

Os Imageboards e os seus Recursos Compartilhados: Um Estudo de Caso em Síntese de Autoria Forense

Luiz H. Valle-Nunes

Universidade do Porto, Portugal

https://doi.org/10.21747/21833745/lanlaw/7_1_2a11

Abstract. *The development of the internet and new information and communication technologies has led to new enunciative practices which occur via emerging textual genres, such as forum threads. These new communicative habits, especially mob behaviours, combined with an anonymity that can be real or just felt may result in different cybercriminal activities, including those related to language. Forensic Linguistics has focused on such practices through studies in forensic authorship analysis with the aim of authorship synthesis, either based on individuals or communities, such as that of imageboards, collaborating in tasks associated with undercover investigations. The current analysis is a case study in forensic authorship synthesis applied to imageboards in Brazilian Portuguese, focusing on the category of resources shared by the community of practice, which is part of the resources and restrictions model presented by Grant and MacLeod (2018, 2020a). The corpus consists of 20 forum threads collected between October and November 2019, with a total of 306 posts and 24,587 tokens. The approach results in a holistic description of linguistic markers found at the microstructural level. It is based on a study of the n-grams obtained through the ‘Corpógrafo’ corpora management tool and a manual analysis of a sub corpus comprised of 10 threads, which point out to phenomena such as code-switching and othering manifestations.*

Keywords: *Undercover investigations, forum threads, code-switching.*

Resumo. *O avanço da internet e de novas tecnologias de informação e comunicação tem ocasionado o surgimento de novas práticas enunciativas, realizadas por meio de gêneros textuais emergentes, como os fios de discussão. Estes novos comportamentos comunicativos, dentre eles os mobilizadores, aliados a um anonimato que pode ser real ou apenas sentido, resultam em diferentes crimes cibernéticos, dentre eles os relacionados à linguagem. A Linguística Forense tem se debruçado sobre tais práticas por meio de estudos em análise de autoria forense com o intuito de síntese de autoria, seja de indivíduos ou seja de comunidades, tais como a dos imageboards, colaborando em tarefas periciais associadas a investigações*

infiltradas. A presente análise trata-se de um estudo de caso em síntese de autoria forense acerca de imageboards em português brasileiro, utilizando a categoria de recursos compartilhados em comunidades de prática, presente no modelo de recursos e restrições proposto por Grant and MacLeod (2018, 2020a). Foram coletados 20 fios de discussão entre outubro e novembro de 2019, compondo um corpus de 306 publicações e 24.587 átomos. A abordagem resulta numa descrição holística de marcadores a nível microestrutural, localizados a partir de um estudo de n-gramas obtidos por meio da ferramenta de gestão de corpora “Corpógrafo” e de uma análise manual de um subcorpus de 10 fios, de modo a evidenciar fenômenos como a alternância de códigos e a alterização.

Palavras-chave: *Investigações infiltradas, fios de discussão, alternância de códigos.*

Introdução

O desenvolvimento de novas tecnologias digitais, aliado à democratização do acesso à internet, possibilitou a criação de plataformas digitais cada vez mais multimodais e focadas na coletividade, permitindo o surgimento de práticas enunciativas (Fiorin, 2008: 4) e gêneros textuais emergentes altamente dinâmicos (Marcuschi, 2005: 13), tais como os tweets, os reels, os fios (*threads*), entre outros. Esta evolução trouxe grandes mudanças para a Comunicação Mediada por Computador (CMC) por contribuir para uma nova relação interativa entre os usuários e as redes. Isto se reflete, num sentido amplo, em diferentes ferramentas e estratégias adotadas pelos usuários para interagirem entre si e em suas percepções de pertença a comunidades de prática (Bucholtz and Hall, 2005), o que influencia as suas projeções enquanto *personas* digitais (Rashid *et al.*, 2013). Associado a tais mudanças, o advento de um efeito de despersonalização (KhosraviNik and Esposito, 2018: 49) teve como resultado a materialização de discursos que potencialmente não se realizariam em interações pessoais (Sousa-Silva, 2018: 119). Com efeito, ocorreu uma escalada na disseminação dos crimes cibernéticos relacionados à linguagem nos mais diferentes espaços da rede, dentre eles os discursos de ódio contra minorias, bem como comportamentos mobilizadores (Bernstein *et al.*, 2011; Fontanella, 2010) de usuários que compartilham de ideologias radicais.

Espaços que começaram a receber grande atenção a partir da década de 2010 são os fóruns de discussão online, que não só utilizam a sensação de anonimato proporcionada pelas plataformas, por conta da possibilidade de criação de contas sem a necessidade da inserção de informações reais dos utilizadores, como também possibilitam o recurso a ferramentas de criação e edição de conteúdo (Lisecki, 2013), que podem aumentar a efemeridade da comunicação materializada (Bernstein *et al.*, 2011), o que dificulta as investigações acerca das publicações dos utilizadores. Para além deste anonimato sentido, outro desafio está relacionado à criação de ferramentas de encriptação da rede, que propiciam um anonimato real por proteger os dados do usuário, de modo a permitir uma comunicação não rastreável em diferentes níveis, fazendo com que a linguagem se torne uma peça-chave para se identificar autores de crimes virtuais e se obter informações sobre as atividades de grupos organizados.

No âmbito dos fóruns de discussão que cultivam o anonimato destacam-se os *imageboards*, que são espaços inerentemente hipertextuais (Xavier, 2002: 29-33) e colaborativos (Koch, 2007: 35) a ponto de constituírem uma cultura online própria (Thibault, 2015). Popularmente conhecidos como *chans*, trata-se de ambientes nos quais o anoni-

mato é requerido e em que diferentes estratégias para a sua manutenção são empregadas (Bergman, 2001: 1), desde a possibilidade do uso de redes privadas virtuais (VPNs) e do roteamento cebola (Okazaki *et al.*, 2015: 371), por meio de navegadores anônimos como o Tor, até a codificação da própria linguagem, que opera por meio de informações e códigos compartilhados, estabelecidos por meio do exercício do poder ideológico (Van Dijk, 2006) dos membros mais integrados à comunidade. Este tipo de fórum já foi alvo de diferentes investigações, como em trabalhos que discutem o compartilhamento de informações (Manivannan, 2012; Richoux, 2016) e a articulação de grupos radicalizados, seja no que diz respeito aos discursos de ódio (Manivannan, 2013; Ruocco, 2020), ou seja na prática do ciberterrorismo (Weimann, 2016) e de ataques online orquestrados (Fontanella, 2010), entre outros. No entanto, trata-se ainda de um objeto de estudo novo que não recebeu grande atenção de trabalhos em português, principalmente numa perspectiva da Linguística Forense, que pode contribuir para que se tenha uma visão mais global acerca da linguagem em fóruns, de modo a identificar, a partir de tarefas periciais, estratégias de codificação importantes para o auxílio a outros profissionais das Ciências Forenses na solução de crimes cibernéticos, além de se possibilitar eventuais atividades de monitoramento.

No quadro da linguagem como prova, a Linguística Forense tem se dedicado a problematizar as questões relativas à identidade linguística (Coulthard and Johnson, 2010), discutindo-a a partir de noções importantes como a do idioleto (Coulthard, 2004), a do estilo idioletal (Turell, 2010) e a da performance linguística (Grant and MacLeod, 2018), com o intuito de se aprimorar o entendimento acerca das recorrências linguísticas e assim avançar na eficácia das tarefas periciais, sejam manuais ou seja também na implementação de softwares para as análises automatizadas sobre grandes quantidades de dados. Desde os anos 2000, não só as discussões se estenderam como também novas tarefas periciais relacionadas aos gêneros textuais emergentes começaram a ser requisitadas aos linguistas forenses, como é o caso da síntese de autoria forense (Grant and MacLeod, 2020b), um desdobramento da análise de autoria forense.

A análise de autoria forense prevê a identificação de marcadores linguísticos recorrentes materializados na comunicação de um ou mais determinados autores (Grant, 2008; Sousa-Silva and Coulthard, 2016). Já a síntese de autoria forense ocorre quando um linguista perito, a partir das ocorrências consistentemente localizadas, constitui um modelo de *persona* linguística a ser assumido por um investigador infiltrado, que poderá simular a autoria de uma vítima ou de um membro de um grupo organizado para que seja possível se obter informações acerca do caso investigado e possibilitar o seu avanço (Grant and MacLeod, 2020a). Para isto, o eventual infiltrado não só precisará compreender com profundidade a performance de uma *persona* linguística (Grant and MacLeod, 2020b: 82), como evitar ao máximo o vazamento de informações que revelem a sua própria identidade por meio da materialidade linguística, o que pode comprometer a sua investigação.

As primeiras experiências de aplicação da síntese da autoria forense foram testadas num programa de treinamento para investigadores infiltrados intitulado *Pilgrim* no Reino Unido, ao longo da década de 2010 (MacLeod and Grant, 2017: 159). Com a participação de linguistas forenses, as tarefas periciais tiveram como intuito adquirir inteligência relativamente as atividades de aliciamento de menores para fins sexuais, a partir de interações entre investigadores infiltrados e potenciais pedófilos em salas de bate-papo online (Grant and MacLeod, 2020b: 90). Partindo desta experiência, Grant and MacLeod

(2018, 2020a) elaboraram uma abordagem para as *personas* linguísticas que chamaram de *modelo de recursos e restrições*, que parte de domínios propostos pela Análise do Discurso Mediado por Computador (ADMC), como o estrutural, o do significado, o interacional e o sociocomportamental (Herring, 2004: 360), utilizando-os como base para uma taxonomia de marcadores linguísticos relevantes para a síntese de autoria (MacLeod and Grant, 2012).

O modelo de recursos e restrições leva em conta o fato de as performances linguísticas estarem inerentemente associadas a performances de identidades, dotadas de processos de extração de recursos para a interação que podem conter traços mais estáveis ou mais dinâmicos (Grant and MacLeod, 2018: 92). Segundo os autores, os recursos para a interação estão disponíveis em diferentes níveis, nomeadamente o do histórico sociolinguístico do autor, de sua fisicalidade, do contexto interacional e da relação do falante com a audiência (Grant and MacLeod, 2018: 93-94). Estes níveis podem atuar como categorias de marcadores linguísticos com maior ou menor produtividade a depender dos contextos, dos gêneros textuais e da persona linguística que o autor pretende assumir.

O presente estudo leva em consideração a quarta categoria proposta no modelo de recursos e restrições, nomeadamente a das informações compartilhadas pelos membros da comunidade de prática, com o intuito de se identificar marcadores persistentes e relevantes para a síntese de autoria forense de *imageboards* em português brasileiro. Com isto, visa-se obter uma visão geral acerca das características compartilhadas de autoria neste ciberespaço, alargando o entendimento sobre as variedades linguísticas da rede em português, principalmente no que diz respeito as suas relações com a integração e o sentimento de pertença em comunidades de prática associadas ao cibercrime. Ao se localizarem as características partilhadas pelo grupo torna-se mais eficaz a prevenção de potenciais vazamentos identitários dos investigadores infiltrados, principalmente os ocasionados por conta de marcadores a nível estrutural, que são tipicamente os mais detectáveis por parte dos membros do grupo, por revelarem o nível de integração do locutor. Para além disto, a identificação de recorrências na materialidade linguística das publicações possibilita a criação de bancos de dados acerca deste tipo de comunidade online para que, por meio de ferramentas de *software*, seja possível implementar diferentes tipos de sistemas para o auxílio de investigações infiltradas, como glossários, bem como maneiras de se gerenciar diferentes perfis sintetizados dos fóruns e assim realizar comparações entre eles.

Com estes objetivos em consideração, nas secções seguintes serão apresentados os dados e a metodologia de análise escolhida para este estudo de caso em síntese de autoria forense.

Dados

Foram coletados 20 fios de discussão (*threads*) compostos por 306 postagens, com um total de 24.587 átomos contabilizados pela ferramenta de gestão de corpora “Corpógrafo”, provenientes de categorias do *imageboard* brasileiro 55chan. As categorias escolhidas para a coleta foram /b/, /pol/, /escoria/, /lit/, /an/ e /esp/, que correspondem respectivamente às temáticas: aleatoriedade, política, questões politicamente incorretas, literatura, cultura japonesa e esportes. As amostras foram coletadas entre outubro e novembro de 2019, levando em consideração as respostas síncronas e assíncronas relativamente a publicação iniciadora.

Método

Em primeiro lugar, vale ressaltar que a abordagem adotada para a obtenção, o armazenamento e o subsequente tratamento das amostras coletadas, teve como premissa o princípio da confidencialidade, de modo a salvaguardar a privacidade dos dados de usuário e o anonimato das publicações, sem que para isto fosse perdida a integridade do material linguístico. Para este efeito, uma vez realizadas as capturas de tela de cada publicação, foram removidos os seus elementos paratextuais, como a identificação (ID), os números de hiperligação e a localização temporal. No entanto, o corpo das publicações foi mantido de maneira integral, com o espaçamento, a tabulação, as travas maiúsculas/minúsculas e a acentuação originais. Uma vez que isto pode se traduzir em questões como a falta de espaço entre os caracteres e até mesmo entre frases, a medição escolhida para o volume de dados se deu em átomos e não em palavras. Responsável pela contagem dos átomos, a ferramenta selecionada para a criação e a gestão do corpus foi o “Corpógrafo” (Sarmiento *et al.*, 2004).

Após a coleta, dividiu-se os 20 fios em dois subcorpora a partir de suas temáticas, com um recorte de 216 publicações compostas por 12.945 átomos para o subcorpus intitulado “A” e 102 publicações com 11.642 átomos para o subcorpus “B”. Ambas as partes foram inseridas no Corpógrafo para a componente quantitativa da análise, que se baseou num estudo de uni, bi e trigramas, com o intuito de se localizar escolhas linguísticas recorrentes a nível microestrutural, como alterações morfossintáticas ou da ordem das frases, além de se quantificar ocorrências com nomes e sintagmas nominais. Além da ferramenta automatizada, para uma segunda análise, manual e de natureza qualitativa, optou-se pelo subcorpus A, no qual se inserem as categorias /b/, /pol/ e /escoria/ e localizam-se os textos cujas questões iniciadoras e as respostas dos usuários estão relacionados aos discursos de ódio, que são tipicamente o foco das investigações infiltradas.

Realizada a análise de autoria forense do grupo em questão, optou-se por alocar as ocorrências mais relevantes para a tarefa pericial de síntese nas quatro categorias de recursos e restrições propostas por Grant and MacLeod (2018: 87-88), nomeadamente a do histórico sociolinguístico, a da fisicalidade, a do contexto da interação e a das informações compartilhadas para a integração da comunidade de prática. Uma vez que o escopo deste estudo diz respeito às informações partilhadas entre os membros da comunidade, o que corresponde unicamente à quarta categoria, recorreu-se maioritariamente a trabalhos que abordam o Discurso Mediado por Computador como o de Androutsopoulos (2007) e de Danet and Herring (2007), entre outros; que se debruçam sobre a comunicação multilíngue como o de Dor (2004), Mozzillo (2009), Seargeant and Tagg (2011), entre outros; que tratam dos discursos de ódio e fenômenos de alterização como o de Schneider (2004), Van Dijk (1984, 2006, 2018) e Wodak (2001); que discutem hegemonias culturais em ciberculturas como Ging (2019) e Salter (2018); além de trabalhos sobre o sistema linguístico do português como é o caso dos trabalhos de Alves (2007) e Villalva (1994).

Assim, nas secções a seguir, serão apresentados os resultados obtidos para a quarta categoria do modelo de recursos e restrições, de modo a apontar para as hegemonias ideológicas presentes nos *imageboards* em português brasileiro.

Resultados

A alternância de códigos

Fatores socioculturais e políticos têm contribuído para a hegemonia de algumas línguas para a comunicação online em detrimento de outras (Danet and Herring, 2007: 17). Este é o caso do inglês, que se estabeleceu como língua franca e passou a influenciar as escolhas linguísticas de usuários nativos de línguas como o português, fazendo parte das características das *personas* digitais criadas nas redes sociais e servindo como pilar para a identificação e a integração de comunidades de práticas de natureza multilíngue (Seargeant *et al.*, 2012: 528-529). Como resultado do processo de anglicização (Dor, 2004), os usuários passaram a alternar segmentos de sua língua materna com segmentos da segunda língua (L2), tanto em comunicações síncronas quanto assíncronas (Androutsopoulos, 2007: 357-359), num fenômeno apontado como alternância de códigos (Seargeant and Tagg, 2011). O processo de alternância influencia diretamente o domínio da estrutura, que já havia sido apontado por MacLeod and Grant (2017: 167-168) como o fator que mais causa o vazamento identitário no caso da síntese de autoria forense. Tendo isto em consideração, daremos atenção à maneira como este fenômeno materializa-se no corpus analisado em diferentes níveis, com base na proposta de Mozzillo (2009), que destaca uma distinção entre ocorrências de natureza intra e intersegmental, além das materializações por meio de acrônimos.

Em relação às ocorrências de natureza intrassegmental, cabe destacar que elas tipicamente se materializam quando elementos internos de um enunciado são influenciados a nível morfossintático de modo a gerar mudanças que têm como base uma segunda língua. Como aponta Mozzillo (2009: 189), este tipo de interferência pode afetar desde a raiz lexical de nomes até o nível da ordem da frase, sendo os tipos mais comuns a inserção e estratégias de sufixação e composição, bem como o recurso a traduções literais. Outros trabalhos como o de Danet and Herring (2007) apontaram o fenômeno da alternância intrassegmental como prototípico da formação identitária de comunidades, uma vez que pressupõem conhecimentos compartilhados das circunstâncias por meio das quais as trocas podem ocorrer de modo a serem aceitas e interpretadas pela audiência.

Ao se analisar o corpus, verificou-se a presença dos três tipos de materializações apontadas por Mozzillo (2009), sendo que as inserções e as traduções literais foram integradas a categoria das realizações intrassegmentais. Com isto, para que fosse possível se obter uma visão geral acerca desta categoria, considerou-se o seguinte esquema das ocorrências encontradas de acordo com a sua tipologia:

Com base nos fios de discussão coletados, como mostra a figura 1, o que se localizou foi uma preferência por realizações intrassegmentais, com um total absoluto de 71 ocorrências, face a apenas 2 intersegmentais. Os acrônimos, com 14 ocorrências localizadas, foram identificados sobretudo na referência entre os usuários integrados a comunidade da prática, com materializações como “OP” (*Original Poster*) para o usuário iniciador do fio e NEET (*Not in Education, Employment or Training*) para os indivíduos com poucas relações sociais fora das redes.

Dada a sua presença maioritária, procedeu-se a uma análise mais pormenorizada das ocorrências intrassegmentais, levando em consideração três categorias, nomeadamente a das inserções diretas, das criações lexicais ou neologismos, e das traduções literais do inglês. A distribuição destas categorias foi organizada em forma de gráfico para a sua melhor visualização:

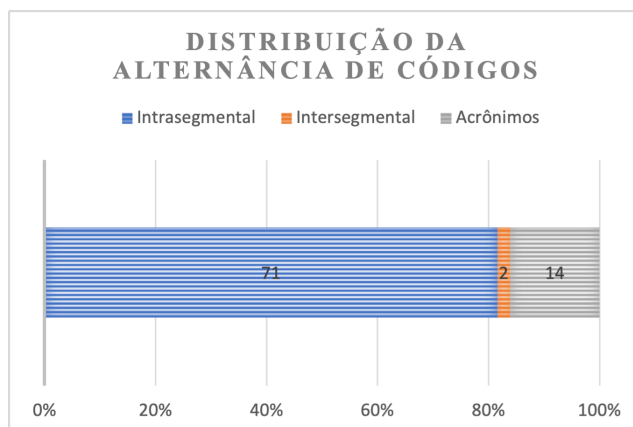


Figura 1. Distribuição da alternância de códigos por meio de anglicização.

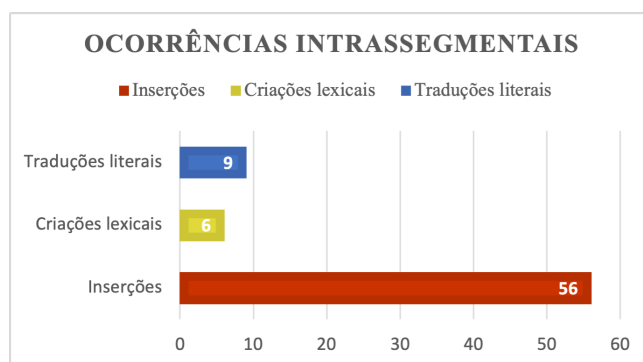


Figura 2. Distribuição das ocorrências intrasegmentais por tipo.

No caso do *imageboard* analisado, as amostras apontam para uma preferência pela utilização de inserções diretas de palavras de língua inglesa, com um valor absoluto de 56 ocorrências, em contraste às traduções literais e as criações lexicais, cujas ocorrências materializaram-se, respectivamente, em apenas 9 e 6 casos. Para uma melhor compreensão de cada tipo de realização da alternância intrasegmental, pode-se recorrer a exemplos de entradas do corpus como:

- (1) “Mas os desenhos do Liefeld têm tantos problemas técnicos que se for citar todos a **thread** entra em **bump limit** [...]” [TB3_R3].
- (2) “[...] logo **razoa por** termos puramente profanos, pois de fato, é a única coisa que você tem, sendo um endemoniado que é; ou talvez não, talvez você seja um **literal demônio**, nunca se sabe” [T9_R4].
- (3) “Não sou o **quotado**, mas ele está certo” [T11_R49].
- (4) “Nem estou nessa discussão mas tive vergonha alheia de você sendo **samefag, cri-ançafag**” [T11_R48].

No exemplo (1), a alternância intrassegmental ocorre por meio de inserções diretas de sintagmas nominais provenientes do inglês, os quais ocorrem de maneira integral, sem interferência na ordem da frase em português brasileiro. Já no exemplo (2), o que se verifica no segmento em destaque é uma tradução literal que influencia diretamente a ordem da frase, mais especificamente no argumento interno do sintagma verbal “ser literalmente um demônio”. Também no exemplo (2), há um outro tipo de anglicização por meio de tradução literal a alterar o sintagma verbal, sendo o foco da alternância o verbo “pensar”, que é realizado como “razoar por”, numa alusão a “*reason through*”. O exemplo seguinte (3) demonstra como a inserção pode sofrer derivações morfológicas do português (Alves, 2007), sendo que em “quotado” o processo de nominalização incide sobre o empréstimo “quote” no lugar de “citar”. Por fim, os processos de alteração morfológica podem ocorrer por meio da composição (Villalva, 1994) inglês+inglês, português+inglês ou ainda inglês+português, tal como no último exemplo (4), em que “*same*” e “criança” são adicionados a “*fag*”, um modo dos usuários se referirem uns aos outros como integrantes dos *boards*.

A questão das traduções literais foi apontada anteriormente por autores como Pairedes *et al.* (2016: 194), os quais ao se debruçarem sobre culturas da rede como a comunidade dos *gamers*, identificaram um nível de anglicização rapidamente assimilado e interpretado no âmbito do processamento da leitura pelos demais participantes da interação. A influência na ordem da frase é um tipo de alternância de códigos que aparece no corpus analisado principalmente nos sintagmas nominais, sendo o movimento sintático tipicamente realizado pelos complementos e não pelo núcleo, tal como nos exemplos:

(5) “GAME DEV / **DESENVOLVIMENTO DE JOGOS FIO** [...]” [TB1_OP].

(6) “[...] Apesar da história ser meio bobinha, é realmente uma obra de arte que deve ser apreciada por todos os **animesamigos** [...]” [TB4_R4].

Nos exemplos em destaque, os núcleos sintagmáticos “fio” e “amigos” mantêm-se a direita dos seus complementos, sob a influência da ordem em inglês para “*game dev thread*” e “*anime friends*”. Uma característica presente nos dois exemplos, (5) e (6), diz respeito a ausência da preposição “de” entre os núcleos e os complementos, embora o exemplo (5) ainda a preserve dentro de seu complemento, em “desenvolvimento [de] jogos”. Já no caso do exemplo (6), a ausência da preposição “de” pode relacionar-se ao fato de se tratar de uma composição, sendo que o movimento sintático escolhido relativamente a escolha da posição do complemento também apareceu anteriormente no exemplo (4) com “criança**fag**” como escolha linguística em detrimento de “*fagcriança*”. Finalmente, cabe dizer que a alteração da ordem de palavras, embora ocorra tipicamente na estrutura interna dos sintagmas nominais, não se materializa a nível dos sintagmas verbais e no caso de expressões idiomáticas:

(7) “[...] mas como eu já **havia dumpado** todos os meus jogos de Wii, não teve jeito. [...]” [TB2_R7].

(8) “[...] >literatura brasileira contemporânea **numa casca de noz**” [TB10_R4].

Ambos os exemplos (7) e (8) mantêm a ordem do português, sendo que no caso de (8) a expressão “numa casca de noz” como tradução de “in a nutshell” não apresentou no corpus nenhum tipo de ocorrência invertida. Uma hipótese pode estar relacionada a natureza semântica do sintagma “casca de noz” em relação, por exemplo, a uma ocorrência como “desenvolvimento de jogos fio”, uma vez que a primeira aponta para uma tipologia e segunda para um fim.

A alternância de códigos com ocorrências intrassegmentais, para além das traduções literais e de inserções no âmbito do predicado, ocorre também nos nomes escolhidos para a referenciação. As escolhas linguísticas localizadas dizem respeito tanto ao grupo de pertença quanto às pessoas que não participam do *imageboard*. Nas próximas secções, apresentam-se estas escolhas de modo a apontar para a hegemonia ideológica da comunidade.

A referenciação entre membros

As escolhas feitas pelos usuários para se referirem uns aos outros e a si próprios configura um aspecto importante relativamente à representação dos atores sociais envolvidos e aos níveis hierárquicos da comunidade. Os utilizadores atribuem diferentes valores àqueles que fazem parte do grupo de pertença face a quem não participa de suas práticas. Há uma relação direta entre os valores considerados positivos e o grupo de pertença, além de uma relação entre valores negativos e os Outros, tal como tipicamente ocorre quando há uma associação ideológica forte prevista para a integração (Wodak, 2001). Dar-se-á início pelos valores associados aos autores do grupo de pertença, de modo a explorar a lexicalização e a significação das formas materializadas.

Em primeiro lugar, verificou-se que as formas utilizadas para referenciação aos atores sociais, para além de apresentarem inserções do inglês e se referirem sempre aos homens, remetem para relações sociais que podem ser divididas entre aquelas associadas ao universo do trabalho, as interpessoais e a fisicalidade dos integrantes:

(i) **Fisicalidade:** “*chad*”;

(ii) **Interpessoais:** “mago”, “falho”, “escravoceta”;

(iii) **Laborais:** “*ex-NEET*”, “*wageslave*”.

Relativamente à fisicalidade, o uso de “*Chad*” para se referir aos homens considerados dentro do padrão de beleza da sociedade já havia sido foco de atenção de autores como Ging (2019) ao estudar o espectro de misoginia – a que chama de *Manosphere* – nos fóruns online. Ao analisar as relações entre a cultura *incel* e a performance da masculinidade nas redes, a autora concluiu que há um chamado “dilema da masculinidade alfa-beta”, em que, por um lado, os usuários enaltecem a figura do macho alfa, isto é, do homem dentro dos padrões normativos da sociedade, mantedor da hegemonia pa-

triarcal e dotado de valores tradicionais e conservadores; e, por outro lado, abraçam os aspectos autodepreciativos da figura do macho beta, que não recebe a mesma atenção e prestígio social, sendo marginalizado pelos seus valores e *hobbies* (Ging, 2019: 648-650). Esta masculinidade híbrida recorre a diferentes formas de referência para as características alfa, dentre elas “*chads*”, “*normies*” e “*fratboys*” (Ging, 2019: 650). Embora não tenham sido encontrados no corpus exemplos destes últimos dois nomes, as ocorrências com *chad* se encaixam no mesmo paradigma apresentado pela autora, apontando assim para a presença do dilema alfa-beta nesta comunidade:

(9) “O pior e **o que mais me irrita** nem é **o fato das depósitos escolherem chad** e chads, blá-blá-blá” [T10_R5].

(10) “**O chad que atenta contra a integridade física e emocional de um falho** é o mesmo que namora a depósito descolada que posta estórias no instagado disseminando correntes de amor ao próximo” [T10_R6].

(11) “**O estereótipo de chad valentão que agride o mais fraco** é mesmo só **um estereótipo falho e deturpado**” [T10_R6].

Há uma dualidade de valores associados a *Chad*, que é aceito pela sociedade e é bem sucedido com as mulheres (exemplo 9), facto que é visto como positivo pelos usuários por conta da manutenção da dominância masculina, mas que também pode receber valores negativos uma vez que este indivíduo “agride o mais fraco” (11), referido também como “falho” (10 e 11), evidenciando que a forma híbrida alfa-beta, um meio-termo, é de facto a mais almejada. É importante destacar que embora *Chad* pertença ao espectro relacionado a fisicalidade, por estar associado a uma cultura popularmente conhecida por *jock*, i.e., de indivíduos que frequentam o ginásio e dão atenção redobrada à sua aparência, “falho” não se encaixa na mesma categoria pelo facto que a fraqueza mencionada não está somente relacionada ao físico, mas às relações que os chamados machos betas têm no seu dia a dia com a sociedade.

Ao se verificar as ocorrências com “falho”, foi possível relacionar a própria noção de falha ao modo como estes indivíduos percebem a sua aceitação pelas pessoas de fora do grupo. Isto é evidenciado no fio identificado no corpus como [TB7], no qual um indivíduo expõe que quer deixar de ser falho, pedindo “ajuda para o seu desenvolvimento” [TB7_OP]. Aqui é utilizado o adjetivo “falho”, associado ao valor negativo de “largado”:

(12) “[...] quero dizer que essa minha versão **largada e falha** vai deixar de existir. Irei excluir minhas redes sociais, só ficarei entrando no chan para relatar o progresso e contribuir com coisas que aprendo nessa jornada [...]” [TB7_OP].

Inicialmente, embora o fio em questão seja dedicado a falha no sentido da fisicalidade, outros usuários começam a expor diferentes valores associados ao nome “falho” que dizem respeito ao intelecto e às interações sociais. As ocorrências a seguir expan-

dem o conceito de “falho” adotado pelos membros da comunidade de prática, além de apontar para uma vontade de se encontrar uma solução:

(13) “[...] Estudarei interações pessoais e oratória para melhorar essa forma ruim que tenho para me comunicar com as pessoas. Achei uns exercícios para os músculos da face na Internet, vou tentar fazer [...]” [TB7_OP].

Valor associado: dificuldades articulatórias (oratória);

(14) “[...] estou tentando me tornar uma pessoa mais interessante, por isso estou lendo mais [...]” [TB7_R1].

Valor: falta de arcabouço de conhecimentos;

(15) “[...] eu digo que vale a pena, estou me sentindo mais confiante para falar com as pessoas e estou só no começo” [TB7_R6].

Valor: dificuldade na comunicação.

Os exemplos apontam para discursos em que as “interações pessoais” a que os usuários se referem têm uma relação direta com “ser mais interessante” (14) e “ser mais confiante” (15), ou seja, são exemplos em que a falha está associada à capacidade de se relacionar com o mundo exterior e ser compreendido por ele. É pautado neste tipo de discurso que o fórum possui uma categoria intitulada “falha e aleatoriedade”, na qual são postados fios relacionados aos problemas da vida cotidiana, tipicamente relacionados ao facto destes indivíduos não se encaixarem em padrões sociais. Ao conhecimento de saber transitar pelas situações do mundo real, os usuários deram o nome de “magia”, o que explica o facto de se referirem uns aos outros como “magos”. O uso deste nome no corpus evidencia uma outra faceta da relação entre os usuários e o mundo exterior, nomeadamente a das relações laborais, como no exemplo:

(16) “O Brasil consegue ser o pior lugar para ser um **magô**, pois na maioria das vezes, se você é um inepto social, nem os empregos mais inóspitos sobram pra você” [T10_R5].

Valor: dificuldades nas interações laborais.

No exemplo (16), o que está em causa é o despreparo daquele que ainda não detém as habilidades interpessoais para o trabalho. É preciso lembrar, no entanto, da dualidade apontada por Ging (2019), já que este despreparo é ao mesmo tempo posto em causa e celebrado. Os usuários escrevem sobre o problema do desemprego, mas também o enaltecem por estarem longe da sociedade tóxica que acreditam existir. Para falarem sobre o mal do trabalho, por exemplo, os usuários recorrem a uma composição de valor negativo que foi importada dos fóruns em inglês, nomeadamente “*wageslave*”, como no exemplo:

(17) “Atualmente sou um *ex-NEET*, *wageslave* sem amigos que mora com os pais , mais alguns meses e eu planejo sair daqui e cortar completamente os laços com a minha família [...]” [T4_R7].

A temática dos utilizadores como incompreendidos pela sociedade reaparece quando o usuário se refere a “cortar os laços” (17) com a família. Esta dificuldade de integração, tanto do ponto de vista interpessoal quanto econômico, foi discutida por autores como Uchida and Norasakkunkit (2015: 2) ao estudarem o fenômeno dos “*Hikikomori*”, que se popularizou no ocidente sob o acrônimo NEET para “*Not in Education, Employment or Training*” (Smith and Wright, 2015: 402-403) e que pode ser associado à cultura *incel* como um culto ao isolamento. Ruocco (2020: 27) já havia afirmado que parece existir uma chamada “cultura NEET” em *imageboards* internacionais como o 4chan, em que a reclusão é vista como uma característica a ser reforçada. Tal cultura parece existir no *imageboard* brasileiro analisado, o que explica o porquê de *ex-NEET* estar associado ao uso de *wageslave*, como alguém que é explorado somente pelo dinheiro, pois odeia o trabalho. Outra hipótese levantada por Ruocco (2020: 27) é a de que esta cultura seja apenas de fachada, parte de uma performance que condiz apenas a *persona* digital destes indivíduos. Mesmo que este seja o caso, deve-se atentar ao fato que uma síntese de autoria forense é baseada na performance linguística materializada, já que o intuito é o de realizar uma performance específica ao contexto do fórum.

Obtidos os resultados acerca da referenciação entre membros, a próxima secção tem por enfoque a referenciação às pessoas que não fazem parte da comunidade de prática, ou seja, os Outros (Wodak, 2001). Apresentam-se, portanto, os apontamentos relativamente às estratégias de alterização (Schneider, 2004), sem deixar de lado a presença da alternância de códigos, persistente característica compartilhada entre as *personas* linguísticas.

A representação dos Outros

Uma das estratégias utilizadas por grupos de pertença para a manutenção de sua hegemonia e para o exercício de poder ideológico é a representação positiva dos membros e a negativa dos Outros, em discursos de persuasão que apontam para comportamentos potencialmente agressivos àqueles que são considerados antagonistas, tipicamente os grupos minoritários (Van Dijk, 1984: 4). A representação de si a partir de características consideradas negativas no Outro constitui um processo de criação identitária chamado de alterização (Schneider, 2004), no qual as práticas discursivas, materializadas por meio de uma gramática pautada no contraste e no conflito (Baumann and Gingrich, 2004), podem realizar marcadores linguísticos em diferentes níveis do texto, como na escolha linguística dos nomes e na manifestação de tópicos prototípicos (Wodak, 2001: 73), que podem compreender características relevantes à síntese de autoria, uma vez que focalizam as ideologias (Van Dijk, 2018) mais pungentes do grupo de pertença.

No que diz respeito às estratégias de alterização, destacam-se no corpus as representações das mulheres e dos negros, sobretudo nas escolhas de nomes pejorativos, que, junto aos adjetivos axiologicamente negativos e os deíticos pessoais, posicionam estes grupos como antagonísticos aos membros da comunidade. Exemplos que evidenciam este antagonismo podem ser localizados em sintagmas como:

- a) **Mulheres**: “uma mãe solteira”, “as mães solteiras”, “msol”, “vadia maluca”, “pitanga”, “depósitos”, “a depósito descolada”, “fêmeas”, “diabolher”, “merdalher”;
- b) **Negros**: “pedaços de carne ambulantes”, “precto”, “esses macacos”, “jovens negros suburbanos”, “animais humanos”.

Efetivamente, os sintagmas apontam para a materialização da identidade masculina excludente e opressora descrita por autores como Salter (2018: 256) e Ging (2019: 639) para os fóruns anônimos, na medida em que as ocorrências coisificam a mulher de forma misógina, reduzindo-a a um papel de total submissão e objetificação sexual. Para a manutenção deste tópico recorrem, por exemplo, a traduções literais do inglês, como é o caso de “a vadia maluca” (*crazy bitch*) e “a depósito” (*dumpster*). Há também casos como “as mães solteiras”, também materializado como “msol”, que adaptam do inglês o acrônimo “MILF” (*Mom I’d Like To Fuck*). Estas escolhas revelam uma interdiscursividade presente na organização ideológica do grupo, fundamental para a sedimentação de suas práticas sociais (Van Dijk, 2006: 117), resgatando discursos de ódio produzidos em outros *imageboards* internacionais, como o 4chan. No entanto, embora a alternância de códigos esteja presente aqui novamente, ela não impede a emergência de novas materializações específicas do português. Como exemplo, a representação do grupo “mulher” materializa-se também em composições português+português com valores negativos, em nomes como “diabolher” e “merdalher”.

O único contexto que aponta para uma diferenciação no que tange a referenciação às mulheres ocorreu com o nome “pitanga”, que se refere às companheiras dos utilizadores. Neste caso, há uma tentativa de colocá-las num patamar menos pejorativo que outras mulheres, mas ainda assim objetificado:

(18) “Quero muito arrumar **uma pitanga** , tentar ter uma vida normal [...]” [T7_R14].

No exemplo, o nome “pitanga” aparece associado a valores positivos, tais como o de “ter uma vida normal”, isto é, estar num relacionamento. Não foram encontradas no corpus ocorrências com os nomes “namorada”, “esposa”, “amante”, “ficante”, “companheira” ou “mina”. A questão da objetificação, no entanto, não se restringe às mulheres, mas afeta ainda mais pejorativamente os negros.

Os negros são referenciados de forma coisificada e animalizada, por meio de tópicos de sobrecarga (Wodak, 2001: 73), em ocorrências como “pedaços de carne” e “animais humanos”. Este uso, entretanto, não parece estar restrito apenas a locutores que se identificam como brancos, como pode ser visto no exemplo:

(19) “Eu me sinto trancado com esses macacos , estou perdendo minha sanidade” [T4_R9].

Em (19), o locutor se refere à própria família com o nome “macacos”, distanciando-se por meio do uso do deíctico. Este tipo de ocorrência já havia sido apontado por Ruocco (2020: 27), que salientou a utilização não esperada de determinados nomes por parte de locutores que são parte do próprio grupo alvo, dada a cultura violenta cultivada nos *imageboards*. Assim, ao se fazer suposições acerca do perfil sociolinguístico destes usuários,

é preciso se ter uma atenção redobrada à reprodução das ideologias da cibercultura, que pode se materializar até mesmo por parte da própria minoria a frequentar estes espaços.

A partir destas considerações, é possível dizer que no âmbito da integração da comunidade de prática, a alternância de códigos incide sobre a alterização, embora não seja o único tipo de recurso a que os usuários recorrem. Verifica-se também o uso de nomes que remetem a valores ideológicos possivelmente compartilhados nos *imageboards* em inglês, constituindo assim uma ligação interdiscursiva entre esses espaços, o que pode ser objetivo de futuros estudos comparativos. Com isto em mente, na próxima secção serão apresentadas algumas considerações finais, bem como outras aberturas de pesquisa numa perspectiva da Linguística Forense.

Considerações finais

O estudo analisou fios de discussão de fóruns do tipo *imageboard* em português brasileiro, com o intuito de se identificar marcadores recorrentes na performance linguística da comunidade de prática, mais especificamente no âmbito dos recursos compartilhados para a sua formação e integração, categoria presente no modelo de recursos e restrições (Grant and MacLeod, 2018). A análise visa chamar a atenção às características linguísticas compartilhadas, de modo a auxiliar na criação de uma inteligência para a mitigação do vazamento identitário a nível estrutural de policiais infiltrados durante tarefas periciais de síntese de autoria forense.

A partir da análise foi possível detectar a presença acentuada da alternância de códigos por meio de processos de anglicização, que ocorrem tipicamente de maneira intrassegmental. A alternância atua sobre o nível morfossintático, sendo os sintagmas nominais os mais afetados, embora tenha-se verificado alterações da ordem da frase em português, principalmente por conta de traduções literais. A anglicização também foi encontrada nas estratégias de referência intragrupo e relativamente aos Outros, nos processos de alterização, em que os valores associados aos nomes escolhidos são negativos e em que a hegemonia ideológica da comunidade se materializa em discursos de ódio.

Dada a perspectiva de uma visão geral adotada nesta pesquisa relativamente aos recursos a nível da integração da comunidade, trabalhos futuros podem debruçar-se sobre uma tipologia das ocorrências de anglicização, bem como apontar para outras línguas potencialmente utilizadas como base para a alternância de códigos no contexto dos *imageboards* brasileiros. Este tipo de identificação dá abertura a potenciais aplicações forenses computacionais tais como a criação de bancos de dados de referência em português para que investigadores infiltrados possam ter acesso mais facilmente a marcadores compartilhados por este tipo de grupo. Possibilita também a elaboração de glossários no que diz respeito aos neologismos e outras inovações de natureza morfossintática como as composições e os truncamentos mais comuns, para que a decodificação dos textos escritos seja mais eficaz. Investigações mais exaustivas acerca deste fenômeno podem resultar, por exemplo, num glossário de valores associados aos nomes, a ser implementado em *softwares*, de modo a permitir que outros profissionais das Ciências Forenses possam aceder mais rapidamente a buscas específicas acerca do léxico do texto em causa durante as investigações infiltradas. A partir da criação de um banco de dados, investigações sobre outros fóruns nesta perspectiva possibilitariam também, por um lado, a realização de análises comparativas relativamente aos marcadores compartilhados nas redes e, por

outro lado, a realização de análises computacionais supervisionadas em aprendizagem automática. Todas estas possibilidades poderão ser utilizadas não só para a resolução de crimes cibernéticos específicos, mas também para o potencial recurso a ferramentas de monitoramento destes ciberespaços.

Nota

Este artigo baseia-se na dissertação de Mestrado em Linguística, aprovada pela Faculdade de Letras da Universidade do Porto e publicada no Repositório Aberto: Valle-Nunes, L.H. (2020). *O cibercrime e as estratégias linguísticas do anonimato: A síntese de autoria forense aplicada aos imageboards em português brasileiro*. Dissertação de Mestrado. Porto: Faculdade de Letras da Universidade do Porto. Disponível em: <https://hdl.handle.net/10216/129313>. Acesso: 16/01/2021.

Referências

- Alves, I. (2007). *Neologismo: criação lexical*. Ática: São Paulo, 3 ed.
- Androutopoulos, J. (2007). Language Choice and Code Switching in German-Based Diasporic Web Forums. In B. Danet and S. Herring, Eds., *The Multilingual Internet: Language, Culture, and Communication Online*. New York: Oxford University Press.
- Baumann, G. and Gingrich, A. (2004). *Grammars of Identity/Alterity - A Structural Approach*. New York: Bergham Books.
- Bergman, M. K. (2001). White Paper: The Deep Web: Surfacing Hidden Value. *The Journal of Electronic Publishing*, 7(1).
- Bernstein, M., Monroy-Hernández, A., Harry, D., André, P., Panovich, K. and Vargas, G. (2011). 4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community. In *Proceedings of the Fifth International AAAI Conference on Web and Social Media, North America*.
- Bucholtz, M. and Hall, K. (2005). Identity and interaction: a sociocultural linguistic approach. *Discourse Studies*, 7(4-5), 585–614.
- Coulthard, M. (2004). Author Identification, Idiolect and Linguistic Uniqueness. *Applied Linguistics*, 25(4), 431–447.
- Coulthard, M. and Johnson, A. (2010). *The Routledge Handbook of Forensic Linguistics*.
- Danet, B. and Herring, S. (2007). *The Multilingual Internet: Language, Culture, and Communication Online*. New York: Oxford University Press.
- Dor, D. (2004). From Englishization to Imposed Multilingualism: Globalization, the Internet, and the Political Economy of the Linguistic Code. *Public Culture*, 16(1), 97–118.
- Fiorin, J. (2008). A Internet vai acabar com a língua portuguesa? *Texto Livre: Linguagem e Tecnologia*, 1(1), 2–9.
- Fontanella, F. (2010). Nós somos Anonymous: anonimato, trolls e a subcultura dos imageboards. In *XXXIII Congresso Brasileiro de Ciências da Comunicação*.
- Ging, D. (2019). Alphas, Betas, and Incels: Theorizing the Masculinities of the Mänosphere. *Men and Masculinities*, 22(4), 638–657.
- Grant, T. (2008). Approaching questions in forensic authorship analysis. In J. Gibbons and M. Turell, Eds., *Dimensions of Forensic Linguistics*. Amsterdam-Philadelphia: John Benjamins.
- Grant, T. and MacLeod, N. (2018). Resources and constraints in linguistic identity performance: a theory of authorship. *Language and Law/Linguagem e Direito*, 5(2), 80–96.

- Grant, T. and MacLeod, N. (2020a). *Language and online identities: The Undercover Policing of Internet Sexual Crime*. Cambridge: Cambridge University Press.
- Grant, T. and MacLeod, N. (2020b). Recursos e restrições na Manutenção de Identidades Linguísticas: uma Teoria de Autoria. In D. Almeida, M. Coulthard and R. Sousa-Silva, Eds., *Perspectivas em Linguística Forense*. Campinas: IEL-UNICAMP, 76–94.
- Herring, S. C. (2004). Computer-Mediated Discourse Analysis: An Approach to Researching Online Behavior. In S. A. Barab, R. Kling and J. H. Gray, Eds., *Designing for Virtual Communities in the Service of Learning*. Cambridge: Cambridge University Press, 338–376.
- KhosraviNik, M. and Esposito, E. (2018). Online hate, digital discourse and critique: Exploring digitally mediated discursive practices of gender-based hostility. *Lodz Papers in Pragmatics*, 14(1), 45–68.
- Koch, I. (2007). Hipertexto e construção do sentido. *Alfa*, 51(1), 23–38.
- Lisecki, M. (2013). Discourse features of Internet message board on the example of 4chan. *Media i Społeczeństwo*, 3, 21–40.
- MacLeod, N. and Grant, T. (2012). Whose Tweet? Authorship analysis of micro-blogs and other short-form messages. In *Proceedings of the International Association of Forensic Linguists' 10th Biennial Conference*, 210–224, Birmingham, UK: Aston University.
- MacLeod, N. and Grant, T. (2017). “go on cam but dnt be dirty”: linguistic levels of identity assumption in undercover online operations against child sex abusers. *Language and Law/Linguagem e Direito*, 4(2), 157–175.
- Manivannan, V. (2012). Attaining the Ninth Square: Cybertextuality, Gamification, and Institutional Memory on 4chan. *Enculturation*, Online.
- Manivannan, V. (2013). Tits or GTFO - The logics of misogyny on 4chan's Random - /b/. *The Fibreculture Journal*, 22(FCJ-158), 109–132.
- Marcuschi, L. (2005). Gêneros textuais emergentes no contexto da tecnologia digital. In L. Marcuschi and A. Xavier, Eds., *Hipertexto e gêneros digitais: novas formas de construção de sentido*. Rio de Janeiro: Editora Lucerna, 2 ed., 13–67.
- Mozzillo, I. (2009). O Code-switching: fenômeno inerente ao falante bilíngue. *PAPIA: Revista Brasileira de Estudos do Contato Linguístico*, 19, 185–200.
- Okazaki, N., Toyoda, K., Yokoyama, E., So, H., Katayama, T. and Park, M. (2015). Countermeasure against fingerprinting attack in Tor by separated contents retrieval. *IEICE Communications Express*, 4(12), 370–375.
- Paredes, L., Ferreira, G. and Gomes, N. (2016). A linguagem dos jogos de vídeo game e sua influência na formação de neologismos: reflexões preliminares sob a perspectiva da Linguística. *Revista Philologus*, 22(66), 191–200.
- Rashid, A., Baron, A., Rayson, P., May-Chahal, C., Greenwood, P. and Walkerdine, J. (2013). Who Am I? Analyzing Digital Personas in Cybercrime Investigations. *Computer*, 46(4), 54–61.
- Richoux, N. (2016). *Computer Mediated Communication: Enregisterment of Gamerspeak and Intertextual Borrowings by 4chan Users*. , Virginia Polytechnic Institute and State University, Blacksburg, Virginia.
- Ruocco, J. (2020). Cómo la extrema derecha se apoderó de 4chan. *Nueva Sociedad (NUSO)*, 286, 25–34.
- Salter, M. (2018). From geek masculinity to Gamergate: the technological rationality of online abuse. *Crime, Media, Culture: An International Journal*, 14(2), 247–264.
- Sarmiento, L., Maia, B. and Santos, D. (2004). *The Corp{ó}grafo - a Web-based environment for corpora research*.

- Schneider, J. (2004). Discursos simbólicos e símbolos discursivos: considerações sobre a etnografia da identidade nacional. *Mana*, 10(1), 97–129.
- Seargeant, P. and Tagg, C. (2011). English on the internet and a ‘post-varieties’ approach to language. *World Englishes*, 30(4), 496–514.
- Seargeant, P., Tagg, C. and Ngampramuan, W. (2012). Language choice and addressivity strategies in Thai-English social network interactions. *Journal of Sociolinguistics*, 16(4), 510–531.
- Smith, R. and Wright, V. (2015). The possibilities of re-engagement: cultures of literacy education and so-called NEETs. *Research in Post-Compulsory Education*, 20(4), 400–418.
- Sousa-Silva, R. (2018). Computational Forensic Linguistics: An Overview of Computational Applications in Forensic Contexts. *Language and Law / Linguagem e Direito*, 5(2), 118–143.
- Sousa-Silva, R. and Coulthard, M. (2016). Linguística Forense. In R. J. Dinis-Oliveira and T. Magalhães, Eds., *O que são as Ciências Forenses? – Conceitos, Abrangência e Perspetivas Futuras*. Lisboa: Pactor, chapter Forensic L.
- Thibault, M. (2015). “Do not talk about Anonymous” Censura, autocensura e anonimato nelle periferie del web. *Lexia - Rivista di semiotica*, 21(22 (dezembro)), 237–254.
- Turell, M. T. (2010). The use of textual, grammatical and sociolinguistic evidence in forensic text comparison. *The International Journal of Speech, Language and the Law*, 17(2), 211–250.
- Uchida, Y. and Norasakkunkit, V. (2015). The NEET and Hikikomori spectrum: Assessing the risks and consequences of becoming culturally marginalized. *Frontiers in Psychology*, 6.
- Van Dijk, T. (1984). *Prejudice in discourse*. Amsterdam: John Benjamins.
- Van Dijk, T. (2018). Discourse and Migration. In R. Zapata-Barrero and E. Yalaz, Eds., *Qualitative Research in European Migration Studies. IMISCOE Research Series*. Cham: Springer, 227–245.
- Van Dijk, T. A. (2006). Ideology and discourse analysis. *Journal of Political Ideologies*, 11(2), 115–140.
- Villalva, A. (1994). *Estruturas morfológicas, unidades e hierarquias nas palavras do português*. , Faculdade de Letras da Universidade de Lisboa.
- Weimann, G. (2016). Terrorist migration to the Dark Web. *Perspectives on Terrorism*, 10(3).
- Wodak, R. (2001). The discourse-historical approach. In R. Wodak and M. Meyer, Eds., *Methods of Critical Discourse Analysis*. London: Sage, 63–94.
- Xavier, A. (2002). *O hipertexto na sociedade da informação: a constituição do modo de enunciação digital*. , Universidade Estadual de Campinas.