Human Computer Interaction : Report Project 2

Andrea LETIZIA & Bastien ROURE

January 2026

Contents

1 Introduction

The AZERTY keyboard, like its American counterpart QWERTY, was designed for use with mechanical typewriters. This system separated frequently used letter pairs to minimize mechanical jams caused by the collision of hammers when typing quickly. This keyboard was therefore not designed for efficient text entry, but rather to adapt to the constraints of the machine. However, despite the transition to the digital, our computer and phone keyboards have retained this layout. This persistence is attributed to socio-technical factors such as user familiarity, standardization of the device, and the high cost of switching to another device.

Some keyboards, such as Dvorak and Bépo, were designed to be optimized for physical computer keyboards. However, text input practices have evolved significantly in recent years with the widespread use of smartphones. These devices have become the primary means of textual interaction, meaning that their digital keyboards now play an important role in everyday text input tasks. This change raises a question about the applicability of keyboard optimization principles initially developed for physical keyboards to the virtual keyboards of touchscreen devices.

Therefore, while the optimized layout of Bépo aims to improve typing efficiency on a physical keyboard, it is not certain that this translates to the context of smartphones. The lack of literature motivated this work, which examines the evolution of typing speed on the Bépo keyboard in contemporary scenarios of text entry on smartphone keypads for French-speaking users.

2 Related Work

2.1 Keyboard optimization

To counter the inefficiency of the QWERTY keyboard, Dvorak proposed a theoretically optimal keyboard in 1940 based on his studies presented a few years earlier [Dvorak et al., 1936]. In this work, he formalized several criteria, the first of which was the minimization of finger movement. He placed the most frequently used letters on the center row of the keyboard, which he considered to be the resting position for the fingers, thus minimizing vertical movement. Another criterion was to maximize the alternation of both hands, to allow for greater fluidity and avoid downtime caused by using only one hand. And finally, distributing the workload between the fingers, in other words, the strong fingers will be used more than the weak fingers. The Bépo keyboard `https://bepo.fr/wiki/Présentation` was subsequently developed for French-speaking users following the same principles, taking into account the structure of the French language.

2.2 Swiching cost and metrics

The transition from a familiar key layout to a new one creates what is known as a **switching cost**, which results in a decrease in typing speed and an increase in error rates. This phenomenon is linked to what is known as the **negative transfer** phenomenon.

According to [MacKenzie and Zhang, 1999], one way to evaluate typing performance on a keyboard is by word typing speed and accuracy. This paper introduces metrics such as **Words Per Minute (WPM)** and various error rates, which highlight the importance of observing the evolution of these metrics over time in order to capture learning effects. This study shows that initial performance on a keyboard is not necessarily representative of long-term use, especially when the keyboard is unfamiliar.

These observations are reinforced by [Kershaw, 2006], which shows that prior keyboard learning can greatly interfere with learning a new keyboard, leading to a temporary decline in performance. This negative transfer is even more pronounced when the two layouts have partial similarities, leading to errors due to automatic responses. These results therefore suggest that experienced AZERTY keyboard users may encounter significant difficulties when switching to another key layout such as Bépo.

To capture these effects, researchers looked at the learning curve, which allows them to analyze performance levels on tests spread out over different sessions spaced out over time. This curve is a key indicator of the cost of change, showing whether performance exceeded the initial level over the study period.

2.3 Text input on smartphones

The article [Li et al., 2019] compares telephone keypads with complex sentences. To do this, in addition to speed and error rate, it measures galvanic skin response, heart rate, and perceived ease of use, which may be other relevant metrics. Sentence complexity is an indication of the number and type of characters used in a sentence. In our case, and particularly for the French language, we can think of words with accents that are particularly common, as well as other special characters such as “ç” or “œ”. With special characters encoded in Bépo, we can imagine that, theoretically, it may be possible to maximize overall performance on these types of texts.

3 Presentation of the solutions

3.1 Selected metrics and experimental support

In order to assess the impact of the Bépo layout on text entry on smartphones, we selected two quantitative metrics, which constitute our dependent variables: Typing speed (WPM – Words Per Minute) and Accuracy: the proportion of words entered correctly, calculated by dividing the number of words with errors by the total number of words entered.

In addition to these measures, we also collected qualitative data to assess user feedback. We asked users to rate typing comfort (score from 0 to 5).
A free comment field also allowed participants to express specific remarks or describe specific situations encountered during the test sessions.

The typing tests were conducted using the website `https://10fastfingers.com/typing-test/french`, via its French typing test. This platform generates random word sequences and automatically provides WPM and accuracy metrics, ensuring consistent measurement for all participants.

3.2 Participant profiles

The study was conducted with five participants, both men and women, belonging to a relatively homogeneous age group (20–30 years old). All participants were regular users of the AZERTY keyboard and reported having no prior knowledge or experience with the Bépo layout.

We asked each participant to estimate how often they write on their phone each day. The more time they spend writing, the more training they will have between tests, which will affect the results. In our panel, daily usage ranges from 15 minutes to over an hour.

The words are written on each person’s personal phone, so that everyone already has their habits and it’s quite simple to change the keyboards in these devices.

3.3 Experimental protocol

The experiment took place over a period of three weeks.
During the first session, we helped participants install and activate the Bépo keyboard on their smartphones and explained the rules and objectives of the study. Before using the Bépo keyboard,

a baseline measurement was taken using each participant's usual AZERTY keyboard. This phase provided an initial point of comparison for typing performance. It also allowed us to present the website.

Participants then committed to using the Bépo keyboard exclusively on a daily basis for all text input on their smartphones for the duration of the study.
The performance tests consisted of copying a sequence of random words generated by the platform for one minute. The instructions were to type the text as quickly as possible with as few errors as possible.
Each participant had to take a test every two days and send a screenshot of the results (WPM, accuracy, and feedback) after each session. Invalid tests (interruptions, distractions, technical bugs, etc.) were not to be taken into account.
In addition, they also had the opportunity to provide additional comments at any time, even outside of the days dedicated to the questionnaire.

4 Results

The results of the three weeks of experiments are shown in Figure 1, where we can observe the evolution of each participant's typing speed on the Bépo keyboard during each test session. For each person, the reference value obtained on AZERTY is shown to indicate whether this threshold has been exceeded. The general trend is an overall increase in typing speed throughout the experiment. However, we see that the results differ between participants: three out of five people exceeded their reference value at the end of the experiment. Among these three, we have an increase ranging from +13% for the largest increase to +5% for the smallest. We note that these three participants are those with the lowest reference time and who also reported spending the least amount of time on the phone. The largest increase was in the profile reporting the lowest estimated screen time, and the smallest increase was in the profile that was the heaviest user of the three. Consequently, participants who reported spending more time writing on their phones each day had a higher average writing score, but they were also the ones who failed to exceed their initial score during the experiment. This brings us back to the concept of negative transfer. When a frequent user has too much prior knowledge, it becomes more difficult for the brain to create new habits and reach a level similar to the previous one.

The accuracy of the participants can be seen in Figure 2a, where we can see different profiles. On average, the errors observed are slightly higher (approximately 1 percentage point) for profiles that spend the most time on their phones. Participants with the lowest average typing speed in AZERTY, who also reported spending less time typing on their phones, tend to be more accurate in their typing. It should also be noted that this accuracy varies between each test. We can see that in the first test, everyone is 100% accurate, which is reflected in the very low scores observed Figure 1 on the first test. There are also behaviors that affected several participants, notably on day 6, when $P1$ and $P4$ set their records with comfortable feedback, and the next day, according to their comments, the text was more complex, which led to a net increase in errors on test 7. Based on these results, in Figure 2b, we can see that the people who feel most comfortable with the keyboard are also the ones who make the most errors in the end. In addition, we observed that as the study progressed over time, our participants felt more comfortable with the exercise, in a fairly linear fashion, with a few understandable exceptions that could be interpreted thanks to the comments left by some participants. A classic pattern was observed during a test deemed more difficult: feelings were obviously affected, and candidates who felt fairly comfortable with the previous tests lowered their confidence levels. This pattern is particularly noticeable with candidates 4 during the 7^{th} test.

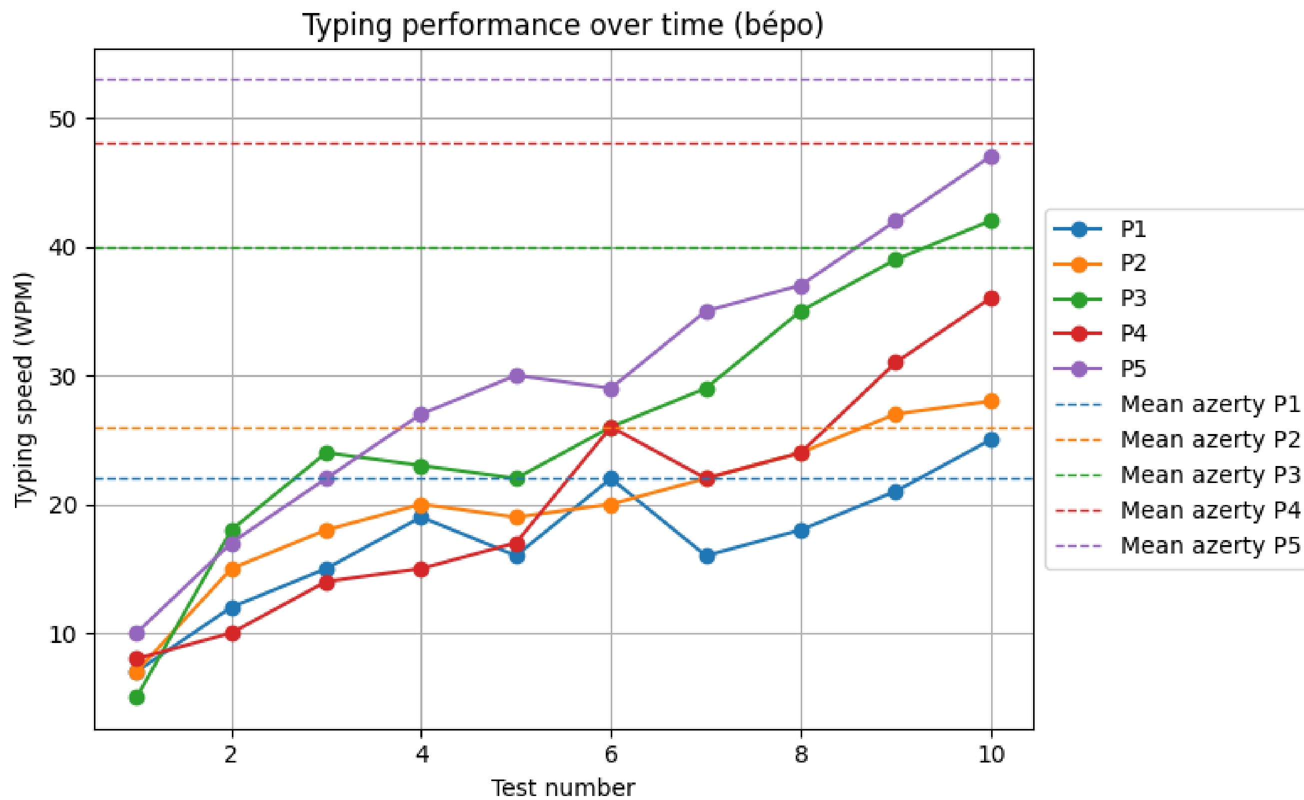
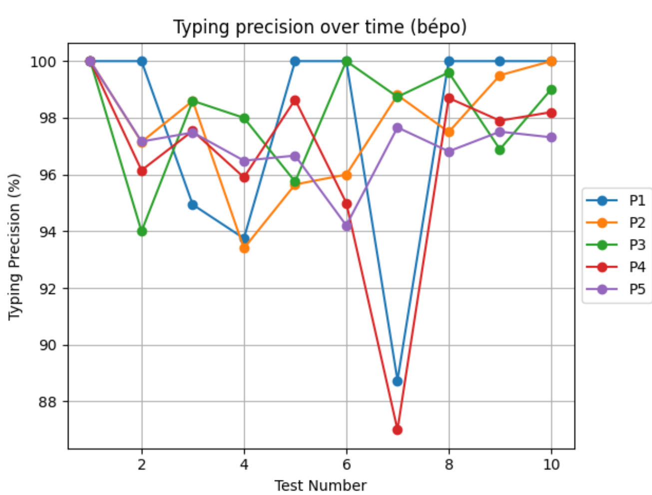


Figure 1: Results of 10 test sessions for 5 participants, with performance (WPM) on the y-axis and the dotted line representing initial typing speed on a azerty keyboard.



(a) Typing precision over time

(b) Confidence score over time

Figure 2: Performance and confidence evolution across the tests

5 General discussion and conclusion

This study focuses on users' adaptation to an optimized keyboard, Bépo, in the context of text entry on a smartphone, a medium for which this type of keyboard has been little studied. Through a three-week longitudinal experiment, we analyzed changes in typing speed, accuracy, and subjective perception among five French-speaking users with varying amounts of daily typing experience, all of whom were new to Bépo and accustomed to the AZERTY keyboard.

The results show that gradual learning of the Bépo keyboard layout can be observed on smartphones, revealing a general trend of improvement in typing speed over the course of the sessions. However, this improvement remains highly dependent on user profiles. Participants who use smartphones less frequently for typing seem to benefit more from the change in layout, even exceeding their initial performance on the AZERTY keyboard. Conversely, users who use smartphones frequently experience a higher transition cost and struggle to return to their baseline level, illustrating the negative transfer phenomenon described in the literature.

The accuracy analysis highlights significant variability between sessions and profiles. Users who are subjectively most comfortable with the Bépo keyboard are not necessarily the most accurate, suggesting a trade-off between speed, perceived comfort, and motor control. These results confirm that typing performance cannot be assessed using a single indicator and that an approach combining quantitative and qualitative metrics is essential, but also that the difficulty of the text has a significant impact on participant's performance.

In conclusion, this study highlights that Bépo has potential for improvement in the medium term. Its adoption depends heavily on pre-existing habits and the context of use. These observations open up avenues for future research, particularly on the influence of ergonomic factors specific to smartphones (posture, how the phone is held, screen size) and on adapting optimized keyboard layouts to new mobile device formats. Results could also be more relevant with longer experience times and texts written in advance with constant or progressive difficulty.

References

[Dvorak et al., 1936] Dvorak, A., Merrick, N. L., Dealey, W. L., and Ford, G. C. (1936). *Typewriting Behavior: Psychology Applied to Teaching and Learning Typewriting*. American Book Company, New York.

[Kershaw, 2006] Kershaw, T. C. (2006). Negative transfer in the learning of typing tasks. In *Proceedings of the Annual Meeting of the Cognitive Science Society*. Paper 2526.

[Li et al., 2019] Li, Y., You, F., You, X., and Ji, M. (2019). Smartphone text input: Effects of experience and phrase complexity on user performance, physiological reaction, and perceived usability. *Applied Ergonomics*, 80:200–208.

[MacKenzie and Zhang, 1999] MacKenzie, I. S. and Zhang, S. X. (1999). The design and evaluation of a high-performance soft keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*, pages 25–31. ACM.