

**UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS E LINGUÍSTICA**

Vinícius Vargas Vieira dos Santos

**BIG DATA, MEIO E LINGUAGEM
Novas tecnologias e práticas linguísticas**

Goiânia

2016

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS (TEDE) NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: **Dissertação** **Tese**

2. Identificação da Tese ou Dissertação

Autor (a):	Vinícius Vargas Vieira dos Santos		
E-mail:	vargasvinicius2@gmail.com		
Seu e-mail pode ser disponibilizado na página?	<input type="checkbox"/> Sim	<input checked="" type="checkbox"/> Não	
Vínculo empregatício do autor			
Agência de fomento: CAPES		Sigla:	
País:	Brasil	UF:	GO
		CNPJ:	00889834/0001-08
Título:	BIG DATA, MEIO E LINGUAGEM Novas tecnologias e práticas linguísticas		
Palavras-chave:	big data, meio, contexto, affordance, performatividade		
Título em outra língua:	BIG DATA , MEDIUM AND LANGUAGE New technologies and linguistic practices		
Palavras-chave em outra língua:	big data, medium, context, affordance, performativity		
Área de concentração:	Estudos Linguísticos		
Data defesa: (dd/mm/aaaa)	29/04/2016		
Programa de Pós-Graduação:	Programa de Pós Graduação em Letras e Linguística		
Orientador (a):	Joana Plaza Pinto		
E-mail:	joplazapinto@gmail.com		
Co-orientador (a):*			
E-mail:			

*Necessita do CPF quando não constar no SisPG

3. Informações de acesso ao documento:

Concorda com a liberação total do documento SIM NÃO¹

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF ou DOC da tese ou dissertação.

O sistema da Biblioteca Digital de Teses e Dissertações garante aos autores, que os arquivos contendo eletronicamente as teses e ou dissertações, antes de sua disponibilização, receberão procedimentos de segurança, criptografia (para não permitir cópia e extração de conteúdo, permitindo apenas impressão fraca) usando o padrão do Acrobat.

Assinatura do (a) autor (a)

Data: ____ / ____ / ____

¹ Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

**UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS E LINGUÍSTICA**

Vinícius Vargas Vieira dos Santos

BIG DATA, MEIO E LINGUAGEM

Novas tecnologias e práticas linguísticas

Dissertação apresentada ao Programa de Pós-Graduação em Letras e Linguística da Universidade Federal de Goiás, como requisito parcial à obtenção do título de Mestre em Letras e Linguística.

Área de concentração: Estudos Linguísticos.

Orientadora: Profa. Dra. Joana Plaza Pinto

Goiânia

2016

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Santos, Vinícius Vargas Vieira dos
BIG DATA, MEIO E LINGUAGEM [manuscrito] : Novas
tecnologias e práticas linguísticas / Vinícius Vargas Vieira dos Santos.
2016.
V, 92 f.: il.

Orientador: Profa. Dra. Joana Plaza Pinto.
Dissertação (Mestrado) - Universidade Federal de Goiás,
Faculdade de Letras (FL), Programa de Pós-Graduação em Letras e
Linguística, Goiânia, 2016.
Bibliografia. Anexos.
Inclui lista de figuras.

1. big data. 2. meio. 3. contexto. 4. affordance. 5.
performatividade. I. Plaza Pinto, Joana , orient. II. Título.

CDU 81




ATA Nº 12/2016

**ATA DA SESSÃO DE JULGAMENTO DA DISSERTAÇÃO DE MESTRADO DO
ALUNO VINÍCIUS VARGAS VIEIRA DOS SANTOS**

Aos vinte e nove dias do mês de abril do ano de dois mil e dezesseis, a partir das 9h na Sala 33 da Faculdade de Letras, realizou-se a sessão pública da Defesa de Dissertação intitulada “Big data, meio e linguagem: novas tecnologias e práticas linguísticas”. Os trabalhos foram instalados pela Orientadora, Professora Doutora Joana Plaza Pinto com a participação dos demais Membros da Banca Examinadora: Professor Doutor Kanavillil Rajagopalan (UNICAMP) e Professor Doutor André Marques do Nascimento (Faculdade de Letras/UFG). A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Dissertação, tendo sido o candidato aprovado pelos seus membros. Proclamados os resultados pela Professora Doutora Joana Plaza Pinto, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que vai assinada pelos Membros da Banca Examinadora e visada pelo Vice-Coordenador do Programa de Pós-Graduação em Letras e Linguística. Goiânia, aos vinte e nove dias do mês de abril do ano de dois mil e dezesseis.


Profa. Dra. Joana Plaza Pinto- Presidente


Prof. Dr. Kanavillil Rajagopalan


Prof. Dr. André Marques do Nascimento

Visto:


Prof. Dr. Sinval Martins de Sousa Filho

Prof. Dr. Sinval Martins de Sousa Filho
Vice-Coordenador do Programa de
Pós-Graduação em Letras e Linguística/UFG
Mat. 256700-1

SUMÁRIO

Resumo.....	04
Abstract.....	05
Introdução.....	06
Capítulo 1: Mídia, Meio e Canal.....	12
Capítulo 2: Designs Computacionais e Affordance	23
Capítulo 3: Colapso de Contexto e Teorias da Linguagem.....	36
Capítulo 4: Práticas Linguísticas em Big Data.....	57
Considerações finais.....	75
Referências	79
Anexos.....	83

RESUMO

Big data, Meio e Linguagem: Novas tecnologias e práticas linguísticas objetiva assimilar 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. Objetos contemporâneos complexos, como big data, nos remetem à consequente necessidade de conceber metodologias que correspondam a suas naturezas superdiversas. Por conseguinte, teve-se em vista, na presente pesquisa, a necessidade de se expandir fronteiras disciplinares, buscando em estudiosos das tecnologias subsídios teóricos para compreensão da natureza dos novos suportes midiáticos. Aparelhos como computadores e celulares com acesso à *World Wide Web* estão aceleradamente transformando o panorama das trocas linguísticas, possibilitando que práticas comunicacionais, a cada dia mais, realizem-se através dos mesmos. É neste ponto que se compreende a própria estrutura computacional como o meio (mídia) através do qual se efetiva a linguagem, a partir de então suas características próprias de design (*affordances*) passam a estimular a ancoragem semântica e a performatividade linguística. 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, meio, contexto, affordance, performatividade

ABSTRACT

Big data, Meio e Linguagem: Novas tecnologias e práticas linguísticas aims to assimilate possible relationships between new digital media and certain conceptual aspects of language, such as meaning and performativity. Big data is a term that refers to digital data gathering, which characterized mass communication media in the past two decades, and it is directly related to the current configuration of Web 2.0 technology services platform. Complex contemporary objects such as big data call for methodological development to meet their super diverse natures. Therefore, it was the goal of this research to expand disciplinary boundaries, searching for theoretical basis in technology studies in the purpose to understand the nature of new media supports. Devices such as computers and mobile phones, with access to the World Wide Web, are increasingly transforming the landscape of linguistic exchanges, enabling communication practices to take place through them. This is why we understand the very computational structure as the medium (media) through which language happens, from then on this structure's design features (affordances) are stimulating semantic anchoring and linguistic performativity. The scales of excessive volume and variety of digital data and its high speed, which characterize the big data, change the social context settings, and thus causing updates on language. After all, contexts in virtual environments collapse because in assuming the characteristics of the medium they reveal themselves as super diverse, simultaneous, fragmented, unstructured, missing family markers, exceeding traditional scales of time, space and social reach.

Keywords: big data, medium, context, affordance, performativity

INTRODUÇÃO

*“quand la data se manifeste,
cela fait à Dada”¹*
Albertine Meunier

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 escalas 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.

A concepção desta pesquisa se deu também como necessidade de desdobramento teórico para produção autoral de arte em novas mídias, mirando-se de fato, epistemologicamente, sobre traços linguísticos do contato entre linguagem e tecnologias digitais, em particular aspectos de significado e performatividade. No entanto, esta pesquisa, apesar de ter o seu foco sobre aspectos linguísticos, se pretende em futuro próximo como contribuição à pesquisa e produção de arte em novas mídias.

Novas tecnologias, dentre outros aspectos, revelam-se como meios de amplas possibilidades para a produção de arte na contemporaneidade, um meio ainda pouco “codificado”, para utilizar o termo de Rubens Gerchman (1973) ao se referir à pintura convencional: “trata-se de um espaço já codificado, tradicional e para se dizer algo que realmente interesse, dentro dele, é muito difícil” (GERCHMAN, 1973, p.163). O ambiente de *data* deflagra-se tanto para possibilidades de invenção e intervenção em arte como para as práticas da linguagem cotidiana.

Ressaltamos que intencionalmente não se delinea neste trabalho a fronteira entre o que os novos meios sugerem ou necessariamente implementam nas práticas linguísticas, podendo gerar níveis de desconforto às exigências de aplicação objetiva entre teoria e práticas exemplificáveis no uso da língua. A referida conduta, para

¹ Anexo 1.

além do foco sobre aspectos epistemológicos, tem suporte metodológico no que se pretende como contribuição para pluralidade das instâncias da linguagem, das artes à pesquisa linguística, dentre outras realizações possíveis pelas faculdades da linguagem humana.

A produção autoral de arte e tecnologia acima mencionada realizou-se, nos últimos anos, a partir de duas vertentes prevaletentes: o uso de dados digitais ordinários disponíveis em rede como matéria para a produção de arte – “o artista deriva da tecnologia uma concepção ativa do material: este passa a ser, ele próprio, um elemento da arte” (CAMPOS, 2015, p. 302) – e experimentações sobre desempenhos não usuais de softwares, com a finalidade de se revelar novas estéticas e linguagens.

Considerando Jacques Ellul, Augusto de Campos (2015) aponta que:

Desligado de suas tradições e sobrecarregado por um mundo estético que é demasiado rico, o artista – devido à tecnologia – encontra-se no pior tipo de situação para fazer uso adequado daquela liberdade que a tecnologia lhe proporciona. O primeiro passo a tentar deve ser, inevitavelmente, de natureza intelectual. Para sair dessa difícil posição, o artista precisa começar a articular uma teoria do que deve ou pode ser feito (CAMPOS, 2015, p. 304).

A natureza diversa e volumosa de dados digitais levou à necessidade de melhor compreender o acúmulo dos mesmos, e em que medida tal acúmulo relaciona-se a aspectos da linguagem; sendo portanto esse o ponto medular que conduziu a proposição da presente pesquisa.

Nas últimas décadas, o fluxo de linguagem, progressivamente, passou a se realizar via interfaces midiáticas. Práticas linguísticas estão sendo mediadas por hardwares e softwares computacionais que, através de seus próprios designs, influenciam modos e escalas² da linguagem e ação humanas.

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

² “scale - the spatial scope, temporal durability, social reach” (BLOMMAERT; RAMPTON, 2011, p.9). Tendo por referência a utilização do termo por Blommaert e Rampton (2011), no presente trabalho “escala” refere-se a variações de amplitude espacial (espaços ou possibilidades de realização da performatividade linguística), durabilidade temporal e alcance social.

armazenamento de computação em nuvem (O'REILLY, 2007). A massa de dados contabilizada em bytes³, 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 termo "big data" é de modo pouco usual traduzido por "megadados" no português. Optamos por utilizar o termo em inglês não apenas por sua maior difusão (mesmo em documentos produzidos em língua portuguesa), mas também porque "data", enquanto termo de massa (como será discutido adiante), carrega de modo metalinguístico as próprias características do meio digital a que se refere.

Na última década, o termo "big data" circulou pelos mais diferentes espaços sociais, de corridas de Fórmula 1 (NEEDHAM, 2013), onde se buscou acumular dados em massa sobre as mais diversas características do veículo para a otimização da performance, a estudos acadêmicos, das artes digitais a análises e estratégias políticas (WU et al. 2013). Segundo o linguista Geoffrey Nunberg (2012), o big data está relacionado com a vitória política de Barack Obama em sua reeleição para presidente, com nossas ansiedades frente à intrusão da tecnologia em nossas privacidades, auxilia epidemiologistas a identificarem focos de gripe e é utilizado para ações de estratégias militares.

A presente pesquisa relaciona, a partir de recorte teórico, conceitos de linguagem como signo, significado, performatividade, contexto, a características apontadas por pesquisadores em tecnologia sobre a natureza dos atuais meios digitais (big data), mediadores de trocas culturais (linguísticas) na atualidade.

Temas contemporâneos como big data apresentam-se como um "tecido junto, isto é, o complexo" (MORIN, 2003, p.14) e são, portanto, um desafio aos modelos disciplinares e metodológicos tradicionais que tendem a reduzir o objeto investigado. Pensar o signo linguístico, agora operado também através de um sistema virtual de linguagem, pode atualizar nosso olhar sobre o mesmo e conseqüentemente sobre desdobramentos de significado e performatividade linguística. Tendo em vista a necessidade de uma compreensão mais ampla dos

³ De acordo com a definição do dicionário da computação on-line, Tech Terms, byte é uma unidade de medida muito utilizada na medição de dados em larga escala. Um único byte corresponde a 8 bits, foi inicialmente desenvolvido para armazenar um único caractere por vez, pois 256 valores são suficientes para representar as letras (minúsculas e maiúsculas) de determinada língua. A unidade byte foi originalmente criada para medição de caracteres, no entanto é atualmente a unidade fundamental de medida para quaisquer padrões de armazenamento de dados.

estudos da linguagem, a investigação linguística do fenômeno aqui proposto deve atuar de modo integrado, contribuindo e se valendo de outras áreas, como pesquisas em tecnologia, sociologia, dentre outras.

Objetos atuais complexos como o big data nos colocam frente a um desafio metodológico e nos levam a reavaliar não apenas as formas de se realizar a pesquisa linguística, mas conseqüentemente as estruturas disciplinares sobre as quais pesquisadores orientam suas linhas de investigação. O cenário de disciplinas compartimentadas faz-se cada vez mais inadequado à pesquisa de temas contemporâneos. Os processos de globalização das últimas décadas têm revelado fenômenos que exigem do pesquisador a capacidade de operar para além das fronteiras tradicionais que delimitam os campos de uma área específica do conhecimento. Segundo Edgar Morin,

há inadequação cada vez mais ampla, profunda e grave entre os saberes separados, fragmentados, compartimentados entre disciplinas, e, por outro lado, realidades ou problemas cada vez mais polidisciplinares, transversais, multidimensionais, transnacionais, globais, planetários. (MORIN, 2003, p. 13)

A relevância da presente pesquisa apoia-se na necessidade de melhor compreensão dos fenômenos midiáticos contemporâneos e de que modo esses interferem nas práticas de linguagem. A revolução desses meios ocorreu com velocidade surpreendente, modificando o panorama das comunicações de massa na passagem do século XX e apresentando, conseqüentemente, desafios a ser enfrentados pelos estudos linguísticos. Como aponta Rajagopalan (1996):

uma área de estudo qualquer - como a Linguística - deve ter alguns princípios doutrinários inegociáveis, sob pena de não poder garantir sua própria sobrevivência institucional. Ora, para facilitar a discussão, seja tal princípio o de que a Linguística jamais pode e deve abdicar seu perfil de uma ciência da linguagem (RAJAGOPALAN, 1996, p. 114)

Big data se trata de um objeto recente e em constantes transformações, portanto não se pretende aqui desenvolver respostas definitivas e conclusas aos desafios que este nos apresenta. Aspira-se, de modo mais coerente com o objeto em questão, a elaboração de direcionamentos frente a determinados aspectos que nos são revelados ao confrontarmos suas características próprias aos conceitos de linguagem acima descritos.

Como poderíamos pensar conceitualmente linguagem e big data sem considerar a utilização do mesmo por diferentes áreas sociais ou para o que pesquisadores em tecnologia nos tem a dizer sobre? A análise do acúmulo de dados a partir de conceitos dos estudos da linguagem pode se realizar tendo como princípio o fato de que dados digitais são (ou convertem-se em) signos de linguagem. Postagens em redes sociais, vídeos, imagens, mapas, áudios, anúncios e tantos outros dados anexados diariamente na rede são objetos virtuais significantes, consideravelmente presente nas práticas da linguagem cotidiana na última década.

Aprender o processo de significação de um suposto sistema linguístico big data, ancorado em relações de valor entre signos e contextos de situação, nos leva à necessidade de reportarmos a uma metodologia que busque antes uma ampliação do objeto do que um reducionismo passível de análise. Uma possível ampliação do olhar sobre o objeto não se refere obviamente a gerar uma quantidade cada vez maior de dados, mas sim ao enlarguemento dos campos de conhecimento aplicados à pesquisa. É de suma importância compreender o que nos tem a dizer tecnólogos sobre dados digitais, quando se propõem a analisar espaços na web, por exemplo.

Presenciamos atualmente um discurso acadêmico de valorização da interdisciplinaridade, no entanto, quando nos deparamos com as metodologias aplicadas, percebemos um contraste entre os dois pontos. A pesquisa de caráter interdisciplinar está voltada para a investigação de temas contemporâneos de maior complexidade, objetos que não mais se limitam às estruturas de uma disciplina específica. O mesmo ocorre com a delimitação de um corpus afinado para objetos que são absolutamente plurais e que se desfazem ao se analisar apenas uma parte do mesmo.

Para atingir os objetivos desta pesquisa, no primeiro capítulo, Mídia, Meio e Canal, discuto a radicalidade da interpretação de Marshall McLuhan⁴ (1964/2005) sobre meio e mensagem. O teórico da comunicação canadense, de certo modo, antecipou a midiaticização digital das práticas comunicativas da contemporaneidade, ao apontar o **meio**, através do qual se efetiva a linguagem, como elemento fundamental para a compreensão da natureza da mensagem. Para o autor, as

⁴ Herbert Marshall McLuhan (Edmonton, 21 de julho de 1911 - Toronto, 31 de dezembro de 1980).

tecnologias, enquanto extensões do sujeito, atuariam de modo significativo, afetando o complexo psíquico e social.

No segundo capítulo, *Designs Computacionais e Affordance*, aponto para os atuais modelos de fluxo de dados digitais, através dos quais práticas linguísticas têm se realizado em nosso cotidiano. Exponho, a partir de informações de tecnólogos, a natureza conceitual do big data e exemplos empíricos da utilização do mesmo em diferentes esferas da vida social. Destaco apologias e críticas aos designs fornecidos por empresas de computação, considerando suas ideologias e consequentes aspirações econômicas, ao determinarem níveis e recursos das interfaces de mídias digitais. O *affordance*⁵ digital, em outras palavras, como as tecnologias orientam ou permitem modos de interação, é apresentado como ferramenta conceitual para que seja possível precisar o objetivo principal desta pesquisa.

No terceiro capítulo, *Colapso de Contexto*⁶ e *Teorias da Linguagem*, argumento sobre a imprecisão de contextos no ciberespaço, que consequentemente reflete sobre significados e práticas de performatividade linguística. Para tal, trago conceitos tradicionais de significado e valor linguísticos, performatividade linguística, dentre outros, contrastando-os com a natureza dos designs computacionais de big data.

No quarto capítulo, *Práticas Linguísticas em Big Data*, busco apresentar uma malha relacional dos assuntos abordados ao longo do texto com a finalidade de que os padrões associados nos revelem pontos pertinentes sobre a relação aqui explorada entre mídias de big data e linguagem, mais precisamente sobre aspectos concernentes a significado e performatividade linguística. Ainda é relatada experiência autoral de arte e tecnologia, produzida durante a presente pesquisa, como elemento simbólico dos apontamentos linguísticos abordados.

⁵ Termo cunhado por Gibson (1986), “affordance” refere-se à capacidade que o design de determinado ambiente tem em estimular ações e comportamentos específicos.

⁶ “Colapso de contexto”(MARWICK & BOYD, 2010; VARIS, 2014) refere-se à complexidade contextual em ambientes on-line.

Capítulo 1. Mídia, Meio e Canal

Para Roman Jakobson (1960/2007), o conceito de “canal” é visto como elemento físico de propagação da mensagem, que se limita à capacidade de funcionar ou não, possibilitando assim a prática comunicativa. “Para ser eficaz, a mensagem requer (...) finalmente um CONTACTO, um canal físico e uma conexão psicológica entre o remetente e o destinatário, que os capacite a ambos a entrarem e permanecerem em comunicação.” (JAKOBSON, 1960/2007, p.121). Salvo exceções, “canal” se manteve para os estudos linguísticos como apenas mais um dos elementos formadores do contexto, superficialmente explorado frente a outros aspectos contextuais (mais profundamente investigados) como o lugar social ocupado pelo usuário, o próprio usuário (enunciador) ou ainda o momento histórico da enunciação (MACHADO, 2006).

Na presente pesquisa se dispensa atenção especial ao canal, ou meio através do qual se performatizam enunciados. Parte-se da premissa apontada por Marshall McLuhan (1964), em que “o meio é a mensagem, porque é o meio que configura e controla a escala e a forma das ações e associações humanas⁷” (MCLUHAN, 1964, p.9). Em tempos de tecnologia massificada, em que as interfaces digitais passam a atuar como intermédios rotineiros da comunicação humana, a lógica apresentada por McLuhan (1964) parece estar ainda mais evidente, pois a interface torna-se, além do meio através do qual a comunicação se efetiva, uma extensão do próprio sujeito, delimitando ou determinando escalas e recursos semióticos para a realização do ato comunicativo. Logo, pode-se considerar que a comunicação humana, ao se efetivar cada vez mais através de interfaces e softwares pré programados, como por exemplo plataformas de redes sociais, está comprometida com as limitações ou aberturas proporcionadas por esses, não apenas no quesito de recurso e escala, mas consequentemente e também no âmbito semântico e pragmático.

A proposição de McLuhan (1964/2005) considera as consequências psíquicas e sociais do que ele chama de modelos computacionais, que seriam hoje de fato as interfaces (hardwares e softwares). Podemos fazer uma analogia com os

⁷ The medium is the message because it is the medium that shapes and controls the scale and form of human association and action.

próprios modelos científicos, que são o meio através do qual algo passa a ser reconhecido como válido. A compreensão científicista do mundo é a própria mensagem; o significado de “raio”, por exemplo, que em outros tempos poderia ser a ação de uma divindade, hoje está ancorado na carga elétrica presente na nuvem e sua correspondência elétrica com o solo. Do mesmo modo, pensar em termos de big data pode revelar significados diferentes sobre a realidade das coisas. O que estamos, no entanto, realmente explicitando aqui é o fato de que o meio é de certo modo a própria mensagem. Partindo desse princípio, em tempos de comunicação a cada dia mais efetivada via tecnologias computacionais, pode-se afirmar que não é possível apreender significado ou atos de fala, nas sociedades de massa, sem considerar os meios, através dos quais fluem diariamente milhares de enunciados, e suas características determinantes. Tudo, ou quase, é data, ou passível de ser. Traduzimos dados culturais em data, como fotografias antigas, discos de música, programas de televisão, porque o conteúdo de um meio é sempre o meio anterior (MCLUHAN, 1964/2005).

Hoje, (...), na era da eletrônica, a classificação dos dados cede ao reconhecimento de estruturas e padrões. (...). Quando os dados se alteram rapidamente, a classificação é por demais fragmentária. Para dar conta dos dados em velocidade elétrica e em situações características de “sobrecarga da informação”, os homens recorrem ao estudo das configurações, como o marinheiro do Maestrom, de Edgar Allan Poe. (MCLUHAN, 1964/2005, p. 11)

Poe (1841/2015) escreve sobre um marinheiro que vai pescar com seu irmão em alto mar, onde correntes de água se cruzavam lhes rendendo grande quantidade de peixes. No entanto, durante uma tempestade, o barco é tomado por um vórtice de água (Anexo 2), sendo levados gradualmente, em movimentos circulares pelo natural giro do vórtice, ao fundo do mesmo. O irmão é arrastado para o fundo do vórtice junto do barco, e o marinheiro consegue se livrar, retornando à superfície. Para conseguir tal realização, o personagem se propõe a observar o modo como outros objetos, fragmentos de navios, blocos de traves, troncos de árvores, caixotes quebrados, barris, dentre outros, se comportam no interior do vórtice. O marinheiro percebe que alguns objetos cilíndricos ofereciam maior resistência à descida ao fundo do turbilhão de água, ele então abandona o barco e se amarra a um barril, conseguindo assim retornar à superfície quando o vórtice se abrandava.

O conto de Poe (1841/2015) é utilizado por McLuhan (1964/2005) de modo metafórico para ilustrar a relação do sujeito com meios tecnológicos de comunicação. Segundo McLuhan (1964/2005), a tecnologia nos coloca diante de “vórtices”, pois anestesia funções e percepções humanas ao se tornarem extensões das mesmas. Ferramentas têm a capacidade de estender os sentidos, o corpo e a mente, como as roupas estendem a pele, os óculos estendem a visão, a cadeira estende o quadril e a coluna, e os talheres estendem as mãos. Assim, artefatos humanos de linguagem, como computadores, seriam extensões do corpo e da mente para práticas de linguagem. Como o personagem de Poe (1841/2015), apenas ao percebermos padrões pertencentes a esses novos meios podemos conceber maneiras de transgredir os efeitos anestésicos que eles nos implicam; “ao estudar o padrão das consequências desse enorme vórtice de energia em que estamos envolvidos, pode ser possível planejar uma estratégia de fuga e sobrevivência” (MCLUHAN, 2002, 4 min.). O desafio está, portanto, em conseguir identificar padrões e estruturas dos meios nos quais estamos imersos, pois um meio dominante estaria sempre além dos domínios da percepção dos que nele se encontram.

Ao efeito anestésico que ferramentas e tecnologias implicam em nossos corpos, sentidos e funções, McLuhan (1964/2005) associou a ideia de “massagem”, como trocadilho com o termo “mensagem”. Para o autor, os novos meios atuam de modo a massagear a mente humana levando o sujeito, imerso como em um vórtice, a não identificar padrões e estruturas ali presentes. Atualmente, tablets, celulares e computadores tornaram-se extensões para milhares de pessoas, acumulando ao menos duas importantes funções da mente humana, a comunicação e o armazenamento de informações. Até que ponto a mente humana poderia mesmo se deixar embalar pelas funções artificiais dos novos meios tecnológicos? Para o autor, o que nós temos feito com nossas tecnologias é colocar o sistema nervoso fora de nós (MCLUHAN, 1964/2005).

Se McLuhan (1964/2005) tomou a direção correta, estaríamos, em certas escalas, anestesiando nossas funções acima descritas a favor de que a máquina as realize por nós. Aqui talvez corretores ortográficos de softwares de editores de texto sejam um bom exemplo do modo como, não apenas o usuário se abstém dia a dia da capacidade de identificar padrões referentes à escrita convencional, mas também como essa mesma escrita convencional é reafirmada, com o auxílio da máquina,

como o padrão de uso de vocabulário, concordâncias, dentre outras características da variação linguística normativa e institucional.

McLuhan (1964/2005) inicialmente parece prever a lógica da tecnologia em nossas vidas pós digitalização dos meios de comunicação. Um olhar mais profundo em sua obra, entretanto, nos leva a perceber que a atualidade do pensamento mcluhiano está no modo radical como compreende a influência das transformações tecnológicas sobre as práticas cotidianas das sociedades em que tais eventos se realizam. Segundo ele, após mais de um século de tecnologia elétrica, delinea-se uma perspectiva globalizante onde tempo e espaço foram transformados, projetando modificações na própria consciência humana, pois extensões tecnológicas sobre o sujeito levam a modificações sobre o complexo psíquico e social, sendo deste modo imprescindível compreender os efeitos dessas transformações.

McLuhan (1964/2005) apresenta um contraste entre a era mecânica, do maquinário industrial, e a era da tecnologia da automação. No primeiro caso tem-se a lógica da fragmentação aliada à centralidade, que representa a natureza da tecnologia da máquina sobre o trabalho humano; no segundo caso, teríamos o oposto, a **integralidade** aliada à **descentralização** determinando os rumos das relações humanas. Seria, pois, essa a lógica de raciocínio que permite que o pensamento de McLuhan permaneça atual para compreendermos de que modo a digitalização da cultura, e mais especificamente da comunicação humana, realiza desdobramentos sobre diversas instâncias da vida social. Para o autor, “a ‘mensagem’ de qualquer meio ou tecnologia é a mudança de escala, cadência ou padrão que esse meio ou tecnologia introduz nas coisas humanas” (MCLUHAN, 1964/2005, p. 22). O autor cita como exemplo a estrada de ferro, que permitiu o surgimento de novos modelos de cidades, lazer e trabalho. A tecnologia da aviação, por outro lado, aumentando a velocidade do transporte, viria a romper, por si só, a lógica da ferrovia e conseqüentemente da política, dentre outras instâncias sociais, independente de seus propósitos de utilização.

Aqui, vale ressaltar, temos metáforas de como McLuhan compreendeu, radicalmente, o meio como a própria mensagem; não seria pois o conteúdo transportado pelo veículo ferroviário ou de aviação o motor profundo de transformações sociais, mas o próprio meio através do qual esse conteúdo se transportava.

Obviamente houve nesse sentido um alargamento do que se compreende por mensagem, o que o levou a receber críticas como a de Umberto Eco (1965/1984). Segundo Eco (1965/1984), McLuhan (1964/2005) teria confundido conceitos básicos da linguagem ao afirmar que o meio é a mensagem; “na verdade, todo o raciocínio de McLuhan é presidido por uma série de erros graves para um teórico da comunicação, para o qual não há diferenças entre o canal de comunicação, código e mensagem⁸” (ECO, 1965/1984, p. 398). Para Eco (1965/1984), o procedimento aplicado por McLuhan é defeituoso, porque ele coloca, ao lado de cada afirmação, outras afirmações contraditórias, assimilando ambas como congruentes e, ainda, não demonstra preocupação com a veracidade dos argumentos, “se contenta em que sejam. Então, o que do nosso ponto de vista pode parecer contraditório, para o dele é simplesmente presença simultânea⁹” (ECO, 1965/1984, p. 394). Eco (1965/1984) diz que no livro de McLuhan, *The Medium is the Message*, o assunto se desenvolve de modo a fundir palavras e imagens, destruindo-se o encadeamento lógico “a favor de uma exposição sincrônica, visivo-verbal, de dados não fundamentados, dispostos como um torvelinho frente a inteligência do leitor¹⁰” (ECO, 1965/1984, p. 395/396). A crítica de Eco, no entanto, atinge patamares de argumentação desarrazoados ao dizer que “há algo de bom em McLuhan, como há nos fumadores de drogas e nos hippies. Veremos o que eles ainda são capazes de fazer¹¹” (ECO, 1965/1984, p. 401).

Parte da crítica ao procedimento de McLuhan reside sobre sua “incapacidade” de apresentar com clareza linear seu raciocínio. É como se estivesse, a todo momento, buscando, sem sucesso definitivo, provar sua teoria de que o meio é a mensagem. Seu texto em modo de mosaico opera a partir de inúmeros e desconexos exemplos e parece nunca encerrar-se com objetividade. “O autor parece querer afogar-nos em uma enorme quantidade de dados, (...), porém a

⁸ En realidad, todo el razonamiento de McLuhan está presidido por una serie de equívocos gravísimos en un teórico de la comunicación, para el cual no se establecen diferencias entre el canal de comunicación, el código y el mensaje.

⁹ se contenta con que sean. Lo que desde nuestro punto de vista podría parecer contradicción, desde el suyo es simplemente presencia simultánea.

¹⁰ a favor de una exposición sincrónica, visivo-verbal, de datos no razonados, dispuestos como un torbellino ante la inteligencia del lector

¹¹ Hay algo bueno en McLuhan, como lo hay en los fumadores de drogas y en los "hippies". Veremos lo que son aún capaces de hacer.

informação central derivada de tudo isso é apenas uma: o meio é a mensagem¹² (ECO, 1965/1984, p. 400/401).

De acordo com Barbosa (2011), críticas ao autor se desdobram desde o estilo e método narrativo e de investigação a questões relacionadas ao determinismo tecnológico.

No primeiro caso, as críticas se apoiariam sobre três aspectos principais: erudição desmedida do autor, pelo fato de que McLuhan integrava discussões de autores de vias distintas “como historiadores, antropólogos, psicólogos, poetas, críticos literários entre outros” (BARBOSA, 2011, p.3), o que de algum modo viria a colaborar para sua ausência de objetividade aparente; a amplitude histórica e conseqüentemente reducionista de seu trabalho, a exemplos do desenvolvimento da escrita e da prensa de Gutenberg¹³; e o uso exagerado de trocadilhos, metáforas e jogos de palavras. Seriam esses três pontos que exigiriam grande esforço do leitor para assimilar a sua linha argumentativa.

Segundo Barbosa (2011), seria a oposição ao conceito de “ponto de vista fixo” que levaria McLuhan a desenvolver um padrão de escrita plural, denominado por ele mesmo como “método mosaico”. Ele haveria compreendido que a escrita, diferentemente da oralidade (meio áudio-tátil, em que diversas instâncias simbólicas independentes da audição estão envolvidas), exige praticamente o sentido da visão para se realizar. O ponto de vista fixo, para McLuhan, estaria vinculado à prática da escrita alfabética e à reprodução da mesma pela prensa de Gutenberg. A exemplo do vórtice de Poe (1841/2015), McLuhan (1964/2005) vislumbrou os novos meios tecnológicos como estruturas complexas e plurais que apenas a partir de um método de linguagem também complexo e plural, que incorporasse na própria escrita a natureza dos meios, poderia render conclusões pertinentes sobre os mesmos.

McLuhan diz abertamente que ele não explica, mas explora o real a partir de suas sondas (probes). O método mosaico consiste em não se prender em um ponto de vista fixo utilizando sondagens sobre o real. McLuhan acredita assim ser possível a partir de uma coleção de inúmeros exemplos, aforismos, metáforas entre outros, fazer o processo de justaposição, que permite dentro de um enorme universo de exemplos, ou um mosaico de exemplos, perceber padrões e relações significativas. (...). Para George P. Elliot

¹² Sí, porque el autor parece querer ahogar- nos en un cúmulo enorme de datos, (...), pero la información central que deriva de todo ello es sólo una: el medio es el mensaje.

¹³ Para McLuhan (1972), a prensa móvel de Gutenberg promoveu, a partir de unidades uniformes, contínuas e repetíveis indefinidamente, o cálculo infinitesimal.

(Hot&Cool, 1968, p. 89) é impossível fazer um resumo racional das ideias de McLuhan, pois sua escrita é anti-lógica, circular, repetitiva, inadequada, aforística e ultrajante. (BARBOSA, 2011, p.4/5)

No segundo caso (BARBOSA, 2011), McLuhan foi alvo de críticas a respeito de seu determinismo tecnológico, conceito muitas vezes carregado de atmosfera negativa e ligado à crença de que as transformações tecnológicas são imperativas e irrefutáveis nas dinâmicas da sociedade. O caráter negativo da crítica reside na ideia de que a partir do instante em que aceitamos que tecnologias condicionam de modo determinante a vida social não há pontos de fuga, restando ao sujeito se prostrar frente às condições a ele afiguradas. Aqui se pode contrastar tais críticas ao conto de Poe (1841/2015) utilizado por McLuhan (1964/2005), em que apesar de se encontrar envolto pelo vórtice de água, o marinheiro, a partir da compreensão dos padrões de comportamento dos objetos, que ali também se encontram, é capaz de conceber uma estratégia para se desvencilhar daquela situação.

A este aspecto metodológico de McLuhan, podemos relacionar a teoria do método contra-indutivo de Paul Feyerabend (1977). Para o filósofo austríaco, a ciência é em si um procedimento anárquico onde “tudo vale” (FEYERABEND, 1977, p. 9), no qual as violações de regras epistemológicas são necessárias ao avanço científico. O autor se opõe à formulação prévia das ideias e sua consequente expressão formal posterior, apontando que ambas as partes do processo devem se desenvolver de modo simultâneo a partir de experimentações, para tal ele utiliza a alegoria do modo como crianças desenvolvem o conhecimento:

Usam palavras, combinam essas palavras, com elas brincam até que apreendem um significado que se havia mantido para além de seu alcance. E a atividade lúdica inicial é requisito básico do ato final de compreensão. Não há razão para supor que esse mecanismo deixe de agir na pessoa adulta. Cabe esperar, por exemplo, que a ideia de liberdade só se faça clara por meio das mesmas ações que supostamente criaram a liberdade. Criação de uma coisa e geração associada à compreensão de uma ideia correta dessa coisa são, muitas vezes, partes de um único e indivisível processo, partes que não podem separar-se, sob pena de interromper o processo. (FEYERABEND, 1977, p. 32)

A radicalidade mcluhiana, apesar de causar incômodos óbvios sobre a compreensão e distinção tradicional de canal e mensagem linguísticos, ou também na simples lógica, ao refletirmos sobre suas metáforas, de que o conteúdo

transportado por veículos também afeta a atividade coletiva, foi o que permitiu um olhar profundo sobre como formas de tecnologias atingem de modo peculiar, por si só e drasticamente, diversas instâncias da vida social. Ainda utilizando das metáforas do autor para melhor esclarecimento da lógica de seu pensamento, as transformações profundas causadas pela luz elétrica vão além de sua capacidade de auxiliar uma cirurgia cerebral ou iluminar uma partida noturna de beisebol, estão ligadas a transformações que extrapolam a sequenciação típica da maquinária industrial, possibilitando um sentido de simultaneidade e instantaneidade nas relações humanas (como explicitado na citação abaixo), afetando assim percepções de tempo, espaço e, conseqüentemente, relações linguísticas e sociais.

A velocidade elétrica, aglutinando todas as funções sociais e políticas numa súbita implosão, elevou a consciência humana de responsabilidade a um grau dos mais intensos. É este fator implosivo que altera a posição do negro, do adolescente e de outros grupos. Eles já não podem ser contidos, no sentido político de associação limitada. Eles agora estão envolvidos em nossas vidas, como nós na deles — graças aos meios elétricos. (MCLUHAN, 1964/2005, p 19)

Para McLuhan (1964/2005), o formato dos meios é determinante para produção da mensagem e seus significados, eis o motivo de sua posição radical ao afirmar que o meio é própria mensagem. Em verdade, o autor propõe uma distinção entre mensagem e conteúdo da mensagem. O tema da linguagem cubista na pintura é exemplar a esse respeito:

O cubismo substitui o “ponto de vista”, ou faceta da ilusão perspectivista, por todas as facetas do objeto apresentadas simultaneamente (...). O cubismo, exibindo o dentro e o fora, o acima e o abaixo, a frente, as costas e tudo o mais, em duas dimensões, desfaz a ilusão da perspectiva em favor da apreensão sensória instantânea do todo. Ao propiciar a apreensão total instantânea, o cubismo como que de repente anunciou que o meio é a mensagem. (MCLUHAN, 1964/2005, p 27)

A temática metalinguística do cubismo, que se difere neste aspecto de referencialidades não metalinguísticas (a exemplo da temática surrealista, onde os referentes estão localizados no universo onírico), operando sobre a própria estrutura, deflagra e transforma a gramática precedente da pintura em planos. A mídia (meio ou suporte) determina a gramática, o sentido e conseqüentemente seu efeito social. Seguindo a lógica de McLuhan, mensagem e conteúdo se contrastam porque o conteúdo seria, por exemplo, uma mulher que chora em *Mulher Chorando*

ou as mazelas da Guerra Civil Espanhola e o bombardeio de aviões alemães em *Guernica* (obras de Picasso), mas a mensagem seria o próprio desfazer da ilusão perspectiva “em favor da apreensão sensória instantânea do todo” (MCLUHAN, 2005, p.27).

Segundo McLuhan (1964/2005), a mensagem poderia então se vincular ou não a um determinado conteúdo. A luz elétrica, por exemplo, seria a própria mensagem, podendo estar vinculada a um conteúdo: “a luz elétrica é informação pura. É algo assim como um meio sem mensagem, a menos que seja usada para explicitar algum anúncio verbal ou algum nome” (MCLUHAN, 1964/2005, p.22). A respeito do exemplo da eletricidade, o autor aponta que por se tratar de um meio, a princípio sem conteúdo, não a percebemos de imediato como um meio de comunicação, pois a mensagem da luz elétrica é por natureza difusa e descentralizada. Ele ilustra sua linha de raciocínio com o exemplo da carne que o assaltante utiliza para distrair o cão de guarda; o aparente conteúdo de um meio, ou seja, o que muitas vezes pressupomos ser o conteúdo, seria o objeto de distração, (o conteúdo expresso pelos significados de uma enunciação, por exemplo), enquanto seu real conteúdo estaria efetivamente latente em outro meio.

Assim, para o autor, essa seria a lógica que permeia todos os veículos de comunicação, o conteúdo de um meio seria, portanto, sempre outro meio antecedente, o conteúdo da escrita não seria, pois, o conteúdo que nos traz as sentenças redigidas, mas a fala. Assim, o conteúdo da palavra impressa seria a própria escrita. “Se alguém perguntar, ‘Qual é o conteúdo da fala?’, necessário se torna dizer: ‘É um processo de pensamento, real, não-verbal em si mesmo’” (MCLUHAN, 1964/2005, p. 22).

Tomando a máxima de McLuhan para um exemplo atual, a diferença entre um post em rede social que defende determinado posicionamento político de outro que aparentemente reforça um posicionamento contrário não é algo tão relevante quanto sua semelhança. Afinal, estão ambos operando a partir de estruturas e padrões de linguagem similares. É neste ponto que se revela a academicamente mal compreendida lógica mcluhiana, na qual menos considerável é o conteúdo aparente, quando mensagens se realizam de modo semelhante. Elas expressam, em alguma medida, a mesma mensagem, a mensagem do próprio meio em que se vinculam. A crítica de McLuhan (1964/2005) está ancorada sobre a perspectiva de que o meio não é um elemento “neutro, um mero instrumento, nem

somente o transmissor” (SOUSA, CURVELLO, RUSSI, 2012) da prática comunicacional, ele guarda padrões de uma nova estrutura, de um novo ambiente.

Antes da velocidade elétrica e do campo integral ou unificado, que o meio fosse a mensagem era algo que não tinha nada de óbvio. Parecia então que a mensagem era o “conteúdo”, como costumavam dizer as pessoas ao perguntarem sobre o que significava um quadro, ou de que uma coisa se tratava. Nunca se lembravam de perguntar do que se tratava uma melodia ou uma casa ou um vestido. (McLuhan, 1964/2005, p. 28).

O linguista Kanavillil Rajagopalan (2012) utiliza-se da linha de pensamento sobre meios e mensagens de McLuhan ao tratar dos desafios que a nova linguagem presente na internet, o internetês, apresenta à Linguística. O internetês é caracterizado, dentre outros aspectos, pelo abreviar de palavras, uso de caracteres imagético-simbólicos (emoticons) e redução do uso de sinais de pontuação. O internetês carrega o uso criativo da escrita, se apartando das regras institucionais da gramática normativa. Rajagopalan (2012), traçando paralelo entre internetês e telegrafês (linguagem utilizada em telégrafos, também caracterizada pela brevidade da escrita), esclarece que ambos os casos exemplificam “o fenômeno do meio moldar a mensagem, na forma que Marshall McLuhan vislumbrou” (RAJAGOPALAN, 2012, p.41), em especial no que diz respeito à rapidez da comunicação, o limite sobre o número de palavras (como no Twitter, por exemplo) e, conseqüentemente, aspectos de redução da forma, como abreviações.

Apoiando-se em dois pontos-chave da discussão mcluhiana, de que o meio “impacta nossas vidas não pelo conteúdo que ele consegue transmitir, mas em virtude de suas próprias características” (RAJAGOPALAN, 2012, p.42) e de que o conteúdo de um meio se revela sempre em outro meio antecedente, Rajagopalan (2012) propõe que o conteúdo do internetês não é a mensagem por ele transmitida, mas a escrita convencional; “a ‘matéria prima’ - por assim dizer – que alimenta (...) o internetês é a própria escrita convencional (de uma língua padrão), e não uma ideia ou mensagem” (RAJAGOPALAN, 2012, p.43). Para o autor, o internetês, apesar de se desenvolver a partir de novas tecnologias, atuando em um meio artificial, é fruto do trabalho coletivo de milhares de usuários e neste ponto está mais próximo das línguas naturais do que de línguas artificiais criadas a partir de projetos pré-orientados.

Alvo de críticas sobre diversas instâncias de seu trabalho (ver BARBOSA, 2011), McLuhan é hoje revisto sobre as dimensões dos novos meios. Para Rajagopalan (2012), apesar do referido autor não ter vivido tempo suficiente para presenciar a “verdadeira eclosão da internet em larga escala, (...) os sinais da revolução a acontecer estavam claramente em seu radar” (RAJAGOPALAN, 2012, p.42).

McLuhan (1964/2005), na busca de melhor definir os padrões e estruturas que se constroem a partir das transformações tecnológicas (representativas de novos meios), apresenta a estratégica distinção conceitual entre mensagem e conteúdo, sendo muitas vezes mal compreendido como tendo confundido conceitos básicos como canal e mensagem, como explicitado acima na crítica de Umberto Eco (1965/1984). No entanto, em tempos de mídias digitais a cada dia mais operadas como suportes da comunicação cotidiana, a teoria mcluhiana do “meio como mensagem” faz-se atual para compreendermos de que modo tais mídias, atuando como intermédio das práticas linguísticas, moldam as escalas e recursos das trocas de linguagem na contemporaneidade.

Ao olharmos para os meios de tecnologia atuais, através dos quais se efetivam práticas linguísticas, nos deparamos com interfaces como tablets, aparelhos celulares e computadores que possuem designs previamente configurados por empresas do setor computacional. No capítulo seguinte, trataremos das características inerentes aos contemporâneos modelos computacionais, suas propriedades e ideologias, em especial ao modelo em uso da Web 2.0.

Capítulo 2. Designs computacionais e Affordance

Na última década, uma torrente de dados digitais se desencadeou nas redes, através de diversas interfaces. A comunicação humana passou a se realizar por intermédios midiáticos de modo sem precedentes, suscitando mesmo atos públicos em ambientes off-line, mobilizados em rede, ou a utilização de mídias que frequentemente compartilham o foco com interações sociais off-line. A informação foi gerada e replicada, formando agrupamentos de textos relacionáveis, que se hibridizam em hipertextos frente aos usuários, interlocutores de uma massa digital em constante expansão, que estão a todo instante também colaborando para o acúmulo e replicagem de dados. A partir de então, temos oportunidades de estudar processos e dinâmicas sociais de novas maneiras, torna-se possível pesquisar sobre atributos diversos de centenas de milhões de pessoas: “podemos ver as imagens e os vídeos que eles criam e comentar, monitorar suas conversas, ler os seus posts e tweets, navegar por seus mapas, ouvir suas faixas de áudio e seguir as suas trajetórias no espaço físico”¹⁴ (MANOVICH, 2011, p.1).

Pesquisa realizada pela Universidade da Califórnia (HILBERT; LOPEZ, 2011) estimou a competência tecnológica do mundo para acumular, transferir e computar informações. Sessenta tecnologias analógicas e digitais foram monitoradas entre 1986 e 2007, segundo os pesquisadores, no ano de 2007 estima-se que a humanidade armazenou cerca de 295 trilhões de megabytes (o que corresponde a 295 hexabytes), transmitiu cerca de 2 quatrilhões de megabytes (o que corresponde a 2 zettabytes) e realizou 6.4 milhões de instruções por segundo (MIPS) em computadores de uso geral.

Os números são importantes para mensurarmos como a informação tem sido armazenada em formato digital nas últimas décadas. Estima-se, segundo Hilbert e Lopez (2011), que o ano de 2002 foi o marco onde o mundo passou a armazenar mais informação digital do que analógica pela primeira vez. A mudança neste ponto ocorreu com velocidade surpreendente, dado o fato que a memória tecnológica digital ocupou espaço de 94% em 2007, em relação ao armazenamento tecnológico analógico, ocupando os 6% restantes.

¹⁴ We can see the images and the videos they create and comment on, monitor the conversations they are engaged in, read their blog posts and tweets, navigate their maps, listen to their track lists, and follow their trajectories in physical space.

O armazenamento de dados digitais 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.

Como um termo genérico, "big data" pode ser bastante nebuloso, da mesma forma que o termo "nuvem" abrange diversas tecnologias. [...], logs de servidores web, sensores de fluxo de tráfego, imagens de satélite, transmissões de áudio, transações bancárias, MP3s de música (...), o conteúdo das páginas web, varreduras de documentos governamentais, trilhas de GPS, telemetria de automóveis, dados do mercado financeiro, a lista continua. Todos eles são realmente a mesma coisa? Para esclarecer as coisas, os três Vs do volume, velocidade e variedade são comumente utilizados para caracterizar os diferentes aspectos do big data. Eles são uma lente útil para se visualizar e compreender a natureza dos dados e das plataformas de software disponíveis para a exploração destes¹⁵ (THAKARE; DESHMUKH, 2012, p.2).

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.

Muitos são os exemplos que podemos listar sobre a utilização do big data por diferentes organizações e intenções. Um dos mais comentados casos foi a utilização de dados em massa pelos estrategistas na campanha de reeleição do

¹⁵ As a catch-all term, "big data" can be pretty nebulous, in the same way that the term "cloud" covers diverse technologies. [...], web server logs, traffic flow sensors, satellite imagery, broadcast audio streams, banking transactions, MP3s of rock music, the content of web pages, scans of government documents, GPS trails, telemetry from automobiles, financial market data, the list goes on. Are these all really the same thing? To clarify matters, the three Vs of volume, velocity and variety are commonly used to characterize different aspects of big data. They're a helpful lens through which to view and understand the nature of the data and the software platforms available to exploit them.¹⁵

presidente americano Barack Obama. Um extenso banco de dados foi estruturado com a intenção de angariar não apenas votos, mas também apoio financeiro à campanha, levando em conta aspectos individuais de cada possível apoiador.

Imagine uma mãe de duas crianças que vive em uma cidadezinha no estado de Ohio, no centro-oeste dos Estados Unidos [...]. Seus filhos estudam em escola pública. Ela costuma tuitar sobre o meio ambiente e ainda mantém uma página no Facebook sobre comida orgânica. Tudo isso ficaria registrado no banco de dados da equipe de Obama. “A campanha mandaria para ela e-mails de Michelle Obama sobre as políticas ambientais planejadas pelo presidente e sobre educação pública”, afirma o consultor Andrew Rasiej (MORAES, 2005).

Como exemplo prático de êxito na aquisição de big data em uma área específica do conhecimento científico, Cukier (2010) cita o Sloan Digital Sky Survey, telescópio ativado em 2000, localizado em Apache Point Observatory, Novo México, EUA. O autor esclarece que o telescópio digital coletou mais dados em suas primeiras semanas de funcionamento do que todos os dados coletados pela história da astronomia. Previa-se, para 2010, uma quantidade convertida em bytes de informação de 140 terabytes; quantidade que se estima coletar a cada cinco dias com o sucessor Synoptic Survey Telescope Grande, com instalação pretendida para 2016, no Chile. O autor aponta ainda para a decodificação do genoma humano com 30 bilhões de pares, que levou 10 anos para ser concluído no ano de 2003, poderia ser realizado nos dias de hoje no período de apenas uma semana.

Todos esses exemplos contam a mesma história: que o mundo contém uma inimaginavelmente vasta quantidade de informação digital que está ficando cada vez mais vasta e cada vez mais rapidamente. Isso faz com que se possa fazer muitas coisas que antes não poderiam ser feitas.¹⁶ (CUKIER, 2010)

Torna-se possível, a partir do uso em massa de mídias digitais, por exemplo, recolher dados culturais em larga escala. O mesmo tem sido uma tendência lucrativa para empresas de tecnologia detentoras de dados pessoais como veremos adiante; “as pessoas podem rastrear hábitos alimentares, exercícios,

¹⁶ All these examples tell the same story: that the world contains an unimaginably vast amount of digital information which is getting ever vaster ever more rapidly. This makes it possible to do many things that previously could not be done.

tempo de permanência on-line, atividade sexual, ciclos mensais, sono, humor e finanças on-line"¹⁷ (YAU, 2009, p.1).

Segundo Yau (2009), recolher dados da vida cotidiana não é um fenômeno dos tempos atuais. Mass Observation, grupo de pesquisa social da Grã-Bretanha, realizava, na década de 1930, coleta de dados rotineiros da sociedade local, desde aparência física como cortes de cabelo e padrões de barbas a gestos, formas de comunicação e comportamento social. Nos tempos atuais, no entanto, há a ampliação em larga medida das possibilidades de coletas de dados da vida social, obviamente proporcionada pelas interfaces digitais presentes no cotidiano das sociedades. Coleta de dados pessoais já é realidade para muitas pessoas que têm seu dia a dia, trabalho, relações familiares, compra de alimentos e produtos domésticos, dentre outros, guiados por softwares que em resumo prometem otimização do tempo e melhor qualidade para as tarefas realizadas.

Gilles Babinet (2014), *digital champions*¹⁸ do governo francês, considera que sistemas digitais de larga escala poderiam ser extremamente mais eficientes que os atuais, para que os governos possam cumprir funções essenciais como segurança, redução das desigualdades sociais, organização do mercado de trabalho, funcionamento da saúde, planejamento urbano, dentre outros. O autor suscita previsões, como a do matemático John von Neumann, a "Singularidade"¹⁹, onde máquinas passariam a assumir parte tão importante sobre o curso da história humana que não mais haveria possibilidades de retorno, e a famosa Lei de Moore, que 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.

A atual capacidade de captura e análise, frente ao armazenamento de dados ocorrido nos últimos anos, ainda é ínfima. O ponto chave está na eficácia de operar dados não estruturados, que no dia a dia têm se acumulado em ritmo cada vez mais acelerado. Este é o desafio que tem orientado pesquisas e empresas do

¹⁷ People can track eating habits, exercise, time spent online, sexual activity, monthly cycles, sleep, mood, and finances online.

¹⁸ Digital champions são designados por cada Estado-Membro da União Europeia para ajudar os Estados e a Comissão Europeia a promover benefícios da abertura digital à sociedade.

¹⁹ Singularidade refere-se à ideologia tecnológica de que chegará o tempo onde as inteligências artificiais irão superar a natureza humana, assumindo o controle definitivamente ou em parte da vida social.

setor. Segundo Babinet (2015), as descobertas das oportunidades são praticamente ilimitadas, uma vez que os dados podem ser reunidos e manipulados em massa.

De acordo com Babinet (2015), a partir da revolução digital, a humanidade experimenta uma nova quebra de paradigmas, pois a informação se tornou de fácil acesso a custos populares. Tal facilidade, no entanto, gerou outra questão a ser considerada, a capacidade de utilização: “geralmente, nós não sabemos ainda otimizar a distribuição desta informação para ser apropriada para as necessidades de todos²⁰” (BABINET, 2015, p. 1). Parte do desafio da utilização de big data está mesmo no desenvolvimento de formas e possibilidades de operar a massa digital para finalidades de bem social. Como exemplo, o autor aponta que a informação, armazenada em nuvens, possibilitará estatísticas indicadoras de determinadas doenças, então milhares de indivíduos poderão ser diagnosticados em massa, com as propensões a essa ou àquela doença anteriormente indicadas, orientando assim tratamentos e alimentos para redução de riscos.

A democratização da gestão de informação está diretamente relacionada ao acesso de diferentes camadas sociais aos meios de telecomunicação (HILBERT, 2013). Nas últimas décadas, assistimos à diversificação de aquisição desses meios por diferentes grupos; faixas etárias e econômicas diversas se apropriaram de tecnologias com maior ou menor intensidade. Diferentes modelos de contrato fornecidos por empresas do ramo disponibilizaram também diferentes níveis de acesso, tornando-se comum o uso em massa das tecnologias de informação e também se desdobrando uma gama de dados de entrada e saída, relacionada em muitas ocasiões ao nível econômico do contratante.

Na era analógica de 1986, a maioria das assinaturas de telecomunicações foram de telefones de linha fixa, e todos eles tinham a mesma performance. Isto resultou em uma relação bastante linear entre o número de assinaturas e a capacidade de tráfego médio. Vinte anos mais tarde, há uma miríade de diferentes assinaturas de telecomunicações com a mais diversa gama de performances. Isso resulta em uma diversidade bidimensional entre sociedades com mais ou menos inscrições, e com maior ou menor capacidade de telecomunicações.²¹ (Hilbert, 2013, p. 16)

²⁰ mais plus généralement nous ne savons pas encore optimiser la distribution de cette information pour qu'elle soit appropriée aux besoins de chacun.

²¹ In the analog age of 1986, the vast majority of telecom subscriptions were fixed-line phones, and all of them had the same performance. This resulted in a quite linear relation between the number of subscriptions and the average traffic capacity. Twenty years later, there's a myriad of different telecom

De acordo com Hilbert (2013), “palavras” são um dos dados mais disponíveis, de baixo custo e estruturados dentre os dados digitais. A ideia apontada pelo autor é o uso de palavras-relacionadas para que o big data seja uma ferramenta capaz, dentre outros, de realizar previsões sobre diversas instâncias da vida social. Aqui podemos apontar coleta e análise de comentários, posts em redes sociais, palavras utilizadas em buscadores web como *Google Search*, mensagens on-line em aplicativos de aparelhos móveis, dentre outros. Como exemplo, o autor cita o *Google Dengue*, onde se realiza anualmente previsões sobre índices de dengue no planeta para o ano seguinte, se baseando nas buscas de palavras relacionadas à doença em sua plataforma. Curiosamente, suas previsões correspondem aos dados informados posteriormente pelos governos dos países afetados pela doença no ano anterior.

Encontramos (...) uma relação estreita entre a quantidade de pessoas que pesquisam tópicos relacionados à dengue e a quantidade de pessoas que realmente têm sintomas da doença. É claro que nem todo mundo que pesquisa por "dengue" está realmente doente, mas um padrão aparece quando todas as consultas de pesquisa relacionadas à dengue são adicionadas ao mesmo tempo. (...). Ao contar a frequência com que vemos essas consultas de pesquisa, podemos fazer uma estimativa da circulação da dengue em diferentes países e regiões ao redor do mundo. (GOOGLE, 2015)

A desigualdade tecnológica, no campo de domínio, coleta e análise de dados em massa, desdobra-se em outras desigualdades sociais, quando determinada comunidade, em maior ou menor escala, pode lançar mão dessa tecnologia a favor de programas sociais, analisar riscos de investimentos, levantar estatísticas em larga escala, previsões de impactos ambientais, econômicos, dentre outros.

Impulsionado pelas novas possibilidades abertas na operacionalização de big data, Chris Anderson escreve em 2008 o controverso “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”. Impulsionando ideologias tecnológicas como a Singularidade, o autor expõe que após a possibilidade de operar dados em massa, modelos explicativos da realidade tornam-se

subscriptions with the most diverse range of performances. This results in a two dimensional diversity among societies with more or less subscriptions, and with more or less telecommunication capacity²¹.

desnecessários; métodos capazes de nos informar sobre o mundo a nossa volta, a partir de recortes significativos de dados, tornam-se obsoletos à medida que se faz possível o uso de tecnologias de big data. Ou seja, modelos que buscavam cientificar a complexidade, nos assegurando dados sobre a realidade das coisas, se tornariam dispensáveis quando é possível lidar com velocidade, sobre variedade em grande volume.

Na escala pentabyte, a informação exige uma abordagem totalmente diferente, que nos obriga a perder a corda de dados como algo que pode ser visualizado em sua totalidade. Força-nos a ver os dados matematicamente primeiro e estabelecer um contexto para isso mais tarde.²² (ANDERSON, 2008).

Como aponta Anderson (2008), o Google pode traduzir idiomas sem necessariamente conhecê-los, sem qualquer análise semântica. A ferramenta Google Translate opera varrendo big data de documentos já traduzidos na rede, comparando padrões e fornecendo resultados, em segundos (GOOGLE, 2010). O autor indica que, antes, tradicionais modelos científicos são construídos a partir de hipóteses testáveis, a experiência então confirma ou descarta a suposta teoria sobre determinada particularidade do mundo; a partir de determinado modelo, dados, até então incertos, nos garantem valores de verdade. Diferentemente, a exemplo do Google Translate (em que se reconhecem padrões de tradução em big data), agora seria então possível analisar dados, mesmo sem lançarmos hipóteses, apenas jogar uma grande massa de informação digitalizada e deixar que algoritmos estatísticos encontrem, neles, padrões significativos.

Como exemplo, Anderson (2008) também cita a experiência de J. Craig Venter e o sequenciamento de genes por supercomputadores que analisam estatisticamente os dados produzidos em determinado ecossistema; “em 2005 ele começou a sequenciar o ar. No processo, ele descobriu milhares de espécies previamente desconhecidas de bactérias e outras formas de vida²³” (ANDERSON, 2008). A lógica da causalidade é substituída pela do nexos (ANDERSON, 2008), as possibilidades de relações entre dados armazenados em massa estão além dos

²² At the petabyte scale, information (...) calls for an entirely different approach, one that requires us to lose the tether of data as something that can be visualized in its totality. It forces us to view data mathematically first and establish a context for it later.

²³ in 2005 he started sequencing the air. In the process, he discovered thousands of previously unknown species of bacteria and other life-forms.

parâmetros científicos tradicionais, podem apresentar resultados não previamente supostos, e a ciência tende a avançar neste aspecto mesmo sem modelos previamente determinados, teorias unificadas ou explicações mecanicistas. Como aponta Alexander Szalay (2012):

Todas as ciências, tanto físicas e sociais, bem como algumas ciências humanas, são cada vez mais intensivas ao uso de dados. Como conjuntos de dados crescem cada vez mais, a capacidade de analisar as informações escondidas neles se torna cada vez mais assustadora. Desenvolver a "Ciência do Big Data" é uma tarefa integrada ao contrário de qualquer outra na ciência anterior. Ela exige a colaboração de estatísticos, cientistas da computação e especialistas disciplinares que podem fundir suas habilidades para criar novas ferramentas (hardware e software), novas técnicas (algoritmos e estatísticas), e uma nova metodologia que produz descobertas através de dados orientados ao invés de inquirir hipóteses orientadas²⁴ (SZALAY, 2012, p.1)

Em caminho divergente dos otimistas da nova realidade virtual de dados acumulados e suas possibilidades de aplicação extra-humanas, o cientista da computação Jaron Lanier (2010), em *You are not a Gadget*, tece uma crítica à ideologia da Singularidade, ao atual modelo da Web 2.0 e conseqüentemente à sua natureza de big data, alegando que estamos lidando com um design computacional previamente configurado para a alienação do sujeito, tornando-o cada vez mais parte do sistema de informação ao invés de autônomo frente à tecnologia. O autor defende que a opção por softwares já programados tornam as pessoas limitadas frente às possibilidades que poderiam se abrir para estruturas pré-programáveis pelo próprio usuário. Os atuais modelos de interface são anteriormente elaborados levando em conta estratégias para obtenção de lucro para as próprias empresas de tecnologia como o Google. Uma interface para compartilhamento de música, por exemplo, poderia estar configurada para facilitar a troca de capital direto entre artistas e público, ao invés disso o que temos é um fluxo indeterminado de dados digitais, onde os músicos acabam por não receber qualquer contrapartida financeira,

²⁴ All sciences, both physical and social, as well as some humanities, are increasingly data-intensive. As datasets grow ever larger, the ability to analyze the information hidden in them becomes increasingly daunting. Developing the "Science of Big Data" is an integrative task unlike any in science before. It requires the collaboration of statisticians, computer scientists, and disciplinary experts who can merge their skills to create new tools (computer hardware and software), new techniques (algorithms and statistics), and a new methodology that yields discovery through data-driven rather than hypothesis-driven inquiry.

enquanto empresas proprietárias dos softwares utilizados angariam grandes quantias.

Segundo o autor, softwares estão sujeitos a “aprisionamento tecnológico” na medida em que programas e seus consequentes designs tornam-se um padrão para uma série de outros softwares que dele se desdobram, em outras palavras, quando programas são projetados a partir de um outro já existente. Nas palavras de Lanier (2010),

A delicada natureza da maturação de programas de computador pode fazer designs digitais ficarem paralisados por um processo chamado “aprisionamento tecnológico”. Isso acontece quando muitos programas são projetados para funcionar com um software existente. O processo de alterar significativamente o software em uma situação na qual muitos outros programas dependem dele é extremamente difícil. Dessa forma, ele quase nunca é realizado (LANIER, 2010, p. 22).

Forma-se com isso uma massa de estruturas padronizadas, que consequentemente têm efeitos sobre a vida social. A ideologia de acesso livre a dados, segundo Lainer (2010), guarda uma ironia, pois a liberdade na web estaria promovendo maior liberdade às máquinas do que às próprias pessoas, à medida que indivíduos passam a atuar anonimamente na formação de uma inteligência artificial comum e limitadora.

A partir de então, práticas de comunicação impessoal estariam rebaixando a interação pessoal em um processo de desumanização acarretado por opções de design computacional previamente determinados. Lanier (2010) aponta para a responsabilidade social das empresas computacionais, esclarecendo que o trabalho em tecnologia da informação deve estar amparado por uma “engenharia social”, pois diferentes designs de mídia estimulam diferentes potenciais da natureza humana. De acordo com o autor, o design da web atual não foi proposto ao acaso, pois no início da década de noventa (1990) havia dezenas de formatos viáveis para a troca de informações digitalizadas em rede. Como consequência, a produtividade voluntária e anônima tornou-se uma commodity lucrativa para empresas do ramo.

Lanier (2010) esclarece que pequenas diferenças no design de determinado software podem se desdobrar em mudanças significativas na relação do sujeito com a interface e consequentemente nas relações interpessoais, sendo que o trabalho da criação em tecnologia da computação deve estar aliado a um compromisso social.

O autor enxerga que existe uma identidade entre o usuário de uma interface digital e a própria máquina, pois ao se relacionar com o computador o usuário passa a fazer parte, em alguma medida, da própria desenvoltura tecnológica, pois características de design estimulam diferentes possibilidades de interação. “Extensões se tornam as estruturas por meio das quais você se conecta ao mundo e a outras pessoas. Essas estruturas, por sua vez, podem mudar a forma como você vê a si mesmo e ao mundo” (LANIER, 2010, p. 20).

Desse modo, discussões sobre como a tecnologia pode atuar nas práticas das relações cotidianas deveriam ser pauta entre grupos de desenvolvedores e usuários das interfaces. Como dito anteriormente, o atual modelo de massa de dados livres foi uma opção dentre várias, quando designers formularam os padrões da Web 2.0 que lidamos nos dias de hoje. De acordo com Lanier

Como os computadores estão ficando mais poderosos em uma velocidade exponencial, os designers e programadores da tecnologia devem ser extremamente cautelosos ao fazer escolhas de design. As consequências de decisões pequenas e inicialmente irrelevantes muitas vezes se estendem para torná-las regras definidoras e imutáveis na vida (LANIER, 2010, p. 24).

Lanier (2010) esclarece que assim como as nuances das notas musicais foram solapadas pelos sons MIDI, tornando-as estruturas rígidas e sem complexidade, modelos impessoais de comunicação digital (como a Web 2.0) podem eliminar ambiguidades de pensamento, engessando significados e práticas sociais. Um bom exemplo seria o modo como as pessoas demarcam seus status de relacionamento em redes sociais como o Facebook, a partir de escolhas previamente determinadas pelo design do referido software, pode haver, por consequência, padronizações na forma como pessoas passam a enxergar seus relacionamentos fora do meio digital. Outro exemplo seria ainda o modo como as pessoas interagem com postagens dessa mesma rede social, marcando ou não se gostou do que lhes foi apresentado, induzindo deste modo a eliminações de ambiguidades ou flexibilidade de opiniões.

O atual modelo de fragmentação da informação em arquivos e troca deliberada dos mesmos também foi uma opção de design computacional que propiciou o acúmulo de dados em massa. Segundo Lanier (2010), além do fato de que arquivos são estruturas aprisionadas, no sentido de que necessitam sempre corresponder a um software compatível, filosoficamente expressam a ideia de que a

representação humana opera a partir de blocos divisíveis, passíveis de organização em um outro nível abstrato. O aprisionamento tecnológico suscitado pelo modelo de compartimento em arquivos é tamanho que a princípio parece impossível prospectar um novo modelo de partilhar informações em rede.

as filosofias correspondentes sobre como os seres humanos podem expressar significado estão tão arraigadas nos designs de software tecnologicamente aprisionados da Internet que podemos nunca ser capazes de nos livrar totalmente delas, ou mesmo nos lembrar de que as coisas poderiam ter sido diferentes. (LANIER, 2010, p. 29)

Para Lanier (2010), portanto, o design computacional, vinculado a aprisionamentos tecnológicos dos próprios designs precedentes, está diretamente relacionado às capacidades de se expressar em ambientes digitais; enquanto o aprisionamento tecnológico torna-se um empecilho para que se vislumbre novas propostas de configuração desses mesmos ambientes. Analisando a Web 2.0, sua natureza de livre compartilhamento de dados, Lanier (2010) aponta para o fato de que enquanto usuários transferem livremente dados em larga escala, a partir de modelos pré projetados para esse mesmo fim, conglomerados empresariais, com domínio sobre o fluxo da informação, a transforma em commodities. O usuário de redes sociais como Facebook, Twitter, Youtube, dentre outros centros digitais para compartilhamento de dados, operando livremente a partir de designs pré-concebidos para fluxo de dados em larga escala, torna-se o próprio produto ou o livre fornecedor da mercadoria que será comercializada: seus gostos estéticos, padrões de comportamento etc.

Enquanto os grandes conglomerados empresariais lucram com a livre transferência de dados, usuários trocam informações em volume e velocidade crescentes. Exemplo do que Lanier (2010) coloca em questão: conforme apontado anteriormente, a produção cultural em seus mais diversos campos, escritores, fotógrafos, músicos dentre tantos outros poderiam se beneficiar economicamente com o fluxo direto de dados autorais, mas se veem entregues a uma prática de livre transferência de dados anônimos, a gerar lucros massivos para empresas detentoras da informação em rede.

Interesses comerciais promoveram a ampla adoção de designs padronizados como os blogs, e esses designs incentivaram o uso de pseudônimos em pelo menos alguns aspectos de seus designs, como os comentários, em vez da orgulhosa extroversão que

caracterizou a primeira onda da cultura da web. Em vez de pessoas sendo tratadas como as fontes da própria criatividade, sites de abstração e agregação comercial apresentavam fragmentos anônimos de criatividade como produtos que poderiam ter caído do céu ou ter sido desenterrados do chão, ocultando as verdadeiras fontes. (LANIER, 2010, p. 33).

Portanto, para Lanier (2010), o design da Web 2.0, previamente determinado por empresas computacionais, tem efeitos causais sobre o modo como usuários passam a se relacionar com as demais esferas da vida. O termo “*affordance*”, originalmente utilizado por Gibson (1986), nos remete exatamente à influência que o design de determinado ambiente tem sobre o sujeito. “*Affordance*” advém do verbo em inglês “*to afford*”, ou seja, “permitir”, e está relacionado, segundo esse autor, à capacidade de um meio, em função de sua natureza física, orientar ou permitir específicas relações com o mesmo. Podemos dizer, em outras palavras, que “*affordance*” é o termo que designa de que modo características de um meio ou objeto influenciam, por si só, formas de interação. Como exemplo, Gibson (1986) nos remete à mais básica das estruturas, o solo: se uma superfície terrestre é quase horizontal (em vez de inclinada), quase plana (em vez de convexa ou côncava), suficientemente ampla (em relação ao tamanho do sujeito), e se o seu conteúdo é rígido (em relação ao peso do sujeito), conseqüentemente proporciona a ação de estar de pé. No entanto, se uma superfície de apoio, também com as quatro propriedades acima mencionadas, encontra-se à altura dos joelhos e acima do solo, proporciona ao sujeito a possibilidade de se sentar. Desse modo, segundo o autor, “diferentes layouts proporcionam comportamentos diferentes²⁵” (GIBSON, 1986, p. 128).

Gibson (1986) aborda a questão exemplificando ações no mundo animal. Ecologistas se referem ao meio em que determinado animal se encontra como “nicho”. Para Gibson, nichos são conjuntos de *affordances* orientando diferentes modos de ação para cada animal. “Em ecologia um nicho é uma configuração de características ambientais que são adequadas para um animal, no qual ele se encaixa, metaforicamente²⁶” (GIBSON, 1986, p. 129). De acordo com o autor, faz-se importante compreender que *affordances* estão relacionados a sentidos físicos, ao

²⁵ Different layouts afford different behaviors.

²⁶ In ecology a niche is a setting of environmental features that are suitable for an animal, into which it fits metaphorically.

contrário de valores e significados, os quais se supõe ser dados mentais e subjetivos.

Nos últimos mil anos, como todo mundo agora percebe, a própria face da terra foi modificada pelo ser humano. O layout das superfícies foi alterado, cortando, limpando, a partir de nivelamento, pavimentação e construção. Desertos naturais e montanhas, pântanos e rios, florestas e planícies ainda existem, mas estão sendo invadidas e remodeladas por layouts feitos pelo ser humano. Além disso, as substâncias do ambiente foram parcialmente convertidas, a partir dos materiais naturais da terra, em vários tipos de materiais artificiais, como (...) o concreto²⁷. (GIBSON, 1986, p. 129)

Assim como houve uma transformação no meio ambiente natural, modificando os níveis de relação do sujeito com o espaço e mesmo entre os próprios indivíduos, muitas vezes nos levando mesmo a desconsiderar de que modo o espaço é desenvolvido a determinar ações específicas, assistimos nas últimas décadas a uma transformação dos meios através dos quais se dão as mais básicas relações da linguagem. Assim como paisagens naturais foram gradualmente substituídas por nivelamentos e concretos, a fala, elemento da comunicação humana, tem sido substituída (em muitos e consideráveis momentos) por meios artificiais (tecnológicos) por onde se materializa a linguagem.

²⁷ In the last few thousand years, as everybody now realizes, the very face of the earth has been modified by man. The layout of surfaces has been changed, by cutting, clearing, leveling, paving, and building. Natural deserts and mountains, swamps and rivers, forests and plains still exist, but they are being encroached upon and reshaped by man-made layouts. Moreover, the substances of the environment have been partly converted from the natural materials of the earth into various kinds of artificial materials such as (...) concrete.

Capítulo 3. Colapso de Contexto e Teorias da Linguagem

De acordo com Rajagopalan (2012), inovações tecnológicas costumam implicar reações em duas medidas: entusiastas que as vislumbram de modo a enxergar suas positivities e benefícios, e céticos que anunciam os novos meios com olhares apocalípticos, como se responsáveis por grandes mudanças catastróficas para a humanidade.

Assim como no conto de Poe (1841/2015) que ilustra a teoria dos meios de McLuhan (1964/2005), para que possamos compreender de que modo novas mídias (caracterizadas pelos fluxos de big data) estimulam a linguagem, é oportuno compreender as características próprias do meio digital. As críticas aos apocalípticos e luditas, que enxergam nas novas tecnologias os sinais de catástrofes sociais, não deve de modo oposto nos direcionar a um entusiasmo acrítico sobre as mesmas. A CMC (comunicação mediada por computador) traz desafios aos estudos da linguagem que exigem, hoje, abordagem ambivalente, que não nos leve a compreender o fenômeno de modo apocalíptico, mas também que nos faça indagar sobre a efetividade das ideologias de liberdade e autonomia do usuário que permeiam a Web 2.0. Neste ponto, o aprisionamento tecnológico apontado por Lanier (2010) e o efeito anestésico das tecnologias sobre o sujeito em McLuhan (1964/2005) parecem estar em acordo.

Segundo Rajagopalan (2012), o internetês tem exacerbado opiniões que veem a nova linguagem como modelo precário da prática comunicacional, como se suas experimentações e inovações de certo modo representassem padrões de linguagem inferiores às línguas “consolidadas” e tradicionais. Opondo-se ao pessimismo dos tradicionalistas da linguagem, o linguista demonstra que a nova linguagem da internet, assim como o Português ou Espanhol, também se configura pela participação social com passar do tempo. O autor aponta para o internetês como elemento sintomático dos novos tempos, que carrega características próprias dos meios em que se realiza, como a velocidade no ato da comunicação. Como bem resume Rajagopalan (2012, p. 39), “objeto de vilipêndio ou, no mínimo, desconfiança, o internetês é, no entanto, nada mais que uma síntese de tudo aquilo que está acontecendo na esfera da comunicação”. De acordo com Rajagopalan

(2012), a questão que se coloca é se o internetês aponta para transformações significativas sobre o que temos conceituado por “língua”.

Questão interessante sobre a nova linguagem dos meios digitais é apontada por José Luiz Fiorin:

é notável você se debruçar sobre as abreviaturas: por que ‘BLZ’ dá ideia de beleza? Isso me remete à origem da escrita alfabética. Quando a escrita alfabética aparece entre os fenícios e depois em outras línguas semíticas, não havia vogais. As vogais são desnecessárias nas línguas semíticas dado o fato de um fenômeno chamado alternância vocálica, as vogais são colocadas conforme a classe de palavras, o radical etc, então elas são muito previsíveis, então eu não preciso de vogal. Quando um menino escreve beleza (BLZ), ele tem a intuição linguística que presidiu a primeira intuição da escrita alfabética: ‘eu abrevio com consoante’. (FIORIN, 2011, 10:15 min.)

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 anteriormente mencionada, possibilitou a utilização de mídias que se entropõem de modo significativo nas práticas comunicacionais. Um detalhe a ser considerado é que os computadores, dentre outras mídias digitais, incorporaram meios como a televisão e o rádio, com uma notável diferença, o usuário deixa de ser apenas o receptor passivo para assumir lugar ativo e dinâmico na utilização do meio. O que essa pesquisa questiona, a partir das teorias de McLuhan (1964/2005) e Lanier (2010), é até que ponto o ativismo do usuário é realmente exercido ou uma ideologia amplamente propagada e agregada às novas tecnologias.

Além de apontar que o conteúdo do internetês não é o conteúdo da mensagem que este carrega, mas a escrita convencional, Rajagopalan (2012) também exclui a difundida ideia de que o internetês reduz-se e abrevia-se como se refletisse as reduções e celeridade da fala. Fiorin (2011), à maneira de Rajagopalan, compreende o fenômeno do internetês como consequência natural das práticas de

linguagem na contemporaneidade, sem conservadorismos e purismos sobre formas linguísticas a ele antecedentes; no entanto, no que compete à compreensão de reduções e abreviações do internetês, opinam de modo divergente. Fiorin (2011) compartilha a divulgada ideia de que as características do internetês remetem às características da fala: “hoje eu tenho visto críticas muito severas ao internetês, ‘vai acabar com a linguagem’. (...). Eles escrevem abreviadamente, mas isso é a característica da instantaneidade do oral presente na escrita” (FIORIN, 2011, 09:30 min.). No entanto, de acordo com Rajagopalan (2012) “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, 2012, p. 40). A ideia que Crystal (2001) e Rajagopalan (2012) 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, 2012, p. 40).

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 mcluhiana, 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, enquanto elemento contextual indispensável, modifica a própria engrenagem contextual na qual se insere, 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, como veremos adiante.

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 significado”²⁸ (FREGE, 1960, p. 71). O contexto de situação, onde se consideram itens não linguísticos a exemplo de 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 (1916/1995), em que “o valor de qualquer termo que seja está determinado por aquilo que o rodeia” (SAUSSURE, 1916/ 1995, p. 135), sendo então compreendido também a partir do contexto de utilização da linguagem.

A exemplo de como significado atrelou-se ao conceito de contexto de enunciação, podemos ainda citar a visão de linguagem performativa de Austin (1962/1998). Ele contrasta inicialmente enunciados constativos (declarativos) a enunciados performativos; o primeiro vinculado à ideia de verdadeiro ou falso (conceitos ligados especialmente à filosofia analítica, da qual Frege é precursor, onde 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) estabelece três categorias para os atos de fala: locucionário, ilocucionário e perlocucionário. O primeiro, ato locucionário, referindo-se à produção de sons de determinado vocabulário, articulação de aspectos semânticos e sintáticos; o segundo referindo-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 referindo-se ao efeito do ato ilocucionário sobre um provável interlocutor. O seguinte exemplo de Austin (1962) ilustra a distinção acima apontada: locucionário (ele me disse “atire nela”); ilocucionário (ele me pediu para atirar nela); perlocucionário (ele me convenceu ou

²⁸ only in a proposition have the words really a meaning.

me levou a atirar nela). Para que os exemplos de Austin (1962) sejam evidentes sobre a distinção dos níveis dos atos de fala é necessário nos inclinar à distinção de “disse”, “pediu”, “ me convenceu, me levou”, onde o primeiro alude apenas ao fato do enunciador ter dito algo; o segundo nos leva a perceber que o ato do dizer possui uma intenção de que alguém realize determinada ação; e o terceiro de que o dizer realmente provocou uma ação por parte do interlocutor. No entanto, para que a produção e efeitos dos performativos se realizem, considera-se que haja determinantes contextuais.

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), onde uma mesma enunciação é direcionada a variados públicos on-line. 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 on-line, 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 etéreo”²⁹ (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. Ainda que o texto de Porter date de 1997, o anonimato na rede ainda é característica fundamental do meio, o autor argumenta sobre experiências semióticas de ambiguidade e equívocos de interpretação; “é a resposta coletiva a essa experiência de ambiguidade, o processo gradual de adaptação ao universo semiótico de álibis eletrônicos de livre flutuação que constitui a cultura da Internet³⁰” (PORTER, 1997, p. XII).

Percebe-se, portanto, duas perspectivas, podendo mesmo ser aparentemente contraditórias, sobre marcação de contexto no meio virtual. A primeira de Marwick & Boyd (2010) e Varis (2014), marcada pela superdiversidade³¹, onde 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. Duas noções que, entretanto, revelam a própria complexidade do elemento contextual em meios de big data, em outras palavras, como 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

²⁹ . 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.

³⁰ It is the collective response to this experience of ambiguity, the gradual process of adaptation to the semiotic universe of free-floating electronic alibis that constitutes the unique culture of the Internet.

³¹ 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.

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-line³². (PORTER, 1997, p. XII).

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

³² 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.

a uma sucessão interminável de exasperantes fantasias sedutoras evocativas do próprio ser³³ (PORTER, 1997, p. XIII).

Assim, as recentes mídias digitais de interação social revelam novas dimensões para a produção e compreensão da linguagem. Tomar o big data como objeto de investigação linguística torna-se, como dito anteriormente, um desafio metodológico para os modelos de pesquisa tradicionais que muitas vezes tendem a reduzir o objeto a partir de um corpus analisável. Isso não apenas pelo fato de se tratar de um objeto de grande volume e variedade de categorias, ou pela velocidade em que este se amplia no tempo, mas também pelo fato de que a análise linguística de um fenômeno está diretamente relacionada à compreensão de seu contexto de situação. Segundo Morin (2003), “todos os problemas particulares só podem ser posicionados e pensados corretamente em seus contextos; e o próprio contexto desses problemas deve ser posicionado, cada vez mais, no contexto planetário” (MORIN, 2003, p. 14).

A configuração textual em um ambiente digital realiza-se em modo de hipertexto; 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, 1916/ 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?

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 onde entram em campo padrões anormais de comunidade.

³³ 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.

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 melhor representada pelo prefixo “super”, indicando níveis mais avolumados de heterogeneidade cultural. Segundo o autor, superdiversidade é utilizado para que seja possível resumir em um termo fatores diferenciais variáveis entre imigrantes, como direitos e restrições legais, experiências de mercado de trabalho, distinções de gênero, perfis em geral, distribuição espacial, dentre outros; ainda para indicar novas conjunturas e interações das variáveis significativas ocorridas com os novos padrões migratórios para o Reino Unido nas últimas décadas (VERTOVEC, 2007).

De acordo com Blommaert (2012), a superdiversidade se revela a partir de duas forças diferentes e relativamente concomitantes. A primeira estaria vinculada ao fim dos processos políticos da Guerra Fria, que possibilitou maiores níveis de interação entre pessoas dos distintos blocos econômicos. Esse período se caracteriza por alterações nos padrões de mobilidade entre pessoas de diversos locais diferentes, um impulso de migrações que decorreu a partir de então. A segunda força reveladora de uma nova e mais ampla escala de diversidade estaria vinculada às transformações tecnológicas dos últimos anos, e o autor define, de modo direto, a internet como paradigma. 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-line³⁴. (VARIS; WANG, 2011, p. 72).

³⁴ 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

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; semelhantemente se torna nossa auto concepção, pois a vida faz-se mais complexa em sua organização e distribuição em ambientes on-line. Segundo o autor, a superdiversidade neste aspecto pode ser definida por: mobilidade, complexidade e imprevisibilidade, tornando-se um fenômeno social paradigmático, onde a linguagem se torna uma ferramenta importante para detecção de sua natureza difusa.

De acordo com Blommaert (2012), o conceito de “comunidades de fala”, onde se supõe que seus agentes linguísticos 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 independentes, estabelecendo relações estáveis e lineares (BLOMMAERT, 2012). No segundo caso, entretanto, as relações entre essas instâncias tornam-se dinâmicas e mesmo imprevisíveis; a estabilidade e linearidade entre as instâncias acima mencionadas cedem lugar a inúmeras exceções que revelam, deste modo, seus contextos móveis de produção de linguagem. Segundo Blommaert (2012), ambientes de superdiversidade (on-line e off-line) proporcionam, ao falante, recursos linguísticos também superdiversos, realizando-se a prática linguística a partir de contextos e formas semióticas de grande complexidade. A partir de então, 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.

Em *Superdiverse Repertoires and the Individual* (2012), Blommaert e Backus exploram o conceito de “repertórios linguísticos”; peça central do vocabulário da “sociolinguística da globalização” (BLOMMAERT, 2010), mas que, segundo os autores, é ainda pouco contemplado pelos estudos da linguagem. Em primeira instância, repertórios formaram a tríade dos recursos linguísticos, juntamente de competência e da própria comunidade falante. De acordo com Blommaert e Backus

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.

(2012), contextos de superdiversidade levam o agente da linguagem a uma série de trajetórias e tecnologias influenciando seus próprios repertórios linguísticos, os quais passam a se atualizar de modo fragmentado em retalhos de competências. Repertórios possuem origens biográficas a partir dos contextos vivenciados pelo(a) falante; sendo portanto itens válidos para a compreensão das subjetividades na contemporaneidade (BLOMMAERT; BACKUS, 2012). Para os autores, repertórios se definem pela totalidade de recursos linguísticos acessíveis ao falante em determinada comunidade, no entanto, os autores associam o sentido de repertórios à visão da etnografia tradicional, em que o conceito de comunidades de fala se realiza de modo isolado para fins de pesquisa, carecendo de atualizações em tempos de superdiversidade.

Para Blommaert (2012), contextos de superdiversidade nos orientam a questionar quais recursos estão disponíveis aos falantes de determinada língua ou línguas. O que é ou não compartilhado em sociedades pós Guerra Fria, em ritmos acelerados de migração e deslocamento, já não está tão claro como se supunha ao adotar conceitos fixos como “comunidades de fala”. Para o autor, aos processos migratórios crescentes ainda se associam as mídias de comunicação digital em massa, colaborando ainda mais para processos de indefinição dos recursos linguísticos disponíveis aos agentes da linguagem e conseqüentemente seus repertórios linguísticos. A internet nos últimos anos seria portanto responsável por elevados índices de diversidade, problematizando termos tradicionalmente utilizados com certa segurança pelos estudiosos da linguagem como “comunidade”, “minoría linguística” e o próprio por ele abordado no texto “repertório”. Padrões até então normalizados pelas práticas sociais tornaram-se anormais e fragmentados por experiências cotidianas com tecnologias móveis e de interação em massa.

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, 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*), possivelmente caótica (conhecido o fato de se compor em maioria de massa de dados não estruturados), e, portanto, hipóteses podem ser levantadas: estamos diante de um sistema de linguagens (ou um antissistema dado sua natureza

caótica); nesse sistema se produzem 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 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, 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). E ainda sobre “deslocamento”, conceito de Ernesto Laclau, Hall diz que “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 deste vasto campo de dados, poderemos encontrar provas da fragmentação do sujeito apontada por estes teóricos da pós modernidade? Para Varis e Wang (2011), a World Wide Web viabiliza novos canais de comunicação, possibilitando, a partir de então, 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 deste tecido complexo, big data, afetam os tradicionais padrões da linguagem não apenas no que corresponde à referenciação, mas também, e conseqüentemente, questões de valor linguístico, significado e performatividade.

Assim como dos contemporâneos espaços de superdiversidade cultural apresentam 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. 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’.³⁵” (NUNBERG, 2012).

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 como (...) água, manteiga, gás, ar, etc³⁶” (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 onde 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

³⁵ data isn't a plural noun like 'pebbles'. It's a mass noun like 'dust.

³⁶ 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.

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.

O termo “big data” deflagra uma exorbitante cadeia de referentes (dados digitais) em constante transformação e ampliação, como dito acima, 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.

Em Saussure (1916/ 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 descarta o fato da realidade extra-linguística influenciar a língua, a qual “sofre sem cessar a influência de todos [indivíduos]” (SAUSSURE, 1916/ 1995, p.88).

Logo, trata-se de pontos diferentes: a arbitrariedade linguística entre signo e referente, ficando as relações de valor estabelecidas de modo sistêmico; e a influência do extra-linguístico, tanto para a mutabilidade quanto para a imutabilidade da língua: “a língua forma um todo com a vida da massa social e esta, sendo naturalmente inerte, aparece antes de tudo como um fator de conservação” (SAUSSURE, 1916/ 1995, p. 88). Assim, mesmo que o princípio da arbitrariedade seja condição necessária para a estrutura linguística saussuriana, o linguista não exclui o papel da sociedade para as alterações ocorrentes na Língua.

Se se tomasse a língua no tempo, sem a massa falante – suponha-se o indivíduo isolado que vivesse durante vários séculos – não se registraria talvez nenhuma alteração; o tempo não agiria sobre ela. (SAUSSURE, 1916/ 1995, p. 92)

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, 1916/ 1995, p. 133). A língua seria, pois, um “sistema de valores” (SAUSSURE, 1916/ 1995, p. 130), e a massa amorfa do pensamento seria capaz de distinguir ideias de “modo claro e constante” (SAUSSURE, 1916/ 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, 1916/ 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, 1916/ 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. E como dito na citação de 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, 1916/ 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 (1964/2005), 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, hipertexto³⁷, o modo como práticas linguísticas têm se constituído on e off-line 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, e como visto anteriormente, 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, 1916/ 1995), extensões tecnológicas estão a todo instante interferindo nas realizações linguísticas de nosso tempo.

Segundo Saussure (1916/ 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, que ordenadas a partir de “relações baseadas no caráter linear da língua, que exclui a possibilidade de pronunciar dois elementos ao mesmo tempo” (SAUSSURE, 1916/ 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, 1916/ 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 este 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. A cena não é incomum nos dias de hoje:

³⁷ 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).

indivíduos frente a interfaces, conectados a determinada rede social, movendo publicações e conseqüentemente absorvendo informações a níveis acelerados e superficiais.

As segundas relações explicitadas por Saussure (1916/ 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, 1916/ 1995, p. 143). As relações de ordem associativa, segundo o autor, não se baseariam na extensão linear – como ocorreria nas associações sintagmáticas – essas 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, 1916/ 1995, p. 143).

Assim como as novas mídias atualizam a tradicional linearidade da fala, relações sintagmáticas, em função de hipertextos, as relações associativas mnemônicas também podem ser afetadas por estímulos decorrentes das tecnologias vigentes. 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.

Saussure (1916/ 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, 1916/ 1995, p. 145/146).

A inteligência artificial, ao expandir em volume, variação e velocidade dos índices mentais do sujeito, poderia, portanto, interferir nas associações mentais

realizadas pelo indivíduo, colaborando a partir de então nas relações de significado. A partir do momento em que 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 associações possíveis a interferir nas práticas de produção e compreensão semântica e conseqüentemente na performatividade linguística.

Enquanto um sintagma suscita em seguida a idéia 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, 1916/ 1995, p. 146).

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 conseqüências 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, conseqüentemente 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.

A questão da materialidade do enunciado é estruturada por Foucault a partir da distinção entre enunciado e enunciação. Uma enunciação se realiza a partir do momento em que o indivíduo emite um conjunto de signos, e quaisquer repetições deste mesmo enunciado tratará portando de uma nova enunciação. Ou seja, um mesmo enunciado pode ser repetido várias vezes que se tratará de enunciações únicas, pois o momento da enunciação nunca será o mesmo, trata-se, portanto, da questão da materialidade. A identidade de um enunciado dependerá de sua localização em um campo institucional; “a instituição constitui a materialidade do que é dito e, por isso, não pode ser ignorada pela análise arqueológica” (MACHADO, 2006, p. 152).

A compreensão da própria instituição torna-se vaga e indefinida em ambientes de superdiversidade digital. A localização institucional já não pode sequer ser determinada a partir de replicagens de enunciados em escalas que fogem ao controle de qualquer observador dos fatos linguísticos. É preciso portanto realizar revisões das teorias, paradigmas e conceitos que nos auxiliam na compreensão dos efeitos da linguagem. As transformações das duas últimas décadas não serão um fato isolado, mas sim uma constante que tende a se desdobrar cada vez mais com o passar do tempo. Como dito anteriormente, a proporção de aumento da carga digital

tem operado em escala geométrica de dois para um, Corrigan (2013), a considerar a relação tamanho por tempo transcorrido.

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 conseqüentemente o modo de percepção, apropriação e produção do significado; como bem coloca McLuhan (1964/2005), a mídia é a mensagem, e aqui apontamos que conseqüentemente atualizam-se repertórios e a performance via linguagem.

É válido para os estudos da linguagem que indaguemos 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, 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.

Se Jaron Lanier (2010) estiver na direção correta, a facilidade de acesso a dados em escala imensurável pode guardar um cavalo de troia para o presente e futuro da comunicação. Como explicitado anteriormente, big data tem se tornado a nova revolução tecno-científica, transformando o modo de operar pesquisas e a possibilidade de resultados. A partir do momento em que o presidente de uma das maiores potências econômicas mundiais se elege com apoio substancial de estruturas midiáticas governadas e pré-determinadas por empresas privadas de grande capital econômico, como anteriormente citada a reeleição de Barack Obama, e que empresas genéricas passam a comprar dados de empresas de tecnologia como estratégias de atuação nos mercados financeiros, devemos questionar

criticamente se os modelos de interface estão realmente configurados com propósitos de uma “feliz heterogeneidade” (VARIS & WANG, 2011).

De acordo com Zizek (1996), mercado e meios de comunicação estão dialeticamente interligados, fato diretamente relacionado à teoria de Guy Debord (1997) a respeito da “sociedade do espetáculo”. Segundo o filósofo e cientista social esloveno é o espaço onde a mídia estrutura antecipadamente nossa percepção da realidade, tornando-a indistinguível de sua própria imagem espetacularmente estetizada (ZIZEK, 1996). As mídias contemporâneas digitais, caracterizadas pela velocidade, volume e variação em proporção crescente, não fogem à lógica apontada por Zizek, e ainda operam em escalas de amplitude sem igual na recente história das sociedades de massa.

Cientistas da linguagem, especialmente advindos de países subdesenvolvidos ou em desenvolvimento, devem se questionar sobre formação discursiva e capacidade de atuação pragmática e mesmo possibilidades de alienação ideológica de indivíduos que a todo instante estão absorvendo uma grande quantidade de informação digitalizada e em constante transformação.

Capítulo 4. Práticas Linguísticas em Big Data

A lógica de McLuhan, 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.

Como tratado anteriormente, dados não estruturados predominam a grande massa digital acumulada nas últimas décadas. 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, um outro conteúdo. A estrutura computacional, hardware/software, conseqüentemente 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, 2012) e o cinema o conteúdo da televisão (MCLUHAN, 1964/2005). 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 conseqüências profundas à linguagem.

Assim como empresas dos mais diversos seguimentos se interessam por estruturar dados para diferentes finalidades, a exemplo da estratégia eleitoral de Barack Obama ou da otimização da performance de carros da Fórmula 1, a estruturação da massa de dados, obviamente em níveis consideravelmente menores, é prática comum ao usuário que navega pela malha de dados significantes do big data. Como no vórtice de Allan Poe, o suposto usuário, massageado pelo meio tecnológico, a princípio, não apreende de modo nítido e amplo a estrutura midiática em que se encontra submerso; o meio, enquanto extensão, nos cega para sua própria natureza (MCLUHAN, 1964/2005).

Efeitos das recentes mídias digitais sobre a linguagem podem ser percebidos por exemplo no internetês, como abordado anteriormente, possivelmente outros, também linguísticos, desdobramentos sociais, ideológicos, estéticos, ainda se revelarão no tempo e no próprio uso das novas configurações das linguagens. Entretanto, também a exemplo do vórtice, podemos reconhecer certos padrões a partir dos quais se faz possível apreender atualizações em níveis escalares da linguagem e assim determinar “estratégias de fuga”, modos para que as restrições ou possibilidades estimuladas pelo meio sejam sobrepujadas.

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 on-line.

O atual contexto de mobilidade e globalização rompe com as dinâmicas lineares e estáveis que caracterizavam comunidades de fala (BLOMMAERT, 2012a), dando lugar a espaços dinâmicos e mesmo imprevisíveis. A ancoragem semântica torna-se complexa e conseqüentemente os atos performativos. Resta-nos, a esse respeito, compreender de que modo tais instâncias da linguagem passam a corresponder à complexidade dos novos meios. A atual pesquisa não tem a ousadia de delinear ou definir tais modos, dado a amplitude do fenômeno, mas antes

deflagrar esses novos desafios aos estudos da linguagem e indicar possíveis direções para tal empenho.

A superdiversidade on-line se materializa digitalmente em uma massa estratificada, atingindo além de repertórios linguísticos, como indica Blommaert (2012a), 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 de então 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 consequentes associações. A complexidade dos “retalhos” de competência, de que nos fala Blommaert (2012), pode se aplicar à lógica da ancoragem semântica e performatividade. 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, onde “o valor de qualquer termo que seja está determinado por aquilo que o rodeia” (SAUSSURE, 1916/ 1995, p. 135), ou pelo viés pragmático, onde 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, onde 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 destes novos meios.

A supervariabilidade 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 on-line, estará diretamente vinculado a uma série de possibilidades 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 on-line), 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. Poderá mesmo, ao acaso, inferir, em uma simples conversa intermediada, palavras, assuntos e informações que lhe surgiram ao “acaso”, ou seja, pela lógica do uso de dados não estruturados que compõem a World Wide Web. Uma conversa sobre política em redes sociais pode, por exemplo, se enveredar para assuntos relacionados pelo próprio buscador on-line, como podemos ver no seguinte diálogo:

- (1) E vc viu o que foi falado ontem na câmara dos deputados?
- (2) Não, manda o link.
- (1) Aqui <https://www.youtube.com/watch?v=L0rEmiPsUMd>
- (2) Nossa, ela rasgou o verbo né? Vc já ouviu essa música? Lançaram ontem <https://www.youtube.com/watch?v=COOnSeCTeTUrAD?>
- (1) Ainda não tinha ouvido. Mas o que tem a ver com a fala da deputada?
- (2) Nada, é que apareceu o link sugerido aqui..
- (1) Ah tá... curti não.. perai que vou te mandar outra música dele que eu gosto muito..

Observe como as estruturas da própria mídia, opções de design previamente estabelecidas, interferem na prática comunicacional dos supostos usuários. A conversa que estava pautada em um determinado assunto é desviada por motivos externos, relacionados ao próprio design da interface por meio da qual a linguagem se configura. As mídias passam a atuar não apenas nas intermediações entre usuários, mas também no próprio ato de produção e leitura realizadas on-line.

A partir do exemplo acima, não seria estranho imaginar um leitor que ao buscar textos na internet sobre o vírus da dengue, em pouco minutos se encontre lendo algum artigo sobre a legalização das drogas ou uma receita de bolo de chocolate.

Como já explicitado, dados digitais não estruturados compõem a maior parte do conteúdo lançado em rede, ou nas palavras da artista da computação francesa Albertine Meunier “quand la data se manifeste, cela fait à Dada³⁸” (MEUNIER, 2014), referindo-se à linguagem anárquica da vanguarda dadaísta do início do século XX. 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?

No entanto, se considerarmos McLuhan (1964/2005), as implicações sociais e linguísticas do “mover-se” de usuários sobre uma malha indefinível e em constante expansão de dados não estruturados não estão claramente dispostas à nossa compreensão. Seguindo a lógica do autor, apenas o reconhecimento de padrões específicos dos meios poderá nos sugerir respostas, como a exemplo das considerações de David Porter (1997), sobre o modo como o usuário tende a preencher experimentalmente, como exercício semântico, as lacunas abertas pela fugacidade de contexto.

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 on-line.

A performance on-line, 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, 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 off-line, projeções do eu, possibilidades de se

³⁸ Quando dados se manifestam, eles fazem à maneira Dada.

performar identidades falsas, que ainda assim se moldam a partir de *affordances* próprios do meio.

A instantaneidade do meio digital aliada à constante permanência on-line (por via de celulares ligados à rede) e aos colapsos de contexto nos levam a inevitáveis ambiguidades de sentido, que interferem diretamente nas relações sociais em diversas instâncias. Nos dias de hoje, não por acaso, tornam-se comuns guias de etiqueta da linguagem on-line, sobre como evitar, por exemplo, mal-entendidos em e-mails profissionais. 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 estimulam 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 on-line está diretamente vinculada aos designs de softwares e conseqüentemente 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 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. As postagens em redes sociais, por exemplo, revelam fugacidade semelhante para quaisquer assuntos, dos mais superficiais aos mais profundos e urgentes. Qual a real diferença de conteúdos distintos como a denúncia de tráfico de animais silvestres, uma nova tradução do poeta Vladimir Maïakovisk e a publicidade de um novo restaurante japonês em sua cidade, quando todos eles são igualmente modelados com a mesma efemeridade e opções interativas de “curtir” ou “compartilhar” previstas no Facebook por exemplo? O que o Facebook, rede social de maior alcance social nessa década, realiza, é o solapar de conteúdos dos mais diversos temas, em alta velocidade, revelando a premissa de McLuhan (1964/2005) que tanto incomodou teóricos da comunicação como Umberto Eco (1965/1984): a mensagem mediada pela tecnologia é o meio e não o seu conteúdo vinculado.

A abordagem de Lanier (2010) no entanto não se configura de modo ludita ou semelhante. O autor parece mais um saudosista de outras formas de se conceber e implementar a atual web 2.0 do que um pessimista das transformações tecnológicas:

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, 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 ao 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 pelo uso da deep web³⁹, 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 a atos criminosos de racismo até a exposição de conteúdo íntimo pessoal, 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

³⁹ 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.

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. O coletivo incógnito estaria de algum modo mais próximo de supostas verdades supra-humanas, tornando perspectivas individuais dispensáveis; “textos sagrados tradicionais funcionam exatamente da mesma forma e apresentam muitos dos mesmos problemas” (LANIER, 2010, p. 38). Lanier (2010) parece buscar garantias para a escassez de marcadores contextuais, enquanto Porter (1997) parece mesmo interessado em como os usuários corresponderão a esse fenômeno.

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. Pois 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? Como no conto de Allan Poe, do marinheiro que se vê entre a vida e a morte no interior de um vórtice marítimo, apenas a compreensão dos padrões e estruturas do meio podem realmente nos retirar da cômoda massagem para uma superfície além e mais profícua. 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 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. Não existe mais um ou alguns contextos de produção. A supervariada 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 (1916/ 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”, (a exemplo do ritual de casamento, pois um juiz ou padre, mesmo com autoridade para efetivar a ação do casamento, não o pode realizar a qualquer momento, em qualquer ambiente), como esses conceitos passam a se aplicar em condições de anonimato e superdiversidade?

Qual relação fazer entre a impessoalidade da Web 2.0 enfatizada por Lanier (2010) e a performatividade dos atos de fala? 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 off-line 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, 1964/2005), 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.

O famoso experimento de Allan Turing, descrito por Lanier (2010), nos leva a indagações sobre as possibilidades de performatividade via linguagem em ambientes virtuais. Turing foi um matemático inglês, um dos pioneiros no surgimento da lógica computacional, considerado por muitos como o primeiro *cracker* da história, pois ele foi o responsável por decodificar o padrão de mensagens em código das forças militares alemãs na Segunda Guerra Mundial produzido pela Enigma, máquina electro-mecânica de criptografia. O famoso teste de Turing (Lanier, 2010) consiste em um jogo onde um humano e uma máquina são isolados para que um terceiro participante (humano) seja capaz de distinguir qual é qual, a partir de perguntas e respostas em texto. O teste de Turing, apesar de datar do ano de 1950, metaforiza uma das grandes questões da Linguística Computacional na atualidade, a tentativa de equivaler a capacidade da máquina para que essa torne-se apta às flexibilidade dos jogos da linguagem humana.

A linguagem de processamento natural, Natural Language Processing (NLP), termo da inteligência artificial e Linguística Computacional, se refere à tentativa de implementar em sistemas computacionais a capacidade de processar linguagens naturais (humanas). Parte considerável desse desafio tecnológico está voltada para a interação entre máquinas e humanos, ou seja, a busca de interfaces que sejam compatíveis à capacidade humana de processar dados (informação) e se comunicar. NLP está relacionada à faculdade de reconhecimento sintático e semântico, ao processamento de signos, reconhecimento de fala, capacidade de fazer referências e reconhecer contextos, interpretar e produzir enunciados.

A inteligência artificial abre novas possibilidades para pesquisas em performatividade linguística, como questões relacionadas a atos performáticos ilocucionários e perlocucionários. Quais os efeitos do falar pré-programado das

máquinas sobre nós? A cada dia mais a inteligência artificial faz-se presente na vida das sociedades de massa. *Chats* de ajuda na internet, por exemplo, são muitas vezes permeados de usuários artificiais programados para perguntas e respostas, a partir de um grande banco de dados de frases programadas. Outro exemplo de inteligência artificial presente no dia a dia de muitos é o aplicativo Siri da Apple, que utiliza de processamento de linguagem natural para funções diversas, como realizar perguntas, recomendar serviços e realizar ações.

A inteligência artificial, em especial ao que tange aos estudos de NLP, pode revelar interessantes aspectos da relação entre sujeito e linguagem. NLP, de acordo com Chaudhari et al (2015), inclui, em um sentido abrangente, o processamento de palavras, reconhecimento de fala, questões acerca de referências de contexto, planejamento de discursos e produção, além de questões estritamente de sintaxe e semântica. Para os autores, dizer que um sistema é capaz de compreender linguagem natural não quer dizer que o mesmo é capaz de produzi-la, mas exclusivamente de interpretá-la. Os problemas típicos da NLP no entanto se revelam também nas práticas linguísticas humanas em meios digitais. Chaudhari et al (2015) diz que ao utilizar o termo “linguagem natural”, ele se refere à linguagem humana em toda a sua confusão e uso variado.

Possivelmente são esses aspectos da linguagem que levaram Austin a refletir sobre a natureza da linguagem ordinária, repleta de infelicidades e ambiguidades em sua prática cotidiana. A lógica “confusa e variada” das línguas naturais, que são o grande desafio para a construção da inteligência artificial, relaciona-se com a ausência dos mais básicos elementos formadores de contexto, afinal máquinas, a priori, não registram situações de tempo ou espaço.

Segundo Chaudhari et al (2015), para Searle computadores não são, ou serão, capazes de realmente compreender a linguagem humana. Segundo Searle (1980), a lógica da inteligência artificial opera a partir de dois seguimentos definidos por ele como: inteligência artificial fraca e inteligência artificial forte, onde a primeira estaria relacionada ao computador como ferramenta aliada a mente e a segunda em que a máquina, a partir de um software adequado é, ela própria, uma mente capaz de cognições de linguagem. Para Searle (1980), na Inteligência artificial forte “os programas não são meras ferramentas que nos permitem testar explicações psicológicas, em vez disso, os programas são, eles próprios, as explicações” (SEARLE, 1980). Searle (1980) esclarece que não faz objeções à inteligência

artificial fraca, mas busca demonstrar que computadores serão sempre incapazes de operar a linguagem humana no sentido da inteligência artificial forte, porque apesar de dominarem a lógica da sintaxe por programação, nunca seriam capazes de operar os aspectos semânticos da língua.

O computador, (...), tem uma sintaxe, mas não semântica. Assim, se você digitar no computador "2 mais 2 igual?" será como digitar "4". Mas ele não tem idéia de que "4" significa 4 ou que isso significa alguma coisa. E a questão não é que falta-lhe algumas informações de segunda ordem sobre a interpretação de seus símbolos de primeira ordem, mas sim que seus símbolos de primeira ordem não têm qualquer interpretação⁴⁰. (SEARLE, 1980, p. 354)

Searle (1980) ilustra sua linha de pensamento, de que a máquina é incapaz de processar linguagem natural com a primazia humana, com o exemplo fictício do Quarto Chinês. Um falante de inglês no interior de um quarto, junto a ele uma caixa com vários escritos em chinês a ele incompreensíveis e um manual de instruções em inglês. Um segundo personagem, falante de chinês, do lado de fora do quarto, passa mensagens em chinês, para que ele, com o auxílio do manual, seja capaz de respondê-las (em chinês) e enviá-las de volta. O personagem situado fora do quarto pode, portanto, supor que o outro realmente está se comunicando em chinês, enquanto o conhecimento é apenas simulado.

No entanto, independente de inteligência artificial forte ou fraca, Searle (1980) não descarta a máquina como ferramenta para o "auxílio" da linguagem, que podemos em alguma medida comparar às extensões de McLuhan. Interfaces de comunicação digital tornam-se, em maior ou menor grau, parte considerável das práticas linguísticas das sociedades de massa. O Google Translate, por exemplo, é uma ferramenta que tem permitido inovadoras escalas de leitura e conversação nos ambientes virtuais.

A experiência linguística⁴¹ tratada a seguir, executada com o Google Translate, é uma produção autoral de arte e tecnologia realizada durante a pesquisa aqui apresentada. Buscou-se operar de modo simbólico o uso de linguagem em

⁴⁰ The computer, to repeat, has a syntax but no semantics. Thus, if you type into the computer "2 plus 2 equals?" it will type out "4." But it has no idea that "4" means 4 or that it means anything at all. And the point is not that it lacks some second-order information about the interpretation of its first-order symbols, but rather that its first-order symbols don't have any interpretations.

⁴¹ A experiência de multitradições foi realizada como trabalho final da disciplina Arte e Tecnologia. Trata-se de uma dentre oito outras peças produzidas. A referida disciplina foi cursada pelo Programa de Pós-Graduação em Cultura Visual, sob docência de Edgar Franco, e integra os créditos necessários para a conclusão do mestrado no Programa de Pós-graduação em Letras e Linguística da Faculdade de Letras da UFG a que se vincula a presente pesquisa.

mídias de big data, através de ferramenta on-line de tradução. A experiência possui caráter artístico e nos interessa antes por sua lógica operacional do que por sua validade estética, que não está sendo tratada aqui. Operou-se da seguinte maneira: um texto de entrada em língua portuguesa, a conhecida Oração do Pai Nosso, na plataforma de tradução *Google Translate*, seguido de um processo de multitradições, indo do alemão ao iorubá, passando, nesse intercurso, por 11 línguas diferentes, e um texto de saída, também em língua portuguesa, como resultado⁴². O processo completo da experiência pode ser conferido no Anexo 3 deste documento.

O texto de saída apresenta completa distorção do texto de entrada, ambos em língua portuguesa, gerando enunciados absolutamente diversos do primeiro. O resultado da experiência revela antes a fragilidade ou maleabilidade da linguagem (mesmo quando operada via máquina) do que um erro no próprio software, mesmo porque a multitradição consecutiva de um mesmo texto não é o propósito de uso da ferramenta em questão. O texto de saída, após inúmeras traduções entre línguas, apresentou a seguinte configuração:

Suporta e no ar fica implícito.
 Benvindo ao reino de seu pai. No entanto, no ar, em terra.
 Para nós hoje, para ser.
 Perdona-nos os nossos pecados, e que você tenha criado.
 Não vamos permitir o tribunal. Somos livres do mal.
 Artistas

Mais uma vez o que nos interessa aqui não é a criação do objeto em si, mas simbolicamente sua metaforização do uso de linguagem em meios digitais de superdiversidade. O texto de saída releva de modo obviamente simbólico e radical dois aspectos aqui apontados sobre linguagem e mídias de big data. Primeiro que um enunciado linguístico digital, dada sua natureza de registro junto à massa de dados, está sujeito a possibilidades de manipulação nos mais diversos níveis de intervenção, como sequentes replicagens, reedições e mesclagens. Seu efeito performativo está sempre sujeito a desdobrar-se massivamente em uma malha de enunciações e, conseqüentemente, outros efeitos desconhecidos. Segundo, que a ancoragem do seu significado, considerando Porter (2012), realiza-se

⁴² A seqüência de idiomas utilizados foi: português, alemão, tailandês, hebraico, japonês, hindi, coreano, maori, hauçá, tailandês (novamente), russo, indonésio, iorubá, português (novamente).

contextualmente escassa, levando o leitor a identificar experimentalmente a unidade do texto a partir de outros marcadores pessoais. Consequentemente também seu efeito perlocucionário estará sujeito a determinantes de contexto experimentalmente pessoais. Nas palavras de Porter (2012) a “leitura criativa”, situação em que um texto contextualmente fendido nos leva a preencher suas lacunas com referências experimentais, pois não há marcadores de tempo ou de espaço, nem mesmo um sujeito enunciador. A experiência metaforiza, em nível escalar exacerbado, as lacunas de significado, que nos aponta Porter (1997), e porque não dizer também a impessoalidade da máquina que inquieta Lanier (2010).

Seja “ferramenta”, enquanto inteligência artificial fraca (SEARLE, 1980) ou mídias como “extensões da mente” (MCLUHAN 1964/2005), não há razões para supor que softwares de inteligência artificial não estejam a cada dia mais presentes na vida de usuários nas próximas décadas.

Se considerarmos que o colapso de contexto se faz presente na ancoragem do significado, também se fará na execução de atos de fala. 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 nem sequer imaginados antes.

A postagem em rede social Facebook do Anexo 4, do cartunista Ivan Cabral (2016), exemplifica os resultados da pesquisa de que o meio digital não apenas é a mensagem, mas é a ancoragem e a performatividade da linguagem em rede. Perceba o modo como as permissividades do meio, através de edição de imagem e

difusão em massa do enunciado, levam a uma ambiguidade sobre o desejo ilocucionário e desdobramentos de efeitos perlocucionários. O enunciador diz “alguém simplesmente pegou uma charge minha e fez as alterações que quis e deixou minha assinatura”. Observe agora, no Anexo 5, uma postagem em mesma rede social, apenas da imagem adulterada, vinculada pela página oficial de uma jornalista. A imagem, no entanto, foi replicada, em uma primeira onda, mais de 70 mil vezes (dados obtidos no dia 20 de março de 2016). A postagem original do autor original, na mesma data, não ultrapassava 100 replicagens, mesmo após a denúncia divulgada pelo mesmo. O cerne da questão é: a escala de alcance social da imagem adulterada, que além da assinatura carrega o traço como elemento de identidade do artista, é consideravelmente maior.

Assim, o que se diz on-line está, de diferentes maneiras, sujeito a forças multiescalares do próprio meio, e o cartunista é colocado na posição de ter dito o que de fato não disse, em outras palavras, de vincular socialmente sua autoria a efeitos perlocucionários contrários ao desejo ilocucionário. Assim que a imagem é potencialmente replicada nos meios, não há mais controle sobre as resultantes performativas da mesma.

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 hacker 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? Talvez os hackers sejam um bom exemplo contemporâneo para a metáfora mcluhiana do marinheiro de Allan Poe. Parece lógico pressupor que os três níveis apontados por Austin (1962/1998), para o ato performativo, podem refletir os meios de big data. O locucionário, como a exemplo do internetês e suas características de celeridade e abreviações, e também os níveis ilocucionário e perlocucionário.

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, 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 consequentes *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), consequentemente a depender das distintas camadas e alcance social que esses venham a atingir. Assim, os efeitos de um ato performativo on-line podem se mostrar radicalmente distintos com base no que *affordances* digitais dispõem ao falante/usuário.

Novamente a metáfora mcluhiana do vórtice de Allan Poe parece pertinente, pois, considerando a linha de pensamento acima expressa, níveis de ilocução estão vinculados à performatividade on-line desempenhada pelo falante. Assim como o marinheiro de Allan Poe necessita compreender a lógica do comportamento dos objetos no interior do vórtice para sobrepujá-lo, em tempos de mídias de massa, o falante usuário pode precisar o efeito de seus atos performativos com base em suas próprias desenvolvimentos on-line; como a exemplo do hacker há pouco mencionado. Conclui-se, a partir de 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 (1963/2005) 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.

Considerações finais

Na presente pesquisa investigou-se conceitualmente a relação entre linguagem e big data, considerando referenciais teóricos de áreas distintas do conhecimento. Afinal, o objeto em questão demanda abordagem em diversas direções, especialmente atenção para o que nos tem a dizer teóricos dos estudos de tecnologia; a contribuição desses autores foi substancial para atingir os objetivos pretendidos. Não se pretendeu a uma construção metodológica linear, o que não significa que não se operou a partir de uma metodologia consciente. O objeto pesquisado, big data, dada sua natureza superdiversa, nos moveu a recortes referenciais também diversos. Informações de diferentes orientações foram apresentadas, propondo-se, na medida em que se revelaram, associá-las coerentemente e de modo recorrente, objetivando deflagrar a relação entre novas mídias e determinados conceitos da linguagem. Buscou-se antes partir do objeto analisado em direção a uma gradual construção metodológica, ao invés de empreender uma metodologia precedente, para que o objeto pudesse de algum modo adequar-se a um modelo tradicional de investigação.

Como já mencionado, a presente pesquisa, para além de seu compromisso epistemológico com os estudos linguísticos, pretende-se como contribuição teórica para pesquisa e produção de arte em novas mídias. “O que” ou “como fazer” via linguagem artística nos novos meios são questões que se desdobrarão à presente investigação. Os meios digitais poderiam guardar para o artista, que agora se vê conectado através de big data, novas possibilidades e funções? Como se relaciona a produção de arte em meios de “fragmentação” de contexto? Segundo Augusto de Campos (2015), considerando McLuhan:

o papel consciente das artes em nosso tempo é explorar e criar a consciência do novo contexto criado pelas novas tecnologias, (...), o circuito elétrico cria um campo simultâneo que é totalmente envolvente e invisível no que tange a sua operação psíquica. (...) um mundo em que tudo acontece ao mesmo tempo, em que não há sequência ou sucessão de eventos. (...), o artista deve responder evidenciando que esse contexto simultâneo se tornou, ele próprio, uma obra de arte. (CAMPOS, 2015, p.307).

Inclinar-se aos colapsos de contexto evidenciados pelos meios de big data pareceu um viés pertinente para que de algum modo fosse possível relacionar teoricamente escalas tecnológicas e escalas linguísticas. A atualização das escalas

de contexto geradas pelas novas tecnologias, o colapso dos modelos tradicionais de situação podem afetar diferentes realizações da linguagem, sendo provável que os níveis de interferência das novas mídias ocorram em camadas tão amplas quanto a própria natureza de volume desmedido e em crescimento constante.

O big data, ao alterar toda a perspectiva contextual, nos coloca frente a um desafio: medir seus efeitos sobre a linguagem, dada a natureza desmedida e superdiversa dos novos meios. As escalas de tempo, espaço e amplitude social estão em níveis amplos e inacessíveis ao usuário comum, no entanto, torna-se indispensável indagar a respeito de suas dinâmicas e consequentes efeitos sobre a linguagem. Ao final da pesquisa, considera-se que a alteração de escalas de contexto, decorrentes das próprias características e padrões dos meios, deflagram atualizações no modo como compreendemos a natureza da linguagem. No que compete à ancoragem do significado, o usuário tende a preencher as ausências de marcadores contextuais de modo experimental, e no que diz respeito à performatividade linguística em ambientes virtuais, o meio, enquanto suporte para a produção da linguagem, está diretamente relacionado ao “o que se faz ao dizer on-line”. Apontou-se para a contemporânea figura do hacker como exemplo de que níveis de ilocução estão relacionados à desenvoltura do usuário e suas consequentes capacidades de ultrapassar as permissividades dos *affordances* computacionais.

Quais padrões seriam reconhecíveis ao se coletar quantidades massivas de dados linguísticos digitalizados, à maneira da própria ciência do big data? Poderia a língua nos revelar aspectos não reconhecíveis na análise metodológica tradicional? Empreendimentos dessa ordem obviamente carecem de investimentos de aparelhamento de hardwares, a exemplo de supercomputadores ou saídas economicamente mais viáveis como computadores agrupados (clusters⁴³). Segundo Hilbert e Lopez (2011), atualmente agrupamentos de consoles de vídeo game são utilizados como substitutos de supercomputadores, para fins científicos e outras tarefas que exigem o uso de dados computacionais em grandes proporções.

O desafio de compreensão do objeto big data coloca em cheque nossos tradicionais modelos de investigação, onde a experimentação metodológica, sua livre construção ao longo do processo, parece a saída mais pertinente. Afinal, como

⁴³ Cluster é o termo utilizado para se referir a agrupamento de computadores com finalidade de processar dados em grande quantidade.

insiste em afirmar Feyerabend (1977), o princípio que não inibe o avanço científico é: “tudo vale” (FEYERABEND, 1977, p. 27). A posição radical de Feyerabend (1977) é relevante para a pesquisa de objetos que estão no centro da recente implosão das mídias digitais, apenas porque ainda não sabemos como lidar com sua natureza complexa. A análise dos objetos das sociedades marcadas pela superdiversidade e acelerada transformação tecnológica exige assim a atualização de modelos de pesquisa.

A intensidade do referido fenômeno social, no breve percorrer de duas décadas, estimula alternativas experimentais para o empreendimento da investigação. A partir de então, a aproximação de características apresentadas por teóricos das novas mídias a conceitos próprios dos estudos da linguagem pareceu um procedimento sensato para avançarmos na compreensão dos efeitos da recente digitalização das práticas de linguagem. Ainda foram utilizados autores que afinam a relação das duas instâncias, como Blommaert (2012/2013), Rajagopalan (2013), Varis e Wang (2011), dentre outros.

A estrutura do big data, se é que assim podemos considerá-la, mostra-se como uma malha em nível escalar de zettabytes (10^{21}) com elementos em diferentes níveis de desestruturação; é neste aspecto que se metaforiza a expressão “navegar” para o uso diário em redes de informação. Uma estrutura (ou desestrutura) de dimensões desmesuradas, porém de natureza clástica – massiva como são massivos os sacos de grãos; dados em massa são anexados em rede a cada segundo, de modo não estruturado, ocasionando contextos em colapso.

Não se pretendeu, todavia, afirmar que a malha desestruturada de dados elimine os efeitos performativos ou irresponsabilize o sujeito pelo que se diz, apesar de que o anonimato, fruto dos 3Vs do big data, diariamente é utilizado para que se amore (nas mais distintas situações) a responsabilidade do ato de dizer em ambientes virtuais. Pretende-se antes reconhecer que enunciados on-line estão sujeitos às forças multiescalares do próprio meio, que o efeito do ato performativo on-line está em certa medida atrelado às condições que o meio nos apresenta. Em outras palavras, que a performatividade da linguagem pode-se realizar em escalas de big data, dado o fato de que os contextos virtuais também incorporam as características do meio.

Tornamo-nos de certo modo usuários andróides à medida em que a máquina, ou mais precisamente seus designs, nos moldam as escalas contextuais

da linguagem, como apontado por McLuhan (1964/2005) a tecnologia torna-se uma extensão da mente. É o ponto em que a própria língua passa a assumir características próprias do meio artificial, como no caso do internetês, e conseqüentemente aspectos semânticos e de performatividade também estarão sujeitos a causas e formas dos meios. A atualização multiescalar das novas mídias incide em atualizações nas escalas da linguagem. O meio tornou-se multimidiático e isso deverá trazer conseqüências significativas nas formas e modelos de interação. O *affordance* dos meios digitais pode moldar escalas da linguagem nos três níveis da performatividade austiniana, tanto para as práticas da língua, quanto para a linguagem das artes em novas mídias.

Referências

- ABREU, Giovanna; NICOLAU, Marcos. A estética do anonimato na Deep Web: a metáfora das máscaras e do homem invisível aplicada ao “submundo” da internet. *Cultura Midiática* Ano VII, n. 12. 2014. Disponível em: <http://periodicos.ufpb.br/index.php/cm/article/view/19746/10908>. Acesso em: 02 out. 2015.
- ANDERSON, Chris. *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, 2008. Disponível em: http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory. Acesso em: 12 jun. 2015.
- AUSTIN, John. *How to do things with words*. 1. ed. Oxford/New York: Oxford University Press, 1962.
- AUSTIN, John. Performativo-constativo. In: OTTONI, Paulo Roberto. *Visão performativa da linguagem*. Campinas : Editora da UNICAMP, 1998. (Viagens da Voz). (Apêndice). p. 107-144.
- BABINET, Gilles. L'ère numérique, un nouvel âge pour l'humanité. Disponível em: <http://www.gqmagazine.fr/uploads/freepages/Ere-numerique.pdf>. Acesso em: 13 mai. 2015.
- BARBOSA, Miranda. *McLuhan e as críticas*. 2011. Disponível em: <http://www.intercom.org.br/papers/nacionais/2011/resumos/R6-2386-1.pdf>. Acesso: 05 dez 2015.
- BAUDRILLARD, Jean, *Simulacros e Simulação*, trad. Maria João da Costa Pereira, Lisboa: Relógio D'água, 1991.
- BLOMMAERT, Jan. *The sociolinguistics of globalization*. Cambridge: Cambridge University Press, 2010.
- BLOMMAERT, Jan. *Ethnography, superdiversity and linguistic landscapes: chronicles of complexity*. Bristol: Multilingual Matters, 2012.
- BLOMMAERT, Jan; BACKUS, Ad. *Superdiverse Repertoires and the Individual*. Tilburg Papers in Culture Studies, paper 24. Tilburg University, 2012.
- BLOMMAERT, Jan. *Citizenship, Language, and Superdiversity: Towards Complexity*. *Journal of Language, Identity & Education*, 12:3, 193-196, 2013
- CAMPOS, Augusto de. *Poesia Antipoesia Antropofagia & Cia*. 1ª ed. São Paulo. Companhia das Letras. 2015.
- CHAUDHARI, Rajashri et al. *Conversion of Natural Language into Java Programme*. In: *International Journal of Advanced Research in Computer Science and Software Engineering*. Volume 5. Jalgaon, 2015.
- CORRIGAN, David. *Integrating and governing big data*. New York: IBM Corporation. 2013
- CRYSTAL, David. *Language and the Internet*. Cambridge: Cambridge University Press. 2001.

CUKIER, Kenneth. Data, data everywhere. *The Economist*, 2010. Disponível em: <http://www.economist.com/node/15557443>. Acesso em: 12 nov. 2015.

DEBORD, Guy. A sociedade do espetáculo: comentários sobre a sociedade do espetáculo. Rio de Janeiro: Contraponto, 1997.

ECO, Umberto. *Apocalípticos e Integrados*. Tradução: Andrés Boglar. Espanha. Editora Lumen. 1984.

FEYERABEND, Paul. Contra o método. Rio de Janeiro. Livraria Francisco Alves Editora S.A. 1977.

FREGE, Gottlob. *The foundations of Arithmetic*. Editora Basil Blackwell & Mott. 1960.

FIORIN, José Luiz. Norma Culta e Variedade Linguística. 2011. Disponível em: <https://www.youtube.com/watch?v=pWvuF0U9zv4>. Acesso em 08 ago. 2015

GIBSON, J.J. *The ecological approach to visual perception*. New Jersey: Lawrence Erlbaum, 1986.

GOOGLE. *Dengue Trends*. Disponível em: <http://www.google.org/denguetrends/about/how.html>>. Acesso em: 22 jun 2015.

GOOGLE. *Inside Google Translate. 2010*. Disponível em: https://www.youtube.com/watch?v=_GdSC1Z1Kzs>. Acesso em: 10 jan. 2015.

GERCHMAN, Rubens. *Uma arte Brasileiro/Latino-Americana*. In: GULLAR, Ferreira. PEDROSA, Mario. (Orgs.). *Arte Brasileira Hoje*. Rio de Janeiro: Paz e Terra, 1973, p. 163-166

HALL, Stuart. *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*. 1991. Disponível em: <http://annabellelukin.edublogs.org/files/2013/08/Halliday-1991-Context-in-language-education-yc7ta4.pdf>> Acesso em: 27 fev. 2015.

HILBERT, Martin; LOPEZ, Priscila. The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*. 2011.

HILBERT, Martin. *Big Data for Development: From Information to Knowledge Societies* (January 15, 2013). Disponível em: <http://ssrn.com/abstract=2205145>>. Acesso em: 19 abr. 2015.

JAKOBSON, Roman. *Linguística e Comunicação*. Tradução: Izidoro Blikstein e José Paulo Paes. Editora Cultrix. 2007.

JESPERSEN , Otto. *The philosophy of grammar*. New York. Norton. 1965

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

LAYCOCK, Henry. *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.

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

MACHADO , Roberto. *Foucault, a ciência e o saber*. Rio de Janeiro. Ed. Jorge Zahar. 2006

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

MANOVICH , Lev. *Trending: The Promises and the Challenges of Big Social Data*, 2011. Disponível em: <<http://manovich.net/content/04-projects/067-trending-the-promises-and-the-challenges-of-big-social-data/64-article-2011.pdf>>. Acesso em: 03 ag. 2015

MARWICK, Alice; BOYD, Danah. *I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience*. New Media & Society 13 (1), 114-133. 2010.

MCLUHAN, Marshall. *Understanding Media: The Extensions of Man*. New York: McGraw-Hill Book Company, 1964.

MCLUHAN, Marshall. *A galáxia de Gutenberg: a formação do homem tipográfico*. São Paulo, Editora Nacional, Editora da USP, 1972.

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

MCLUHAN, Marshall. *O despertar de McLuhan*. Dir.: Kevin McMahon. 2002. Disponível em: <https://www.youtube.com/watch?v=BkmTYGGq0cQ>. Acesso em 10 nov. 2015.

MEUNIER, Albertine. Manifeste Datadada #1. Datadada. 2014. Disponível em: <http://www.datadada.net/>. Acesso em: 8 dez. 2015.

MORAES, Mauricio. *Big Data Ajudou Obama a ganhar as eleições*, 2005. Disponível em: <<http://info.abril.com.br/noticias/ti/big-data-ajudou-obama-a-ganhar-eleicoes-15012013-25.shl>>. Acesso em 12 mai. 2015.

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

NEEDHAM, Jeffrey. *Disruptive Possibilities: How Big Data Changes Everything*. Sebastopol: O'Reilly Media. 2013

NUNBERG, Geoffrey. *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, Tim. *What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. O'Reilly Media. 2007

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

POE, Edgar. *The Complete Works of Edgar Allan Poe: a descent into the maelstrom*. Disponível em: <http://www.unilibRARY.com/ebooks/Poe,%20Edgar%20Allan%20-%20Complete%20Works.pdf>. Acesso em 10 nov. 2015.

PORTER, David. 1996a. Introduction. In Porter (ed.), xi–xviii. *Internet culture*. New York and London: Routledge. 1997.

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

RAJAGOPALAN, Kanavillil. *O Austin do qual a linguística não tomou conhecimento e a linguística com a qual Austin sonhou*. Cad.Est.Ling., Campinas, (30):105-115, Jan./Jun. 1996

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

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

SEARLE, John. *Minds, brains and programs*: in *The behavioural and Brain Sciences*. Cambridge University Press, 1980.

SOUSA, Janara; CURVELLO, João; RUSSI, Pedro. *100 anos de McLuhan*. Organizadores: Janara Sousa, João Curvello, Pedro Russi. Casa das Musas, Brasília, 2012.

SZALAY, Alexander. *Data-intensive discovery in science: the fourth paradigm*. 2012. Disponível em: <http://www.cra.org/ccc-old/nitrdsymposium-szalay.php>. Acesso em 08 ago. 2015

THAKARE, Girish; DESHMUKH, Shilpa. *Big Data: An Analysis of large data*. International Journal of Innovative Research in Science & Engineering, 2012.

VARIS, Piia, WANG, Xuan. *Superdiversity on the Internet: A Case from China*. *Diversities* Vol. 13, No. 2. University of Tilburg, 2011.

VARIS, Piia. *Digital Ethnography*: in *The Routledge Handbook of Language and Digital Communication*, 2014.

VERTOVEC, Steven. *Super-diversity and its implications*. *Ethnic and Racial Studies*, v.30, n. 6, 1024-1054, 2007.

WU et al. *Data Mining with Big Data*. Department of Computer Science. University of Vermont. 2012

YAU, Nathan. *Beautiful Data: The Stories Behind Elegant Data Solutions*. Organização de SEGARAN, Toby; HAMMERBACHER, Jeff, 2009. Disponível em: <http://vample.com/ebooks/Oreilly.Beautiful.Data.Jul.2009.pdf>. Acesso em 29 set. 2015.

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

ZIZEK, Slavoj. *Um mapa da ideologia*. Tradução: Vera Ribeiro. Rio de Janeiro. Editora Contraponto. 1996

Anexo 1

MANIFESTE DATADADA #1

(Albertine Meunier)

c'est ma data là
mais ceci n'est pas de votre ressort
la data c'est GAFA
la datum c'est Badaboum
ma data est une poule aux oeufs d'or
Lucky la Data qui ne peut être vendue
Pâquerettes à Gogo
quand la data se manifeste,
cela fait à Dada
à hue et à dia!
quand la data se vide de sens
elle change de direction
je répète
la data c'est GAFA
la datum c'est Badaboum
la data, c'est comme un ange
Unisex et pur esprit
c'est la data là
c'est la data là
qu'on se le dise !

*GAFA: Google, Apple, Facebook, Amazon

Anexo 2. Vórtice



Figura 1: Ilustração de Harry Clarke para o conto de Allan Poe (1919).

Fonte: <http://50watts.com/Harry-Clarke-Illustrations-for-E-A-Poe>

Anexo 3. Experiência em Google Translate. nº02.

Texto Original: Oração do Pai Nosso.

Português:

português inglês espanhol Detectar idioma ▾

Pai nosso, que estais no céu, santificado seja o vosso nome.
Venha a nós o vosso reino, seja feita a vossa vontade, assim na terra como no céu.
O pão nosso de cada dia nos dai hoje.
Perdoai-nos as nossas ofensas assim como nós perdoamos a quem nos tem ofendido.
E não nos deixeis cair em tentação, mas livrai-nos do mal.

Amém

🔊 🗨️ ▾ 🔊

Alemão:

português inglês espanhol alemão - detectado ▾

Vater unser, der du bist im Himmel, geheiligt werde dein Name.
Kommen Sie zu dein Reich, dein Wille geschehe auf Erden wie im Himmel.
Das tägliche Brot gib uns heute.
Vergib uns unsere Schuld, wie auch wir vergeben unserm Schuldigern.
Und lassen Sie uns nicht in Versuchung, sondern erlöse uns von dem Bösen.

amen

🔊 🗨️ ▾ 🔊

Tailandês:

português inglês espanhol tailandês - detectado ▾

พระบิดาของเราผู้สถิตในสวรรค์ที่สักการบูชาเป็นชื่อของเจ้า
มาถึงอาณาจักรของเจ้าเจ้าจะทำบนโลกในสวรรค์
ขนมปังในชีวิตประจำวันให้เราวันนี้
ยกโทษให้เราละเมิดของเราในขณะที่เรายกโทษให้ลูกหนี้ของเรา
และให้เราไม่ได้เข้าไปในการทดลอง แต่ให้พ้นจากความชั่วร้าย

สาธุ

🔊 🗨️ ▾ 🔊

Hebraico:

português inglês espanhol hebraico - detectado ↕

אבה, שבשמיים, יתקדש שמך.
 תבוא מלכותך, "עשה הרצון על פני כדור הארץ, בגן העדן.
 לחם לתת לנו היום.
 סלח לנו על חטאים כפי שאנו סולחים לחיבנו.
 ותן לנו לא לידי ניסיון. אבל לגאול אותנו מכל רע.

אמן

Japonês:

português inglês japonês Detectar idioma ↕

アバ、天の父よ、汝の名である神聖。
 あなたの王国来るが、天国で、地球上で行われます。
 パンは私たちに、今日を与える。
 我々は債務者を許すようにわたしたちの罪を赦し。
 とではない誘惑に私たちをみましよう。しかし、悪からお救いください。

アーティスト

あ

Hindi:

japonês português inglês hindi - detectado ↕

अब्बा, स्वर्ग में पिता, पवित्र तेरा नाम है।
 तेरा राज्य आए, लेकिन स्वर्ग में, यह पृथ्वी पर किया जाएगा।
 हमारे लिए, आज रोटी दे।
 हम हमारे पापों को क्षमा और देनदार की अनुमति है।
 प्रलोभन में हमें एक और नहीं है चलो। हालांकि, बुराई से बचा कर दीजिये।

कलाकार

Ä

Coreano:

japonés portugués inglés coreano - detectado

아바, 하늘에 계신 아버지, 거룩한 네 이름입니다.
 아버지의 나라가 올하지만, 하늘에, 그것은 지상에서 수행됩니다.
 우리를 위해, 오늘 빵을 제공합니다.
 우리는 우리의 죄 지은자를 용서하고 사용할 수있다.
 우리없는 유혹에 보자. 그러나, 악에서 구 하옵소서하시기 바랍니다.

예술가

Maori:

japonés portugués inglés maori - detectado

E Apa, e i roto i te rangi, ko to koutou ingoa tapu.
 Haere mai ki te kingitanga o tona papa, engari i roto i te rangi, kei te mahi i taua mea i te whenua.
 Hoki a matou, i tenei ra, ki te whakarato i ngā taro.
 Murua o matou hara, me te tangata kia hanga.
 E kore matou e tukua te whakamatautauranga. Heoi, whakaorangia matou i te kino.

Artist

Haucá:

japonés portugués inglés haucá - detectado

Tallafi, da kuma a cikin sama, ka mai tsarki suna.
 Barka da zuwa mulkin mahaifinsa, amma a cikin iska, yin ta a cikin kasa.
 Domin mu, a wannan rana, don samar da abinci.
 Gāfarta mana zunubanmu, kuma ya gina.
 Ba za mu bari fitina. Amma kubutar da mu daga sharri.

Artist

Tailandés

japonés portugués inglés tailandés - detectado

สนับสนุนและในท้องฟ้าคุณเป็นของบริสุทธิ์
ยินดีต้อนรับเข้าสู่อาณาจักรของบิดาของเขา แต่ในอากาศบนบก
สำหรับเราในวันนี้สำหรับอาหาร
ทรงโปรดยกบาปของเราและท่านได้สร้าง
เราจะไม่ยอมให้การพิจารณาคดี แต่เราส่งมอบให้จากความชั่วร้าย

ศิลปิน

Indonésio

japonés portugués inglés indonésio - detectado

Mendukung dan di langit, Anda adalah kudus.
Selamat datang di kerajaan ayahnya. Namun, di udara, di darat
Bagi kami hari ini, untuk makan.
Mengampuni segala dosa kita, dan telah Anda buat.
Kami tidak akan membiarkan pengadilan. Kami memberikan dari kejahatan.

artis

Iorubá

japonés portugués inglés ioruba - detectado

Se atileyin ati ninu awon orun, ti o ba wa ni mimo.
Kaabo si ijoba baba re. Sugbon, ni air, lori ile
Fun wa loni, lati je.
Dariji wa ese wa, ati awon ti o ba ti da.
A yoo ko gba laaye kootu. A fi jise lati buburu.

olorin

Português (final)

indonésio português ioruba ▼ Traduzir

Suporta e no ar, fica implícito.
Bem-vindo ao reino de seu pai. No entanto, no ar, em terra
Para nós hoje, para ser.
Perdoa-nos os nossos pecados, e que você tenha criado.
Não vamos permitir que o tribunal. Somos livres do mal.

Artistas

☆ ☰ ✎ 🔊 ✓

Anexo 4

Ivan Cabral adicionou 2 novas fotos.

Ontem às 00:06 · 🌐

Essa web é mesmo uma terra sem lei. O pessoal acha que pode dizer tudo, fazer tudo, que não dá em nada... Alguém simplesmente pegou uma charge minha e fez as alterações que quis e deixou minha assinatura. Abaixo as charges original e adulterada:



Anexo 5

**Rachel Sheherazade**

18 de março às 22:17 · 🌐

👍 Curtir Página

Pessoal acabaram de jogar um artefato que dispersou uma multidão que fazia uma manifestação a favor do Lula e Dilma. Segue foto tirada no momento exato desse ato que ainda não foi apurado o numero de vítimas.



👍 Curtir 💬 Comentar ➦ Compartilhar

👍 😄 ❤️ 91 mil

Principais comentários ▾

79.700 compartilhamentos

3,4 mil comentários

Fonte:

<https://www.facebook.com/RachelSheherazade01/photos/a.404251809679414.1073741831.304120173025912/829676117136979/?type=3&pnref=story>