467	$\beta_{1,4}$	0.0367	0.002
		$\beta_{2,4}$	0.0289	0.010
5	0.0466	$\beta_{1,5}$	0.0453	0.000
		$\beta_{2,5}$	-0.0111	0.318
6	0.0296	$\beta_{1,6}$	0.0282	0.018
		$\beta_{2,6}$	-0.0090	0.420
7	0.0216	$\beta_{1,7}$	0.0190	0.108
		$\beta_{2,7}$	-0.0101	0.369
8	0.0299	$\beta_{1,8}$	0.0133	0.262
		$\beta_{2,8}$	-0.0268	0.018
9	0.0310	$\beta_{1,9}$	0.0234	0.050
		$\beta_{2,9}$	0.0203	0.069
10	0.0083	$\beta_{1,10}$	0.0036	0.762
		$\beta_{2,10}$	-0.0075	0.504

A Implementation Details of APS

A full version of the APS algorithm is presented in Algorithm 3. In contrast to the abbreviated pseudocode presented above in Algorithm 2, it details the process of determining the search window W , the linear regression-based hill validation, and the final period selection strategy.

B Amplitudes and Corresponding p -Values of HR Modeling

To verify the periodicity of information identified by APS, we fit HR models using the hints and valid periods returned by APS as the scaling factors, and examine the significance of the harmonic terms. We simplify the experimental setting of HR modeling, picking 10 terms ($K = 10$) in Equation (2), and fitting the model on the entire document set of WSJ. The amplitudes A_k and the corresponding p -values of the coefficients $\beta_{1,k}$ and $\beta_{2,k}$ are reported in Tables 6 and 7.

C Full Results of Human vs. Generated Texts

We provide the full results of the detection of periods in human-written vs. LLM-generated texts in Table 8, including the counts and ratios of documents with non-empty period hints ($|P_1|$) and valid periods ($|P_2|$) identified by APS in three corpora. For FACE, we choose the News portion of the

dataset in both English and Chinese, and select the texts generated by LLaMA3-70B and Qwen2-72B as the LLM-generated texts, respectively. For EvoBench, we choose the XSum portion of the dataset, and select the subset generated by GPT-4o as the LLM-generated texts.

D A Complete Example of APS on a WSJ Document

For a text sample from document #2380 from PTB-WSJ (presented in Figure 6), we apply APS to detect the periods of surprisal. In the token surprisal curve, shown in Figure 7, we mark the intervals of paragraphs with dashed red lines. The average length of the paragraphs is 61.76 tokens. The periodogram of the surprisal curve is shown in Figure 8, where the dashed lines represent different confidence levels, and the period hints above the 99% confidence level are marked with crosses. The candidate periods are 60, 99, 101, 374, and 403 tokens. The ACF curves at the period hints of 61, 99, and 375 tokens (the best splits within the windows) are shown in Figure 9a, Figure 9b and Figure 9c respectively. We use ACF curves to filter and refine the period hints. For this example, the period hint of 403 tokens is filtered out, and the period hint of 99 tokens is refined to 101 tokens. The final detected valid periods are 61, 101, and 374 tokens.

Algorithm 3 ACF Filtering (full version)

```
1: function ACFFILTERING( $\mathbf{x}$ , periodHints)
2:    $N \leftarrow \mathbf{x}.length$ 
3:    $ACF \leftarrow \text{GETACF}(\mathbf{x})$ 
4:    $M \leftarrow \text{periodHints}.length$ 
5:    $refinedPeriods \leftarrow \{\}$ 
6:   for  $i = 1$  to  $M$  do
7:      $\tau \leftarrow \text{periodHints}[i]$ 
8:      $\tau \leftarrow N/k \leftarrow \text{periodHints}[i]$ 
9:                                      $\triangleright$  Determine search window using adjacent periods
10:     $\tau_{prev} \leftarrow N/(k - 1)$ 
11:     $\tau_{next} \leftarrow N/(k + 1)$ 
12:     $W \leftarrow [\frac{\tau + \tau_{next}}{2} - 1, \frac{\tau + \tau_{prev}}{2} + 1]$ 
13:                                      $\triangleright$  Find optimal split point using linear regression
14:     $\epsilon_{min} \leftarrow \infty$ 
15:     $t_{best} \leftarrow -1$ 
16:     $slope_L \leftarrow 0, slope_R \leftarrow 0$ 
17:    for each  $t$  in  $W$  do
18:      Split  $W$  at  $t$  into  $L$  and  $R$ 
19:       $slope_L^t, \epsilon_L^t \leftarrow \text{LINEARREGRESSION}(L, ACF[L])$ 
20:       $slope_R^t, \epsilon_R^t \leftarrow \text{LINEARREGRESSION}(R, ACF[R])$ 
21:      if  $\epsilon_L^t + \epsilon_R^t < \epsilon_{min}$  then
22:         $\epsilon_{min} \leftarrow \epsilon_L^t + \epsilon_R^t$ 
23:         $t_{best} \leftarrow t$ 
24:         $slope_L \leftarrow slope_L^t, slope_R \leftarrow slope_R^t$ 
25:      end if
26:    end for
27:                                      $\triangleright$  Validate if the split forms a hill
28:     $\theta_L \leftarrow \arctan(slope_L)/(\pi/2)$ 
29:     $\theta_R \leftarrow \arctan(slope_R)/(\pi/2)$ 
30:     $\Delta\theta \leftarrow |\theta_L - \theta_R|$ 
31:     $isHill \leftarrow (slope_L > slope_R) \wedge (\Delta\theta > 0.01)$ 
32:    if  $isHill$  then
33:       $\tau_{refined} \leftarrow \text{FINDPEAK}(ACF, W, t_{best})$ 
34:       $refinedPeriods.ADD(\tau_{refined})$ 
35:    end if
36:  end for
37:  return  $refinedPeriods$ 
38: end function
```

Table 8: Counts and ratios of documents with non-empty period hints ($|P_1|$) and valid periods ($|P_2|$) identified by APS in human-written vs. LLM-generated texts across three corpora.

	EvoBench _{en}		FACE _{en}		FACE _{zh}	
	Human	GPT-4o	Human	LLaMA3-70B	Human	Qwen2-72B
$ \Sigma $	150	150	5000	5000	5000	5000
$ P_1 $	9	19	1032	1882	1209	1592
$ P_2 $	5	11	740	1503	943	1231
$ P_1 / \Sigma $	6.00%	12.67%	20.64%	37.64%	24.18%	31.84%
$ P_2 / P_1 $	55.56%	57.89%	71.71%	79.86%	78.00%	77.32%
$ P_2 / \Sigma $	3.33%	7.33%	14.80%	30.06%	18.86%	24.62%

Friday , October 13 , 1989

The key U.S. and foreign annual interest rates below are a guide to general levels but do n't always represent actual transactions .

PRIME RATE : 10 1/2 % . The base rate on corporate loans at large U.S. money center commercial banks .

FEDERAL FUNDS : 8 13/16 % high , 8 1/2 % low , 8 5/8 % near closing bid , 8 3/4 % offered . Reserves traded among commercial banks for overnight use in amounts of \$ 1 million or more . Source : Fulton Prebon (U.S.A .) Inc .

DISCOUNT RATE : 7 % . The charge on loans to depository institutions by the New York Federal Reserve Bank .

...

FEDERAL HOME LOAN MORTGAGE CORP . (Freddie Mac) : Posted yields on 30-year mortgage commitments for delivery within 30 days . 9.91 % , standard conventional fixedrate mortgages ; 7.875 % , 2 % rate capped one-year adjustable rate mortgages . Source : Telerate Systems Inc .

FEDERAL NATIONAL MORTGAGE ASSOCIATION (Fannie Mae) : Posted yields on 30 year mortgage commitments for delivery within 30 days (priced at par) 9.86 % , standard conventional fixed-rate mortgages ; 8.85 % , 6/2 rate capped one-year adjustable rate mortgages . Source : Telerate Systems Inc .

MERRILL LYNCH READY ASSETS TRUST : 8.33 % . Annualized average rate of return after expenses for the past 30 days ; not a forecast of future returns .

Figure 6: Sample text from document #2380 from the PTB-WSJ corpus.

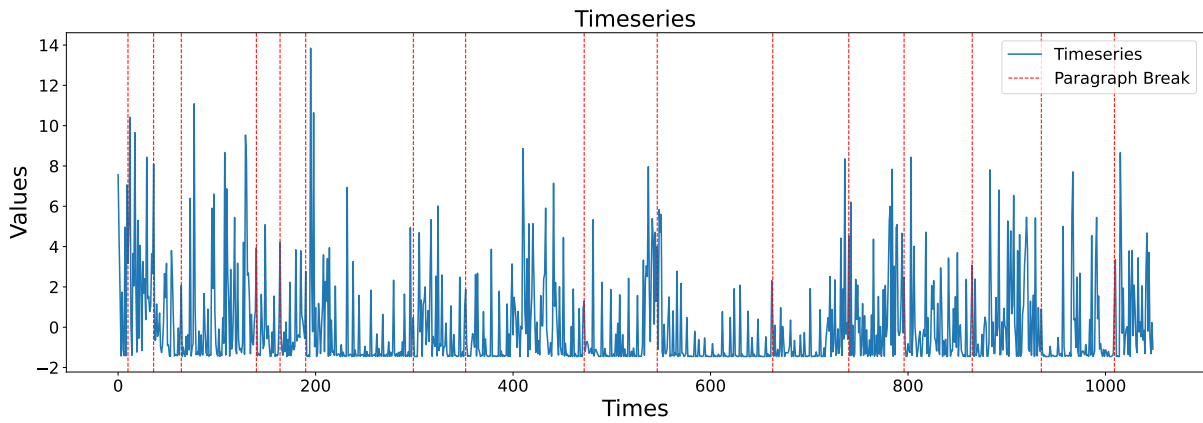


Figure 7: The token surprisal curve of document the #2380 from the PTB-Wsj corpus shown in Figure 6. The dashed red lines indicate the intervals of paragraphs. The average length of the paragraphs is 61.76 tokens.

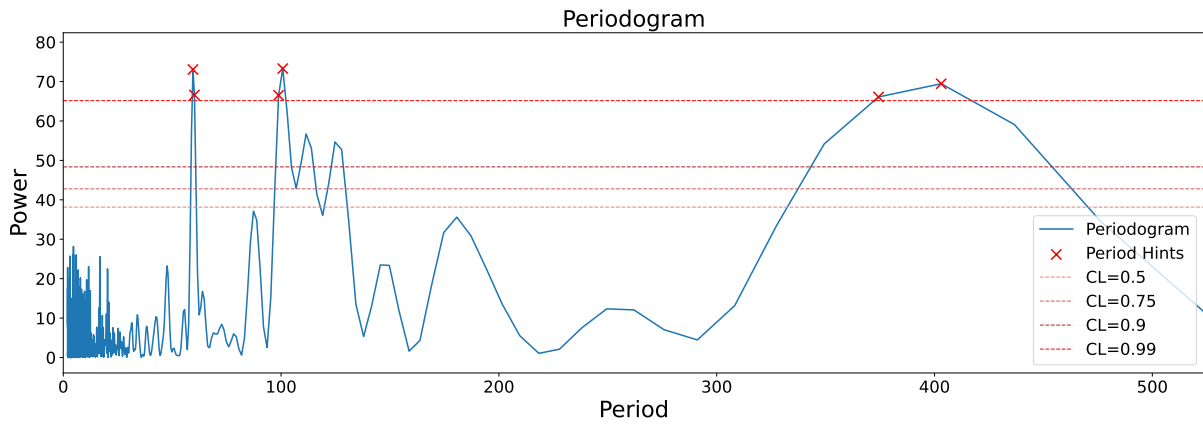
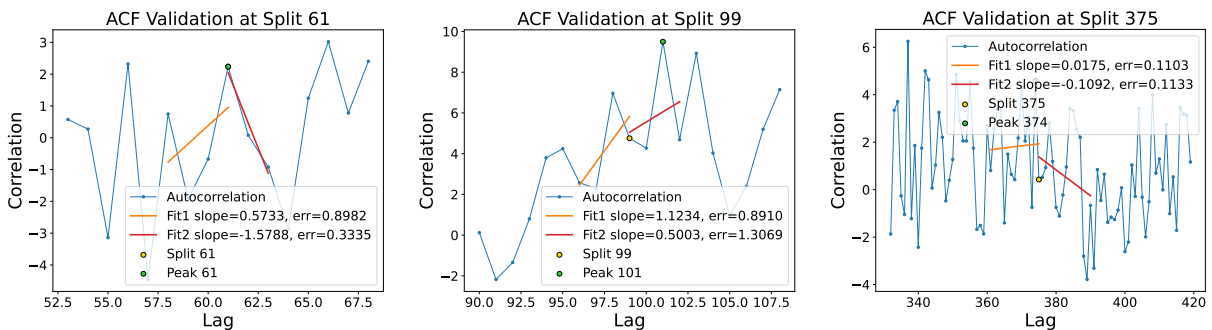


Figure 8: The periodogram of the surprisal curve in Figure 7. The dashed lines represent different confidence levels. The period hints above the 99% confidence level are marked with crosses.



(a) ACF at the period hint of 61 tokens. (b) ACF at the period hint of 99 tokens. (c) ACF at the period hint of 375 tokens.

Figure 9: Examples of the ACF curves at the period hints 61, 99, and 375 returned by APS. The yellow split points are the optimal split points determined by the linear regression-based hill validation, the green peak points are the peaks within the search windows.