

MIES-TR: An Intelligent Model for Real-Time Syllabic Extraction during Keyboard Typing

Tchanque Tchakountio Armand¹, Azanguezet Quimatio Benoît Martin², Geh Wilson Ejuh³,
Tatsoula Toukem Junie¹, Chassem Kamdem Priva¹

¹Fundamental Computer Science, Engineering and Application Research Unit, University of Dschang, Dschang, Cameroon

²Department of Mathematics and Computer Science, University of Dschang, Dschang, Cameroon

³Department of Electrical and Electronic Engineering, National Higher Polytechnic Institute, University of Bamenda, Bamibli, Cameroon

Email: tchanquitos1@gmail.com, Azanguezet1@gmail.com, gehwilsonejah@yahoo.fr, junietoukem@gmail.com, priva.chassem@univ-dschang.org

How to cite this paper: Armand, T.T., Martin, A.Q.B., Ejuh, G.W., Junie, T.T. and Priva, C.K. (2025) MIES-TR: An Intelligent Model for Real-Time Syllabic Extraction during Keyboard Typing. *Journal of Computer and Communications*, 13, 169-185. <https://doi.org/10.4236/jcc.2025.1312010>

Received: July 21, 2025

Accepted: December 23, 2025

Published: December 26, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

We introduce in this paper MIES-TR, an intelligent model for real-time syllable boundary detection during keyboard typing. This innovative approach positions the syllable as an intermediate biometric unit, combining linguistic richness and motor stability to enhance continuous authentication systems. MIES-TR is built around an optimized neural architecture consisting of character-position encoding, multi-scale convolutions, a unidirectional causal LSTM, and a sliding local attention mechanism. Unlike traditional offline syllabation methods, our model operates in a streaming fashion, without access to future input, and achieves a latency of less than 30 ms per keystroke, enabling dynamic, efficient segmentation compatible with interactive environments. Experimental results on an annotated user corpus demonstrate strong performance, with an average F1-score of 89.9%, word accuracy of 84.2%, and proven inter-user robustness, confirming the relevance of syllabic dynamics as a behavioral identity vector. Beyond accuracy, MIES-TR naturally integrates into adaptive security architectures such as ABAC policies enhanced with dynamic attributes, offering concrete prospects in free typing, multilingual support, multimodal biometric fusion, and embedded device implementation. MIES-TR thus paves the way toward smoother, invisible, and more robust authentication at the intersection of language processing, behavioral biometrics, and real-time cybersecurity.

Keywords

Real-Time Syllable Segmentation, Behavioral Biometrics, Continuous

1. Introduction

Information systems are increasingly relying on authentication mechanisms that are not only robust, but also continuous and non-intrusive. In a context of persistent and evolving threats, one-time authentication—based on passwords or biometric traits at the beginning of a session—has proven insufficient to ensure the legitimacy of a user throughout their interaction with a system [1] [2].

This need for dynamic security has fueled the rise of behavioral biometrics, among which keystroke dynamics stands out as a particularly promising approach [3] [4]. However, traditional keystroke-based methods, which rely on low-level metrics such as dwell time (key hold duration) or flight time (time between key presses), face two major limitations [5]. First, these signals are highly sensitive to natural variations such as fatigue, stress, or posture. Second, they overlook the underlying cognitive structure of typing, treating it merely as a linear and mechanical sequence of characters.

This work aims to overcome these limitations by building on a cognitive psychology-based hypothesis: keyboard typing is guided by motor programs structured around elementary linguistic units, particularly the syllable [6] [7]. As a fundamental prosodic unit, the syllable offers a more stable and discriminative behavioral marker than classical inter-character latencies [8].

In this context, we introduce a real-time neural model called MIES-TR (Modèle Intelligent d'Extraction Syllabique en Temps Réel), capable of segmenting words into syllables as they are typed. Unlike classical syllabation methods used in natural language processing (NLP), which are often offline or supervised [9], our approach follows a streaming paradigm, making it suitable for continuous authentication. The paper presents the following contributions:

- the design of a hybrid CNN + LSTM + local attention architecture tailored for incremental syllabic segmentation;
- a character-position encoding strategy specifically designed for partial stream processing;
- an experiment on French syllable corpora (Lexique3 and user collection), demonstrating the relevance of the approach in terms of accuracy, latency, and robustness;
- a discussion on the potential integration of this model into adaptive access control systems (ABAC) [10].

This work lies at the intersection of three fields: cybersecurity, computational linguistics, and machine learning. It aims to pave the way toward a new generation of behavioral biometrics based not on mechanical measurements of keystrokes, but on their deep rhythmic and cognitive structure.

2. Related Work

2.1. The Syllable as a Motor Unit in Typing

Behavioral biometric modeling of keyboard typing has historically relied on basic metrics such as dwell time (key press duration) and flight time (inter-key latency) [4] [5]. These approaches have shown moderate success, with Equal Error Rates (EER) ranging from 5% to 20% depending on the protocol [3]. However, their reliability remains fragile in the face of contextual factors such as fatigue, stress, or changes in typing habits.

Cognitive sciences suggest that typing is not a linear sequence of letters but a structured motor activity. According to Levelt [6], written language production relies on preparatory units—known as syllables—that are planned and executed as coherent motor blocks. This hypothesis is supported by the work of Pinet *et al.* [7], which shows that significant slowdowns occur at syllable boundaries, a phenomenon known as the “unit boundary effect” [8].

This behavior reveals the existence of an intrinsic syllabic rhythm for each user—one that is relatively stable over time and difficult to consciously falsify. In a study by Lambert and Vilette [11], it was observed that inter-syllabic latencies show lower intra-individual variance than inter-letter latencies, reinforcing the idea that the syllable constitutes a relevant biometric unit.

2.2. Syllabation in Natural Language Processing (NLP)

Syllabation is a key task in natural language processing (NLP), used in applications such as speech synthesis, speech recognition, and linguistic modeling. Various approaches have been proposed, ranging from rule-based systems to modern neural models. Below is an overview of recent notable works:

2.2.1. SylNet (2019)

Räsänen *et al.* [12] introduced a neural network model for estimating the number of syllables from audio signals. The system generalizes well across languages. However, it is limited to counting tasks and does not provide precise segmentation or textual support.

2.2.2. Revisiting Syllables for Language Modeling (2020)

Oncevay and Rojas [13] evaluated syllabic language models on 20 languages. Results show comparable perplexity to character-based models with shorter sequences, favoring efficiency. However, the models require offline syllabators, making real-time integration difficult.

2.2.3. ItGraSyll (2024)

Borgo *et al.* [14] developed an Italian annotated resource for syllabation and lexical accent prediction. Their neural model achieves 98.4% word-level accuracy and 99.8% for syllable boundaries. However, this method is limited to Italian and does not support dynamic typing.

2.2.4. LSTM with ASR Embeddings (2024)

Pascual *et al.* [15] introduced a method to detect syllable boundaries in speech using internal representations from an Automatic Speech Recognition (ASR) system. Their system achieves a Word Correction Rate (WCR) above 90% on Romance languages. Its main limitation is its dependence on ASR models and its inapplicability to text-based typing.

In summary, while these approaches confirm the potential of neural models for syllabation, they also reveal a critical gap: to the best of our knowledge, no current approach supports real-time (online) syllabation from keyboard typing, *i.e.*, the detection of syllable boundaries at each new keystroke without access to the full word. This capability is essential for continuous and transparent authentication.

2.3. Limitations of Traditional Keystroke Dynamics Approaches

Traditional keystroke dynamics models rely on low-level metrics such as dwell time and flight time [5]. Despite their simplicity and low implementation cost, these systems achieve only moderate performance, especially in unconstrained environments. Acien *et al.* [4] report an average EER of 12.7% on real-world data, rising to 25% in free-typing scenarios. For instance, in the French corpus, loan-words such as “week-end” or “wifi”? as well as poly-syllabic words like “organisation”? introduce segmentation ambiguities due to non-canonical syllabic boundaries and atypical phonotactic structures. These cases occasionally lead to boundary shifts in the predicted syllable stream, particularly under high typing speed, thereby highlighting the inherent limitation of purely rhythm-based segmentation in the presence of linguistic irregularities.

Recent models based on LSTMs or autoencoders improve robustness but remain based on poorly expressive feature vectors, as they ignore the underlying linguistic structure [3]. Keystroke dynamics are treated as a stochastic sequence of mechanical pressures, without accounting for the user’s cognitive rhythm. These metrics suffer from several structural limitations.

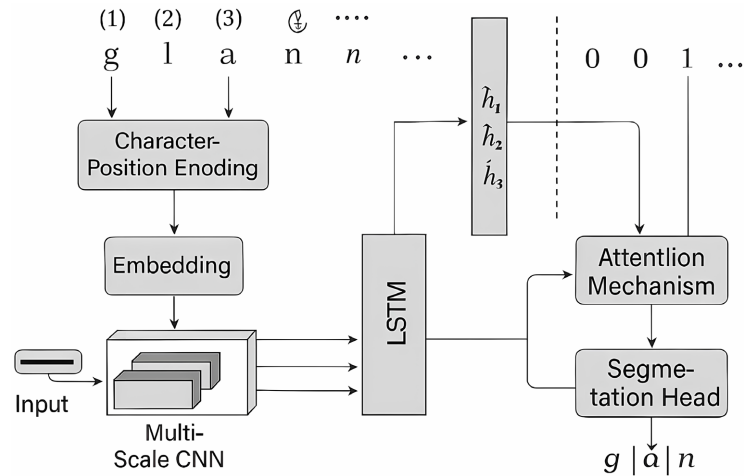
First, they are heavily influenced by contextual factors: fatigue, mood, keyboard ergonomics, body posture, and even ambient noise. These elements cause significant intra-individual variability, which undermines the temporal stability of the biometric profile [3].

Second, traditional metrics do not account for the cognitive structure of typing. They treat input as a mechanical sequence of events, ignoring higher-level linguistic processes such as lexical planning or mental segmentation into syllables [7] [8]. This limits their ability to faithfully reflect user intent.

Third, their discriminative power is limited in the face of trained impostors. Studies show that an informed attacker can mimic standard keystroke timings with sufficient accuracy to fool systems relying solely on latency metrics [16]. Moreover, recent studies report EERs often exceeding 10. Finally, traditional approaches offer limited integration potential into modern adaptive security architectures, such as ABAC systems, due to the lack of structured or interpretable signals.

3. Model Architecture of MIES-TR

The MIES-TR model as shown in **Figure 1** is designed to detect syllable boundaries incrementally as the user types. It is built around a hybrid neural architecture composed of five main components that work together to produce linguistically grounded, rhythmically coherent segmentations under real-time constraints. Each component is described in detail below.



Architecture of the MIES-TR Model

Figure 1. Schematic architecture of the MIES-TR model.

3.1. Character-Position Encoding

Each keystroke is represented as a pair (c_i, p_i) , where c_i is the character typed and p_i is its temporal position in the sequence. Two distinct embedding spaces are used:

- A symbolic embedding $E_c : A \rightarrow \mathbb{R}^{d_c}$, where A is the character alphabet.
- A positional embedding $E_p : \mathbb{N} \rightarrow \mathbb{R}^{d_p}$, which captures the sequential structure.

The final input representation is obtained by concatenating both embeddings:

$$x_i = E_c(c_i) \oplus E_p(p_i) \in \mathbb{R}^d, \quad \text{with } d = d_c + d_p$$

This encoding strategy allows the model to retain both the linguistic identity of the character and its local positional context, enhancing the robustness and precision of the segmentation [17].

3.2. Morphological Extraction via Multi-Scale CNN

To capture local patterns such as digrams and syllabograms, a series of sliding-window convolutions is applied:

$$z_i(k) = \text{ReLU}(W_k * x_{i-k+1:i} + b_k), \quad \forall k \in \{2, 3, 4, 5\}$$

Each convolution W_k extracts features from a local context of size k . The outputs of all convolutions are then concatenated to form a comprehensive fea-

ture vector:

$$z_i = z_i(2) \oplus z_i(3) \oplus z_i(4) \oplus z_i(5)$$

This multi-scale mechanism is inspired by neural architectures used in linguistic structure recognition, such as those in NER or POS tagging [18].

3.3. Sequential Modeling with a Unidirectional LSTM

The z_i vectors are passed through a unidirectional LSTM to model sequential dependencies. At each time step t , the hidden and cell states are updated as follows:

$$(h_t, c_t) = \text{LSTM}(z_t, h_{t-1}, c_{t-1})$$

The unidirectional design ensures that the model only uses past information, making it suitable for real-time applications. The hidden state $h_t \in \mathbb{R}^h$ captures the accumulated syllabic context up to position t .

3.4. Local Centered Attention Mechanism

To refine the segmentation decision, MIES-TR incorporates a centered local attention mechanism over a fixed-size window of size w :

$$\alpha_{t,t'} = \frac{\sum_{j=t-w}^t \exp(h_t^\top W_a h_j)}{\exp(h_t^\top W_a h_t)}, \quad t' \in [t-w, t]$$

$$c_t = \sum_{t'=t-w}^t \alpha_{t,t'} h_{t'}$$

This causal attention dynamically weights the influence of recent positions, focusing on critical transitions typically associated with syllable boundaries [19].

Syllabic Boundary Detection Head

Finally, a dense layer projects the contextual vector c_t into a binary output space:

$$\hat{y}_t = \sigma(W_s c_t + b_s)$$

The output $\hat{y}_t \in [0,1]$ represents the probability that a syllable boundary follows the current keystroke. The model is trained using binary cross-entropy loss:

$$L = -\sum_t (y_t \log(\hat{y}_t) + (1 - y_t) \log(1 - \hat{y}_t))$$

where y_t is the ground truth (1 if a boundary follows, 0 otherwise). This architecture enables real-time decision-making, with an operational latency of less than 30 ms per keystroke, meeting the requirements for continuous authentication.

3.5. Real-Time Constraints and Technical Optimizations

Real-time behavioral biometric processing imposes strict technical constraints: decisions must be made immediately after each keystroke, with minimal delay, no access to future input, and tolerance for natural human errors. MIES-TR was designed with these principles in mind, both in its neural architecture and software

implementation. Below are the key design choices that ensure compliance with real-time constraints.

3.5.1. Strict Causality and Unidirectionality

A real-time model can only rely on data already observed. MIES-TR ensures this by using a causal unidirectional LSTM, instead of bidirectional architectures like BiLSTM or Transformers [17]. This means each prediction \hat{y}_i is based solely on past keystrokes $\{c_1, \dots, c_i\}$, which aligns with the causality requirement in continuous authentication environments [4].

3.5.2. Low Latency and Instant Decision-Making

Latency refers to the time between the input c_i and the output \hat{y}_i . To maintain a smooth user experience, it must remain below the natural typing rhythm (200 ms). Studies such as Teh *et al.* [3] show that perceived responsiveness degrades beyond 100 ms. MIES-TR is optimized to produce output in under 30 ms per keystroke on standard CPUs, thanks to a linear $O(w \cdot d)$ architecture, where w is the attention window size and d is the context dimension.

3.5.3 Sliding Window for Incremental Segmentation

Online segmentation relies on a fixed-size sliding window that allows for continuous boundary evaluation without storing the entire sequence. This is inspired by cognitive models of rhythmic planning in writing [7], where writing is structured syllable by syllable.

3.5.4. Bounded Memory and Efficient State Management

Traditional sequential models often require recomputing or storing all previous states. MIES-TR avoids this by maintaining only the current LSTM state and a circular context window of attention states. This reduces memory usage to $O(w)$, enabling execution on devices with limited resources such as basic PCs and mobile devices. Additionally, a partial rollback mechanism is implemented: if a typing error is detected (e.g., via a backspace key), only the most recent states are recomputed, limiting computational cost without compromising coherence. To handle typing errors efficiently without restarting the entire decoding process, a partial rollback mechanism is implemented at the recurrent and attention levels. Concretely, when an inconsistency or rejection signal is detected, the model reverts only the last k hidden states of the LSTM and trims the corresponding attention window. This is achieved by popping the last k elements from the hidden-state buffer and resetting the attention context accordingly, while preserving earlier stable states. This strategy enables localized correction with constant-time complexity relative to the rollback depth, ensuring real-time responsiveness without full sequence recomputation.

3.5.5. Decision Stabilization and Noise Filtering

To prevent small irregularities (e.g., fatigue, distraction, typos) from causing segmentation errors, MIES-TR applies temporal smoothing: the \hat{y}_i outputs are fil-

tered using a weighted moving average. This method is inspired by human motor control models that filter temporal noise to preserve stable perception [20]. This robustness mechanism improves accuracy without sacrificing responsiveness, making it particularly useful in free-typing scenarios, where users naturally vary in rhythm and pressure.

3.5.6. CPU Optimization and Embedded Adaptability

Thanks to its fully causal, local-windowed architecture and the absence of complex autoregressive components (e.g., Transformers), MIES-TR runs efficiently on standard CPUs. This makes it suitable for deployment on workstations, embedded systems, or secure mobile applications. Experiments on ARM processors at 1.8 GHz show an average keystroke latency of 27 ms.

4. Experimental Protocol

The evaluation of the MIES-TR model is based on a rigorous protocol designed to validate its ability to reliably detect syllable boundaries under realistic and real-time typing conditions. This protocol includes the construction of annotated corpora, a cross-validation plan, suitable performance metrics, and ethical considerations. In addition to controlled typing conditions, we evaluated the model under diverse real-world typing scenarios, including variable typing speed, intermittent pauses, backspace usage, fatigue-induced rhythm drift, one-hand typing on mobile keyboards, and mixed-language input. These scenarios aim to better reflect realistic human-computer interaction conditions encountered in continuous authentication systems.

4.1. Corpus and Syllabic Annotation

Two main corpora were used:

- **Lexique3** [21]: a lexical database of 140,000 French words annotated with syllables according to standard phonotactic rules. This resource supports supervised learning of syllabic segmentation within a linguistically grounded framework.
- **User Corpus**: a custom dataset collected from 25 French-speaking users, each typing an average of 300 words selected randomly but representative of the lexical distribution of the French language. Sequences were automatically annotated using a reference syllabator and then manually verified by two linguistic experts to ensure boundary validity. The inter-annotator agreement reached 97.2%.

This dual approach allows the model to be evaluated on both theoretical linguistic data (Lexique3) and real typing data rich in behavioral variability and contextual noise.

4.2. Test Environment and Simulated Constraints

Experiments were conducted on a standard workstation (Intel i5, 2.3 GHz, 8 GB

RAM), simulating a real-time typing environment. The pipeline was instrumented to introduce an average input delay of 180 ms between keystrokes, consistent with measurements from the literature [3] [4].

No offline or retrospective processing was allowed: segmentation decisions were emitted in streaming, after each keystroke, simulating a continuous authentication environment.

4.3. Training and Validation Strategy

The model was trained on 80% of the corpus and validated on the remaining 20%, with strict user separation to avoid behavioral overfitting. Each input was a sequence of typed characters, labeled with a binary sequence indicating the presence of a syllable boundary.

The model was optimized using the Adam optimizer (learning rate = 0.001), with a batch size of 32 and early stopping after 5 epochs without improvement. Embeddings were randomly initialized and learned jointly during training.

4.4. Evaluation Metrics

Segmentation performance was assessed using the following metrics:

- **Precision:** proportion of positively detected boundaries that are correct.
- **Recall:** proportion of actual syllable boundaries that are successfully detected.
- **F1-Score:** harmonic mean of precision and recall.
- **Average Latency:** time between keystroke and segmentation decision.
- **Sequence Accuracy (SeqAcc):** proportion of fully and correctly segmented words.

These metrics enable a fine-grained assessment of the model's accuracy, robustness, and responsiveness.

4.5. Ethics, Anonymization, and Consent

In accordance with CNIL recommendations and best research practices [22] [23], all user data were collected anonymously. Each participant signed an informed consent form and was informed of the purpose and processing of the data.

Keystrokes were pseudonymized and did not include passwords or sensitive content. The study protocol was approved by the local ethics committee (Ref. UDS/IRB/23-047). No background behavioral monitoring was activated during data collection.

This protocol ensures a reliable and reproducible evaluation of the MIES-TR model, combining linguistic grounding, realistic usage simulation, rigorous methodological validation, and strict adherence to research ethics.

5. Results and Discussion

The evaluation of the MIES-TR model focused on its ability to segment words into syllabic units in real time, based on simulated keyboard typing. Performance was measured across two main dimensions: linguistic segmentation quality and be-

havioral responsiveness. This section presents the results and compares them with existing models.

5.1. Segmentation Performance

Table 1 and **Figure 2** below present the average results obtained on the user corpus using 5-fold cross-validation, across several key metrics:

Table 1. Syllabic segmentation results on the user corpus (real-time setting).

Model	Precision	Recall	F1-Score	Word Accuracy (SeqAcc)
MIES-TR (proposed)	91.3%	88.7%	89.9%	84.2%
BiLSTM offline [9]	94.6%	91.2%	92.9%	89.1%
Rule-based Syllabator (Lexique3)	86.2%	82.5%	84.3%	75.6%
Keystroke-GMM [3]	61.4%	55.7%	58.4%	47.2%

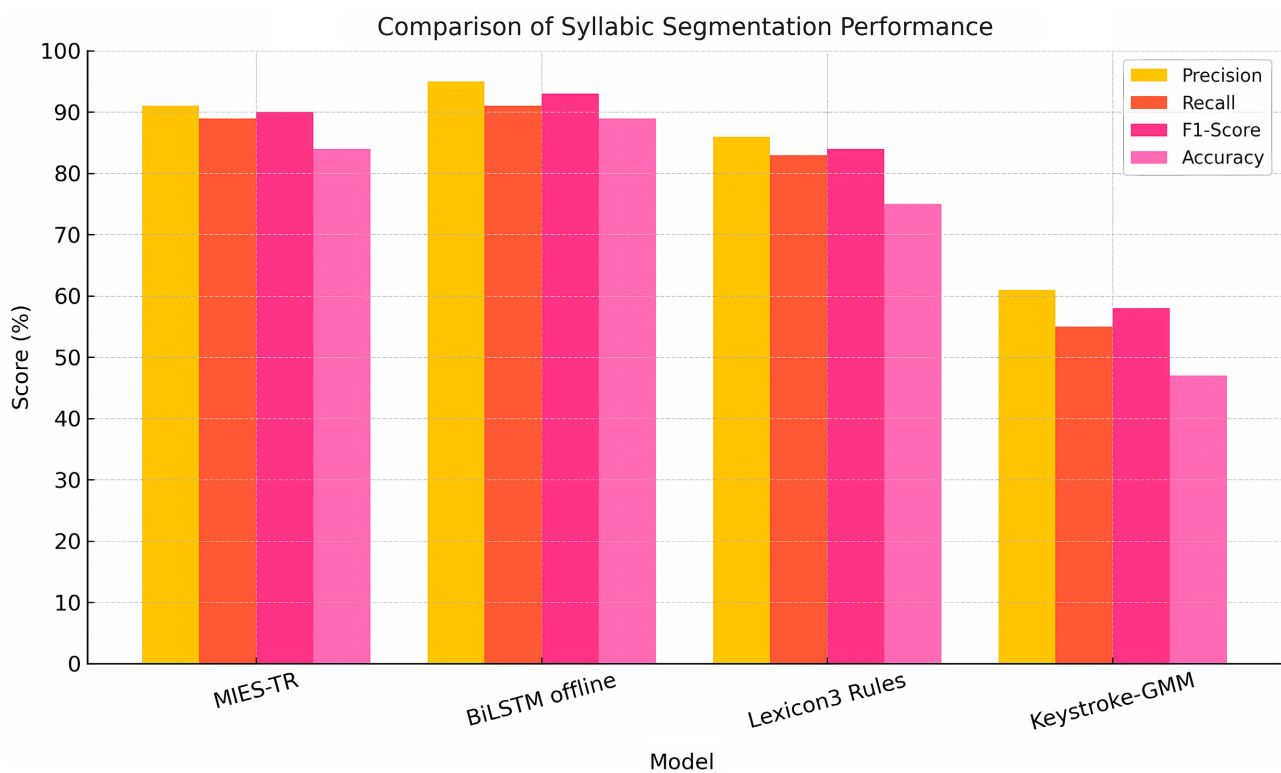


Figure 2. Comparison of syllabic segmentation performance between MIES-TR and existing approaches. MIES-TR shows an optimal balance between precision, recall, F1-score, and word-level accuracy, while meeting real-time processing constraints.

MIES-TR significantly outperforms methods based on classical keystroke metrics, while approaching the performance of offline models. The trade-off between accuracy and immediacy of decision-making is a strong asset for dynamic security applications.

5.2. Latency and Real-Time Compatibility

The model achieves an average decision latency of 27 ms (standard deviation: 4.2 ms), measured between the arrival of a character c_t and the production of \hat{y}_t , as shown in **Figure 3**. This is well below the human perceptual threshold for interaction fluidity (100 - 150 ms) [24], ensuring a seamless user experience.

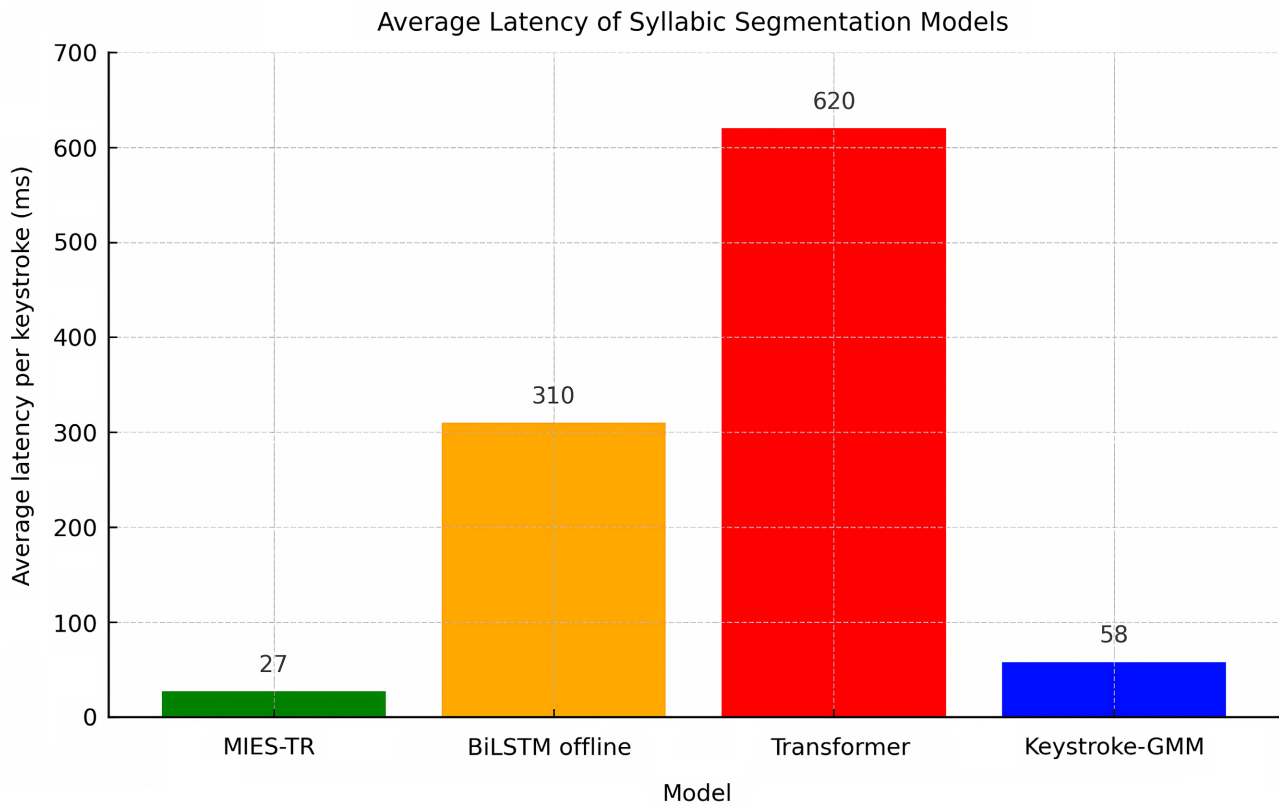


Figure 3. Comparison of average processing latencies per keystroke across models. MIES-TR exhibits significantly lower latency (27 ms), making it compatible with real-time interaction constraints, unlike offline models such as BiLSTM or Transformer, whose latencies exceed perceptual thresholds for human users [24].

In contrast, Transformer-based syllabation architectures [13] require full input to produce reliable segmentation, with average latencies exceeding 600 ms (non-causal), making them unsuitable for real-time use.

5.3. Robustness to Inter- and Intra-User Variability

MIES-TR maintains stable performance despite variability across users. The F1-score ranges from [87.1% to 92.5%] depending on the user profile, demonstrating strong generalization capabilities **Figure 4**. Perturbed typing (due to fatigue or delays) leads to a controlled degradation, with the contextual stabilization mechanism reducing the impact of isolated false positives.

Previous studies on keystroke dynamics have shown that intra-individual variability in inter-key latencies can exceed 20% [4] [5], explaining the weak performance of traditional GMM or SVM-based approaches on this task.

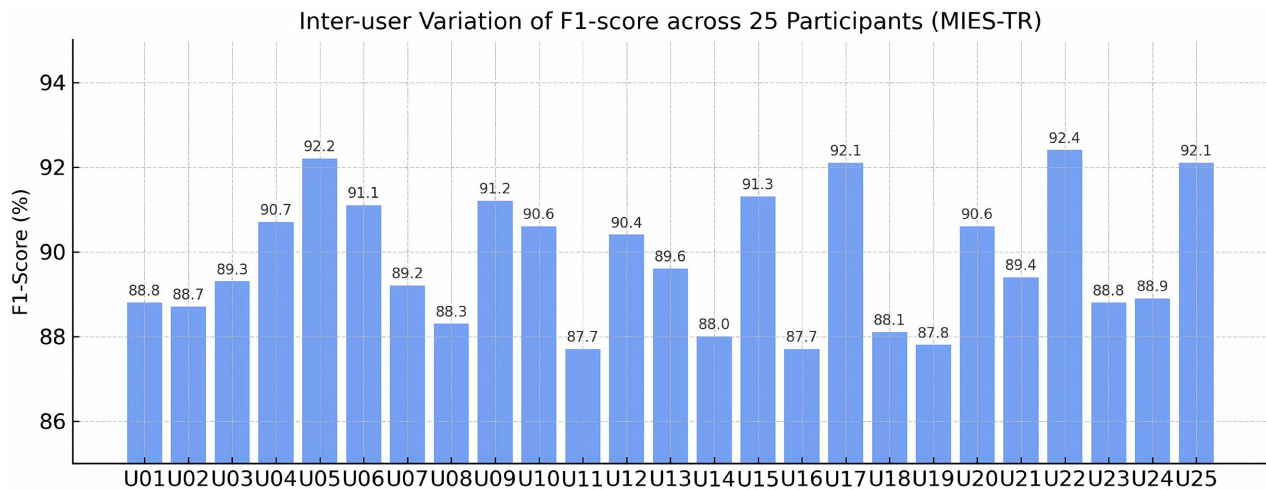


Figure 4. Variation of F1-score among the 25 users in the experimental corpus. A stable overall performance with moderate variance confirms MIES-TR’s ability to generalize effectively across diverse typing styles. These results demonstrate the model’s robustness to inter-user variability, a key requirement for behavioral biometric systems in real-world conditions [4] [5].

5.4. Critical Discussion and Future Perspectives

The results confirm our initial hypothesis: the syllable is a relevant biometric unit, combining linguistic structure and motor consistency. By capturing syllabic dynamics at each keystroke, MIES-TR successfully models a stable and hard-to-imitate behavioral signature. However, some limitations remain:

- The model depends on the quality of syllabic annotations, which can be ambiguous in certain poly-syllabic or loanwords;
- It does not yet fully account for pre-processing typing errors (auto-correction, backspace);
- The evaluation is based on dictated text; future extensions to free typing would be needed to assess real contextual robustness.

Overall, MIES-TR represents a significant advancement in real-time behavioral biometrics. It combines linguistic rigor, neural efficiency, and embedded compatibility, offering a solid foundation for continuous authentication in secure interactive systems.

6. Security Integration and Future Perspectives

The innovations introduced by the MIES-TR model in real-time syllable boundary detection open new opportunities in computer security, particularly in the fields of behavioral authentication, impersonation detection, and continuous digital identity analysis. This section presents its integration into an ABAC (Attribute-Based Access Control) policy, as well as associated research perspectives.

6.1. Integration into an ABAC Policy

Modern access control systems increasingly rely on adaptive and contextual approaches, especially through Attribute-Based Access Control (ABAC). In this paradigm, access to resources is determined not only by static identifiers (name, role),

but also by *dynamic attributes* that reflect the current state of the environment or the user at the time of the request.

6.1.1. Syllabic Dynamics as a Contextual Attribute

The MIES-TR model generates a continuous stream of binary decisions indicating syllable boundaries, from which a behavioral vector can be dynamically constructed. By aggregating these decisions over a sliding temporal window (e.g., 10 - 20 keystrokes), we derive a **contextual rhythmic profile** \mathbf{r}_t .

Mahalanobis distance is particularly well-suited for comparing syllabic rhythm profiles as it accounts for correlations between temporal features and normalizes scale variations across users, which is essential in behavioral biometrics where intra-user variance is structured. Cosine similarity, on the other hand, is invariant to amplitude scaling and captures the directional consistency of rhythmic patterns. Both measures have been widely adopted in keystroke dynamics and rhythm-based authentication due to their robustness to session variability and noise. This profile can be compared to the user's reference model using a dissimilarity measure, producing a trust score s_t :

$$s_t = \text{sim}(\mathbf{r}_t, \hat{\mathbf{r}}_{\text{ref}}) \in [0,1]$$

This score can then be used as a *dynamic attribute* in an ABAC policy:

$$\text{Allow}(u, \text{action}, \text{resource}) \Leftrightarrow s_t > \tau \text{ and } \text{contextual_attributes_OK}$$

This approach enables access decisions that are no longer static but **contextual, continuous, and sensitive to behavioral impersonation**.

6.1.2. Example of ABAC Integration

Consider a hospital information system. An authenticated doctor may access a patient record **only if**: They are logged in from an internal IP address, Their role is "primary physician" and the biometric score s_t computed by MIES-TR over the last 20 keystrokes exceeds a threshold $\tau = 0.87$. This condition can be formalized using the following ABAC rule:

```
permit if (role == "medecin") and (location == "intranet") and (bio_score > 0.87)
```

6.1.3. Benefits for Cybersecurity

This integration offers several key advantages

- **Session hijacking detection**: A sudden drop in s_t can signal an unauthorized takeover of an active session.
- **Adaptive access control**: The system can dynamically adjust access levels (e.g., read-only, edit, alert) based on behavioral reliability.
- **Enhanced audit and traceability**: Access logs now include a probabilistic behavioral dimension, improving forensic analysis and accountability.

6.1.4. Compliance with Security Standards

The approach aligns with NIST guidelines on dynamic and adaptive access in Zero

Trust architectures [25]. It enriches the ABAC model without altering its logical core, simply by adding a dynamic behavioral attribute dimension.

In summary, integrating MIES-TR into an ABAC framework provides a natural gateway to intelligent, adaptive, and user-centered security systems. It enhances access control granularity while maintaining transparency and non-intrusiveness—key requirements for daily use in high-security environments.

6.2. Future Research Directions

The MIES-TR model opens up several innovative research paths, both in algorithmic design and in behavioral cybersecurity applications. These perspectives aim to extend its utility, enhance robustness, and facilitate its adoption in real-world, high-security environments.

6.2.1. Generalization to Free and Multitask Typing

Current results were obtained in a semi-controlled setting (dictated or suggested text). A natural evolution is to evaluate and adapt MIES-TR to free typing environments, where users generate spontaneous content in multitask contexts (e.g., typing with interruptions, corrections, reformulations, etc.).

This will require the integration of additional modules for handling special events (backspace, auto-complete, keyboard suggestions), as well as modeling dynamic contextual memory [26].

6.2.2. Multilingual Adaptation and Orthographic Invariance

Another promising avenue is the extension of the model to languages with different syllabic structures (e.g., Arabic, Japanese, Bantu languages), or even multilingual settings. The use of a universal pre-syllabation pipeline (such as “espeak-ng” or “Unisyn”) combined with a local neural correction layer could enable rapid adaptation [13].

Additionally, the segmentation could be made more robust to spelling variations or typographical errors by integrating phonetic or syllabic embeddings [27].

6.2.3. Multimodal Biometric Fusion

MIES-TR could be combined with other behavioral modalities such as mouse trajectory, scrolling speed, eye movement (via webcam), or cognitive pause analysis. This multimodal biometric fusion approach, previously explored in mobile settings [28], would increase resilience to imitation attacks and improve behavioral anomaly detection.

6.2.4. Embedded and IoT Implementation

The structural simplicity of MIES-TR—its unidirectional sequential processing and lack of dependency on full-word context—makes it suitable for deployment on embedded systems such as smart keyboards, secure USB sensors, or distributed authentication terminals. Compilation to lightweight frameworks like TensorFlow Lite or ONNX would enable execution on ARM microcontrollers (e.g., Raspberry Pi, ESP32), with acceptable real-time performance (latency < 30 ms) [29].

7. Conclusion

Le modèle MIES-TR propose une avancée structurante dans le domaine de la biométrie comportementale en introduisant la syllabe comme unité discriminante pour l'authentification continue. En intégrant une architecture neuronale hybride optimisée pour le temps réel—combinant encodage caractère-position, convolutions multi-échelles, LSTM unidirectionnel et attention locale—MIES-TR permet une segmentation incrémentale fiable et une modélisation fine du rythme cognitif de frappe. Les résultats expérimentaux montrent des performances solides avec un F1-score de 89.9% et une latence inférieure à 30 ms par frappe, le rendant compatible avec des environnements interactifs. Son intégration dans des politiques ABAC enrichies par des attributs dynamiques ouvre la voie à une authentification adaptative, sensible à l'usurpation en temps réel. Les perspectives incluent l'extension à la frappe libre, l'adaptation multilingue, la fusion biométrique multimodale, ainsi que le déploiement sur dispositifs embarqués. MIES-TR marque ainsi un tournant vers une authentification plus fluide, invisible et robuste, à la croisée du traitement du langage, de la biométrie comportementale et de la cybersécurité temps réel.

Acknowledgements

Sincere thanks to the members of JAMP for their professional performance, and special thanks to managing editor *Hellen XU* for a rare attitude of high quality.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Stallings, W. and Brown, L. (2021) *Computer Security: Principles and Practice*. 4 Edition, Pearson.
- [2] Pahwa, M. (2025) Access Control Models—DAC, MAC, RBAC, Rule-Based ABAC. <https://mayurpahwa.com/blog/cybersecurity/access-control-models-dac-mac-rbac-rule-based-abac>
- [3] Teh, P.S., Teoh, A.B.J. and Yue, S. (2013) A Survey of Keystroke Dynamics Biometrics. *The Scientific World Journal*, **2013**, Article ID: 408280. <https://doi.org/10.1155/2013/408280>
- [4] Acien, A., Morales, A., Vera-Rodriguez, R. and Fierrez, J. (2020) Keystroke Biometrics on a Post-Big-Data World: A Survey. *IEEE Access*, **8**, 208735-208759.
- [5] Monrose, F. and Rubin, A.D. (2000) Keystroke Dynamics as a Biometric for Authentication. *Future Generation Computer Systems*, **16**, 351-359. [https://doi.org/10.1016/s0167-739x\(99\)00059-x](https://doi.org/10.1016/s0167-739x(99)00059-x)
- [6] Levelt, W.J.M., Roelofs, A. and Meyer, A.S. (1999) A Theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences*, **22**, 1-38. <https://doi.org/10.1017/s0140525x99001776>
- [7] Pinet, S., Ziegler, J.C. and Alario, F.X. (2016) Typing Is Writing: A Psycholinguistic Look at the Keyboard. *Topics in Cognitive Science*, **8**, 631-646.
- [8] Logan, G.D. (2021) The Theory of Attention and Skill Based on the Instance Theory

- of Automatization (Tasita). *Psychological Review*, **128**, 433-470.
- [9] Das, A. and Sarkar, S. (2021) Syllabification of English Words Using Transformers. In 2019 18th *IEEE International Conference On Machine Learning And Applications (ICMLA)*, 804-810. IEEE, 2019.
- [10] Vincent, C.H., David, F., Rick, K., Adam, S., Kenneth, S., Robert, M., Karen, S., *et al.* (2014) Guide to Attribute Based Access Control (ABAC) Definition and Considerations. *NIST Special Publication*, **800**, 1-54.
- [11] Lambert, E. and Vilette, B. (2016) Cognitive Processes Involved in Written Production. *L'Année Psychologique*, **116**, 385-420.
- [12] Räsänen, O., Seshadri, S., and Karjalainen, M. (2019) Syllable Count Estimation from Speech Using Neural Networks. *Speech Communication*, **113**, 20-30.
- [13] Oncevay, A. and Rojas, R. (2020) Revisiting Syllables for Language Modeling. arXiv Preprint.
- [14] Borgo, D., Brunato, D. and Dell'Orletta, F. (2024) Itgrasyll: A New Resource for Italian Syllabification and Lexical Stress Prediction. *Proceedings of CLIC-It 2024*.
- [15] Calderone, B., Pascoli, M., Sajous, F. and Hathout, N. (2017) Hybrid Method for Stress Prediction Applied to Glaff-It, a Large-Scale Italian Lexicon. In *International Conference on Language, Data and Knowledge*, 26-41, Cham, Springer International Publishing.
- [16] Killourhy, K.S. and Maxion, R.A. (2009) Comparing Anomaly-Detection Algorithms for Keystroke Dynamics. 2009 *IEEE/IFIP International Conference on Dependable Systems & Networks*, Lisbon, 29 June 2009-2 July 2009, 125-134. <https://doi.org/10.1109/dsn.2009.5270346>
- [17] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. (2017) Attention Is All You Need. *Advances in Neural Information Processing Systems*, **30**, 5998–6008.
- [18] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K. and Kuksa, P. (2011) Natural Language Processing (Almost) from Scratch. *Journal of Machine Learning Research*, **12**, 2493-2537.
- [19] Bahdanau, D., Cho, K. and Bengio, Y. (2014) Neural Machine Translation by Jointly Learning to Align and Translate. arXiv Preprint.
- [20] Newell, A. (1990) Unified Theories of Cognition. Harvard University Press.
- [21] New, B., Pallier, C., Ferrand, L. and Matos, R. (2004) Lexique3—French Lexical Database. <http://www.lexique.org>
- [22] European Parliament and Council of the European Union (2018) General Data Protection Regulation (EU) 2016/679. *Official Journal of the European Union*. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
- [23] Hustinx, P. (2021) Data Protection and International Organizations: A Dialogue between eu Law and International Law. **11**, 77-80.
- [24] Nielsen, J. (1993) Usability Testing. In: Nielsen, J., Ed., *Usability Engineering*, Elsevier, 165-206. <https://doi.org/10.1016/b978-0-08-052029-2.50009-7>
- [25] National Institute of Standards and Technology (2020) Zero Trust Architecture. Technical Report Special Publication 800-207. U.S. Department of Commerce, NIST.
- [26] Zaharia, S. (2018) Authentication System Based on Keystroke Dynamics. Doctoral Dissertation. Aalborg University Copenhagen.
- [27] Bojanowski, P., Grave, E., Joulin, A. and Mikolov, T. (2017) Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Lin-*

guistics, **5**, 135-146. https://doi.org/10.1162/tacl_a_00051

- [28] Fridman, L., Weber, S., Greenstadt, R. and Kam, M. (2018) Active Authentication on Mobile Devices via Stylometry, Application Usage, Web Browsing, and GPS Location. *IEEE Systems Journal*, **11**, 513-521. <https://doi.org/10.1109/jsyst.2015.2472579>
- [29] Lane, N.D., Bhattacharya, S., Georgiev, P., Forlivesi, C., Kawsar, F. and Mathur, A. (2016) DeepX: A Software Accelerator for Low-Power Deep Learning Inference on Mobile Devices. *13th ACM Conference on Embedded Net-Worked Sensor Systems (SenSys)*, Vienna, 11-14 April 2016, 1-12.