

PRÁTICAS LINGUÍSTICAS EM *BIG DATA**

PRACTICAL LANGUAGE IN BIG DATA

Vinícius Vargas Vieira dos Santos
Universidade Federal de Goiás
vargasvinicius2@gmail.com

RESUMO: O presente artigo objetiva tratar possíveis relações entre novas mídias digitais e certos aspectos conceituais da linguagem, como significado e performatividade. *Big data* é o termo que se refere ao acúmulo de dados digitais que caracterizou as mídias de comunicação em massa nas duas últimas décadas e está diretamente relacionado à atual configuração da plataforma de serviços de tecnologia Web 2.0. As escalas de desmedido volume e variedade de dados digitais e altos índices de velocidade que caracterizam o *Big data* modificam as paisagens de contexto social, provocando, conseqüentemente, atualizações nas escalas da linguagem. Afinal, contextos em ambientes virtuais entram em colapso, pois ao assumir as próprias características do meio, revelam-se superdiversos, simultâneos, fragmentados, não estruturados, ausentes de marcadores familiares, excedendo escalas tradicionais de tempo, espaço e alcance social.

PALAVRAS-CHAVE: *Big data*; linguagem; contexto; superdiversidade.

ABSTRACT: This article aims at assimilating possible relationships among new digital media and certain conceptual aspects of language, such as meaning and performativity. Big data is a term that refers to digital data accumulation that characterized the mass communication media in the last two decades and it is directly related to the current configuration of Web 2.0 technology services platform. The scales of excessive quantity and variety of digital data and high speed data which characterize the Big data change the social context of landscapes, causing consequently updates on the scales of language. After all, contexts in virtual environments collapse, because they assume the characteristics of the environment, that is, they show themselves as superdiverse, simultaneous, fragmented, unstructured, missing family markers, exceeding traditional scales of time, space and social reach.

KEYWORDS: Big data; language; context; superdiversity.

1 Introdução

Este trabalho objetiva atualizar certos aspectos conceituais dos estudos da linguagem frente às novas realidades sociais configuradas a partir do uso em massa de tecnologias digitais. O objetivo principal desta pesquisa é evidenciar que novas tecnologias, ao se efetivarem como intermédios das práticas linguísticas, estimulam

* Este artigo resume parte de pesquisa de mestrado em linguística, defendida na Faculdade de Letras da Universidade Federal de Goiás, em 2016, sob orientação da Profa. Dra. Joana Plaza Pinto. A referida pesquisa, "Big Data Meio e Linguagem", recebeu apoio financeiro e avaliação sistemática da Capes.

escalas – “escopo espacial, durabilidade temporal, alcance social” (BLOMMAERT; RAMPTON, 2011, p. 9)¹ – e recursos de linguagem nos usuários e usuárias. Outro objetivo mais específico é apresentar suporte teórico para compreender de que modo as novas mídias (a partir de suas características próprias) podem estimular determinados aspectos das práticas de linguagem na contemporaneidade.

Tecnologias digitais, em especial aquelas através das quais fluem dados de comunicação, como redes sociais, por exemplo, estão suportadas por uma plataforma de serviços de tecnologia: “Web 2.0”, termo cunhado e divulgado por Tim O’Reilly (2007). A Web 2.0, diferentemente da Web 1.0 (plataforma anterior de serviços), se caracterizou, dentre outros aspectos, pelo fluxo de dados em larga escala, incentivo ao livre compartilhamento de conteúdo e coletividade no armazenamento de computação em nuvem (O’REILLY, 2007). A massa de dados contabilizada em *bytes*, segundo registros fornecidos pela empresa IBM (CORRIGAN, 2013), tem seu volume duplicado em progressão geométrica a, aproximadamente, cada dois anos. Ao desmedido acúmulo de dados digitais gerados, dentre outros, pela Web 2.0, nomeou-se popularmente: *big data*.

O *big data* se compreende no que especialistas dos estudos de tecnologia, a exemplo Zikopoulos et al. (2012), têm definido sobre os conceitos de velocidade, volume e variedade, conjuntamente referidos por 3Vs. Conteúdos e formatos diversos têm sido acumulados em grande quantidade, com uma rapidez crescente. O desafio atual para a utilização de dados digitais em larga escala compreende a distinção de dados estruturados e dados não estruturados (ZIKOPOULOS et al., 2012). O primeiro se refere a uma determinada quantidade de informação que já se encontra organizada em um banco de dados específico, já eficiente para análise. O segundo, e aqui se compreende a maior parte do conteúdo lançado em rede, se refere à informação que não obedece a um padrão organizacional, como ocorre em postagens de redes sociais por exemplo. Atualmente, existem empresas especializadas na estruturação de dados para clientes que pretendem, por motivos diversos, ter acesso a dados digitais estruturados.

2 A internet transforma a linguagem

Para o linguista David Crystal (2001), a internet tem propiciado novas formas de comunicação, se distinguindo em determinados aspectos das tradicionais formas de conversação da fala e da escrita. Dentre as características implicadas pelos novos meios, estão: ausência de *feedback* simultâneo, multiplicidade de interações simultâneas, possibilidades de manipulação da mensagem já realizada (“copiar” e “colar” textos digitalizados) e hipertextualidade.

Presenciamos, em escala gradual, o uso da máquina para a intermediação das práticas diversas de linguagem. O barateamento dos equipamentos computacionais de pequeno porte dos últimos anos, cumprindo a *previsão de Moore*², possibilitou a utilização de mídias que se entrepõem de modo significativo nas práticas comunicacionais.

Além de apontar que o conteúdo do internetês não é o conteúdo da mensagem que

- 1 “Scale – the spatial scope, temporal durability, social reach” (BLOMMAERT; RAMPTON, 2011, p. 9). (Todas as traduções de fontes em inglês foram feitas pelo autor para fins exclusivos deste artigo).
- 2 Postula a capacidade de computadores aumentarem sua eficiência frente a um custo cada vez menor de produção, possibilitando assim o uso da interface digital a um número cada vez maior de usuários.

este carrega, mas a escrita convencional, Rajagopalan (2013) também exclui a difundida ideia de que o internetês reduz-se e abrevia-se como se refletisse as reduções e celeridade da fala. De acordo com Rajagopalan (2013):

para alguns estudiosos de internet, o internetês é nada mais que uma forma de transmitir mensagens, utilizando uma escrita reduzida e ‘truncada’ que imita a modalidade da fala. (...). O linguista britânico David Crystal, (...), que cunhou o termo “internet linguistics”, tem uma opinião notavelmente diferente (RAJAGOPALAN, 2013, p. 40).

A ideia que Crystal (2001) e Rajagopalan (2013) compartilham é que o determinante das abreviações e celeridade do internetês é a própria internet: “a internet tem feito com que a linguagem evolua para uma nova forma de comunicação, diferente em aspectos fundamentais das formas conversacionais da fala e da escrita” (CRYSTAL, 2001 apud RAJAGOPALAN, 2013, p. 40).

3 Contextos em colapso

Se o meio tecnológico é mesmo determinante na produção da mensagem e, assim como qualquer outro aspecto contextual, contribui para a produção dos significados e atos performativos, não seria estranho nos lançar sobre o que realmente são os novos meios. A união de *hardwares* e *softwares* configura de modo substancial o que compreendemos por novas mídias. Neste ponto também não seria estranho admitirmos que a ampla compreensão e crítica dos recursos e escalas linguísticos motivados por novas mídias passa necessariamente pela compreensão desses objetos da computação. Se considerarmos a lógica de McLuhan (2005), o aspecto contextual “meio”, determinante, dentre outros, para as práticas de linguagem, pode revelar transformações profundas sobre a própria malha contextual em que se insere. O meio digital, como elemento contextual indispensável, modifica a própria engrenagem contextual, porque em função de sua própria estrutura e padrões se deslocam tempo e espaço da enunciação, além de se reconfigurarem padrões de enunciador e enunciatário.

Na história do pensamento linguístico, “contexto” tornou-se elemento imprescindível para compreensão do significado. Gottlob Frege (1960) anunciou que o significado de uma palavra apenas se revelaria no contexto da sentença, não sendo possível determinar o dado semântico em termos isolados: “apenas em uma sentença as palavras realmente possuem um significado³” (FREGE, 1960, p. 71). O contexto de situação, em que se consideram itens não linguísticos, como lugar, tempo e posição do enunciador, também foi apontado por diversos autores, a exemplo de Malinowski (1923) e Halliday (1991), como princípio determinante para compreensão do significado. Desse modo, significado ganhou amplitude frente às relações de valor inicialmente explicitadas por Saussure (1995), em que “o valor de qualquer termo que seja está determinado por aquilo que o rodeia” (SAUSSURE, 1995, p. 135), sendo então compreendido também a partir do contexto de utilização da linguagem. Assim como significado atrelou-se ao conceito de contexto de enunciação, podemos ainda citar a visão de linguagem

3 “only in a proposition have the words really a meaning”.

performativa de Austin (1962/1998). Ele contrasta inicialmente enunciados constativos (declarativos) a enunciados performativos. O primeiro é vinculado à ideia de verdadeiro ou falso (conceitos ligados especialmente à filosofia analítica, da qual Frege é precursor, em que se pretendia inicialmente estudar o funcionamento da linguagem antes de determinar modelos lógicos para o exercício filosófico), e o segundo operando sobre as possibilidades de felicidade ou infelicidade a depender se o autor de determinado enunciado encontra-se de acordo com as condições necessárias para a realização do mesmo. O enunciado ou ato performativo seria aquele que ao contemplar as necessidades de determinado contexto se realizaria como ação via linguagem (o que se faz ao dizer). Austin (1962/1998) estabelece três categorias para os atos de fala: locucionário, ilocucionário e perlocucionário. O primeiro, ato locucionário, refere-se à produção de sons de determinado vocabulário, articulação de aspectos semânticos e sintáticos; o segundo, ilocucionário, refere-se à produção de uma ação através de um possível enunciado, como, por exemplo, o ato performativo do batismo, quando se diz “eu batizo você” ou algo semelhante; e o terceiro, perlocucionário, refere-se ao efeito do ato ilocucionário sobre um provável interlocutor. O seguinte exemplo de Austin (1962/1998) ilustra a distinção acima apontada: locucionário (ele me disse “atire nela”); ilocucionário (ele me pediu para atirar nela); perlocucionário (ele me convenceu ou me levou a atirar nela).

Para tornar performativo nosso enunciado, e sem equívoco, podemos empregar no lugar da forma explícita, vários expedientes mais primitivos, como a entonação, por exemplo, e o gesto; além disso, e sobretudo, o próprio contexto, no qual são pronunciadas as palavras, pode tornar bastante certa a maneira pela qual se deve tomá-las (AUSTIN, 1962/1998, p. 114).

Paulo Roberto Ottoni (1998), considerando Derrida, aponta que para Austin “tanto a enunciação constativa quanto a performativa são atos de discursos produzidos na situação total em que se encontram os interlocutores” (OTTONI, 1998, p. 52). Assim considerando, o contexto em que se situa o falante torna-se elemento indispensável não apenas para ancoragem semântica, mas também ações produzidas através da linguagem. Assim, “atos de fala requerem permanentemente um valor de contexto” (OTTONI, 1998, p. 52). Contexto seria, pois, considerado em seu aspecto totalizante, controlado e finito, sendo que as condições para que a performatividade se realize venham a se submeter a quem fala e conseqüentemente a outros aspectos contextuais (circunstâncias particulares), como, por exemplo, espaço e tempo da enunciação.

Diferentes enunciações indicam, a partir de então, contextos diferentes. A malha complexa do *big data*, entretanto, tende a deslocar a noção tradicional de tempo, espaço e sujeito enunciador, o qual passa a atuar no processo de comunicação por intermédio da extensão midiática (previamente determinada por *designs* computacionais, como a exemplo da Web 2.0), levando a situações de “colapso de contexto” (MARWICK, BOYD, 2010; VARIS, 2014). As autoras utilizam essa expressão ao exemplificar mídias sociais como espaços de sobreposição de referências contextuais de diferentes segmentos (trabalho, família, amigos, dentre outros possíveis), em que uma mesma enunciação é direcionada a variados públicos *online*. A noção de significado, portanto, torna-se complexa a partir do momento em que o contexto social também se torna complexo.

A respeito da questão contextual em ambientes *online*, David Porter (1997),

considerando encontros anônimos e experimentais possibilitados pela rede, diz que correspondentes (interlocutores) no ciberespaço tornam-se muitas vezes sujeitos sem rosto, sem corpo, sem história, sem inflexões, assinaturas ou gestos, além do que ali se opta por revelar. As palavras muitas vezes “parecem desprovidas de contexto, palavras desesperadamente sobrecarregadas pela falta de marcadores familiares de identidade, neste reino estranho e etéreo⁴” (PORTER, 1997, p. XI). Para David Porter (1997), se a internet pode ser compreendida como um espaço cultural, não seria no sentido elitista do termo ou de esfera cultural homogênea; seria antes um espaço marcado pelas condições peculiares de convivência virtual e coletiva próprias do meio, como encontros anônimos e experimentais. De acordo com o autor, são ambientes permeados por vozes sem corpo e pontos de vista descontextualizados.

Percebe-se, portanto, duas perspectivas, podendo mesmo ser aparentemente contraditórias, sobre marcação de contexto no meio virtual. A primeira, de Marwick e Boyd (2010) e Varis (2014), marcada pela *superdiversidade*⁵, em que uma malha de contextos desconexos torna complexa a noção de significado, e a segunda, de Porter (1997), em que o anonimato, decorrente, dentre outros, da superdiversidade de usuários, resulta na ausência de indícios contextuais, levando a ancoragens experimentais do significado. São duas noções que, entretanto, revelam a própria complexidade do elemento contextual em meios de *big data*. Em outras palavras, o ambiente virtual é capaz de apresentar uma trama de contextos superdiversos, desalinhados e, simultaneamente, insuficientemente definidos. O incentivo que atrai pessoas à internet, de acordo com Porter (1997), é o poder de um novo meio de comunicação para conectar pessoas, a expectativa de trocas e sociabilidade. As interações originam espaços sociais que nos levam a reavaliar o próprio conceito que fazemos de comunidade. O confronto das expectativas de sociabilidade (trocas linguísticas) com as ambiguidades e ausência de identidades reconhecíveis trazem como consequência os contornos culturais (e linguísticos) desse novo espaço.

Como os participantes se ajustam às condições dominantes de anonimato e à experiência potencialmente desconcertante de ser reduzido a uma voz isolada, flutuando em um vazio eletrônico amorfo, tornam-se hábeis em reconstituir as palavras sem rosto à sua volta em corpos, histórias, vidas: um compromisso criativo por meio do qual eles se tornam co-produtores dos mundos virtuais em que habitam, e as fronteiras que distinguem a experiência “real” e “virtual” começam a desaparecer. Atos de leitura criativa, em outras palavras, podem substituir a presença física nestes encontros on-line⁶ (PORTER, 1997, p. XII).

Porter (1997) complementa indicando que os atos de interpretação, desdobramentos frente aos contextos imprecisos de espaços virtuais, não são aleatórios,

- 4 “There are words, but they often seem words stripped of context, words desperately burdened by the lack of the other familiar markers of identity in this strange, ethereal realm”.
- 5 O termo “superdiversidade” é utilizado por Steven Vertovec (2007) no artigo *Super-diversity and its implications*, publicado na revista acadêmica *Ethnic and Racial Studies*, que será retomado adiante.
- 6 “As participants adjust to the prevailing conditions of anonymity and to the potentially disconcerting experience of being reduced to a detached voice floating in an amorphous electronic void, they become adept as well at reconstituting the faceless words around them into bodies, his-tories, lives: an imaginative engagement by which they become fully vested co-producers of the virtual worlds that they inhab-it, and the boundaries distinguishing “real” from “virtual” experience begin to fade. Acts of creative reading, in other words, can and do stand in for physical presence in these online encounters”.

porque se o desejo de interação com o outro é o impulso à utilização dos novos meios, somos então levados à personificação do que se encontra anonimamente presente do outro lado. Ao se deparar com uma malha desestruturada de fragmentos eletrônicos, de *hardwares* e *softwares* advindos da indústria computacional, o sujeito portanto tende a substancializar na mente um “leitor” com marcadores familiares, pelos quais esse pode ser (re)conhecido (PORTER, 1997). Para o autor, em meios com desmedidas possibilidades de personificação do outro, a precisão com fidelidade das projeções imaginárias torna-se uma questão secundária. A exposta linha de pensamento leva Porter (1997) a concluir que

a interação que define a cultura da Internet não está na interface entre o usuário e o computador, mas sim entre o usuário e o imaginário coletivo da vasta audiência virtual para quem se submete a uma sucessão interminável de exasperantes fantasias sedutoras evocativas do próprio ser⁷ (PORTER, 1997, p. XIII).

A configuração textual em um ambiente digital realiza-se em modo de *hipertexto*⁸; uma massa de informação não estruturada, em diferentes formatos de mídia, moldando significados complexos (em razão de sua própria estrutura midiática plural). A lógica da linearidade do significante já não se apresenta como é tradicionalmente compreendida pelos estudos da linguagem. Sendo o valor do signo determinado por aquilo que o rodeia (SAUSSURE, 1995), em um ambiente múltiplo e em constante atualização, como no caso da Web 2.0, contextualmente, o significado (fenômeno social) se tornará de difícil apreensão. Em outras palavras, como garantir a ancoragem do significado quando estamos lidando com a prática da navegação ininterrupta em um desmedido emaranhado de dados não estruturados?

4 Superdiversidade

A superdiversidade de dados exige, pois, uma contextualização global para sua investigação. Segundo Blommaert e Backus (2012), “superdiversidade” é um termo descritivo que indica as novas dimensões da pluralidade sociocultural e linguística que emergem das migrações e mobilidades sociais pós Guerra Fria, bem como do surgimento de sistemas globais de comunicação móvel, como a internet. Estruturas tradicionais como grupos sociais de bairro e outros cedem lugar para novas possibilidades de estruturação social, quando entram em campo padrões anormais de comunidade

Vertovec (2007) utiliza “superdiversidade” para diferenciar a diversidade populacional na Grã-Bretanha, que nas últimas três décadas teria sofrido impacto considerável de imigrantes de origens díspares (socioeconomicamente e juridicamente diferenciados), sem precedentes em sua história; sendo, portanto, a diversidade local

7 “The defining interaction of Internet culture lies not in the interface between the user and the computer, but rather in that between the user and the collective imagination of the vast virtual audience to whom one submits an endless succession of enticing, exasperating, evocative figments of one’s being”.

8 Termo cunhado pelo filósofo e sociólogo americano Ted Nelson para se referir a textos em meios tecnológicos, onde se agregam conjuntos de informação em forma de blocos de texto (palavras), imagens e áudio. Ver Lapuente (2013).

melhor representada pelo prefixo “super”, indicando níveis mais avolumados de heterogeneidade cultural.

Blommaert (2012) explicita que a Web 2.0, juntamente a telefones celulares, forneceu a expansão no intercâmbio de informações a longa distância, evidenciando relações translocais, transformando radicalmente as formas de interação, desdobrando-se no modo como regulamos a vida social. Varis e Wang (2011) apontam que:

superdiversidade (on-line) é controlada e moldada por forças multiescalares. Atender a essas dinâmicas entre liberdade, criatividade e normatividade é crucial para a obtenção de uma compreensão detalhada e diferenciada de superdiversidade na Internet; ainda, a maneira em que tal dinâmica funciona, e, mais fundamentalmente, quais as formas de normatividade estão em jogo e em que medida eles organizam práticas on-line⁹ (VARIS; WANG, 2011, p. 72).

As duas forças por trás da superdiversidade, migrações e tecnologias digitais, levam Blommaert (2012) a lançar duas questões: “quem é o outro?”; “quem somos nós?”. O outro passa a assumir sentido de “fluxo constante” (BLOMMAERT, 2012), se tornando, categoricamente, cada vez mais difícil de delimitação. De modo semelhante, ocorre com nossa autoconcepção, pois a vida faz-se mais complexa em sua organização e distribuição em ambientes *online*. Segundo o autor, a superdiversidade nesse aspecto pode ser definida por: mobilidade, complexidade e imprevisibilidade, tornando-se um fenômeno social paradigmático, em que a linguagem se torna uma ferramenta importante para detecção de sua natureza difusa. De acordo com Blommaert (2012), o conceito de “comunidades de fala”, cujos agentes linguísticos, supõe-se, reproduziam padrões de linguagem ancorados em aspectos de tradição atemporais, contrasta com o atual contexto de mobilidade, globalização e superdiversidade. No primeiro caso, aspectos de ordem social e linguística eram considerados de modo independente, estabelecendo relações estáveis e lineares (BLOMMAERT, 2012) e, no segundo caso, entretanto, as relações entre essas instâncias tornam-se dinâmicas e mesmo imprevisíveis. Assim, a estabilidade e linearidade entre as instâncias acima mencionadas cedem lugar a inúmeras exceções que revelam, desse modo, seus contextos móveis de produção de linguagem. Segundo Blommaert (2012), ambientes de superdiversidade (*online* e *offline*) proporcionam, ao falante, recursos linguísticos também superdiversos, realizando-se a prática linguística a partir de contextos e formas semióticas de grande complexidade. A partir desse contexto, estamos diante de repertórios linguísticos igualmente complexos, dinâmicos e instáveis; superdiversidade se manifesta como um aglomerado de camadas, desafiando, dentre outros, os estudos da linguagem.

Para que os estudos da linguagem assimilem as novas roupagens e relações linguísticas assumidas a partir do uso intensivo de tecnologias digitais, como significados, contexto de uso e atos de fala, dentre outros, é necessário compreender a nova gramática do *big data*. Pressupõe-se, portanto, que uma massa de dados diversa, em constante atualização, possa ser investigada sob aspectos linguísticos. Estamos lidando com uma estrutura de natureza complexa (rede; *web*), composta em maioria de massa de dados

9 “superdiversity (online) is controlled and shaped by multiscalar forces. Attending to these dynamics between freedom, creativity and normativity is crucial for obtaining a detailed and nuanced understanding of superdiversity on the Internet; yet the way in which such dynamics work, and, more fundamentally, what forms of normativity are at play and to what extent they organize online practices”.

não estruturados, um sistema de linguagens (ou um antissistema dado sua natureza caótica). Nesse sistema, produzem-se relações de valor (significação) complexas, a partir do momento em que a vizinhança linguística opera em características de volume, variedade e velocidade, se modificando a cada momento sob situações de colapso contextual.

Stuart Hall (2002), em *A Identidade Cultural na Pós-Modernidade*, diz que “as sociedades modernas são, [...], por definição, sociedades de mudança constante, rápida e permanente. Esta é a principal distinção entre as sociedades ‘tradicionais’ e as ‘modernas’” (HALL, 2002, p. 14). Os meios digitais estão relacionados com tais características de mudança constante atribuídas às sociedades modernas; dados digitais são, a todo instante, produzidos, reproduzidos (integralmente ou em partes), mesclados e mesmo descartados. O geógrafo David Harvey, segundo Stuart Hall (2002), fala da modernidade como “caracterizada por um processo sem-fim de rupturas e fragmentações internas no seu próprio interior” (HARVEY, 1989 apud HALL, 2002, p. 12); discorre, ainda, sobre “deslocamento”, conceito de Ernesto Laclau, que Hall (2002) assim explica: “uma estrutura deslocada é aquela cujo centro é deslocado, não sendo substituído por outro, mas por ‘uma pluralidade de centros de poder’” (HALL, 2002, p. 16).

De acordo com Piia Varis e Xuan Wang (2011), a internet é o espaço da superdiversidade por excelência. O paradigma pós-moderno da diversidade cultural (HALL, 2002) parece se confirmar no ambiente digital. A partir desse vasto campo de dados, poderemos encontrar provas da fragmentação do sujeito apontada por esses teóricos da pós-modernidade? Para Varis e Wang (2011), a *World Wide Web* viabiliza novos canais de comunicação, possibilitando, a partir de seu surgimento, novas formas linguísticas e culturais. As autoras apontam que não se pode ignorar esse novo padrão comunicacional quando se pretende explicar o mundo de hoje, sendo necessário, portanto, que se leve em conta a capacidade da internet em complexificar a natureza da vida cultural nos mais diferentes aspectos (VARIS; WANG, 2011). As características linguísticas desse tecido complexo, *big data*, afetam os tradicionais padrões da linguagem não apenas no que corresponde à referência, mas também, e conseqüentemente, questões de valor linguístico, significado e performatividade. Assim como dos contemporâneos espaços de superdiversidade cultural, apresentam-se padrões variados, distintos e híbridos, o contexto de superdiversidade digital a que se vincula o signo “*big data*” pode nos possibilitar novas percepções a respeito da natureza linguística do signo em tempos de novas mídias.

5 Termos de massa

Geoffrey Nunberg (2012), ao abordar o tema *big data*, reflete metalinguisticamente a respeito, apontando para o fato de que “data não é um substantivo plural como ‘pedrinhas’. É um substantivo de massa como ‘poeira’¹⁰” (NUNBERG, 2012, s/p). O termo “massa”, para adjetivar palavras, no que se refere a características semânticas, registra-se na obra de Otto Jespersen (1924/1963): “existem muitas palavras que não invocam a ideia de coisa definida, com forma determinada ou limites precisos. Eu as chamo de ‘masswords... denotam alguma substância em si mesma, independente da forma, tal

10 “data isn’t a plural noun like ‘pebbles’. It’s a mass noun like ‘dust’”.

como [...] água, manteiga, gás, ar, etc'¹¹" (JESPERSEN, 1963, p. 198). Jespersen abordou palavras de massa tanto para termos concretos quanto abstratos, como "satisfação" (JESPERSEN, 1963, p. 198). Segundo Henry Laycock (2004), teóricos posteriores, no entanto, trataram substantivos de massa exclusivamente no âmbito de termos concretos. Quine (1964/2010), ao elaborar uma ontogênese da referência linguística, aborda termos de massa explicitando que tais termos possuem a propriedade semântica de referência cumulativa. "Água" e "açúcar", por exemplo, estabelecem referência de massa porque qualquer quantidade que se acumule dos mencionados referentes, esses continuam a ser água e açúcar. Tais termos, segundo Quine (2010), têm característica gramatical de palavras singulares, por resistirem à pluralização, (apenas em casos bastante específicos utiliza-se "águas" ou "açúcares"), possuem características semânticas de palavras singulares, no tocante a não dividirem suas referências, mas se diferenciam dessas mesmas (singulares) ao não nomearem uma única unidade. Desse modo, Quine (2010) aponta para o caráter semântico em que se distinguem os termos de massa, pois seriam aqueles que acumulam a referência ao invés de dividi-la. Nesse aspecto, termos de massa tornam-se bastante específicos frente às categorias nominativas de singular e plural, o que leva Laycock (2004) a apontar para substantivos de massa como não singulares e não plurais. Qual seria o alvo referente para o termo "*big data*"? Uma massa indeterminada, em expansão constante e progressiva, de números computacionais convertidos em dados de linguagem e de conseqüente impossível apreensão do todo?

O termo "*big data*", a partir de sua massa referencial caracterizada por altos índices de volume, variação e velocidade, deflagra, em si mesmo, uma atualização linguística em tempos de intermediação tecnológica das relações sociais. Isso porque diferente dos termos de massa tradicionalmente compreendidos nessa categoria, "água", "poeira", "areia", "gás", dentre outros, "*big data*" traz ainda uma distinção no que diz respeito a seu volume e variedade (em constante expansão); afinal, as sociedades digitais acumulam desordenadamente dados linguísticos de mídias e conteúdos heterogêneos.

Nesse contexto, o termo "*big data*" deflagra uma exorbitante cadeia de referentes (dados digitais) em constante transformação e ampliação, como dito anteriormente, a se diferenciar no que concerne a mídias e conteúdo. Diferente de palavras coletivas, que fazem referência a ampla quantidade de objetos semelhantes, como gataria (coletivo de gatos) ou plurais mais abrangentes, como o substantivo "ser", que se refere de modo geral ao que possui vida real ou hipotética, abarcando "gatos", "árvores", "dragões" e tantos outros, *big data* possui a peculiaridade de se vincular a uma quantidade indeterminável e interminável de referentes, visto que os bancos de dados estão em ampliação rápida e permanente.

6 Linguística saussuriana e mídias digitais

Em Saussure (1995), a questão do referente traz o princípio da arbitrariedade, já que a nomeação não é motivada pela realidade material. No entanto, mesmo sendo arbitrária a relação entre o signo e aquilo a que ele se refere, e o valor das unidades linguísticas sendo construído entre elas mesmas no interior do sistema, Saussure não

11 "I call these " masswords " ; they may be either material, in which case they denote Borne substance in itself independent of form, such as [...], water, butter, gas, air, etc".

descarta o fato da realidade extra-linguística influenciar a língua, a qual “sofre sem cessar a influência de todos [indivíduos]” (SAUSSURE, 1995, p. 88).

Segundo Saussure, o significado se dá pela relação de valor linguístico entre os próprios signos no interior do sistema de linguagem, “visto ser a língua um sistema em que todos os termos são solidários e o valor de um resulta tão-somente da presença simultânea de outros” (SAUSSURE, 1995, p. 133). A língua seria, pois, um “sistema de valores” (SAUSSURE, 1995, p. 130), e a massa amorfa do pensamento seria capaz de distinguir ideias de “modo claro e constante” (SAUSSURE, 1995, p. 130). De acordo com Saussure, “a coletividade é necessária para estabelecer os valores cuja a única razão de ser está no uso e no consenso geral” (SAUSSURE, 1995, p. 132). Ainda, segundo o referido linguista, a lógica do valor, mesmo fora da língua, seria constituída pelos aspectos: semelhante (possível de comparação) e dessemelhante (suscetível de troca); sendo o conteúdo da unidade linguística determinado pelo que está fora dela ou “o valor de qualquer termo que seja está determinado por aquilo que o rodeia” (SAUSSURE, 1995, p. 135).

As mídias digitais, e todo seu acervo de expansão ininterrupta, são hoje um espaço de linguagens. Imagens, vídeos, textos dos mais variados, objetos linguísticos híbridos são a todo instante atualizados nas telas de diferentes interfaces. O objeto digital enquanto meio de linguagem não nos permite mais a clara distinção entre signo verbal e signo não verbal. Estamos diante de um novo objeto híbrido e complexo a desafiar os modelos tradicionais da pesquisa em Linguística. Segundo Morin (2013), há uma inadequação entre saberes separados e uma realidade polidisciplinar. Consoante Saussure, “na língua, como em todo sistema semiológico, o que distingue um signo é tudo o que o constitui. A diferença é o que faz a característica, como faz o valor e a unidade” (SAUSSURE, 1995, p. 141). Pensar *big data* como um sistema semiológico, ou como extensão artificial dos sistemas semiológicos humanos, implica pensar sobre de que modo as relações de valor poderiam ocorrer a partir de um ambiente em considerável expansão e atualização, ou de que modo características como os 3Vs (volume, variedade e velocidade) poderiam impactar o processo de significação.

Aqui vale retomar a distinção entre conteúdo e mensagem de McLuhan (2005), ao apontar para o fato de que o meio tecnológico configura as escalas e modos das associações e ações humanas. O que nos levaria a questões como a possibilidade do *big data* influenciar, por sua própria natureza (3Vs), ou ainda por sua sintaxe complexa, em hipertexto, o modo como práticas linguísticas têm se constituído *on* e *offline*, ou o modo como a ação é realizada, via linguagem, pelo usuário no ambiente digital. Mídias digitais, ao se interporem nas práticas de linguagem contemporâneas, trazem uma nova configuração para o desempenho linguístico social. Ao operarem como extensões para as atividades comunicativas, passam a configurá-las em função de seu *design* computacional, sendo que *designs* computacionais não são projeções isentas de ideologias. Assim como a memória humana é caráter fundamental da língua, pois é em função dela que se dão as relações associativas (SAUSSURE, 1995), extensões tecnológicas estão a todo instante interferindo nas realizações linguísticas de nosso tempo.

Segundo Saussure (1995), as relações de linguagem se desenvolvem por dois âmbitos: sintagmáticas e associativas. As primeiras estariam relacionadas ao encadeamento dos elementos dos sistemas linguísticos, as palavras, ordenadas a partir

de “relações baseadas no caráter linear da língua, que exclui a possibilidade de pronunciar dois elementos ao mesmo tempo” (SAUSSURE, 1995, p. 142). As relações sintagmáticas estariam suportadas pelas oposições entre os termos na frase, sendo este o princípio de valor linguístico a partir do qual se revelam significados, “a frase é o tipo por excelência de sintagma. Mas ela pertence à fala, e não à língua” (SAUSSURE, 1995, p. 144). No entanto, mídias digitais extrapolam a linearidade da fala ao nos permitir as possibilidades simultâneas do hipertexto. A natureza de velocidade, volume e variação do *big data* levaria esse processo a níveis exagerados, como, por exemplo, as experiências multimidiáticas em redes sociais, com altos índices de informação e atualização aceleradas. Significados passam a se construir a partir de tais índices tecnologicamente configurados, possivelmente levando o sujeito à apreensão nesses mesmos padrões.

As segundas relações explicitadas por Saussure (1995), associativas, não se dariam, pois, através da fala, mas sim no campo da memória do falante: “fora do discurso, as palavras que oferecem algo de comum se associam na memória e assim se formam grupos dentro dos quais imperam relações muito diversas” (SAUSSURE, 1995, p. 143). As relações de ordem associativa, segundo o autor, não se baseariam na extensão linear – como ocorreria nas associações sintagmáticas. Elas se dariam na mente do usuário, mas também realizadas a partir de relações linguísticas de valor. “A relação sintagmática existe in praesentia; repousa em dois ou mais termos igualmente presentes numa série efetiva. Ao contrário, a relação associativa une termos in absentia numa série mnemônica virtual” (SAUSSURE, 1995, p. 143).

Assim como a tradicional linearidade da fala (de relações sintagmáticas) atualiza-se frente às novas mídias, a exemplo do hipertexto; as relações associativas mnemônicas também podem ser estimuladas pelo design dos meios digitais. Ao atuarem como extensões das práticas de linguagem, a inteligência artificial, com sua natureza de dados em massa, pode atualizar índices de associações linguísticas mentais de valor, pois a vizinhança linguística não mais dependerá apenas das associações mentais de semelhança entre termos contextualmente afins, mas estará também sujeita à fragmentação desordenada típica dos vigentes designs computacionais de *big data*. Saussure (1995) utiliza o termo “grupos” para se referir às associações mentais possíveis entre elementos linguísticos. Tais grupos, de acordo com o autor, não se restringem à vinculação de elementos que contenham uma similaridade explícita:

o espírito capta também a natureza das relações que os unem em cada caso e cria com isso tantas séries associativas quantas relações diversas existam, [...], uma palavra qualquer pode sempre evocar tudo quanto seja suscetível de ser-lhe associado de uma maneira ou de outra (SAUSSURE, 1995, p. 145-146).

A inteligência artificial, ao expandir volume, variação e velocidade dos índices associativos do sujeito, poderá, portanto, refletir-se nas ancoragens de significado. Quando nos reportamos ao fato de que associações no ciberespaço se realizam de modo não estruturado, ou em outros casos específicos se dão por ajustes de *design* para fins publicitários, a partir da estruturação de dados do indivíduo por empresas computacionais, (que passam a vender este ou aquele produto a depender das ações realizadas em rede), nos deparamos com uma complexa malha de dados a interferir nas práticas de produção e compreensão semântica e, conseqüentemente, na performatividade linguística.

Enquanto um sintagma suscita em seguida a ideia de uma ordem de sucessão e de um número determinado de elementos, os termos de uma família associativa não se apresentam nem em número definido nem numa ordem determinada. [...]. Um termo dado é como o centro de uma constelação, o ponto para onde convergem outros termos coordenados cuja soma é indefinida (SAUSSURE, 1995, p. 146).

7 Big Data e linguagem: transformações escalares

A configuração textual em um ambiente digital, como dito anteriormente, realiza-se em modo de hipertexto. Blocos de informação em diferentes formatos de mídia se misturam moldando significados complexos. A lógica da linearidade do significante já não se apresenta como é tradicionalmente compreendida pelos estudos da linguagem. Sendo o valor do signo determinado por aquilo que o rodeia, em um ambiente múltiplo e em constante atualização, contextualmente complexo, o significado (fenômeno social) faz-se também multiforme e intrincado. O *big data* e sua natureza de contexto em colapso, sua característica de volume, velocidade e variedade, rompem com os modelos tradicionais de garantia do significado linguístico social. Como garantir a ancoragem do significado, quando estamos lidando com a prática da navegação ininterrupta em um emaranhado de dados não estruturados? Projete por um instante a possibilidade ilusória de se mapear toda a produção linguística digital não estruturada, encontrar padrões recorrentes das mais variadas instâncias da linguagem, compreender como enunciados são criados, replicados, confrontados, encadeados, e assim por diante, até se formarem ideias sobre as coisas do mundo, construção de identidades, atos de fala e suas consequências sociais, e ainda outras realizações da cultura através da linguagem. Estudiosos da linguagem podem se valer da mediação digital para compreender relações em diversas camadas da vida social. No entanto, é necessário se perguntar até onde as interfaces atualizam noções básicas sobre o que é linguagem. Como o uso linguístico se desenvolve a partir de estruturas tecnológicas, *hardwares* e *softwares*, previamente desenhados, consequentemente modulam os mais diversos aspectos da produção linguística. Para se pensar, por exemplo, “como fazer coisas com as palavras”, é preciso considerar possibilidades oferecidas ou negadas pela interface digital. Desdobrando a questão um pouco mais, para compreendermos tais possibilidades para a prática linguística, questionarmos como conceitos linguísticos básicos (através dos quais temos noções sobre o que é e como se realiza a linguagem) se atualizam nestes novos ambientes comunicativos, pode ser um caminho para que possamos responder a questões dessa ordem. Ou seja, para que possamos compreender de que modo as interfaces influenciam a comunicação é preciso indagarmos, por exemplo, se a construção de significados tem operado de modo a considerar o meio em que estes são projetados e ancorados; se a malha discursiva tem se configurado a partir de novos padrões de velocidade e variação; se as escalas de indexicalidade podem refletir a pluralidade rarefeita da marcação de contexto em meios virtuais.

O ciberespaço tornou-se o espaço preenchido pelo *big data*, o qual está a todo instante sendo somado pela massa de usuários e por estes também apreendido linguisticamente de modo hipertextual. A apreensão dessa massa de volume, variedade e velocidade, sem precedentes nas interfaces midiáticas de comunicação, atualiza

consequentemente o modo de percepção, apropriação e produção do significado. Como bem coloca McLuhan (2005), a mídia é a mensagem, e aqui apontamos que consequentemente atualizam-se repertórios e a performance via linguagem. É válido para os estudos da linguagem que indagemos a respeito de como as estruturas da nova comunicação digital podem atualizar a prática linguística. O como fazer coisas com as palavras em ambiente de volume, variedade, velocidade e armazenamento em larga escala poderia sofrer alterações significativas que nos levariam a questionar como a pragmática linguística tem se realizado nesses ambientes mediados por interfaces pré-programadas e desenhadas para determinados fins. Indivíduos estão se permitindo a todo instante que a comunicação seja pré orientada em suas mais profundas possibilidades por modelos anteriormente projetados, não sem quaisquer ideologias determinadas. A partir do momento em que a comunicação se realiza sobre suportes fornecidos por grandes conglomerados econômicos, como Microsoft, Apple e Google, dentre outros, faz-se imprescindível que linguistas se questionem sobre de que modo tais suportes de comunicação estão pré-orientando práticas discursivas e pragmáticas em nossa sociedade.

A lógica de McLuhan (2005), de que a “mensagem” de qualquer “meio” ou “tecnologia” é a mudança de “escala” nas diversas instâncias da vida social, nos leva a questionar sobre as transformações tecnológicas, atualizando escalas da linguagem em sociedades de massa. Se considerarmos que a mídia é a mensagem, a mensagem dos novos meios são as próprias características de volume, velocidade, variedade e não estruturação que o fenômeno do *big data* nos revela. Trata-se de transformações escalares que incidem sobre a língua, pois modificam a amplitude espacial, durabilidade temporal e alcance social da experiência do falante/usuário. O conteúdo de entrada e saída de linguagem em meios digitais opera a partir de padrões escalares de *big data*, pois compartilham da extensa massa digital que vem sendo acumulada em ritmo crescente, de seus espaços transitórios, da superdiversidade de dados que se avizinham em amplas escalas de tempo e espaço e das possibilidades que a estruturação dos mesmos, para qualquer finalidade, representa.

O acúmulo crescente e desordenado é a própria mensagem que nos traz as tecnologias de *big data*. Noutro patamar está o seu conteúdo, não aquele suposto “conteúdo” das imensuráveis e desmedidas imagens, vídeos, mensagens de texto cotidianas, mas o seu metaconteúdo, que é, como previu McLuhan (2005), um outro conteúdo. A estrutura computacional, *hardware/software*, e, consequentemente, a Web 2.0 e o *big data* operam de modo multimidiático sem precedentes. A internet incorporou imagens de segurança, de satélites, as revistas e os jornais, o rádio, a fotografia, o cinema e a televisão, os livros, dentre tantos outros objetos culturais; trata-se, portanto, de uma integradora de mídias em larga escala. É esse o seu conteúdo, assim como é a própria escrita padrão o conteúdo do internetês (RAJAGOPALAN, 2013) e o cinema o conteúdo da televisão (MCLUHAN, 2005). A partir do raciocínio mcluhiano, diferem-se, portanto, três itens sobre o *big data*: sua “mensagem”, seu “conteúdo” e a “mensagem que esse conteúdo vincula”. A mensagem do *big data*, a não estruturação de dados em níveis escalares em alta frequência de volume, velocidade e variação, é que pode trazer consequências profundas à linguagem.

Como apontado por Blommaert (2012), a Web 2.0 expandiu a informação, possibilitando novas relações e transformando formas de interação, se tornando o que ele

chama de “força por detrás da superdiversidade”, superdiversidade que, em ambiente digital, de acordo com Varis e Wang (2011), é controlada e moldada por forças multiescalares. A linguagem de diferentes maneiras, enquanto forma básica de interação, sofre e sofrerá os efeitos da Web 2.0, revelando assim a própria superdiversidade, porque estará sujeita aos moldes do próprio meio de *big data*. As escalas do uso linguístico na internet, conseqüentemente, deverão corresponder em alguma medida às forças multiescalares que moldam o meio *online*. O atual contexto de mobilidade e globalização rompe com as dinâmicas lineares e estáveis que caracterizavam comunidades de fala (BLOMMAERT, 2012), dando lugar a espaços dinâmicos e mesmo imprevisíveis. A ancoragem semântica torna-se complexa e, conseqüentemente, também os atos performativos.

A superdiversidade *online* se materializa digitalmente em uma massa estratificada, atingindo além de repertórios linguísticos, como indica Blommaert (2012), possivelmente modos de produção e leitura. Segundo o autor, a superdiversidade revela variedades de trajetórias, e a tecnologia passa a atuar na apreensão dos mais diferentes repertórios. Podemos, a partir do exposto, sugerir que um mesmo objeto de linguagem poderá compartilhar dos mais diversos hipertextos, atingindo diferentes níveis de alcance social. Os significados tornam-se complexos, porque signos estão sendo replicados em alto teor de mobilidade e conseqüentes associações. Objetos linguísticos, dada a natureza de colapso contextual, podem ser apreendidos em fragmentos, e os efeitos dos atos de fala podem se estilhaçar nas mais diversas direções. O reconhecimento de algumas características dos atuais meios digitais, relacionado a teorias linguísticas referentes à significação e performatividade, pode revelar possibilidades sobre o modo como tais características dos meios interferem sobre aspectos da linguagem. Ao se analisar a produção do significado por meio da tradicional relação estruturalista saussuriana, em que “o valor de qualquer termo que seja está determinado por aquilo que o rodeia” (SAUSSURE, 1995, p. 135), ou pelo viés pragmático, pelo qual o significado se dá pelo seu contexto de utilização, “atos de fala requerem permanentemente um valor de contexto” (OTTONI, 2000, p. 52), ou pelo apontamento de McLuhan (2005), para quem o meio é a mensagem; em todos esses modelos de compreensão do fato semântico, podemos considerar que este, ao se realizar em tempos de superdiversidade digital, sofre atualizações que correspondem exatamente à natureza desses novos meios.

A supervariada de dados é exatamente aquilo que passa a envolver o signo; a vizinhança linguística de um termo não está mais submetida apenas à lógica do sistema linguístico mnemônico. A partir do instante que as relações estão mediadas por interfaces digitais, variedade, velocidade e volume passam a atuar juntamente à mente e às relações semânticas do sistema compartilhado pelos falantes/usuários de mídias. Isso porque as tecnologias da informação tornam-se extensões da mente do usuário. Este, portanto, não contará mais apenas com as dinâmicas fornecidas pela mente, mas, ao produzir ou receber enunciados *online*, estará diretamente vinculado a uma série de possibilidade fornecidas pela máquina. Está ao alcance da mente, com espaços de tempo frequentes para busca – como aponta Crystal (2001), ausência de *feedback* simultâneo é uma das características da linguagem *online* –, enciclopédias, obras literárias, músicas, revistas sobre temas variados, imagens e toda a vasta malha *big data* (impossível de se definir em categorias), uma ampla quantidade de informação caracterizada por dados em disposição não estrutural. Não é incomum a quem se predispõe a práticas comunicativas

em redes sociais, por exemplo, alternar determinada conversa com pesquisas em suportes de busca, como o Google Search. Que efeitos tais práticas podem revelar nas produções de significado e mesmo nos atos performativos? A máquina, ao se tornar extensão da mente dos falantes, colabora para a produção de enunciados suportados por uma série imensurável de imagens, assuntos, referências supervariadas, dentre todas as possibilidades fornecidas pelo universo de dados digitais.

É essa mesma massa não estruturada que assume lugar importante como extensão para a mente do falante/usuário, alargando como consequência os próprios conceitos por via dos quais tentamos compreender o que a língua de fato é. Não seria a língua um depósito mental?

8 Designs computacionais (*affordances*)

Perfilando a teoria aqui sugerida, os três Vs do *big data* (volume, velocidade e variação) estão em relação de causa e consequência com o fenômeno dos contextos em colapso, que, por sua vez, se desdobram na ancoragem experimental dos significados e consequentemente no efeito dos atos performativos produzidos *online*. A performance *online*, como em redes sociais, grupos de discussão, *chats* em vídeo, *posts*, comentários em *sites*, dentre outras possibilidades, em certa medida, contará com outros marcadores contextuais virtuais, como, por exemplo, o avatar de identificação e nome, dentre outros. Uma mesma postagem, entretanto, sendo replicada por centenas de usuários (avatars) diferentes, acompanhada de comentários também diferentes, faz-se plural e multidirecionada. Também, em certa medida, há curiosas encenações de identidades, simulacros (BAUDRILLARD, 1981) que mesclam dados advindos do *offline*, projeções do eu, possibilidades de se performar identidades falsas, que ainda assim se moldam a partir de *affordances*¹² próprios do meio.

Não estamos aqui, entretanto, apontando para o fato de que ambiguidades de sentido sejam características próprias dos novos meios, pois são inerentes às próprias condições da linguagem, como já apontado por Austin (2000). No entanto, não podemos desconsiderar as possibilidades de volume, velocidade e variedade favorecerem ambiguidades de sentido, porque a fugacidade do contexto em colapso e sincronicamente amorfo estimula como um *affordance* a ilocução e perlocução de enunciados.

Como apontado por David Porter (1997), o usuário inevitavelmente buscará completar as fendas contextuais com outras referências encontradas para além do contexto de situação em que o enunciado é produzido. De acordo com Lanier (2010), a expressão de significados *online* está diretamente vinculada aos *designs* de *softwares* e consequentemente sujeita ao aprisionamento tecnológico presente nos mesmos. Há, neste ponto, certa divergência entre Lanier (2010) e Porter (1997), em que o primeiro enxerga uma proposta de impessoalidade ideologicamente fundamentada, diretamente relacionada com interesses econômicos a favor de conglomerados corporativos do setor computacional; o segundo, de modo mais otimista, encontra nos modos de relacionar-se com a linguagem em ambientes virtuais novas possibilidades de ancoragem do

12 Termo cunhado por Gibson (1986), “*affordance*” refere-se à capacidade que o *design* de determinado ambiente tem em estimular ações e comportamentos específicos.

significado a partir da experiência do sujeito. Para Lanier (2010), a flexibilidade do significado é suprimida assim como houve a compressão do áudio através da sintetização do MIDI, eliminando a sinuosidade e transitoriedade de notas a favor da condensação da unidade sonora. O MIDI para o autor é um exemplo de como decisões de *design* se desdobram a partir de aprisionamentos tecnológicos, influenciando a geração de *softwares* posteriores.

Apesar de suas limitações, o MIDI se tornou o esquema padrão para a representação musical em software. Programas musicais e sintetizadores foram projetados para trabalhar com o MIDI, e rapidamente passou a ser impraticável alterar ou descartar todo esse software e hardware. O MIDI criou raízes e, apesar de tentativas hercúleas de reformá-lo em muitas ocasiões por um batalhão de poderosas organizações comerciais, acadêmicas e profissionais internacionais ao longo de várias décadas, ele continua sendo o padrão (LANIER, 2010, p. 23).

Apesar de aparentemente contraditórios, ambos modos de apreender a ancoragem do significado em meio digital parecem coerentes e talvez coexistam de modo pertinente. Para Lanier (2010), a aparente liberdade da Web 2.0 estaria mais relacionada a suas ideologias propagadas do que às práticas que ela nos fornece com modelos comunicativos prontos e delimitados, nos impedindo até mesmo de imaginar outros formatos alternativos de interação social em meios tecnológicos. Para o autor seria essa incapacidade de vislumbrar alternativas um dos pontos críticos do aprisionamento tecnológico, exatamente porque a partir do momento em que determinado *software* apresenta baixos níveis de dinamismo, parece que não há outra coisa a se fazer além de se condescender frente às possibilidades fornecidas.

O design da web, como ela é atualmente, não foi inevitável. No início dos anos 1990, havia talvez dezenas de tentativas viáveis de criar um design para apresentar informações digitais em rede de forma que pudesse atrair uma utilização mais popular. Empresas como a General Magic e a Xanadu desenvolveram designs alternativos com características fundamentalmente diferentes e que nunca chegaram a ser lançados (LANIER, 2010, p. 21).

Trata-se, portanto, de modelos e propostas que cederam espaços para outros formatos (considerados pelo autor como "impessoais"). Ou seja, o sujeito não é o protagonista do meio, apesar de desempenhar o seu papel nele, mas sim a formação de uma inteligência assujeitada e artificial ou as próprias máquinas. Para Lanier (2010), no entanto, de modo semelhante a McLuhan (2005), há modos de se desvencilhar das supostas imposições ou limitações expressas nas mídias contemporâneas. Tendo como conjectura o fato de que os meios digitais substanciados pela Web 2.0 operam a favor da impessoalidade das relações, o autor descreve alguns procedimentos para que usuários possam se esquivar dos *affordances* tecnológicos. Lanier (2010) propõe ao usuário: não postar mensagens anonimamente, a menos que haja qualquer perigo iminente. Podemos ilustrar essa última parte ao conhecido uso do anonimato em rede na onda de protestos e transformações sociais ocorridas em países árabes, emergidas a partir de 2011, a

Primavera Árabe. Considerou-se o uso do anonimato em rede, em especial a *deep web*¹³, como responsável pela divulgação de informações importantes para o resto do mundo e também entre os próprios resistentes (ABREU; NICOLAU, 2014). Lanier, não descartando o óbvio sobre a importância do anonimato para situações emergentes, advoga a favor do incentivo à marcação de identidade pelo usuário, e que os próprios *designs* computacionais sejam favoráveis a tais práticas. Para o autor, um incentivo vazio à cultura do anonimato apresenta riscos de se construir uma inteligência artificial desumanizada, como a exemplo da violência virtual, que no repertório da *web* conceitua-se por “trolagem”, prática recorrente que se estende desde atos criminosos de racismo até a exposição de conteúdo íntimo pessoal e chantagens, dentre outros. Para o autor, a cultura do anonimato é um incentivo à prática recorrente de “trolagem” nos meios digitais. Outra sugestão do autor é direcionada a usuários colaboradores da Wikipedia, que se esforcem para que outros usuários tomem conhecimento dos tópicos que esse contribui. Aqui se tem a crítica ao atual modelo de retenção de informação da Wikipedia, onde, segundo Lanier (2010), autores têm seus pontos de vista eliminados a favor de um conhecimento que se dispõe a ser desautorizado e genérico, promovendo, a partir de então, idealizações de verdades únicas, descartando ambiguidades sobre o objeto.

A questão urgente é então o fato de que a performatividade linguística em meios digitais está diretamente atrelada às opções de atividades determinadas pelo *software* em uso. Isso porque, assim como significados, em função dos colapsos de contexto virtuais desestruturados, absurdamente massivos e simultaneamente clásticos, se ancoram de modo criativo pelo usuário, atos ilocucionários e perlocucionários também estarão, em certa medida, submetidos às fendas contextuais facilitadas pela estrutura computacional. O que realmente incomoda determinado usuário, que ao postar em rede social uma denúncia de fato por ele considerada importante, vê esse mesmo conteúdo ser esquecido em poucos minutos? O que realmente podemos fazer ao dizer, quando o que se diz é rapidamente sobreposto por conteúdos de ordens mais diversas?

9 Considerações finais

A tecnologia tornar-se um espetáculo do capital, por exemplo, ao pesquisar e denunciar em redes sociais por repetidas vezes a situação de descaso do poder público para com os povos indígenas Guarani-Kaiowá no Mato Grosso do Sul, o suposto usuário pode acabar se surpreendendo com uma promoção publicitária de viagem aérea para a capital de Campo Grande. A fragmentação da cultura digital, segundo Lanier (2010), leva a dados digitais descontextualizados; dados esses somados à grande massa de *hexabytes* não estruturados da *World Wide Web*. Considerando que a massa digital opera sob a lógica apontada por Marwick e Boyd (2010) e Varis (2014), de colapso do contexto de produção, não podemos mais determinar objetivamente a qual ambiente o enunciado está ou estará vinculado.

13 *Deep Web* é o termo utilizado ao se referir às camadas ditas profundas da internet. É um ambiente identificado por anonimato, ausência ou drástica redução de controle institucional sobre o conteúdo virtual, é também o espaço caracterizado por conteúdos ilegais, alta disseminação de vírus computacionais e mensagens criptografadas.

Não existe mais um ou alguns contextos de produção. A supervariabilidade digital passa a uma recontextualização constante não controlada pelo falante/usuário; seus enunciados passam a significar de modos diversos, porque signos em sistemas de *big data* possuem variadas e instantaneamente renovadas relações de valor. Como aponta David Porter (1997), na ausência de marcadores de contexto, a leitura criativa de cada interlocutor torna-se a regra, abrindo espaços para experiências semióticas de ambiguidade e equívocos de interpretação. O ambiente virtual é em certa medida amorfo, e os signos, ao atuarem como massa indistinta e descontextualizada, acabam por se assemelhar à “massa amorfa do pensamento” de que nos fala Saussure (1995, p. 130). O usuário reúne material simbólico apreendido, preenchendo lacunas contextuais com suas próprias experiências e referências de linguagem; a esse procedimento facilitam-se pluralidades de leitura, justamente porque contextos superdiversos entram em cena. Se a felicidade ou infelicidade dos atos performativos, que considera se os mesmos são bem ou malsucedidos, estão ancoradas em aspectos contextuais como “quem fala” (autoridade) ou “onde se fala”; como esses conceitos passam a se aplicar em condições de anonimato e superdiversidade? O que se faz ao dizer quando o próprio ato de dizer está engrenado a ferramentas digitais com diferentes níveis escalares de alcance social, durabilidade temporal e amplitude espacial? Os níveis escalares dos atos de fala, locucionário, ilocucionário e perlocucionário, em alguma medida, compreendem aos níveis escalares do *big data*, porque o *big data* altera a malha contextual por via da qual fluem atos performativos da linguagem. Através dos novos meios, somos bombardeados com uma carga de informação em altos níveis de volume, velocidade e variedade. São centenas de enunciados sobre temas diversos com efeitos perlocucionários sobre o sujeito usuário. A internet é o espaço onde as diferenças entre um texto publicitário, a fala do indivíduo desconhecido e níveis escalares familiares de comunicação se retraem, parecendo lógico pensarmos que esse processo, como um *affordance*, contemple meios *offline* e outras instâncias da vida social.

O *design* computacional passa a atuar nos processos de ancoragem semântica e performatividade, justamente porque se tornam extensões do sujeito (MCLUHAN, 2005), e assim “nossas identidades podem ser alteradas pelas idiosincrasias dos gadgets” (LANIER, 2010, p. 19). A identidade na era da digitalização dos mais diversos índices culturais, sexualidade, ideologias políticas e a própria linguagem, também poderá se compreender como elemento passível às reverberações do *big data*. O olhar sobre si e o outro passa pela mediação tecnológica, sendo este o ponto que leva Lanier (2010) a defender uma engenharia computacional com maior comprometimento social.

Ao submeter-se às “fendas contextuais” do meio virtual, o efeito do perlocucionário, através de encontros experimentais com a leitura (PORTER, 1997), pode se desligar por completo dos domínios de seu enunciador, atrelando-se às mais diversas enunciações. O que se “diz” digitalmente, em grande parte, mantém-se registrado. Talvez seja esse um dos principais pontos que diferenciam a enunciação *offline* e *online*. A partir de então, o enunciado torna-se disponível a replicagens sob intenções várias, edições, reedições, ou seja, torna-se, em grande parte das vezes, conteúdo coletivo junto à massa digital e também recorrentemente material comercializado para estruturação de dados com fins econômicos, como no exemplo das políticas on-line pró Guaranis-Kaiowás. O que se faz ao dizer, em outra escala, pode ser contribuir para que uma empresa de aviação lhe ofereça passagens aéreas. Não se pretende aqui desacreditar de mobilizações políticas

ancoradas em ambientes virtuais, mas atentar-se que, em escalas de *big data*, o efeito performativo pode sofrer os mais excêntricos desvios ou ser multiplicado em muitos efeitos sequer previstos.

Obviamente, a lógica austiniana de “o que se faz com as palavras” encontrará novos cursos no meio virtual. Isso porque a atualização de níveis escalares de amplitude espacial e alcance social nos meios informa também variações nos níveis de performatividade (ao próprio uso que o sujeito faz da linguagem). O *hacker* é exemplo bastante útil a essa questão. O que são *hackers* senão indivíduos que, ao deterem “ferramentas” para sobrepor os limites dos meios, tornam-se capazes de exercer níveis de ilocução comumente não realizados em rede?

Parece então lógico supor que os três níveis apontados por Austin (1962/1998), para o ato performativo, podem refletir os meios de *big data*. No aspecto locucionário, a exemplo do internetês e suas características de celeridade e abreviações, e também os níveis ilocucionário e perlocucionário, como anteriormente exposto. Seguindo a lógica mcluhiana, de que o conteúdo de um meio é sempre um meio a ele antecedente, podemos vislumbrar a Web 2.0, ou a internet, como tendo seu conteúdo não nas letras de músicas, roteiros de seriados, números bancários, temáticas de imagens etc., mas o conteúdo da internet seria seus próprios meios anteriores, o livro, o cinema, a fotografia, a televisão, o rádio, dentre outros meios. A internet tem seu conteúdo não em um outro meio, mas em um emaranhado de meios. Seu conteúdo é o multimídia por excelência. O que o falante faz ao dizer na internet? Para responder a essa pergunta seria apropriado antes considerar que se pode dizer de muitas maneiras no superdiverso ambiente virtual. O mesmo enunciado pode ser expresso e veiculado por diversas mídias e diversos espaços virtuais, a partir de ferramentas e conseqüentes *affordances* também diferentes, estando os efeitos de um ato performativo atrelados à diversidade de vias possíveis, ou seja, o meio denuncia níveis escalares desiguais para o próprio efeito do que se diz. A exemplo, podemos presumir que um mesmo enunciado poderá revelar diferentes efeitos caso seja proferido tradicionalmente como uma postagem em rede social, ou se, para essa mesma postagem, o autor optar pelo uso de ferramentas pagas para “impulsionar a publicação”. Também podemos supor que um mesmo enunciado sofrerá diferentes níveis de ilocução, caso seja proferido em formatos diversos de mídia (vídeo, áudio, texto), conseqüentemente a depender das distintas camadas e alcance social que esses venham a atingir.

Conclui-se, então, que as garantias de felicidade ou infelicidade dos atos de fala, em meios de *big data*, estão atreladas à capacidade do usuário falante em manipular as próprias ferramentas que o meio dispõe. Se McLuhan (2005) estiver certo, ao transcorrer por meios de volume, variedade e velocidade massivos e sincronicamente fendidos, a linguagem poderá corresponder apresentando intenções semelhantes. Afinal, se o meio é a mensagem, porque é ele que controla e configura as escalas das associações e ações, ancoragem semântica e performatividade *online* refletem de modo escalar os próprios meios artificiais: *affordances* que modulam, como consequência, as associações (elemento básico para o processo semântico) e ações via linguagem.

Referências

- ABREU, G.; NICOLAU, M. A estética do anonimato na Deep Web: a metáfora das máscaras e do homem invisível aplicada ao “submundo” da internet. *Cultura Midiática*. João Pessoa. Ano VII, n. 12. 2014. Disponível em: <<http://periodicos.ufpb.br/index.php/cm/article/view/19746/10908>>. Acesso em: 02 out. 2015.
- AUSTIN, J. *How to do things with words*. Oxford/New York: Oxford University Press, 1962/1998.
- AUSTIN, J. Performativo-constativo. In: OTTONI, P. R. *Visão performativa da linguagem*. Campinas: Editora da UNICAMP, 1998. (Viagens da Voz). (Apêndice). p. 107- 144.
- BAUDRILLARD, J. *Simulacros e Simulação*. Tradução de Maria João da Costa Pereira, Lisboa: Relógio D’água, 1981.
- BLOMMAERT, J. *Ethnography, superdiversity and linguistic landscapes: chronicles of complexity*. Bristol: Multilingual Matters, 2012.
- BLOMMAERT, J.; BACKUS, A. Superdiverse Repertoires and the Individual. *Tilburg Papers. Culture Studies*, paper 24. Tilburg: Tilburg University, p. 11-32, 2012.
- BLOMMAERT J.; RAMPTOB B. Language and Superdiversity. *Diversities*, v. 13, n. 2, UNESCO & MPI MMG, p. 1-21, 2011. Disponível em: <http://www.mmg.mpg.de/fileadmin/user_upload/Subsites/Diversities/Journals_2011/2011_13-02_art1.pdf>. Acesso em: 13 mai. 2017.
- CORRIGAN, D. *Integrating and governing big data*. New York: IBM Corporation, 2013.
- CRYSTAL, D. *Language and the Internet*. Cambridge: Cambridge University Press, 2001.
- FREGE, G. *The foundations of Arithmetic*. Oxford: Editora Basil Blackwell & Mott, 1960.
- GIBSON, J. J. *The ecological approach to visual perception*. New Jersey: Lawrence Erlbaum, 1986.
- HALL, S. *A identidade cultural na pós-modernidade*. 7ª ed. Rio de Janeiro: DP&A, 2002.
- HALLIDAY, M. A. K. The notion of “context” in language education. In: LE, T.; McCAUSLAND, M. (Eds.) *Language Education: interaction and Development, Proceedings of the International Conference*, Vietnam, April 1991. University of Tasmania (Launceston), 1–26. Disponível em: <<http://annabellelukin.edublogs.org/files/2013/08/Halliday-1991-Context-in-language-education-yc7ta4.pdf>>. Acesso em: 27 fev. 2015.
- JESPERSEN, O. *The philosophy of grammar*. New York. Norton. 1963.

LANIER, J. *Você não é um aplicativo*. Tradução de Cristina Yamagami. São Paulo. Saraiva, 2010.

LAPUENTE, M. *Hipertexto, el nuevo concepto de documento en la cultura de la imagen*. 2013. Disponível em: <<http://www.hipertexto.info/>>. Acesso em: 09 ago. 2015.

LAYCOCK, H. *Mass nouns, count nouns and non-count nouns: Philosophical aspects*. 2004. Disponível em: <<http://semanticsarchive.net/Archive/jRiNmJkM/Mass%20nouns,%20Count%20nouns%20and%20Non-count%20nouns.PDF>>. Acesso em: 29 abr. 2015.

MALINOWSKI, B. The problem of meaning in primitive language. In: OGDEN, C. K.; and RICHARDS, I. A. *The meaning of meaning. A study of the influence of language upon thought and of the science of symbolism*. Supplement I. London, 1923.

MARWICK, A.; BOYD, D. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*. Chicago. 13 (1), p. 114-133, 2010. Disponível em: <<http://journals.sagepub.com/doi/pdf/10.1177/1461444810365313>>. Acesso em: 17 abr. 2017.

MCLUHAN, M. *Os meios de comunicação como extensões do homem: (understanding media)*. Tradução de Décio Pignatari. São Paulo. Ed. Cultrix, 2005.

MORIN, E. *A cabeça bem-feita: repensar a reforma, reformar o pensamento*. Tradução de Eloá Jacobina. 8. ed. Rio de Janeiro: Bertrand Brasil, 2003.

NUNBERG, G. *Forget YOLO: Why 'Big Data' Should Be The Word Of The Year*. 2012. Disponível em: <<http://www.npr.org/2012/12/20/167702665/geoff-nunbergs-word-of-the-year-big-data>>. Acesso em: 02 jan. 2015.

O'REILLY, T. *What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. O'Reilly Media, 2007.

OTTONI, P. *Visão performativa da linguagem*. Campinas. Ed.da Unicamp,1998.

PORTER, D. Introduction. In: PORTER, D. (Ed.). *Internet culture*. New York and London: Routledge, 1997, p. xi-xviii.

QUINE, W. V. O. *Palavra e Objeto*. Petrópolis. Vozes, 2010.

RAJAGOPALAN, K. Como o internetês desafia a linguística. In: SHEPHERD, T. G.; SALIÉS, T. G. (Orgs.). *Linguística da internet*. São Paulo: Contexto, 2013, p. 37-53.

SAUSSURE, F. *Curso de Linguística Geral*. São Paulo. Cultrix, 1995.

VARIS, P.; WANG, X. Superdiversity on the Internet: A Case from China. *Diversities*, Tilburg: University of Tilburg, v. 13, n. 2., p. 71-83, 2011. Disponível em:

<<http://unesdoc.unesco.org/images/0021/002147/214772e.pdf#214784>>. Acesso em: 17 abr. 2017.

VARIS, P. Digital Ethnography. In: GEORGAKOPOULOU, A.; SPILIOTI, T. (Eds.). *The Routledge Handbook of Language and Digital Communication*. London: Routledge, p. 55-68, 2014.

VERTOVEC, S. Super-diversity and its implications. *Ethnic and Racial Studies*, London, v. 30, n. 6, p. 1024-1054, 2007. Disponível em: <http://www.uio.no/studier/emner/sv/sai/SOSANT2525/h14/pensumliste/vertovec_super-diversity.pdf>. Acesso em: 17 abr. 2017.

ZIKOPOULOS, P. et al. *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. McGraw-Hill: Aspen Institute, 2012.

Recebido em 01 setembro de 2016.
Aprovado em 29 de novembro de 2016.