

HIỆN TƯỢNG ALGOSPEAK: CHIẾN LƯỢC NGÔN NGỮ MÃ HOÁ TRONG KHÔNG GIAN THUẬT TOÁN (TRƯỜNG HỢP MẠNG XÃ HỘI TIKTOK)

NGUYỄN PHƯƠNG NGÂN¹

ĐỖ THI XUÂN DUNG²

Abstract: This paper examines the linguistic features of Algospeak and their implications for multimodal discourse in algorithmic spaces. Algospeak is a phenomenon of language adaptation to digital media, where users modify their expression to avoid the platform's strict censorship. Using a combination of quantitative and qualitative analysis, we extract data from 10 trending TikTok videos selected by keyword and popularity, where content creators have coded their message using various multimodal forms to describe violent events without violating algorithmic principles. The findings highlight the evolving nature of language in the digital context and raise important questions for linguistic theory, digital media, and multimodal discourse in the future.

Keywords: *Algospeak, algorithms, social media language, content moderation, multimodal discourse*

1. Đặt vấn đề

1.1. Sự phát triển nhanh chóng của truyền thông kỹ thuật số đã đặt ra những thách thức mới đối với việc sử dụng ngôn ngữ, đặc biệt là trong khía cạnh ràng buộc bởi quy tắc kiểm duyệt theo thuật toán. Các nền tảng mạng xã hội như Facebook, TikTok, Instagram và YouTube sử dụng các hệ thống tự động để lọc hoặc ngăn chặn nội dung được coi là không phù hợp, chẳng hạn như bạo lực, ngôn từ kích động thù địch hoặc nội dung phản cảm. Người dùng - đặc biệt là những người thảo luận về các chủ đề này (câu chuyện vốn chứa đựng các từ ngữ về bạo lực, giới tính, chủng tộc...) buộc phải phát triển những giải pháp diễn đạt thay thế để tránh nội dung họ bị kiểm duyệt bởi thuật toán. Hiện tượng này được gọi là Algospeak, thể hiện chiến lược mã hoá ngôn ngữ nhằm thích nghi với những áp chế của nền tảng công nghệ.

1.2. Algospeak đặt ra những câu hỏi quan trọng về sự tương tác giữa ngôn ngữ, công nghệ và diễn ngôn trên mạng xã hội. Thuật toán ảnh hưởng đến việc sử dụng ngôn ngữ trong giao tiếp trực tuyến như thế nào? Những đặc trưng ngôn ngữ mới nào xuất hiện từ Algospeak? Và hiện tượng này đặt ra những thách thức nào đối với nghiên cứu ngôn ngữ trong tương lai? Bài viết giải quyết các vấn đề trên bằng cách phân tích 10 bài đăng trên Tiktok Việt Nam. Những video này, được chọn dựa trên bộ lọc từ khoá và sự hiện diện của chúng trong các đề xuất hàng đầu, chứng minh cách người dùng cá nhân sử dụng Algospeak để điều hướng các hạn chế về thuật toán trong khi vẫn duy trì hiệu quả giao tiếp về nội dung bạo lực.

2. Giới thuyết chung

2.1. Đặc trưng của thuật toán trong truyền thông kỹ thuật số

Theo nhà nghiên cứu Fry, “thuật toán (algorithm) chỉ đơn giản là một chuỗi các hướng dẫn logic chỉ ra cách thực hiện một nhiệm vụ từ đầu đến cuối” [1, tr.8]. Trong đó, thuật toán thực hiện hai nhiệm vụ cơ bản có khả năng tác động và chi phối đến thông tin là: (1) Xếp thứ tự ưu tiên, tức là lập danh sách có thứ tự, chẳng hạn như trong kết quả tìm kiếm và (2) Phân loại, tức là chọn một số danh mục nhất định như phân loại và xoá nội dung không phù hợp trên các nền tảng đa phương tiện (ví dụ YouTube) [6, tr.102]. Quyền năng của thuật toán không chỉ giới hạn ở việc điều chỉnh nội dung mà còn mở rộng

¹ Nghiên cứu sinh Trường Đại học Khoa học, Đại học Huế

² Đại học Huế (Tác giả liên hệ)

quản lý các hoạt động thảo luận, xác định khả năng hiển thị và tác động đến mức độ tương tác của người dùng. Do đó, thuật toán được xem như “người gác cổng kỹ thuật số”, chúng “ảnh hưởng đến những gì có thể nhìn thấy trực tuyến, do đó định hình những gì được xem là phù hợp lợi ích cộng đồng” [7, tr.29].

Thuật toán kiểm duyệt, nằm trong các chức năng kể trên, sử dụng công cụ tự động để phát hiện và xoá nội dung vi phạm chính sách của nền tảng. Các hệ thống này sử dụng kỹ thuật xử lý ngôn ngữ tự nhiên (natural language processing - NLP) và thị giác máy tính để xác định nội dung có hại, gây hiểu lầm hoặc phản cảm. Đây là một hành động cần thiết để đảm bảo môi trường thông tin văn minh và an toàn. Tuy nhiên, do tính phức tạp của ngôn ngữ và sắc thái văn hoá, việc kiểm duyệt theo thuật toán thường dẫn đến kiểm duyệt quá mức hoặc diễn giải sai ngữ cảnh. Ví dụ, diễn ngôn về chủ đề nhạy cảm như sức khỏe tinh thần, bạo lực có thể bị đánh dấu là không phù hợp ngay cả khi được trình bày trong bối cảnh giáo dục hoặc báo chí. Như Steen, Yurechko và Klug kết luận trong nghiên cứu của mình: “Việc kiểm duyệt nội dung theo thuật toán của Tiktok không có khả năng hiểu được ngữ cảnh cụ thể mà người tham gia thảo luận về các chủ đề lành tính” [6, tr.12]. Cụ thể hơn, một từ ngữ bạo lực được sử dụng không phải lúc nào cũng với mục đích kích động mà có thể nhằm truyền tải về một tin tức xã hội, lấy các trường hợp khảo sát tại mục 4 bài viết là ví dụ.

Như vậy, dù không phải là các chuyên gia công nghệ, người dùng Internet đương đại liên tục tiếp xúc với các hệ thống thuật toán phức tạp và tinh vi phủ rộng khắp các nền tảng. Thuật toán không phải là trung gian thụ động mà luôn có sự tham gia tích cực vào việc định hình diễn ngôn kỹ thuật số. Vai trò của chúng trong việc cấu trúc các tương tác trực tuyến đặt ra những thách thức về cách thức ngôn ngữ được sử dụng và biến đổi. Khi các thuật toán tiếp tục phát triển, sự hiểu biết của chúng ta về ý nghĩa của chúng đối với giao tiếp và diễn ngôn kỹ thuật số cũng phải thay đổi.

2.2. Diễn ngôn đa thức trong không gian mạng xã hội

Không gian mạng xã hội, tương tự như các không gian trực tuyến khác, bị định đoạt và chi phối bởi thuật toán công nghệ nhưng cũng cung cấp hàng loạt tài nguyên biểu đạt mới bên ngoài ngôn ngữ - chẳng hạn hình ảnh, âm thanh, biểu tượng, hiệu ứng... Điều này giúp người dùng đa dạng sự lựa chọn từ hình thức diễn ngôn đơn thức (chỉ có ngôn ngữ) sang đa thức với sự kết hợp của nhiều yếu tố phi ngôn ngữ, tạo nên hiện tượng “diễn ngôn đa thức”. Như nhà nghiên cứu Kress và van Leeuwen nhận xét, trong môi trường truyền thông kỹ thuật số, “tính đa thức đang chuyển sang trung tâm của hành động giao tiếp” [5, tr.45] và hiện tượng này phản ánh “tinh thần thời đại” [5, tr.46].

Phương pháp phân tích diễn ngôn đa thức (multimodal discourse analysis - MDA) trong các bài đăng Facebook, video Tiktok ... do đó không chỉ dừng lại ở việc giải mã ngữ nghĩa văn bản mà còn phải mở rộng sang việc đọc hiểu các yếu tố phi ngôn ngữ - như cấu trúc hình ảnh, tương tác giữa âm thanh và lời nói, vai trò của biểu tượng thị giác, được đặt trong những bối cảnh nhất định. Trong “Cẩm nang phân tích đa phương thức”, nhà nghiên cứu Jewitt nhấn mạnh MDA tập trung vào “hiểu và mô tả các nguồn kí hiệu và nguyên tắc của hệ thống ý nghĩa của chúng để hiểu cách mọi người sử dụng các nguồn này trong bối cảnh xã hội cho các mục đích cụ thể” [3, tr. 33].

Nhìn nhận diễn ngôn dưới khía cạnh đa thức giúp thúc đẩy tìm hiểu sự đa dạng hoá trong cách diễn đạt ngôn ngữ, khiến nhà nghiên cứu chú ý hơn đến vai trò của yếu tố cảm xúc và thị giác trong giao tiếp. Điều này khiến cho ngôn ngữ không còn là một nguồn lực tạo nghĩa độc lập (chẳng hạn như trong một tác phẩm văn chương hay đoạn hội thoại thông thường) mà là một phần của những tập hợp các nguồn lực có khả năng đồng bộ hoá lớn hơn. Như nhà nghiên cứu Jordan nhận định: “Lĩnh vực đa phương tiện mới... loại bỏ sự thống trị của bảng chữ cái, tái tích hợp hoàn toàn hình ảnh nghe nhìn với chữ viết. Không gian mạng nổi lên ở đây như một thành phần có ý nghĩa chiến lược trong những thay đổi này, bởi vì nó vừa giới thiệu các hình thức truyền thông mới vừa cung cấp một ứng dụng đa phương tiện cực kỳ phức tạp” (dẫn theo [2, tr.38]).

2.3. Hiện tượng ngôn ngữ Algospeak

Algospeak (ngôn ngữ mã hoá) là một thuật ngữ mới nổi, ghép từ “algorithm” (thuật toán) và “-speak” (nói). Hiện tượng này xuất hiện như một dạng biến thể hoặc mã hoá trong giao tiếp do người dùng mạng xã hội phát triển với “mục đích chính là sử dụng ngôn ngữ để tránh kiểm duyệt nội dung đặc biệt theo thuật toán” [6, tr.2], cụ thể là khi thảo luận về các vấn đề nhạy cảm hoặc gây tranh cãi. Sự phát triển của Algospeak gắn liền với sự tinh vi ngày càng tăng của các hệ thống kiểm duyệt nội dung. Khi các hệ thống này trở nên tiên tiến hơn, người dùng phải liên tục điều chỉnh ngôn ngữ, phát triển các chiến lược để duy trì sự rõ ràng của thông điệp trong khi tránh sự ngăn chặn của thuật toán.

Hiện tượng Algospeak đã xuất hiện trong các nghiên cứu gần đây về ngôn ngữ học kỹ thuật số và quyền lực công nghệ. Bằng cách phỏng vấn người sáng tạo nội dung về chủ đề đồng tính trên TikTok, Steen cùng cộng sự đã kết luận định nghĩa Algospeak gắn liền với diễn ngôn đa thức khi cho rằng: “Algospeak vượt ra ngoài các yếu tố văn bản thành các biến thể trực quan và nghe nhìn khéo léo”, do đó “Algospeak cần được hiểu là các từ mã và biến thể ngôn ngữ, giao tiếp trực quan và đa phương thức và sự gắn kết nghe nhìn trong tương tác trên mạng xã hội” [6, tr.12]. Đồng thời, các nhà nghiên cứu chỉ rõ nguyên nhân xuất hiện của Algospeak là do hành động quản lý của nền tảng Tiktok gây “cản trở nhu cầu giao tiếp và quyền tự do ngôn luận của người dùng” [6, tr.12]. Một số Algospeak được chỉ ra như thay đổi cách đánh máy (@b0rt!0n thay cho *abortion* tức “phá thai”), dùng biểu tượng (🐱 chỉ bộ phận sinh dục nữ, 🍆 chỉ bộ phận sinh dục nam)...

Tại Việt Nam, trong khả năng khảo sát của người viết, tuy chưa có nghiên cứu nhưng Algospeak đã xuất hiện phổ biến trong các nền tảng mạng xã hội như Facebook, Tiktok... Sự sáng tạo này dễ dàng nhận thấy qua nhiều khía cạnh, chẳng hạn như biến âm (nói chệch, nói lái: *chết* thành *chiết*, *lừa đảo* thành *đào lừa*...), thay đổi tên gọi (*rượu* thành *đồ uống có cồn*, Facebook thành *Phở bò*, *đánh* thành *tác động vật lý*), lỗi chính tả cố ý (*giết* thành *gi.ết*, *hung khí* thành *h.u.n.g k.h.í*). Đặc biệt hơn, Algospeak còn thường xuyên xuất hiện ở các dạng đa phương thức khi sử dụng kết hợp biểu tượng cảm xúc, ngôn ngữ cơ thể, âm thanh, hình ảnh kỹ thuật số... Các hình thức biến đổi này tạo thêm nhiều lớp nghĩa cho diễn ngôn khiến thuật toán khó phát hiện và kiểm duyệt nội dung hơn. Nghiên cứu Algospeak do đó có tiềm năng mở ra những góc nhìn mới cho sự tiến hoá của ngôn ngữ, khả năng đọc viết kỹ thuật số của người dùng.

3. Phương pháp và dữ liệu nghiên cứu



Ảnh 1. Dữ liệu lựa chọn theo bộ lọc Tiktok



Ảnh 2. Video 6. Người dùng tránh sử dụng các từ ngữ bạo lực

Bài viết sử dụng phương pháp phân tích diễn ngôn đa phương thức, do đó quan tâm đến sự tương tác của các phương thức tạo nghĩa từ ngôn ngữ đến các yếu tố phi ngôn ngữ (chẳng hạn cử chỉ, biểu cảm, âm thanh...) để tạo nên Algospeak. Bài viết khảo sát định lượng kết hợp định tính đối với 10 bài đăng dạng video trên nền tảng Tiktok Việt Nam theo từ khoá *nhungvuankinhhoang* (*Những vụ án kinh hoàng*), đạt mức đề xuất và tương tác hàng đầu, đăng tải trong 06 tháng cuối năm 2024. Từ khoá được lựa chọn bởi nó điều hướng đến chủ đề bạo lực, vi phạm nguyên tắc cộng đồng và bị ràng buộc bởi thuật toán. Tuy nhiên, những người sáng tạo nội dung đã khéo léo xây dựng thông tin dễ hiểu, dễ tiếp cận với mức độ lan truyền cao (thể hiện ở số lượng tương tác hàng chục nghìn lượt thích và hàng trăm bình luận, chia sẻ) mà vẫn đánh lừa được thuật toán.

Dữ liệu được thu thập dựa trên hai khía cạnh chính:

(1) Ngôn ngữ: bao gồm văn bản nói - viết, lời chú thích...;

(2) Yếu tố trực quan và âm thanh, hoạt động cận ngôn: cử chỉ, hành động người nói, âm thanh lồng ghép, biểu tượng cảm xúc...

Đối với lời nói, người viết sử dụng phần mềm Capcut để trích xuất phụ đề video. Các hiện tượng ngôn ngữ được biến đổi có chủ đích nhằm né tránh thuật toán được xem là các bất thường cần được phân loại và thống kê tần suất xuất hiện. Phạm vi khảo sát không quan tâm đến ngôn ngữ bị biến đổi do ngôn ngữ địa phương, yếu tố địa lí - khu vực. Đồng thời, do tính chất nhạy cảm của dữ liệu và đảm bảo quyền riêng tư, các video trích dẫn trong bài viết được ẩn danh và xoá mờ hình ảnh người.

4. Kết quả nghiên cứu

Đánh giá sơ bộ, 10 video trích xuất có lượt tương tác cao (hàng nghìn lượt thích) cho thấy nội dung được truyền tải thành công đến công chúng. Ý nghĩa thông điệp được thể hiện dưới nhiều phương thức bao gồm ngôn ngữ, trực quan hình ảnh, hành động, âm thanh kỹ thuật số. Tuy nhiên, dữ liệu xuất hiện những bất thường trong cách sử dụng ngôn ngữ, trong đó người sáng tạo nội dung cố tình sửa đổi từ ngữ bạo lực, nhạy cảm để tránh bị thuật toán phát hiện.

Video	Chủ đề	Lượt thích	Ví dụ Algospeak	Giải nghĩa
1	Đưa tin về dự thảo luật tại Iraq	85,6 nghìn	Trẻ 🤔, cưỡng d***	Trẻ em, cưỡng dâm
2	Sơ lược về anh hùng Mai Thị Nương	29,8 nghìn	Moi liver, tách lego	Moi gan, phân xác
3	Đánh giá phim	40,1 nghìn	T.r.a t.á.n, P!g, *tiếng bíp	Tra tấn, lộn
4	Đưa tin về vụ án cũ	261,9 nghìn	Lột skin, cho đăng xuất	Lột da, giết
5	Đưa tin về vụ án cũ	23,9 nghìn	Thithe, choon, *tiếng bíp, *khẩu hình	Thi thể, chôn
6	Đưa tin về vụ án cũ	10,7 nghìn	Đã** xuấ*, con dờ	Chết, con dao
7	Triển lãm Zombbie (video do trí tuệ nhân tạo thực hiện)	405,9 nghìn	Th.âm h.ọạ, kh.ùng kh.iếp	Thảm họa, khủng khiếp
8	Đưa tin về vụ án cũ	340 nghìn	Tình *hiệu ứng âm thanh làm mất âm chữ sau	Tình dục
9	Đưa tin về vụ án cũ, quảng cáo phim	31,9 nghìn	Hạ s*t	Giết
10	Đưa tin về vụ án	98,7 nghìn	Không sử dụng	

Trong một số video, người sáng tạo nội dung đã lí giải cho chiến lược sử dụng ngôn ngữ trong video nhằm thích nghi thuật toán nền tảng. “Trước khi kể câu chuyện này thì tui cũng hi vọng là mọi người có thể thông cảm cho tui sẽ có những từ ngữ mà tui phải **nói tránh, nói né để tránh gây của TikTok**, nhưng mà không có nghĩa là tui không tôn trọng các vị anh hùng” (Video 2). Trong đó, **gây của TikTok** (Video 2), **bị đánh gây** (Video 6) đề cập đến hình phạt của nền tảng khi người dùng vi phạm

thuật toán. Điều này là minh chứng cho việc họ hiểu rõ cơ chế kiểm duyệt của TikTok và biết rằng nếu dùng từ ngữ trực tiếp liên quan đến bạo lực, video có thể bị xoá hoặc giảm tương tác. Đồng thời, người nói muốn nhấn mạnh việc tách biệt việc né tránh kiểm duyệt với thái độ cá nhân, khẳng định rằng việc sử dụng ngôn ngữ thay thế không đồng nghĩa với việc xem nhẹ nội dung lịch sử hay chính trị.

4.1. Đặc điểm ngôn ngữ trong Algospeak

Dữ liệu từ 10 video TikTok được phân tích trong nghiên cứu này cho thấy hiện tượng Algospeak xuất hiện với tần suất cao, đặc biệt là trong các nội dung liên quan đến bạo lực hoặc vi phạm nguyên tắc cộng đồng. Trong số này, 90% video chứa ít nhất một dạng biến đổi ngôn ngữ nhằm tránh kiểm duyệt với tổng số 83 Algospeak và 80% video kết hợp từ hai chiến lược khác nhau để tối ưu hoá khả năng vượt qua thuật toán. Các chiến lược phổ biến nhất bao gồm biến đổi ngữ âm và chính tả, biến đổi tên gọi, cũng như biến đổi phương thức biểu đạt.

4.1.1. Biến đổi về mặt ngữ âm và sử dụng lỗi chính tả cố ý

Một trong những chiến lược cơ bản và dễ nhận biết nhất (xuất hiện 42 lần trong 8/10 video, chiếm tỉ lệ 50,6% trên tổng số tần suất sử dụng Algospeak) là thay đổi phát âm của từ gốc hoặc cố tình sử dụng lỗi chính tả để đánh lạc hướng thuật toán nhận diện nội dung. Điều này được thực hiện theo nhiều cách khác nhau, bao gồm nói lái, thay đổi ngữ âm, chèn kí tự đặc biệt hoặc dùng lỗi đánh máy cố ý.

Ví dụ, trong dữ liệu khảo sát, từ *giết* được thay thế bằng nhiều biến thể thay đổi ngữ âm (*quyết, iết*). Hoặc trong trường hợp đọc lược bỏ âm đầu như từ *chết* thành *éch*; lược bỏ âm cuối như từ *con dao* thành *con dờ*, *cái xác* thành *cái xờ*. Chơi chữ bằng hình thức nói lái cũng diễn ra phổ biến: *tâm thâm* (tâm thần), *tĩnh hù* (tử hình)...

Ngoài ra, người dùng thường xuyên chèn kí hiệu dấu chấm (.) hoặc dấu sao (*), khoảng trống văn bản, cố ý tạo lỗi gõ máy tính đối với các từ khoá nhạy cảm. Ví dụ: *cưỡng d***, p.h.â.n x.*.c, kh.ửng kh.iếp, S.n.u.f.f, P!g...* Hình thức chữ viết này làm gián đoạn quá trình quét từ khoá của hệ thống kiểm duyệt.

4.1.2. Biến đổi về tên gọi

Chiến lược biến đổi tên gọi được sử dụng 11 lần trong 04/10 video, được hiểu là thay thế thuật ngữ bị hạn chế/cấm thành các từ có ý nghĩa tương tự về mặt văn hoá hoặc ngữ cảnh, đảm bảo truyền đạt ý nghĩa một cách gián tiếp. Đây là một trong những phương pháp hiệu quả nhất để tránh kiểm duyệt, do thuật toán kiểm duyệt nội dung thường dựa vào danh sách từ khoá cố định.


Ví dụ các video sử dụng thuật ngữ *tách Lego* thay vì *phân xác*. Hình ảnh trò chơi Lego gợi lên ý tưởng về sự tháo rời và sắp xếp lại, tạo ra một lớp nghĩa mới ít mang tính bạo lực hơn nhưng vẫn có thể được người xem hiểu rõ. Tương tự, từ *chết* thường được thay bằng các cụm từ như *đăng xuất, offline* - đều là những ẩn dụ tri nhận về thế giới mạng, nơi người dùng kĩ thuật số rời khỏi hệ thống hoặc không còn hoạt động.

Ngoài ra, trong nhiều trường hợp, người dùng TikTok còn trộn mã và chiến lược đa ngôn ngữ khi kết hợp giữa tiếng Việt và tiếng Anh hoặc sử dụng số và chữ cái xen kẽ để tránh bị thuật toán nhận diện. Ví dụ, *4 hình* có thể thay thế cho *tử hình*, hoặc *moi liver* được sử dụng thay vì *moi gan*. Việc trộn mã ngôn ngữ này không chỉ làm giảm khả năng bị phát hiện mà còn tạo ra một phong cách giao tiếp đặc trưng trong cộng đồng mạng, vốn đa ngôn ngữ và đa văn hoá.

4.1.3. Biến đổi về mặt phương thức biểu đạt

Ngoài việc sửa đổi và sáng tạo với hệ thống ngôn ngữ, người dùng TikTok còn tận dụng các yếu tố đa phương thức để truyền tải nội dung mà không cần dùng từ ngữ trực tiếp. Các phương thức này bao gồm sử dụng biểu tượng cảm xúc (*emoji*), hiệu ứng âm thanh (*tắt âm, tiếng bíp, bóp méo âm*

thanh), hành động mô phỏng (*khấu hình, cử chỉ*), hoặc kết hợp hình ảnh minh họa thay vì mô tả bằng lời. Các hình thức này xuất hiện 30 lần (chiếm 36,1% tổng số lần sử dụng Algospeak) trong 07/10 video.

Ví dụ, thay vì sử dụng từ *trẻ em*, người dùng có thể thay thế bằng biểu tượng Trẻ . Tương tự, trong một số video, khi đề cập đến các vấn đề nhạy cảm, người sáng tạo nội dung có thể giả vờ phát âm từ đó nhưng không phát ra âm thanh, dùng cử chỉ, khiến thuật toán không thể nhận diện từ khoá bị cấm nhưng người xem vẫn có thể hiểu thông điệp qua khẩu hình miệng.

Ngoài ra, một số video TikTok còn sử dụng hiệu ứng âm thanh thay thế, chẳng hạn như tiếng bíp hoặc âm thanh gây nhiễu để thay thế cho các từ ngữ bị cấm. Cách làm này không chỉ giúp nội dung vượt qua kiểm duyệt mà còn tạo ra hiệu ứng hài hước hoặc mang tính biểu tượng, làm cho thông điệp trở nên sinh động hơn.

4.1.4. Nhận xét

Những phát hiện trên cho thấy Algospeak không chỉ là việc thay thế ngôn ngữ mang tính tự phát, mà là một tập hợp chiến lược ngôn ngữ mang tính hệ thống, sáng tạo và chủ động nhằm phản ứng với kiểm duyệt tự động. Các dạng biến đổi linh hoạt như biến âm, lỗi chính tả có chủ đích, thay đổi tên gọi và biểu đạt đa phương thức cho thấy người dùng không đơn thuần tránh né mà đang tích cực tái cấu trúc ngữ nghĩa để giữ lại nội dung cốt lõi trong môi trường hạn chế.

Phân tích cũng cho thấy sự kết hợp giữa các cấp độ biểu đạt - từ ngôn ngữ viết, nói, đến icon, hiệu ứng âm thanh - phù hợp với quan niệm về diễn ngôn đa thức Kress và van Leeuwen, nơi quan niệm về ngôn ngữ là phương tiện độc tôn để biểu diễn và giao tiếp là không còn khả thi nữa (xem thêm [5, tr. 111]). Ngoài ra, hiện tượng sử dụng từ thay thế như *đăng xuất* cho *chết* hay *tách Lego* cho *phân xác* là ví dụ cụ thể của ẩn dụ tri nhận trong môi trường số, mở rộng cách người dùng gán nghĩa dựa trên vốn văn hoá mạng và khả năng suy diễn của cộng đồng.

4.2. Ý nghĩa thực hành diễn ngôn của hiện tượng Algospeak

Algospeak như một hành vi ứng phó và thích nghi của người dùng đối với những tác động của thuật toán. Hành vi này đã được Leyla Dogruel đề xuất đối với “cá nhân có thể áp dụng các chiến lược cho phép họ sửa đổi các thiết lập được xác định trước trong các môi trường được quản lý theo thuật toán... để thay đổi đầu ra của thuật toán” [7, tr.77]. Điều này liên quan kiến thức và năng lực sử dụng các phương tiện truyền thông của người dùng. Đối với trường hợp khảo sát của bài viết, người dùng đã có những hiểu biết nhất định về quy tắc thuật toán, dẫn đến việc chuyển mã, chuyển đổi từ ngữ, hành động nhằm đánh lừa thuật toán. Nội dung về bạo lực chắc chắn sẽ bị ngăn chặn hoặc loại bỏ, trong khi đó Algospeak giúp loại trừ nguy cơ này, tức là làm “thay đổi đầu ra của thuật toán”.

Algospeak mở rộng diễn ngôn khỏi phạm vi giao tiếp dựa trên ngôn ngữ, tích hợp với các yếu tố đa phương thức như hình ảnh, biểu tượng cảm xúc, âm thanh... Sự chuyển dịch này phù hợp với quan điểm của Kress và van Leeuwen về diễn ngôn đa thức, khi nhấn mạnh đa thức trở thành trung tâm trong giao tiếp. Algospeak không chỉ là sự thay thế từ ngữ mà còn là sự đồng bộ hoá các nguồn lực biểu đạt (ngôn ngữ, hình ảnh, âm thanh) để tạo ra một hệ thống ý nghĩa phức tạp hơn. Chẳng hạn, việc sử dụng biểu tượng che giấu từ ngữ nhạy cảm cho thấy người dùng tận dụng các tài nguyên kỹ thuật số để tăng cường hiệu quả giao tiếp, đồng thời giảm thiểu nguy cơ bị kiểm duyệt. Bởi đặc trưng đa thức, Algospeak còn xuất hiện dưới nhiều hình thức đa dạng và phức tạp khác như Internet Meme, tệp Gif...

Bên cạnh đó, hiện tượng Algospeak cũng đặt ra câu hỏi quan trọng về mối quan hệ giữa kiểm duyệt thuật toán và ngôn ngữ. Như Steen, Yurechko và Klug đã chỉ ra, hành động kiểm duyệt của TikTok có thể “cản trở nhu cầu giao tiếp và quyền tự do ngôn luận của người dùng” [6, tr.12] thì Algospeak xuất hiện như một phản ứng sáng tạo trước sự kiểm soát này. Khi người dùng buộc phải

thay đổi ngôn ngữ để tránh bị hạn chế giao tiếp, điều này có thể làm ảnh hưởng đến tính rõ ràng và nghiêm túc của thông tin, hàng loạt các từ ngữ không chính thức, phi chuẩn xuất hiện.

4.3. Những thách thức mới về nghiên cứu ngôn ngữ trong không gian của thuật toán

Sự xuất hiện của Algospeak đặt ra nhiều thách thức đối với nghiên cứu ngôn ngữ trong không gian số. Một trong những vấn đề lớn nhất là tính linh hoạt và thay đổi nhanh chóng của ngôn ngữ. Không giống như ngôn ngữ truyền thống, hệ thống từ vựng của Algospeak không cố định mà luôn biến đổi để thích ứng với thuật toán kiểm duyệt. Điều này khiến việc thu thập và phân loại dữ liệu về Algospeak trở nên khó khăn, vì một thuật ngữ có thể bị thay thế chỉ trong một thời gian ngắn.

Đồng thời, người tiếp cận cũng gặp không ít thách thức về khả năng “đọc - hiểu” kỹ thuật số, bởi thông tin đã nằm ngoài phạm vi của bảng chữ cái thông thường. Khi một từ hoặc cụm từ có thể có nhiều biến thể khác nhau, việc giải mã thông tin trở nên phức tạp hơn, đặc biệt đối với những người không quen thuộc với hệ thống mã hoá này. Điều này không chỉ ảnh hưởng đến quá trình truyền thông trên mạng xã hội mà còn đặt ra vấn đề đối với các nhà nghiên cứu ngôn ngữ học trong việc xác định các khuôn mẫu diễn ngôn trong không gian số.

5. Kết luận

Hiện tượng Algospeak - như chính tên gọi của nó, đã làm nổi bật mối quan hệ phức tạp giữa ngôn ngữ và công nghệ. Tại các nền tảng phổ biến như Tiktok, nơi bị kiểm duyệt ngôn từ nghiêm ngặt và giao tiếp bị áp đặt bởi thuật toán, người dùng đã phát triển các chiến lược sáng tạo để điều hướng ngôn ngữ, mở rộng cách thức truyền tải ý nghĩa. Thông qua biến đổi ngữ âm, lỗi chính tả, tên gọi và cả các yếu tố đa phương thức, Algospeak minh hoạ cho bản chất năng động và thích ứng của ngôn ngữ trước văn hoá, xã hội và nay là thuật toán.

Bài viết nhấn mạnh tầm quan trọng của việc hiểu Algospeak như một hiện tượng ngôn ngữ kỹ thuật số và diễn ngôn đa phương thức. Nó phản ánh những bước dịch chuyển quan trọng cũng như thách thức về chuẩn hoá ngôn ngữ trong tương lai. Các nghiên cứu tiếp theo có thể khám phá các dạng Algospeak khác, giúp mở ra những hiểu biết sâu sắc hơn về cách công nghệ định hình ngôn ngữ và giao tiếp trong bối cảnh kỹ thuật số. Quan trọng hơn, đây là hướng đi tất yếu trong một thời đại ngày càng bị công nghệ chi phối.

TÀI LIỆU THAM KHẢO

1. Fry, H. *Hello world. Being human in the age of algorithms*. W. W. Norton & Company, New York & London. 2018.
2. Iedema, R. *Multimodality, resemiotization: extending the analysis of discourse as multi-semiotic practice*. *Visual Communication*, 2(1), pp. 29-57. 2003.
3. Jewitt, C. *Different approaches to multimodality*. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (2nd ed., pp. 28-39). Routledge. New York, NY. 2009.
4. Klug, D., Steen, E. and Yurechko, K. *How algorithm awareness impacts algospeak use on TikTok*. WWW '23: The ACM Web Conference 2023, pp: 1-4. 2023.
5. Kress, G., & Van Leeuwen, T. *Multimodal discourse: The mode and media of contemporary communication*. Arnold. London. 2001.
6. Steen, E., Yurechko, K. and Klug, D. *You can (not) say what you want: using algospeak to contest and evade algorithmic content moderation on TikTok*. *Social Media + Society*, 9(3), pp. 1-17. 2023.
7. Taddicken, M., and Schumann, C. (Eds.). *Algorithms and communication*. Digital Communication Research, Vol. 9. Berlin. 2021.
8. Tiktok, link: <https://www.tiktok.com>