

Độ chuẩn xác của công cụ sửa lỗi Merci App trong dạy viết tiếng Pháp

Đỗ Thị Bích Thủy

Email: thuydtb1976@vnu.edu.vn
Trường Đại học Ngoại ngữ - Đại học Quốc gia Hà Nội
Đường Phạm Văn Đồng, quận Cầu Giấy,
Hà Nội, Việt Nam

TÓM TẮT: Nghiên cứu này có mục đích đo độ chuẩn xác của công cụ sửa lỗi tự động Merci App, phiên bản miễn phí dùng trong dạy viết tiếng Pháp. Dữ liệu nghiên cứu là 30 bài viết trình độ B2 của 30 sinh viên năm hai ngành Ngôn ngữ Pháp. Kết quả cho thấy, số lỗi Merci App nhận chuẩn xác là 82,7% tính trên tổng số lỗi ứng dụng phát hiện ra và giảm xuống 45,3% tính trên tổng số lỗi cấp độ thấp xuất hiện trong bài. Số lỗi Merci App nhận đúng, phân tích và đề xuất sửa đúng là 63,4% tính trên tổng số lỗi ứng dụng phát hiện ra, tỉ lệ này giảm xuống còn 34,7% tính trên tổng số lỗi cấp độ thấp trong bài. Ứng dụng thường nhận sai các lỗi về hợp giống số trong một cụm danh từ; không nhận được nhiều lỗi diễn đạt, từ vựng, hợp giống số, giới từ; phân tích và đề xuất sửa sai các lỗi từ vựng, động từ và hợp giống số. Chúng tôi đề xuất cần hướng dẫn người học sử dụng công cụ sửa lỗi tự động, cung cấp từ vựng siêu ngôn ngữ để các em hiểu được phân tích lỗi và sửa lỗi trong Merci App nhưng giảng viên vẫn cần tập trung sửa lỗi ở cấp độ cao và các lỗi cấp độ thấp mà ứng dụng không nhận được, nhận sai, phân tích sai, sửa sai.

TỪ KHÓA: Sửa lỗi tự động, độ chuẩn xác, dạy viết tiếng nước ngoài, Merci App, tiếng Pháp.

→ Nhận bài 18/4/2024 → Nhận bài đã chỉnh sửa 19/5/2024 → Duyệt đăng 15/6/2024.

DOI: <https://doi.org/10.15625/2615-8957/12410604>

1. Đặt vấn đề

Ứng dụng công nghệ thông tin trong giảng dạy nói chung và dạy viết nói riêng đang là một xu hướng tất yếu trong thời đại số, ở cả bậc học Phổ thông và Đại học. Những năm gần đây, nhiều công cụ chữa lỗi tự động đã được sử dụng rộng rãi trong nhiều lớp học ngoại ngữ như Criterion, Grammarly, Pigai để chữa tiếng Anh; Merci App, Bon Patron, Antidote để chữa tiếng Pháp... Các công cụ này phát hiện lỗi sai, phân loại lỗi, giải thích vì sao sai và thường đưa ra các phương án chữa lỗi để người dùng lựa chọn. Rất nhiều nghiên cứu mới đây nhất khẳng định các công cụ chữa lỗi này có hiệu quả trong việc giúp người học cải thiện độ chính xác khi viết bài bằng tiếng nước ngoài [1], [2], [3].

Hoang (2022) [3] đã tiến hành đo độ chuẩn xác của Criterion sử dụng để chữa lỗi bài viết cho 38 sinh viên năm thứ hai học tiếng Anh bậc Đại học. Tác giả đã mã hoá các lỗi theo Lavolette et al. (2015) [4], chia thành 3 loại lỗi, cụ thể như sau:

- *CC (correct code)*: Là những lỗi được Criterion phát hiện đúng, phân loại lỗi, phân tích lỗi đúng và đề xuất sửa đúng.

- *IC (incorrect code)*: Là những lỗi được phát hiện đúng nhưng phân loại, phân tích sai hay đề xuất sửa lỗi sai.

- *FP (false positive)*: Không phải là lỗi mà Criterion báo lỗi.

Hoang (2022) [3] đã đo độ chuẩn xác bằng cách lấy tổng số correct code chia cho tổng correct code + incorrect code + false positive. Kết quả nghiên cứu cho thấy, 82,2% số lỗi Criterion tìm được là correct code, 6,1% là incorrect code và 11,7% là false positive. Máy thường sai khi báo lỗi dấu phẩy, lỗi câu không hoàn chỉnh, lỗi chính tả với tên riêng.

Lavolette et al. (2015) [4] đã tiến hành nghiên cứu đo độ chuẩn xác của Criterion khi chữa bài viết cho sinh viên nước ngoài học tiếng Anh tại Mỹ. Kết quả cho thấy, 75% số lỗi được phát hiện là chuẩn xác, 14% là phát hiện lỗi đúng nhưng phân tích phân loại lỗi và đưa ra đề xuất sửa sai, 11% là phát hiện lỗi không chuẩn xác, nghĩa là không phải là lỗi. Lavolette et al. (2015) [4] đã tính độ chuẩn xác theo các loại lỗi và nhận thấy Criterion chữa tốt về lỗi viết hoa, lỗi dấu phẩy, từ sai, chia động từ sai (các lỗi này có độ chuẩn xác lên tới 85%). Máy chữa không tốt các lỗi về run-on sentence (được hiểu là câu mà trong đó hai hay nhiều mệnh đề độc lập được viết đằng sau một mệnh đề độc lập khác mà không sử dụng dấu chấm câu hay thành phần kết nối phù hợp), mạo từ, lỗi chính tả với độ chuẩn xác chưa được 50%. Tiếp theo, Lavolette et al. (2015) [4] đã chia các bài luận của sinh viên thành các đơn vị văn bản (T-unit) để xem

Criterion không phát hiện được những lỗi sai nào. Theo [5], một đơn vị văn bản bao gồm một mệnh đề chính và các mệnh đề phụ. Tổng cộng, Lavolette et al. (2015) [4] đã nghiên cứu trên dữ liệu 266 đơn vị văn bản, trong đó có 206 T-unit có chứa ít nhất một lỗi sai. Tuy nhiên, Criterion chỉ phát hiện được 111 đơn vị văn bản có lỗi sai, bỏ sót 46% số T-unit chứa lỗi sai. Theo Lavolette et al. (2015) [4], Criterion hay bỏ sót lỗi về từ vựng, đặc biệt là lỗi dùng sai từ loại, ví dụ khi sinh viên dùng từ character thay vì phải dùng từ, characteristics. Máy cũng không phát hiện ra lỗi giới từ trong 7 đơn vị văn bản, lỗi mạo từ trong 9 đơn vị văn bản.

Grégoire (2021) [6] đã sử dụng công cụ chữa lỗi tiếng Pháp tự động Antidote cho 304 học sinh cấp hai tại Quebec, Ca-na-đa. Các em học sinh đã được chia thành bốn nhóm: Nhóm A viết bài trên giấy, nhóm B dùng chương trình soạn thảo văn bản Word nhưng không dùng tính năng chữa lỗi chính tả ngữ pháp, nhóm C dùng Antidote nhưng không được hướng dẫn cách sử dụng, nhóm D dùng Antidote sau khi được hướng dẫn cách dùng. Kết quả cho thấy, hai nhóm C và D dùng công cụ chữa lỗi tự động đã giảm đáng kể lỗi chính tả nhưng lỗi ngữ pháp và từ vựng thì không giảm, thậm chí lỗi từ vựng còn nhiều hơn so với hai nhóm còn lại. Nhóm B dùng Word là nhóm có kết quả bài kiểm tra đầu ra thấp nhất trong số bốn nhóm và thấp hơn kết quả đầu vào. Tuy nhiên, Grégoire (2021) [6] đã không đo độ chuẩn xác của công cụ sửa lỗi tự động Antidote.

Ứng dụng chữa lỗi tự động cho tiếng Pháp Merci App ra đời vào năm 2020, được coi là anh em sinh đôi của Grammarly. Do rất mới nên Merci App ít được sử dụng để dạy viết trong lớp học ngoại ngữ và chúng tôi chưa tìm thấy nghiên cứu đo độ chuẩn xác của Merci App khi sử dụng trong lớp học tiếng Pháp như tiếng nước ngoài. Đỗ Thị Bích Thủy [7] đã so sánh đánh giá của 40 sinh viên khi sử dụng Merci App và Google Doc để chữa bài viết, tuy nhiên nghiên cứu tập trung đo ở góc độ định lượng ví dụ so sánh tổng lỗi ngữ pháp, chính tả, từ vựng mà mỗi công cụ tìm được trên cùng một bài viết hay so sánh số lỗi được sinh viên sửa theo đề xuất của máy... chứ chưa đi sâu phân tích chất lượng của nhận xét sửa. Một điểm cần lưu ý là do trình độ tiếng còn hạn chế nên sinh viên có thể đánh giá chưa thật đúng độ chuẩn xác của các nhận xét chữa lỗi.

Như vậy, cần tiến hành thêm nghiên cứu để biết được Merci App có thể nhận được, phân tích và đề xuất sửa một cách chính xác bao nhiêu lỗi trong bài viết tiếng Pháp. Câu hỏi chúng tôi đặt ra trong bài báo này là: Công cụ chữa lỗi tự động Merci App có độ chuẩn xác như thế nào khi chữa bài viết trình độ B2 cho sinh viên Việt Nam học tiếng Pháp?

2. Nội dung nghiên cứu

2.1. Phương pháp nghiên cứu

2.1.1. Đối tượng nghiên cứu và thu thập dữ liệu

Dữ liệu sử dụng cho nghiên cứu của chúng tôi là 30 bài viết luận trình độ B2 của 30 sinh viên năm thứ hai khoa Pháp một trường đại học phía Bắc Việt Nam. Các em đều là đầu vào tiếng Anh, mới học tiếng Pháp được ba học kì. Có 27 sinh viên nữ và 3 sinh viên nam. Bài viết của các em được gõ trên máy tính với phần mềm Word không sử dụng tính năng sửa lỗi tự động của Word cũng như bất kì công cụ chữa lỗi tự động nào, sau đó nộp cho chúng tôi và chúng tôi sẽ dán bài vào Merci App.

2.1.2. Xử lí dữ liệu

Tất cả các bài viết của sinh viên đã được mã hoá từ V01 đến V30. Chúng tôi sử dụng công cụ chữa lỗi tự động Merci App, phiên bản miễn phí để phát hiện các lỗi trong bài viết của sinh viên. Các lỗi sai sẽ được Merci App gạch chân, khi kích vào đó công cụ chữa lỗi sẽ cho thông tin phân loại lỗi, đưa ra các đề xuất chữa lỗi, phân tích lỗi sai tại đâu và nhắc lại quy tắc ngữ pháp. Ví dụ, trong câu “*En effet, les réseaux sociaux quy offrent de nombreux avantages aux hommes a transformé profondément notre mode de vie ainsi que nos habitudes quotidiennes.*”, Merci App gạch dưới lỗi “a”, phân loại lỗi là lỗi hợp giống số, đưa ra đề xuất sửa lỗi là “ont”, giải thích là động từ “a transformé” phải hợp ngôi và số với chủ ngữ là nhóm danh từ “Réseaux sociaux” là ngôi thứ ba số nhiều, nhắc lại quy tắc ngữ pháp là động từ phải được chia theo ngôi và số của chủ ngữ.

Chúng tôi đã phân loại lỗi dựa theo các bước trong quy trình sửa bài viết của Butterfield et al. (1996) [8], [9]. Theo đó, sửa bài gồm ba bước: phát hiện lỗi (detection), phân tích lỗi (diagnosis) và chữa lỗi (correction). Tất cả các lỗi mà Merci App nhặt ra và gạch chân báo lỗi sẽ được mã hóa là N (nhặt), các lỗi đó nếu đúng là lỗi sai được mã hóa là NĐ (nhặt đúng), nếu không phải là lỗi được mã hóa là NS (nhặt sai), tương ứng với False Positive [3]. Các lỗi nhặt đúng sẽ được xem xét xem lỗi có được phân tích và đưa ra phương án chữa lỗi chính xác không. Nếu phân tích lỗi sai thì được mã hóa là PTS (phân tích sai), nếu đề xuất sửa lỗi sai thì được mã hóa là SS (sửa sai). Tổng số lượng nhặt đúng và nhặt sai = nhặt. Sau đó, chúng tôi phân loại các lỗi phân tích sai, sửa sai và nhặt sai để xem Merci App có độ chuẩn xác thấp nhất với loại lỗi nào. Cuối cùng, chúng tôi sẽ đo xem Merci App không nhặt được những lỗi nào KN (không nhặt). Tuy nhiên, chúng tôi sẽ chỉ tập trung vào nhặt các lỗi ở cấp độ thấp (cấp độ câu, cụm từ và từ) mà Merci App bỏ sót chứ không tìm các lỗi ở cấp độ cao (viết đúng đề, ý tưởng, cấu trúc). Toàn bộ quá trình

mã hóa và nhật lỗi, phân loại lỗi được tác giả bài báo và một giảng viên đồng nghiệp tiến hành độc lập sau khi xử lý chung ba bài viết. Độ tin cậy giữa hai người mã hóa là Cronbach's Alpha = 0.86.

2.2. Kết quả

2.2.1. Độ chuẩn xác của Merci App về khả năng nhật lỗi

a. Số lượng lỗi nhật đúng, nhật sai và không nhật được

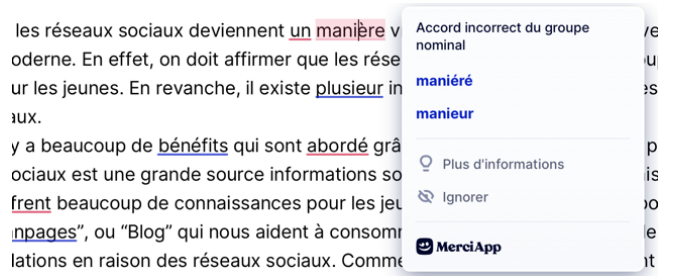
Tổng số lỗi sinh viên mắc trong 30 bài luận là 729 lỗi, trong đó Merci App nhật được 399 lỗi, chiếm 54,7% và không nhật được 330 lỗi, chiếm 45,3%. Trong số 399 lỗi ứng dụng nhật được có 330 là nhật đúng nghĩa là đúng là lỗi, còn 69 lỗi là nhật sai nghĩa là không phải là lỗi. Tỷ lệ các phát hiện lỗi chính xác nhật đúng chiếm 82,7%, tỷ lệ lỗi nhật sai chiếm 17,3% tổng số lỗi máy phát hiện ra. Tuy nhiên, nếu tính trên tổng số lỗi mà sinh viên mắc trong toàn bài thì các phát hiện lỗi chính xác chỉ chiếm 45,3%, các lỗi ứng dụng không nhật được chiếm tỷ lệ tương đương 45,3% và các lỗi ứng dụng nhật sai chiếm 9,5% (xem Bảng 1).

b. Phân loại lỗi nhật sai

Merci App nhật sai những loại lỗi nào? Trong 30 bài luận, Merci App đã nhật sai 69 lỗi, nghĩa là không phải là lỗi nhưng máy báo lỗi. Đứng đầu trong các loại lỗi bị nhật sai là các lỗi hợp giống số chiếm tới 71%. Tiếp theo là các lỗi về chính tả và chia động từ chiếm 8,7% số lỗi bị nhật sai. Đứng vị trí số bốn là lỗi về viết

hoa 5,8%, cuối cùng là lỗi về giới từ, phong cách với tỉ lệ khá thấp dưới 5% (xem Bảng 2).

Lí do của tỉ lệ lỗi hợp giống số nhật sai rất cao là vì trong một cụm danh từ khi bị sai hợp giống số thì Merci App sẽ báo lỗi tất cả các từ trong cụm đó. Ví dụ, khi sinh viên viết “un manière” thì máy sẽ báo lỗi “un” là không hợp giống số với “manière” và báo lỗi “manière” không hợp giống số với “un”, rồi đề xuất một danh từ giống đực có dạng thức chính tả gần với “manière” nhất. Như vậy, thường với một cụm danh từ, máy báo hai hoặc ba lỗi thì sẽ có một lỗi nhật đúng còn các lỗi còn lại là nhật sai (xem Hình 1).



Hình 1: Lỗi nhật sai

Đây cũng chính là yếu điểm của tất cả các công cụ chữa lỗi tự động, không dựa được vào ngữ cảnh bài viết để xác định xem từ nào là hạt nhân trong cụm danh từ và các mạo từ, tính từ đi kèm phải hợp giống số với danh từ hạt nhân đó. Dưới đây là một ví dụ nhật sai lỗi chia động từ. Trong câu “Les réseaux sociaux comme Facebook, Instagram, Twitter... sont appréciés chez nous.”, Merci App báo lỗi động từ “sont” và yêu cầu phải viết hoa là “Sont” do máy nhận biết sai là động từ này bắt đầu một câu.

c. Phân loại lỗi Merci App không nhật được

Tổng số lỗi mà Merci App bỏ sót trong 30 bài viết là 330 lỗi. Bảng 3 cho thấy kết quả phân loại các lỗi sai máy không nhật được. Đứng đầu danh sách chiếm 23,6% các lỗi không nhật được là các lỗi diễn đạt, đặc biệt là lỗi sai khi sinh viên muốn viết câu phức. Dùng được các cấu trúc câu phức đa dạng trong bài viết là một yêu cầu thiết yếu để đạt trình độ B2 nhưng rất nhiều sinh viên chưa nắm chắc mảng kiến thức này. Ví dụ, sự khác biệt giữa pour và pour que, cấu trúc nhấn mạnh c'est que ...

Đứng thứ hai, thứ ba và thứ tư trong một tỉ lệ rất sát nhau từ 16,7% đến 14,5% là lỗi về từ vựng, về hợp

Bảng 1: Số lượng lỗi nhật đúng (ND), nhật sai (NS) và không nhật được (KN)

T	N	ND	NS	KN
729 (N + KN)	399 (ND + NS)	330	69	330
		82,7% N	17,3% N	82,7% N
	54,7% T	45,3% T	9,5% T	45,3% T

(Ghi chú: T: Tổng số lỗi toàn bài; N: Nhật (Tổng lỗi app nhật được); ND: nhật đúng; NS: nhật sai; KN: không nhật)

Bảng 2: Phân loại lỗi nhật sai (NS)

Tổng NS	Hợp giống số	Chính tả	Chia động từ	Viết hoa	Giới từ	Phong cách
69	49	6	6	4	3	1
%	71%	8,7%	8,7%	5,8%	4,3%	1,4%

Bảng 3: Phân loại lỗi cấp độ thấp Merci App không nhật được (KN)

Tổng KN	Diễn đạt	Từ vựng	Hợp giống số	Giới từ	Động từ	Từ loại	Mạo từ	Chấm câu	Đại từ	Lược dấu phẩy trên
330	77	55	54	48	30	28	22	8	6	2
%	23,3%	16,7%	16,4%	14,5%	9%	8,5%	6,7%	2,4%	1,8%	0,6%

giống số và về giới từ. Merci App đặc biệt gặp khó khăn khi nhận lỗi hợp tính từ với danh từ đi kèm, đặc biệt ở những câu mà danh từ không đi liền sát với tính từ. Tiếp theo là các lỗi về động từ, lỗi nhầm loại từ, lỗi mạo từ với tỉ lệ 9%, 8,5% và 6,7%. Một điểm đáng lưu ý là tương tự như Criterion, Merci App không nhận được lỗi sai khi người viết dùng sai từ loại. Ví dụ, dùng tính từ “Raisonnable” thay cho trạng từ “Raisonnement” hay dùng danh từ “Nombre” thay cho tính từ “Nombreux”. Đứng cuối cùng trong danh sách lỗi máy không nhận được là lỗi về chấm câu, đại từ, lược dấu phẩy trên với tỉ lệ rất thấp dưới 3%.

Như vậy, có thể thấy, Merci App không hiệu quả với các loại lỗi đòi hỏi phải hiểu tình huống giao tiếp và hiểu sự kết nối giữa các yếu tố ngôn ngữ không luôn đúng sát cạnh nhau như lỗi diễn đạt, từ vựng, hợp giống số giữa danh từ và tính từ khi đứng cách xa nhau.

2.2.2. Độ chuẩn xác của Merci App về khả năng phân tích lỗi và sửa lỗi

a. Số lượng phân tích lỗi sai và sửa lỗi sai

Trong số 330 các phát hiện lỗi chính xác, có 253 lỗi xếp loại nhận, phân tích, sửa đúng (đúng cả phần phân tích lỗi và phần đề xuất sửa lỗi). Điều này có nghĩa là, số nhận xét lỗi hoàn toàn chính xác và thực sự hữu dụng cho sinh viên chỉ chiếm 63,4% tổng số lỗi Merci App phát hiện ra và chiếm 34,7% tổng lỗi cấp độ thấp toàn bài (xem Bảng 4). Một điểm đáng lưu ý là các lỗi chính tả thường sẽ có nhiều phương án đề xuất chữa lỗi, và người dùng phải đủ năng lực để chọn phương án chữa đúng. Khi mã hóa thì chỉ cần có một phương án đúng trong các đề xuất sửa lỗi Merci App đưa ra thì chúng tôi đều xếp loại sửa đúng nhưng chúng tôi không chắc là trên thực tế sinh viên chọn được phương án sửa lỗi đúng trong số 4-5 đề xuất. Ví dụ, với lỗi chính tả “réaux” (lỗi của bài V03) thì Merci App đưa ra 5 đề xuất sửa lỗi: réseau, réseaux, réaux, ressaut, réaux. Trong 5 đề xuất này chỉ có một đề xuất đúng là “réaux”. Hay với lỗi chính tả “vélunérable” (lỗi của bài V03), Merci App đưa ra 4 đề xuất sửa lỗi: vulnérable, vulnérables, vénérable, vénérables. Trong 4 đề xuất này chỉ có một

Bảng 4: Số lượng phân tích lỗi sai (PTS), sửa lỗi sai (SS)

T	N	ND	NPTSD	PTS	SS
729	399	330	253	57	67
	%	82,7% N	63,4% N	14,3% N	16,8% N
		45,3% T	34,7% T	7,8% T	9,2% T

(Ghi chú: T: Tổng lỗi toàn bài; N: Nhận (Tổng lỗi app nhận được); ND: Nhận đúng; NPTSD: Nhận, phân tích, sửa đúng; PTS: Phân tích sai; SS: Sửa sai)

đề xuất đúng là “vulnérables”. Số lượng phân tích lỗi sai là 57 lỗi, chiếm 14,3% số lỗi máy phát hiện được và chiếm 7,8% tổng lỗi toàn bài. Số lượng đề xuất sửa lỗi sai là 67 lỗi, chiếm 16,8% số lỗi máy phát hiện được và chiếm 9,2% tổng lỗi toàn bài. Như vậy, Merci App phân tích lỗi tốt hơn một chút so với đề xuất phương án sửa lỗi.

b. Phân loại phân tích lỗi sai và đề xuất sửa sai

Bảng 5 và Bảng 6 cho thấy kết quả phân loại các phân tích lỗi sai và các đề xuất sửa sai. Tổng phân tích lỗi bị sai là 57, thấp hơn so với đề xuất sửa sai 10 lỗi. Từ vựng là loại lỗi có tỉ lệ phân tích và đề xuất sửa sai cao nhất, chiếm hơn 40%, tiếp đến là lỗi về động từ hơn 20%, lỗi về hợp giống số hơn 10%. Cuối cùng là các lỗi về giới từ, vị trí đại từ và lỗi chính tả chiếm dưới 10%. Chúng ta thấy, ở Bảng 2 là Merci App không nhận sai lỗi từ vựng nhưng do không hiểu được ý muốn diễn đạt của người viết nên thường đưa ra phương án sửa lỗi từ vựng sai.

Bảng 5: Phân loại phân tích lỗi sai (PTS)

Tổng	Từ vựng	Động từ	Hợp giống số	Giới từ	Vị trí đại từ	Chính tả
57	28	13	7	4	3	2
	49,1%	22,8%	12,3%	7%	5%	3,5%

Bảng 6: Phân loại đề xuất sửa sai (SS)

Tổng	Từ vựng	Động từ	Hợp giống số	Giới từ	Vị trí đại từ	Chính tả
67	31	16	10	4	3	3
	46,3%	23,9%	14,9%	5,9%	4,5%	4,5%

Sau đây là một vài ví dụ phân tích lỗi sai và đề xuất sửa sai. Ví dụ, khi sinh viên dùng sai tính từ “facile”, Merci App đã phân tích đúng là chỗ này cần dùng một động từ nhưng đưa ra đề xuất sửa sai là “fascinant”. Hay khi sinh viên viết: “*Ensuite, ils représentent un moyen de communication*”, Merci App nhận đúng lỗi là “représent” nhưng giải thích sai là vị trí này cần một động từ và đưa ra đề xuất sửa sai là “reprisent”.

2.3. Thảo luận

Để xác định độ chuẩn xác của Merci App trong sửa lỗi bài viết, chúng tôi đã phân tích 30 bài viết B2 của sinh viên năm thứ hai chuyên tiếng Pháp tại một trường đại học ở miền Bắc Việt Nam. Kết quả cho thấy, công cụ sửa lỗi tự động tìm được 399 lỗi, bỏ sót 330 lỗi trong tổng số 729 lỗi sinh viên mắc trong 30 bài viết. Như vậy, Merci App bỏ sót 45% số lỗi cấp độ thấp trong bài, chưa tính tới các lỗi cấp độ cao mà ứng dụng không

nhận xét được và chúng tôi cũng chưa phân tích trong nghiên cứu này. Kết quả này khá tương đồng với nghiên cứu của Lavolette et al. (2015) [4], theo đó Criterion bỏ sót 46% số T-unit chứa lỗi sai; không phát hiện được lỗi về từ loại, lỗi giới từ và mạo từ.

Trong số 399 lỗi ứng dụng tìm được có 330 lỗi thực sự là lỗi chiếm 82,7%, còn 69 phát hiện không phải là lỗi chiếm 17,3%. Tỷ lệ lỗi Merci App nhậ sai cao hơn kết quả nghiên cứu của Lavolette et al. (2015) [4] là 11% và của Hoang (2022) [3] là 11,7%.

Một số nghiên cứu đã phân loại lỗi nhậ sai trong tiếng Anh và thấy Criterion hay nhậ sai lỗi về đoạn văn, giới từ, dấu phẩy [3], mạo từ, chính tả, run on sentence [4], trong khi Merci App nhậ nhiều lỗi sai nhất về hợp giống số rồi đến chính tả, chia động từ. Lỗi nhậ sai về giới từ không nhiều nhưng lỗi không nhậ về giới từ khá cao. Như vậy, phân loại lỗi nhậ sai tùy thuộc vào đặc trưng của từng ngôn ngữ.

Trong số 330 lỗi thực sự là lỗi thì 253 lỗi được phân tích lỗi và đề xuất sửa chất lượng, nghĩa là nhận xét thực sự hữu dụng cho người dùng chỉ chiếm 63,4% tổng số lỗi Merci App tìm được và chiếm 34,7% tổng số lỗi sinh viên mắc khi viết bài. Độ chuẩn xác của công cụ chữa lỗi tự động Merci App chúng tôi tìm thấy trong nghiên cứu này là 82,7%, tương đồng với kết quả đã tìm thấy trong các nghiên cứu trước của Hoang (2022) [3] là 82,2%, cao hơn so với kết quả của Lavolette et al. (2015) [4] là 75%, Dikli & Bleyle (2014) [10] là 72,7%. Cùng với thời gian, các công cụ chữa lỗi tự động đã được hoàn thiện và có độ chuẩn xác cao hơn trước. Tỷ lệ các lỗi nhậ đúng, phân tích đúng, đề xuất sửa đúng là 63,4%, cao hơn tỷ lệ nhận xét được người dùng sửa theo máy trong nghiên cứu của Hoang (2022) [3] là 54%; của Koltovskai (2020) [11] là 57%; tương đương với kết quả của Ranalli et al. (2017) [12] là 55-65%; và thấp hơn kết quả của Lavolette et al. (2015) [4] là 73%.

Tỷ lệ lỗi Merci App phân tích sai là 14,3% và đưa ra đề xuất sửa sai là 16,8%, tương đương với kết quả nghiên cứu của Lavolette et al. (2015) [4] có tỷ lệ phân tích sai và sửa sai là 14%.

3. Kết luận

Các nghiên cứu trước đây thường tính độ chuẩn xác bằng cách lấy số lỗi nhậ chính xác chia cho tổng số lỗi mà máy tìm được. Theo chúng tôi, hiệu quả thực sự của ứng dụng sửa lỗi tự động cần được đo bằng cách lấy số lỗi được nhậ, phân tích và đề xuất sửa chính xác chia cho tổng lỗi sinh viên mắc trong bài (bằng tổng lỗi app tìm được cộng với số lỗi máy không tìm được) với nghiên cứu này là 34,7%. Như vậy, khi dùng Merci App, người viết có thể sửa được hiệu quả khoảng 1/3 số lỗi cấp độ thấp mắc trong bài. Một lần nữa, chúng tôi khẳng định vai trò quan trọng của giảng viên trong việc dạy viết, dù cho có sự hỗ trợ đặc lực của các công cụ chữa lỗi tự động. Chúng tôi khuyến nghị giảng viên cần làm thật kỹ phần phân tích đề, xác định cấu trúc bài viết, các ý chính cần triển khai trong bài viết trước khi cho sinh viên viết nhằm hạn chế tối đa các lỗi cấp độ cao.

Để sử dụng hiệu quả Merci App, nên có buổi hướng dẫn cách sử dụng ứng dụng này, lưu ý những lỗi công cụ hay nhậ sai như báo lỗi hợp giống số cho tất cả các thành phần trong nhóm danh từ. Tiếp theo, giảng viên cần cung cấp từ vựng siêu ngôn ngữ và làm demo trên 2-3 bài viết để sinh viên hiểu được các phân tích lỗi trên Merci App. Cuối cùng, sau khi app chữa xong và sinh viên sửa xong bài trên ứng dụng, giảng viên vẫn phải chữa lại mới đảm bảo được sự tiến bộ của người học. Giảng viên cần tập trung sửa các lỗi cấp độ cao và các lỗi cấp độ thấp mà Merci App thường không nhậ được hay nhậ sai, sửa sai như lỗi diễn đạt, lỗi từ vựng, lỗi hợp giống số và lỗi giới từ.

Điểm hạn chế của bài báo này là số lượng mẫu 30 bài viết còn khiêm tốn. Ở những nghiên cứu sau, có thể xem xét tiến hành trên những mẫu lớn hơn và trong thời gian dài hơn để đo sự tiến bộ của sinh viên khi sử dụng công cụ chữa lỗi tự động. Một hướng nghiên cứu nữa là khảo sát quan niệm và cách giảng viên sử dụng công nghệ để chữa bài viết ngoại ngữ, theo hướng nghiên cứu của Koltovskaia [13].

Tài liệu tham khảo

- [1] Barrot, J. S. (2023), *Using automated written corrective feedback in the writing classrooms: effects on L2 writing accuracy*, Computer Assisted Language Learning, 36(4), 584-607, <https://doi.org/10.1080/09588221.2021.1936071>.
- [2] Guo, Q., Feng, R., & Hua, Y. (2022), *How effectively can EFL students use automated written corrective feedback (AWCF) in research writing?*, Computer Assisted Language Learning, 35(9), 2312-2331, <https://doi.org/10.1080/09588221.2021.1879161>.
- [3] Hoang, T. L. G. (2022), *Feedback precision and learners' response: A study into ETS Criterion automated corrective feedback in EFL writing classrooms*, The JALT CALL Journal, 18(3), 444-467.
- [4] Lavolette E, Polio C & Kahng J, (2015), *The accuracy of computer-assisted feedback and students' responses to it*, Language Learning & Technology, 19(2), 50-68, <http://dx.doi.org/10.125/44417>.
- [5] Hunt, K. W. (1965), *Grammatical structures written at three grade levels*, NCTE Research Report No.3, Champaign, IL: National Council of Teachers of English, Retrieved from <http://files.eric.ed.gov/fulltext/ED113735.pdf>
- [6] Grégoire, P. (2021), *L'utilisation d'un outil numérique d'aide à la révision et à la correction à la fin du secondaire : effets sur la qualité de l'écriture*, Canadian

- Journal of Education / Revue canadienne de l'éducation, 44(3), 788-814, <https://doi.org/10.53967/cje-rce.v44i3.4809>.
- [7] Đỗ Thị Bích Thủy, (2024), *Sử dụng Google Docs và Merci App trong dạy viết tiếng Pháp*, Tạp chí Khoa học Giáo dục, số 2, tr. 25-34.
- [8] Butterfield, E., Hacker, D., & Albertson, L. (1996), *Environmental, cognitive and metacognitive influences on text revision: Assessing the evidence*, Educational Psychology Review, 8(3), 239-260.
- [9] Do, T. B. T. (2023), *Effects of scaffolded peer review training on revision quantity and quality in foreign language writing*, Iranian Journal of Language Teaching Research, 11(2), 55-73, doi: 10.30466/ijltr.2023.121330.
- [10] Dikli, S., & Bleyle, S. (2014), *Automated essay scoring feedback for second language writers: How does it compare to instructor feedback*, *Assessing Writing*, 22, 1-17, <https://doi.org/10.1016/j.asw.2014.03.006>.
- [11] Koltovskaia S. (2020), *Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study*, *Assessing Writing*, 44, <https://doi.org/10.1016/j.asw.2020.100450>.
- [12] Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017), *Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation*, *Educational Psychology*, 37(1), 8-25, <https://doi.org/10.1080/01443410.2015.1136407>.
- [13] Koltovskaia, S. (2023), *Postsecondary L2 writing teachers' use and perceptions of Grammarly as a complement to their feedback*, *ReCALL*, 35(3), 290-304, doi:10.1017/S0958344022000179.

THE ACCURACY OF THE MERCI APP - AN ERROR CORRECTION TOOL - IN FRENCH INSTRUCTION

Do Thi Bich Thuy

Email: thuydtb1976@vnu.edu.vn
University of Languages and International Studies,
Vietnam National University, Hanoi
Pham Van Dong street, Cau Giay district,
Hanoi, Vietnam

ABSTRACT: *This study aims to measure the accuracy of the free version of the Merci App, an automatic error correction tool used in teaching French writing. The research data consists of 30 B2 level essays written by 30 second-year students majoring in French Language. The results show that the number of errors accurately detected by the Merci App accounts for 82.7% of the total errors detected by the application and 45.3% of the total number of local errors made by students in the drafts. The number of errors being well detected, diagnosed, and corrected by the Merci App represents 63.4% of the total errors detected by the application and 34.7% of local errors appearing in the essays. Merci App often incorrectly identifies errors in gender and number agreement in a noun phrase; misses a large number of errors in expression, vocabulary, gender and number agreement, prepositions; wrongly diagnoses and corrects many vocabulary, verbs, gender and number agreement. We recommend that language teachers instruct learners to use automatic error correction tools, provide a meta-linguistic vocabulary to help them understand error diagnoses and error corrections in the Merci App, and focus on correcting high or low level errors that can not be found or inaccurately detected, diagnosed and corrected by this tool.*

KEYWORDS: Automatic error correction, accuracy, foreign language instruction, Merci App, French.