

Reescrita sentencial baseada em traços de personalidade

Personality-dependent sentence rewriting

Georges Basile Stavrakas Neto
Universidade de São Paulo (USP)
georges.stavrakas@gmail.com

Ivandr  Paraboni 
Universidade de S o Paulo (USP)
ivandre@usp.br

Resumo

Sistemas de Gera o de L ngua Natural (GLN) s o centrais para o desenvolvimento de comunica o humano-computador realista e psicologicamente plaus vel que n o recorra ao uso de texto fixo ou predefinido, fazendo uso de uma ampla gama de estrat gias para modelar alguma forma de varia o estil stica. Entre estas estrat gias, o uso de modelos computacionais da personalidade humana emergiu como uma alternativa popular na  rea. Neste contexto, o presente trabalho apresenta um modelo de GLN do tipo texto-para-texto (ou reescrita sentencial) para o portugu s que leva em conta, al m da senten a a ser reescrita, informa es sobre a personalidade de um locutor-alvo de interesse. Mais especificamente, o modelo transforma a senten a de entrada em outra na qual certas formas lexicais s o substituídas por termos mais adequados ao tipo de personalidade-alvo fornecido. Resultados sugerem que as senten as geradas com base em personalidade s o mais pr ximas das que seriam produzidas por um locutor humano com as caracter sticas de personalidade fornecidas do que seria poss vel sem acesso a essa informa o, e abrem assim caminho para futuros estudos de gera o de l ngua natural personalizada em portugu s.

Palavras chave

gera o de l ngua natural, personalidade

Abstract

Natural Language Generation (NLG) systems are central to the development of psychologically plausible human-computer communication that does not rely on canned text, and which makes use of a wide range of strategies to model some stylistic variation. Among these, the use of computational models of human personality has emerged as a popular alternative in the field. In this context, the present work presents a text-to-text (or sentential rewriting) GLN model for Portuguese that takes into account, in addition to the sentence to be rewritten, information about the personality of a target speaker of interest.

More specifically, the model transforms the input sentence into another one in which certain lexical forms are replaced by terms more suited to a certain personality type. Results suggest that personality-based generation produces sentences that are closer to those produced by a human speaker with those personality traits than what would be possible without access to this information, thus paving the way for future studies of speaker-dependent natural language generation in Portuguese.

Keywords

natural language generation, personality

1. Introdu o

Sistemas de gera o de l ngua natural (GLN) produzem texto a partir de uma representa o lingu stica ou n o-lingu stica fornecida como entrada, e s o centrais para o desenvolvimento de comunica o humano-computador realista e psicologicamente plaus vel que n o recorra ao uso de texto fixo ou predefinido. Aplica es de GLN incluem, por exemplo, a gera o de resumos textuais a partir de dados obtidos de unidades de terapia neonatal intensiva (Portet et al., 2009), hist ricos de pacientes e relat rios de enfermagem (Hunter et al., 2012; di Eugenio et al., 2014; Jordan et al., 2014), cartas personalizadas de cessac o do tabagismo (Reiter et al., 2003), boletins de previs es meteorol gicas (Reiter et al., 2005), di logos e textos narrativos (Walker et al., 2011a,b), poesia (Zhang & Lapata, 2014), legendas de imagens (Karpathy & Fei-Fei, 2015; Xu et al., 2015) e muitas outras.

Podemos em princ pio distinguir dois tipos de sistemas de GLN (Reiter & Dale, 2000): sistemas do tipo dados-para-texto, que recebem uma representa o n o lingu stica como entrada (e.g., dados num ricos provenientes de sensores) e produzem texto como sa da (e.g., relat rios do mercado financeiro), e sistemas do tipo texto-para-texto, que recebem como entrada um texto j  redigido em l ngua natural (possivelmente por um autor humano) e produzem uma vers o modifi-



cada deste mesmo texto como saída (e.g., uma versão simplificada, resumida, traduzida, adaptada estilisticamente etc.) No presente trabalho será abordado exclusivamente o segundo tipo de sistema, tratando da questão da reescrita de sentenças preexistentes com base em critérios a serem discutidos a seguir.

Sistemas de GLN podem, em princípio, produzir sempre o mesmo texto de saída para uma determinada entrada. No entanto, sistemas que visam gerar texto de maneira mais natural ou semelhante ao humano geralmente implementam uma ampla gama de estratégias para modelar variação estilística, incluindo o controle de parâmetros que afetam o tamanho ou complexidade das sentenças, uso de pausas, ênfases, pontuação, escolha lexical e muitos outros. Uma visão detalhada destas estratégias é apresentada por Mairesse (2008). Entre estas possibilidades, o uso de modelos computacionais da *personalidade humana* emergiu como uma alternativa popular na área (Mairesse & Walker, 2010).

De especial interesse para o presente trabalho, consideraremos o uso do modelo dos Cinco Grandes Fatores (CGF) da personalidade humana, ou *Big Five* (Goldberg, 1990). O modelo CGF baseia-se no pressuposto de que as diferenças de personalidade são reveladas pela maneira como os indivíduos se expressam em língua natural e, em especial, nas escolhas lexicais realizadas. Dada a sua motivação linguística, o modelo CGF tem sido aplicado a uma ampla gama de estudos tanto na interpretação (Plank & Hovy, 2015; Álvarez-Carmona et al., 2015; González-Gallardo et al., 2015) como na geração de língua natural (Mairesse & Walker, 2011; Herzig et al., 2017) e, embora haja várias outras teorias de personalidade em discussão (em especial, o modelo MBTI de Briggs Myers & Myers (2010)), a quase totalidade dos estudos em Ciência da Computação tende a ser baseado no modelo CGF.

O modelo CGF compreende cinco dimensões fundamentais da personalidade: *Extroversão*, *Agradabilidade*, *Conscienciosidade*, *Neuroticismo* e *Abertura à experiência*. Cada dimensão CGF é associada a um escore que representa a intensidade com que um indivíduo manifesta cada aspecto da personalidade. Assim, por exemplo, um escore abaixo da média (de acordo com uma população de interesse) para Extroversão indica um indivíduo introvertido, enquanto que um escore acima da média indica um indivíduo verdadeiramente extrovertido. Escores de personalidade podem ser obtidos por diversos métodos, incluindo a aplicação de questionários específicos (John et al., 2008).

Para melhor apreciar o papel das dimensões de personalidade CGF na produção humana de língua natural, considere-se a tarefa de produzir uma descrição textual simples de uma determinada cena como na figura 1, extraída da base GAPED de Dan-Glauser & Scherer (2011).



Figura 1: Um exemplo de imagem extraída da base GAPED.

Em uma situação desse tipo, diferentes locutores poderiam produzir um grande número de descrições alternativas do mesmo contexto. Por exemplo, um indivíduo com maior escore para a dimensão *Agradabilidade* pode descrever o personagem da cena como “uma linda criança”, enquanto que um indivíduo de menor escore pode descrevê-lo como “um pirralho sujo”, dentre muitas outras possibilidades. De forma análoga, podemos também conceber aplicações computacionais em que um sistema não apenas descreve uma imagem seguindo um padrão fixo ou pré-definido, mas o faz impondo um estilo específico ditado por uma personalidade-alvo de interesse.

A variação humana em GLN é um tópico recorrente na pesquisa na área (Viethen & Dale, 2010; Mairesse & Walker, 2011; Ferreira & Paraboni, 2014) e será também o tema do presente trabalho. De forma mais específica, propõe-se desenvolver um modelo de reescrita sentencial para o português que leva em conta, além da sentença a ser reescrita, informações sobre a personalidade de um locutor-alvo de interesse. Modelos deste tipo são potencialmente úteis, por exemplo, quando procura-se obter maior engajamento do leitor (ou usuário), como no caso de aplicações em educação, no relacionamento com clientes ou consumidores e outras.

O modelo a ser apresentado objetiva transformar uma sentença fornecida como entrada em outra na qual certas formas lexicais são substituídas por termos mais adequados ao tipo de

personalidade-alvo fornecido¹. Neste contexto, a hipótese a ser investigada é a de que as sentenças geradas com base em personalidade sejam mais próximas das que seriam produzidas por um locutor humano com as características de personalidade fornecidas do que as que seria possível obter utilizando-se um modelo de geração de texto sem acesso à informação de personalidade.

O restante deste artigo está organizado da seguinte forma. Na Seção 2 são discutidos trabalhos relacionados à área de GLN baseada em personalidade. Na Seção 3.2 é descrito o cópús a ser empregado neste trabalho, a anotação semântica realizada e os modelos de lexicalização e reescrita sentencial propostos. Na Seção 4 estes modelos são avaliados tanto sob uma perspectiva de aprendizado de máquina como orientada à tarefa de reescrita propriamente dita. Os resultados destas duas formas de avaliação são apresentados nas Seções 4.1 e 4.2. Finalmente, a Seção 5 apresenta algumas conclusões e trabalhos futuros.

2. Trabalhos relacionados

Estudos de GLN baseado em personalidade são de modo geral escassos, e não foi possível identificar na literatura iniciativas deste tipo (ou sistemas de GLN minimamente completos) voltados para o idioma português. Nesta seção são discutidos assim alguns estudos da área de GLN baseada em personalidade dedicados ao idioma inglês.

Estudos de GLN são fortemente relacionados ao trabalho seminal de Mairesse (2008) e suas extensões. Todos estes estudos abordam a geração de texto em língua inglesa, e nenhum deles trata da tarefa de geração texto-para-texto (como no caso da reescrita sentencial tratada no presente trabalho), mas apenas da geração de dados-para-texto. Estes estudos são brevemente discutidos a seguir.

O estudo de Mairesse & Walker (2010) abordou uma ampla gama de decisões de geração que podem ser influenciadas por um perfil de personalidade-alvo. O trabalho se concentra na geração de língua natural do tipo ponta-a-ponta, ou dados para texto, apresentando um sistema de GLN configurável que produz recomendações textuais de restaurantes. O sistema resultante — chamado PERSONAGE — é treinado com base em dados anotados com informação de personalidade proveniente do modelo CGF, e os textos assim produzidos foram reconhecidos por juízes

¹O presente foco na questão da escolha lexical é motivado pela natureza lexical do próprio modelo CGF, mas outras direções possíveis de pesquisa são discutidas na seção 5.

humanos como refletindo certos traços de personalidade específicos. A lexicalização no sistema PERSONAGE é realizada para cada palavra de conteúdo no texto com base em três parâmetros: a frequência do léxico, o tamanho da palavra e a força verbal (por exemplo, “sugerir” seria mais fraco do que “recomendar”). Esses parâmetros fazem uso do conhecimento obtido a partir de vários recursos lexicais on-line (por exemplo, WordNet e VERBOCEAN) e de contagens de frequência de corpus.

Diversos estudos subsequentes foram desenvolvidos como extensões do sistema PERSONAGE. Estes estudos incluem uma série de melhorias na arquitetura do sistema original, e acrescentam suporte aos outros quatro traços de personalidade do modelo CGF (além de Extroversão) (Mairesse & Walker, 2011). Além disso, a arquitetura PERSONAGE original foi também aplicada a outros domínios, incluindo a geração personalizada de fofocas (Khosmood & Walker, 2010), diálogos sobre jogos de computador (McCormick, 2012), escrita criativa (Lukin et al., 2014), narração de histórias (Bowden et al., 2016), geração de gestos (Aly & Tapus, 2016) e geração de *feedback* a clientes (Herzig et al., 2017).

No caso do idioma português, não foram identificados estudos de GLN baseados em personalidade exceto dois trabalhos prévios relacionados ao projeto atual para tratamento de outras questões de pesquisa, e ambos utilizando o próprio cópús *b5* a ser discutido na próxima seção. Estes dois estudos, detalhados por Paraboni et al. (2017) e Lan & Paraboni (2018), que abordam a questão da relação entre personalidade e a tarefa de seleção de conteúdo na geração de expressões de referência (GER). Esta tarefa, que é um subcomponente da arquitetura de sistemas de GLN, utiliza algoritmos próprios para decidir, nos estágios iniciais da geração de texto, qual o conteúdo semântico a ser expresso na forma de descrições definidas (e.g., “a mão esquerda do menino”). Uma visão geral da área, que está fora do escopo do presente trabalho, é apresentada por van Deemter (2016).

3. Método

Nesta seção são descritos o cópús empregado neste trabalho (Seção 3.1), a preparação de dados realizada (Seção 3.2) e os modelos de lexicalização e reescrita sentencial propostos (Seção 3.3).

3.1. O corpus b5

Neste trabalho será utilizado o corpus *b5* (Ramos et al., 2018) de textos em português brasileiro rotulados com escores de personalidade do modelo *CGF* relativos aos seus autores. O corpus foi utilizado em estudos prévios de geração de texto baseada em personalidade (Paraboni et al., 2017) e inferência de traços de personalidade a partir de textos e caracterização autoral (dos Santos et al., 2017; Hsieh et al., 2018; Silva & Paraboni, 2018b,a). Detalhes desta organização são discutidos por Ramos et al. (2018).

Os textos a serem utilizados são provenientes de duas bases (ou subcorpus) do corpus *b5* denominados *b5-text* e *b5-caption*, em ambos os casos coletados em uma série de experimentos presenciais envolvendo participantes humanos engajados na tarefa de descrição de imagens. Para este fim, foram empregados 10 estímulos visuais extraídos da base GAPED (Dan-Glauser & Scherer, 2011) de imagens classificadas por valência e significância normativa, e designadas de modo a despertar diferentes graus e tipos de emoção. Um exemplo de imagem deste tipo foi apresentado na Figura 1 da Seção 1.

O corpus contempla duas subtarefas de descrição de imagens: em versão detalhada (em texto multi-sentencial) e em versão resumida (na forma de uma sentença única, ao estilo de uma legenda que resume o conteúdo da imagem). Um exemplo de descrição textual resumida para a imagem anterior, tal qual observado em *b5-caption*, poderia ser simplesmente “*Trabalho infantil*”. A versão detalhada desta mesma descrição, tal qual observada em *b5-text*, é apresentada a seguir.

Criança de cerca de 5 anos, trabalhando injustamente, despejando terra num balde por algum motivo para ajudar alguém. Cenário de pobreza no qual ao seu fundo existe uma casa construída de sacos plásticos e alguns tijolos.

O corpus *b5* contém descrições textuais de 10 contextos visuais como o da figura anterior. Todas as imagens foram descritas por um grupo de 151 participantes, havendo portanto 151 versões longas e outras 151 versões curtas de cada uma —todas potencialmente influenciadas pelas diferentes personalidades de seus autores. Todos os 1510 textos longos e 1510 legendas encontram-se rotulados com escores de personalidade *CGF* obtidos por meio de inventários de personalidade dos participantes que os escreveram.

A porção *b5-text* dos dados contém 84463 *tokens* e a porção *b5-caption* contém 4896. Os textos completos da porção *b5-text* são usados para fins de treinamento dos modelos a serem discutidos. A porção *b5-caption* foi utilizada para fins de teste tal qual descrito na Seção 4.

3.2. Preparação dos dados

A geração de texto com variação lexical decorrente de personalidade (ou outros fatores) requer a identificação de conceitos de nível semântico a serem lexicalizados. Assim, o uso do corpus *b5* descrito na seção anterior demanda uma tarefa de anotação semântica com o objetivo de estabelecer um mapeamento entre conceitos e suas lexicalizações possíveis. Esta atividade é descrita a seguir.

Os modelos de reescrita a serem propostos tratam da lexicalização de três tipos de conceitos semânticos representados nas cenas do corpus *b5-text* (Ramos et al., 2018): conceitos a serem realizados como substantivos (como a maior parte das entidades das cenas do corpus, pessoas, objetos etc.), adjetivos (e.g., características concretas e abstratas das entidades da cena, como cores, emoções etc.) e verbos (e.g., ações que as entidades presentes na cena executam). Estes conceitos são aqui denominados CRS, CRA e CRV, respectivamente, em referência à forma de realização superficial (substantivo, adjetivo ou verbo) que cada um assume no texto, e foram escolhidos por estarem entre os mais facilmente observáveis nas imagens do corpus.

Dado que um determinado conceito pode ser lexicalizado de muitas formas, o primeiro passo desse estudo consistiu em computar todos os conceitos referenciados no corpus *b5* e então associar cada conceito a uma lista de formas lexicais possíveis. Para este fim, foi atribuído um rótulo único a cada conceito, aqui designado na forma \$nome, onde ‘nome’ é um termo representado em inglês como forma de distingui-lo de suas formas lexicais em português. A seguir apresentamos um exemplo deste tipo de notação, em que o CRS \$bucket é associado a três formas lexicais possíveis.

\$bucket → {‘balde’, ‘recipiente’, ‘pote’}

O procedimento para cômputo dos mapeamentos de conceitos para lexicalizações difere para cada tipo de conceito (CRS, CRA ou CRV) por motivos discutidos a seguir. Em todos os casos, a anotação dos dados foi realizada por dois avaliadores, e nos poucos casos de ambiguidade observados, foram decididos por um terceiro.

No caso dos CRS, uma vez que os textos do *córpus b5* são fortemente apoiados nos elementos visuais de cada cena, optou-se por modelar apenas a lexicalização de conceitos que representam elementos referenciáveis das cenas, incluindo por exemplo personagens e objetos físicos, entidades representando indivíduos ou grupos, e entidades representando um todo ou suas partes, mas excluindo-se conceitos mais abstratos ou de caráter retórico presentes no discurso (e.g., problemas, questões, diferenças etc.) Assim, criou-se inicialmente uma lista de todos elementos visuais que poderiam ser realizadas na forma de substantivos, atribuindo-se um rótulo identificador a cada um. Por exemplo, para a Figura 1 anterior poderiam ser enumerados, dentre outros, os CRS \$bucket, \$child, \$tent para os objetos balde, criança e barraca, respectivamente.

A seguir, o *córpus* foi analisado sintaticamente com uso da ferramenta PALAVRAS (Bick, 2000), escolhida pela conveniência de uso de sua versão online, e por de modo geral apresentar bom desempenho para o Português brasileiro. Foram selecionados todos os substantivos com um mínimo de cinco ocorrências nos textos, um limite mínimo foi adotado para evitar casos muito infrequentes, e que não seriam passíveis de aprendizado automático. Estes substantivos foram então manualmente associados a cada um dos CRS enumerados no passo anterior examinando-se as imagens quando necessário para fins de desambiguação, objetivando agrupar todos os sinônimos referentes a cada um dos conceitos. Um exemplo de mapeamento de um CRS para uma lista de substantivos foi ilustrado no exemplo acima, para o caso do CRS \$bucket.

Neste processo, alguns novos CRS que não tinham sido originalmente listados foram descobertos a partir dos próprios substantivos, e a lista de CRS foi atualizada de acordo quando pertinente. O resultado desta atividade foi a produção de uma série de mapeamentos de CRS para suas formas lexicais possíveis, contendo 151 CRS associados a 191 substantivos únicos com cinco ou mais ocorrências no *córpus*.

Uma vez definidos os CRS possíveis, o *cômputo* dos CRA foi realizado de forma semi-automática utilizando-se a informação sintática disponível. De forma mais específica, os adjetivos associados a cada substantivo ou pronome² de interesse no *córpus* foram manualmente associados a cada um dos CRS definidos no passo

anterior. Assim, \$black-bucket é o CRA que representa a cor escura de um balde, e \$black-skin é o CRA que representa a cor escura da pele de um indivíduo. A definição de conceitos distintos associados a cada tipo de objeto a que se referem é motivada pelas observações de que seus conjuntos de lexicalizações possíveis podem não coincidir totalmente, e de que mesmo sendo idênticas estas lexicalizações podem ser empregadas de forma diferente para cada conceito por locutores com determinados tipos de personalidade.

Cada grupo de adjetivos assim identificado recebeu um rótulo único representando seu significado e contendo indicação do CRS ao qual se refere. Um exemplo de mapeamento de um CRA para uma lista de adjetivos é ilustrado a seguir, representando a qualidade do objeto balde de possuir uma cor escura.

$$\text{\$black-bucket} \rightarrow \{\text{'escuro'}, \text{'preto'}, \text{'negro'}\}$$

Seguindo-se este procedimento, 114 CRA foram associados a 191 adjetivos únicos com cinco ou mais ocorrências no *córpus*.

No caso dos CRV, que representam ações que podem estar relacionadas a múltiplas entidades, optou-se simplesmente por computar todos os verbos do *córpus* e agrupá-los em conjuntos de sinônimos (ou *synsets*) obtidos a partir do dicionário TeP 2.0 (Maziero et al., 2008). A cada *synset* foi então atribuído um rótulo arbitrário único. Um exemplo de mapeamento de um CRV para uma lista de verbos é ilustrado a seguir.

$$\text{\$v1} \rightarrow \{\text{'cansar'}, \text{'debilitar'}\}$$

Como resultado, foram identificados 198 CRV associados a 352 verbos únicos com cinco ou mais ocorrências no *córpus*.

Finalmente, dado o objetivo de explorar lexicalizações alternativas (i.e., de acordo com uma personalidade-alvo) dos CRS, CRA e CRV identificados, apenas conceitos com mais de uma lexicalização possível foram mantidos. A Tabela 1 sumariza os três conceitos mais frequentes de cada tipo, o número de vezes que suas lexicalizações são utilizadas no *córpus*, e suas possíveis lexicalizações.

Os mapeamentos entre conceitos (CRA, CRS ou CRV) e suas possíveis lexicalizações serão tomados por base para o modelo de reescrita sentencial baseada em personalidade discutido na próxima seção.

²A identificação de antecedentes anafóricos foi feita de forma manual. Para abordagens computacionais, ver por exemplo Paraboni & de Lima (1998); Cuevas & Paraboni (2008).

Tipo	ID	Instâncias	Lexicalizações
CRS	\$child	1399	criança, menino, garoto, bebê
	\$scene	761	imagem, cena, cenário, foto
	\$soil	437	terra, barro, areia
CRA	\$white-human	989	claro, branco
	\$black-human	823	escuro, preto, negro
	\$pale-object	674	pálido, desmaiado, claro
CRV	\$v1	8024	haver, ser, estar, existir
	\$v2	4592	querer, relevar, estar, assentar
	\$v3	3144	ser, servir

Tabela 1: Três conceitos mais frequentes de cada classe (CRS, CRA e CRV).

3.3. Modelo proposto

Nesta seção é apresentado o modelo de reescrita sentencial que é o foco principal deste trabalho. O modelo recebe como entrada uma sentença s em Português e os escores das cinco dimensões de personalidade P de um locutor-alvo de interesse, e produz como saída uma nova sentença s' que é uma variação de s em que certas palavras foram substituídas por sinônimos que seriam tipicamente empregados por um locutor de personalidade P . Neste modelo, a tarefa de escolha lexical é implementada com uso de métodos de aprendizado de máquina supervisionado (ou classificadores). Estes dois componentes —classificadores e o reescrevedor sentencial propriamente dito— são ilustrados na Figura 2, e serão detalhados separadamente nas próximas seções.

3.3.1. Classificadores de escolha lexical

O componente principal do modelo de reescrita é formado por um conjunto de 10 classificadores (um para cada contexto do cópús $b5\text{-text}$) que recebe como entrada um conceito semântico representado por id e escores de uma personalidade-alvo P , e retorna a lexicalização w de id que seria tipicamente empregada por locutores de personalidade P para descrever o conceito id . Por exemplo, dado $id = \$child$ e um perfil P de maior agradabilidade, a classificação pode sugerir a forma lexical $w = \text{'criança'}$. Com menor agradabilidade, por outro lado, o classificador pode sugerir, e.g., $w = \text{'pirralho'}$.

O uso de classificadores individuais para cada contexto do cópús é motivado pela observação de que cada imagem é efetivamente um domínio distinto, e conceitos superficialmente semelhantes (como ‘mulher’) possuem significados e lexicalizações distintas em cada um (i.e., cada imagem representa uma mulher diferente, que em alguns

casos pode ser descrita como ‘moça’ e em outras não). Assim, cada subconjunto do cópús $b5\text{-text}$ precisa ser tratado com um cópús distinto.

Para cada palavra w representando um adjetivo, substantivo ou verbo com 50 ou mais ocorrências no cópús, foram criadas instâncias de aprendizado supervisionado como segue. Primeiramente, cada palavra w é pesquisada nos mapeamentos de conceitos para palavras (cf. seção anterior) e o conceito id correspondente é recuperado. Em paralelo a isso, o autor do texto é pesquisado na base de participantes do cópús $b5$ e seus escores de personalidade P são também recuperados. Assim, a instância de aprendizado é formada pelos atributos id e P , e a classe a ser apreendida é a palavra w . Havendo mais de um mapeamento para a palavra em questão, múltiplas instâncias de aprendizado são criadas. Finalmente, todas as instâncias relativas a conceitos com menos de cinco ocorrências foram descartadas.

Este procedimento resultou em um conjunto de dados contendo 35k instâncias de aprendizado de 734 lexicalizações possíveis para todos os CRA, CRS e CRV com frequência mínima no cópús e que apresentavam alguma variação lexical³. A distribuição geral de classes por tipo de conceito é ilustrada na Tabela 2.

Tipo de Conceito	Instâncias de aprendizado	Lexicalizações possíveis
CRS	10428	191
CRA	5741	191
CRV	18905	352

Tabela 2: Número de instâncias e lexicalizações por tipo de conceito.

³O número de possíveis escolhas lexicais para cada conceito variou entre 2 e 12 alternativas.

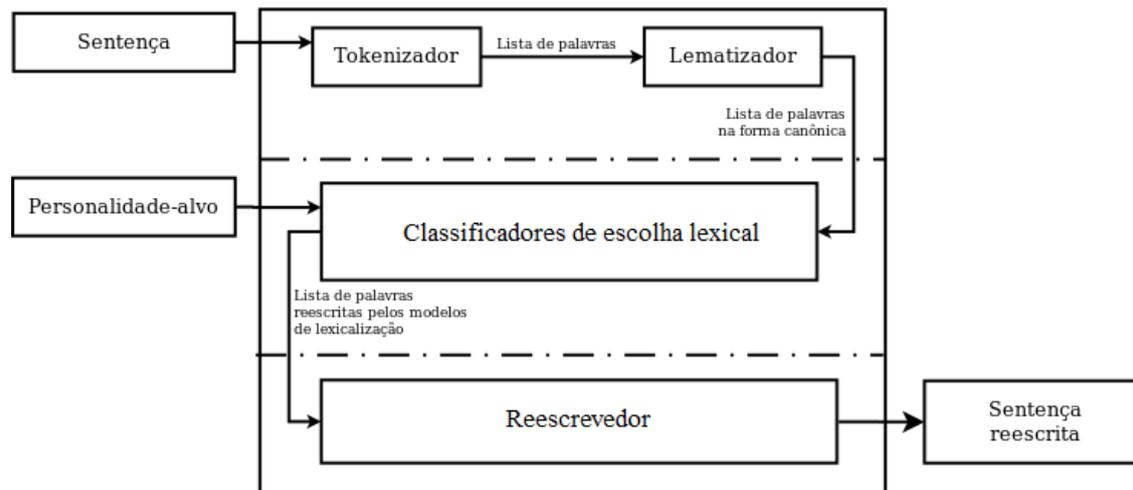


Figura 2: Fases da reescrita sentencial.

Em um experimento-piloto foi avaliado o uso de máquinas de vetor de suporte (SVM), indução de árvore de decisão e classificação Naive Bayes na tarefa de lexicalização. Como resultado, optou-se por construir todos os classificadores utilizando-se SVMs com núcleo de base radial. Os parâmetros ótimos de cada modelo foram obtidos por *grid search*, e os classificadores resultantes foram então incorporados ao modelo de reescrita sentencial propriamente dito.

3.3.2. Reescrevedor sentencial

Os classificadores descritos na seção anterior são combinados em um modelo reescrevedor que recebe como entrada uma sentença em língua natural e efetua substituições lexicais de acordo com um determinado perfil de personalidade fornecido. Assim, é possível converter uma dada sentença em versões alternativas que refletem um maior ou menor grau de extroversão, agradabilidade, conscienciosidade, neuroticismo ou abertura à experiência, tal qual definido no modelo dos Cinco Grandes Fatores (Goldberg, 1990).

O processo de reescrita segue um *pipeline* de três fases onde a sentença fornecida é analisada e modificada conforme o diagrama da Figura 2 anterior. Primeiramente, a sentença de entrada passa por um processo de *tokenização* e lematização dos termos constituintes, gerando assim uma lista de palavras na sua forma canônica. A seguir, os mapeamentos de palavras para conceitos discutidos na Seção 3.2 são utilizados para obter os conceitos que as palavras de entrada representam. Finalmente, os classificadores de escolha lexical são invocados para predizer a forma modificada (baseada em personalidade) de cada

palavra original e, caso a predição resulte em uma forma válida para o conceito fornecido⁴, a substituição lexical é realizada. Nota-se também que, caso a palavra de entrada já esteja na forma adequada para o tipo de personalidade em questão, a lexicalização sugerida pelo classificador pode ser idêntica à palavra de entrada, e neste caso nenhuma substituição é realizada.

3.3.3. Exemplo de funcionamento

Esta seção descreve um exemplo simplificado de funcionamento do modelo proposto. Para este fim, considere-se a tarefa de reescrever a frase de entrada a seguir com base em um conjunto de traços de personalidade-*P*.

Homem maduro demonstrando preocupação.

Inicialmente, é obtido o conceito (CRS, CRV ou CRA) referente a cada substantivo, verbo ou adjetivo da frase fornecida. Neste exemplo seriam obtidos os conceitos \$man (com possíveis lexicalizações como ‘homem’, ‘senhor’ etc.) \$old (‘maduro’, ‘idoso’, ‘velho’ etc.), \$show (‘demonstrar’, ‘mostrar’, ‘revelar’ etc.) e \$negemotion (‘preocupação’, ‘sofrimento’, ‘desespero’ etc.)

Cada conceito, associado a um conjunto de traços *P* fornecido, forma uma instância de teste a ser submetida ao respectivo classificador de escolha lexical (de CRS, CRA ou CRV, conforme o caso) de modo a obter a lexicalização mais adequada (i.e., de acordo com a personalidade-alvo

⁴Ou seja, dado que os classificadores não possuem desempenho perfeito, é possível que façam predições inconsistentes a serem desconsideradas.

fornecida). Supondo-se, respectivamente, escores de Extroversão, Agradabilidade, Conscienciosidade, Neuroticismo e Abertura à experiência como $P = \{3.0, 3.0, 2.4, 2.0, 3.5\}$, um exemplo de instância de aprendizado a ser submetidas ao classificador de lexicalização de CRA seria como segue:

[\$old, 3.0, 3.0, 2.4, 2.0, 3.5]

A seguir, os classificadores de CRS, CRA e CRV são invocados, retornando a lexicalização de cada conceito fornecido. No exemplo, seriam obtidas as lexicalizações ‘homem’, ‘velho’, ‘mostrar’ e ‘desespero’, respectivamente. Finalmente, os termos originais da frase de entrada são substituídos pelas formas lexicais sugeridas pelos classificadores com a devida flexão, resultando na seguinte sentença reescrita:

Homem velho mostrando desespero.

Variando-se os escores de personalidade em P , lexicalizações alternativas podem ser sugeridas pelos classificadores, resultando em sentenças distintas porém de significado aproximado. Por exemplo, um maior grau de Extroversão modificaria a lexicalização do verbo \$show, e um maior grau de Neuroticismo modificaria simultaneamente as lexicalizações do adjetivo \$old e do substantivo \$negemotion. Assim, dado um perfil de personalidade $P = \{4.0, 3.0, 2.4, 4.1, 3, 5\}$, a seguinte sentença seria obtida.

Homem idoso revelando sofrimento.

Cabe destacar, entretanto, que a interação entre os cinco traços de personalidade é de modo geral complexa, e que o efeito de um traço pode ser atenuado ou amplificado por outro. Por exemplo, um escore mais baixo de Conscienciosidade neste exemplo faria com que a lexicalização do adjetivo \$old fosse definida como ‘idoso’ independentemente do grau de Neuroticismo.

4. Avaliação e resultados

A avaliação dos modelos propostos foi realizada de duas formas: considerando-se o desempenho individual dos classificadores de escolha lexical, e considerando-se o uso destes classificadores como parte de um modelo completo de reescrita sentencial. As duas formas de avaliação são detalhadas nas seções a seguir.

Dado que os classificadores, se tomados isoladamente, não possuem função prática, sua avaliação objetiva unicamente ilustrar seu desempenho na tarefa de lexicalização de conceitos individuais sob a perspectiva de aprendizado de

máquina. Para este fim, foi realizada a comparação entre os resultados dos modelos baseados em personalidade com uma alternativa de *baseline* na qual os cinco atributos de aprendizado representando os escores de personalidade do autor-alvo são omitidos. Assim, a escolha lexical neste caso passa a ser feita apenas com base no conceito fornecido como entrada. Todos os classificadores com e sem acesso à informação de personalidade foram avaliados utilizando-se o conjunto de dados completo do *corpus b5-text* com validação cruzada de 10 partições, e computando-se sua medida F_1 média.

No caso da reescrita sentencial, por outro lado, objetiva-se verificar se o uso de informação de personalidade permite gerar sentenças mais próximas das produzidas por humanos (tal qual observadas no *corpus*) do que uma alternativa que não tenha acesso a este tipo de informação (ou seja, construída utilizando-se os classificadores de *baseline* acima.) Para este fim, todos os modelos com e sem acesso à informação de personalidade foram treinados com o conjunto completo de textos do *corpus b5-text*, e aplicados à geração de sentenças do *corpus b5-caption*, que contém legendas das mesmas imagens e escritas pelos mesmos autores dos textos de treinamento. A avaliação neste caso foi realizada seguindo-se prática comum na avaliação de sistemas de GLN, em especial no que diz respeito à tarefa de realização superficial (Reiter & Belz, 2009), comparando-se cada sentença original do *corpus* e sua versão modificada e medindo-se a distância de edição (Damerau, 1964) entre ambas. É importante observar também que os reflexos de diferentes personalidades na produção textual tende a ser sutis (Walker et al., 2011b), o que dificulta o uso de métodos de avaliação mais sofisticados (e possivelmente de maior custo) como o uso de julgamento humano.

4.1. Resultados dos modelos de classificação de escolha lexical

A Tabela 3 apresenta os resultados médios de lexicalização dos CRS, CRA e CRV do *corpus b5-text*.

Para os três tipos de conceitos, observa-se que o uso de informação de personalidade (à direita na tabela) apresenta resultados superiores ao modelo de *baseline* que não tem acesso a esse tipo de informação, e que os resultados da lexicalização na forma de adjetivos são consideravelmente superiores aos demais. Este comportamento era em grande parte esperado, já que o modelo CGF é mais fortemente relacionado ao uso

Tipo	sem personalidade			com personalidade		
	P	R	F ₁	P	R	F ₁
CRS	0,45	0,64	0,52	0,73	0,75	0,70
CRA	0,78	0,79	0,76	0,85	0,82	0,82
CRV	0,76	0,77	0,74	0,78	0,78	0,76

Tabela 3: Médias ponderadas de precisão (P), revocação (R) e medida F₁ do modelo de lexicalização completo. O maior valor de medida F₁ de cada tipo é destacado.

de adjetivos (Goldberg, 1990), e estes seriam assim os elementos que melhor refletem a distinção entre tipos de personalidade.

Dado que estes resultados incluem diversos conceitos com pequeno número de ocorrências, é possível obter uma visão mais clara do desempenho efetivo dos classificadores analisando-se apenas o resultado médio dos 10 conceitos mais frequentes de cada tipo. Os resultados desta análise são apresentados na Tabela 4.

Tipo	sem personalidade			com personalidade		
	P	R	F ₁	P	R	F ₁
CRS	0,53	0,58	0,52	0,6	0,58	0,56
CRA	0,48	0,52	0,48	0,74	0,67	0,68
CRV	0,55	0,52	0,51	0,72	0,56	0,58

Tabela 4: Médias ponderadas de precisão (P), revocação (R) e medida F₁ do modelo de lexicalização para os dez conceitos mais frequentes de cada tipo. O maior valor de medida F₁ de cada tipo é destacado.

Novamente, observa-se que os classificadores com acesso à informação de personalidade apresentam resultados mais próximos do desempenho humano do que os classificadores de *baseline*.

Por fim, como forma de comparar o desempenho dos classificadores com e sem acesso à informação de personalidade por meio de exemplos, as Tabelas 5, 6 e 7 apresentam resultados das alternativas de lexicalização dos conceitos mais frequentes de cada tipo (CRA, CRS e CRV).

No caso dos CRS mais frequentes listados na Tabela 5, observa-se que o uso de personalidade foi superior em seis das onze lexicalizações mais frequentes, havendo também um empate (no grupo criança/jovem) e quatro casos em que é ligeiramente inferior, embora todos concentrados nas classes menos frequentes. Estes resultado era de certa forma esperado dado que substantivos tendem a apresentar menor variação de sinônimos do que adjetivos e verbos.

Substantivos	sem	com
	personalidade	personalidade
foto	0,45	0,53
imagem	0,56	0,62
cena	0,00	0,33
criança	0,64	0,64
jovem	0,00	0,00
casaco	0,69	0,67
jaqueta	0,00	0,33
senhora	0,69	0,68
mulher	0,64	0,62
calça	0,86	0,82
jeans	0,00	0,35

Tabela 5: Medida F₁ média obtida para os cinco CRS mais frequentes, por forma lexical.

Adjetivos	sem	com
	personalidade	personalidade
alto	0,83	0,85
forte	0,65	0,71
bravo	0,69	0,82
simples	0,89	0,89
singelo	0,00	0,44
rural	0,71	0,63
rústico	0,00	0,30
velho	0,63	0,71
idoso	0,74	0,75
preocupado	0,76	0,92
apreensivo	0,00	0,80

Tabela 6: Medida F₁ média obtida para os cinco CRA mais frequentes, por forma lexical.

Verbo	sem	com
	personalidade	personalidade
mostrar	0,73	0,74
dar	0,50	0,60
revelar	0,00	0,67
dividir	0,33	0,75
destacar	0,80	0,81
separar	0,48	0,56
carregar	0,58	0,56
passar	0,61	0,60
tirar	0,93	0,92
separar	0,48	0,56
afastar	0,00	0,20
espalhar	0,56	0,58
dar	0,50	0,60

Tabela 7: Medida F₁ média obtida para os cinco CRV mais frequentes, por forma lexical.

No caso dos CRA a vantagem dos modelos baseados em personalidade é mais expressiva, sendo superiores em 9 dos 11 casos de lexicalização considerados. Finalmente, no caso dos CRV, o uso de informação de personalidade é superior em 10 de 13 casos.

4.2. Resultados de reescrita sentencial

Embora os resultados da avaliação do classificador apontem uma certa vantagem no uso de informação de personalidade na tarefa de lexicalização, resta a questão de quão efetiva seria a tarefa de reescrita sentencial baseada nestes classificadores. A avaliação global do modelo de reescrita é descrita a seguir.

A Tabela 8 apresenta a distância de edição média obtida pelas duas variantes do modelo proposto —com e sem acesso à informação de personalidade— para cada contexto do cópulus.

Contexto	sem personalidade		com personalidade	
	Média	DP	Média	DP
1	3,28	3,01	2,81	3,29
2	3,52	1,93	2,23	2,41
3	3,02	3,69	2,91	3,73
4	0,45	3,28	0,56	3,51
5	1,11	3,39	1,26	3,25
6	1,07	2,65	1,30	3,06
7	2,89	2,70	2,28	3,14
8	0,32	1,47	0,36	1,72
9	3,34	4,22	2,70	3,74
10	2,27	3,61	1,57	2,92
Média	3,37	3,36	2,53	3,36

Tabela 8: Médias e desvios padrão das distâncias de edição por contexto. O melhor resultado para cada contexto é destacado.

Com base nesses resultados, observa-se que a distância de edição média dos modelos com acesso à personalidade é inferior à dos modelos sem acesso a esta informação. Além disso, os modelos baseados em personalidade possuem, em média, menor distância de edição em 6 dos 10 contextos visuais do cópulus, e observa-se ainda que os 4 contextos em que o uso de informação de personalidade foi prejudicial são justamente aqueles em que ocorreu menor volume de substituição lexical (indicado pelo menor grau de distância de edição).

A diferença entre a distância de edição média dos modelos de reescrita sentencial com e sem personalidade é significativa de acordo com um

teste ANOVA de fator único ($F(1, 1624) = 13.23$, $MSE = 11.39$, $p < 0.001$). Esse resultado oferece suporte à hipótese de pesquisa deste estudo, ou seja, de que o uso de informação de personalidade permite a geração de sentenças mais próximas daquelas que seriam geradas por um locutor humano com estas características.

5. Conclusão

Este trabalho apresentou um modelo de reescrita sentencial que leva em conta, além da sentença a ser reescrita, informações sobre a personalidade de um locutor-alvo de interesse. O modelo proposto efetua substituições lexicais com base no modelo de personalidade CGF e é, até onde temos conhecimento, o primeiro deste tipo para o português (pelo menos em sua variante brasileira). Resultados sugerem que levar em conta a personalidade do autor-alvo de interesse produz sentenças mais próximas das que seriam produzidas por um locutor humano com as mesmas características do que um modelo de geração sem acesso a este tipo de informação, e sugerem assim a possibilidade de geração de texto mais natural e/ou psicologicamente plausível.

Uma das possíveis limitações do presente trabalho é o fato de ser baseado em um conjunto de textos rotulados manualmente com informações de conceitos semânticos de interesse para fins de aprendizado de máquina supervisionado. Assim, o uso dos métodos aqui discutidos em outro domínio em princípio exigiria um novo trabalho de anotação de cópulus, possivelmente com avaliação do grau de concordância entre juízes (no presente caso omitido dada a simplicidade do domínio e do pequeno tamanho do cópulus considerado). Alternativas mais generalizáveis, como as baseadas em métodos semi ou não supervisionados de aprendizado, são deixadas como trabalho futuro.

Finalmente, observa-se que o foco do presente trabalho foi a questão da escolha lexical feita por indivíduos com diferentes personalidades, e foi motivado pela própria natureza lexical do modelo CGF. No entanto, para além deste modelo específico, é possível considerar diversas outras formas de variação humana na produção de língua natural. Estas variações incluem o uso de certas construções sintáticas preferenciais, tempos verbais e até mesmo a determinação do conteúdo semântico a ser representado em forma textual. Estas variações, que seriam uma extensão natural do presente trabalho, também são deixadas como trabalho futuro.

Agradecimentos

Este trabalho contou com apoio FAPESP nro. 2016/14223 0.

Referências

- Álvarez-Carmona, Miguel A., A. Pastor López-Monroy, Manuel Montes y Gómez, Luis Villaseñor-Pineda & Hugo Jair Escalante. 2015. INAOE's participation at PAN'15: Author profiling task. Em *Working Notes of the Conference and Labs of the Evaluation Forum (CLEF)*, s.p.
- Aly, Almir & Adriana Tapus. 2016. Towards an intelligent system for generating an adapted verbal and nonverbal combined behavior in human–robot interaction. *Autonomous Robots* 40(2). 193–209. doi 10.1007/s10514-015-9444-1.
- Bick, Eckhard. 2000. *The parsing system PALAVRAS: Automatic grammatical analysis of portuguese in a constraint grammar framework*: Aarhus University. Tese de Doutorado.
- Bowden, Kevin K., Grace Lin, Lena Reed, Jean E. Fox Tree & Marilyn A. Walker. 2016. M2D: Monolog to dialog generation for conversational story telling. Em *9th International Conference on Interactive Digital Storytelling*, 12–24. doi 10.1007/978-3-319-48279-8_2.
- Briggs Myers, Isabel & Peter B. Myers. 2010. *Gifts differing: Understanding personality type*. Davies-Black 2nd edn.
- Cuevas, Ramon Re Moya & Ivandré Paraboni. 2008. A machine learning approach to Portuguese pronoun resolution. Em *Advances in Artificial Intelligence (IBERAMIA)*, 262–271. doi 10.1007/978-3-540-88309-8_27.
- Damerau, Fred J. 1964. A technique for computer detection and correction of spelling errors. *Communications of the ACM* 7(3). 171–176. doi 10.1145/363958.363994.
- Dan-Glauser, Elise S. & Klaus R. Scherer. 2011. The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behavior Research Methods* 43(2). 468–477. doi 10.3758/s13428-011-0064-1.
- van Deemter, Kees. 2016. *Computational models of referring. a study in cognitive science*. Cambridge: MIT Press.
- di Eugenio, Barbara, Andrew Boyd, Camillo Lugaresi, Abhinaya Balasubramanian, Gail Keenan, Mike Burton, Tamara Goncalves Rezende Macieira, Jianrong Li & Yves Lussier. 2014. PatientNarr: Towards generating patient-centric summaries of hospital stays. Em *8th International Natural Language Generation Conference (INLG)*, 6–10. doi 10.3115/v1/W14-4402.
- Ferreira, Thiago Castro & Ivandré Paraboni. 2014. Classification-based referring expression generation. Em *Computational Linguistics and Intelligent Text Processing (CICLing)*, 481–491. doi 10.1007/978-3-642-54906-9_39.
- Goldberg, Lewis R. 1990. An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology* 59. 1216–1229. doi 10.1037/0022-3514.59.6.1216.
- González-Gallardo, Carlos E., Azucena Montes, Gerardo Sierra, J. Antonio Núñez-Juaréz, Adolfo Jonathan Salinas-López & Juan Ek. 2015. Tweets classification using corpus dependent tags, character and POS n-grams. Em *Working notes of the Conference and Labs of the Evaluation Forum (CLEF)*, s.p.
- Herzig, Jonathan, Michal Shmueli-Scheuer, Tommy Sandbank & David Konopnicki. 2017. Neural response generation for customer service based on personality traits. Em *10th International Conference on Natural Language Generation*, 252–256. doi 10.18653/v1/W17-3541.
- Hsieh, Fernando Chiu, Rafael Felipe Sandroni Dias & Ivandré Paraboni. 2018. Author profiling from Facebook corpora. Em *11th International Conference on Language Resources and Evaluation (LREC)*, 2566–2570.
- Hunter, James, Yvonne Freer, Albert Gatt, Ehud Reiter, Somayajulu Sripada & Cindy Sykes. 2012. Automatic generation of natural language nursing shift summaries in neonatal intensive care: BT-Nurse. *Artificial intelligence in medicine* 56(3). 157–172. doi 10.1016/j.artmed.2012.09.002.
- John, Oliver P., Laura P. Naumann & Christopher J. Soto. 2008. Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. Em *Handbook of personality: Theory and research*, 114–158. New York: Guilford Press.
- Jordan, Pamela, Nancy Green, Christopher Thomas & Susan Holm. 2014. TBI-Doc: Generating patient & clinician reports from brain

- imaging data. Em *8th International Natural Language Generation Conference (INLG)*, 143–146. doi 10.3115/v1/W14-4423.
- Karpathy, Andrej & Li Fei-Fei. 2015. Deep visual-semantic alignments for generating image descriptions. Em *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3128–3137. doi 10.1109/CVPR.2015.7298932.
- Khosmood, Foaad & Marilyn Walker. 2010. Grapevine: a gossip generation system. Em *5th International Conference on the Foundations of Digital Games*, 92–99.
- Lan, Alex Gwo Jen & Ivandr  Paraboni. 2018. Definite description lexical choice: taking speaker’s personality into account. Em *11th International Conference on Language Resources and Evaluation (LREC)*, 2999–3004.
- Lukin, Stephanie M., James O. Ryan & Marilyn Walker. 2014. Automating direct speech variations in stories and games. Em *Games and Natural Language Processing Workshop (GAMNLP)*, s.p.
- Mairesse, Fran ois & Marilyn A. Walker. 2010. Towards personality-based user adaptation: psychologically informed stylistic language generation. *User Model. User-Adapt. Interaction* 20(3). 227–278. doi 10.1007/s11257-010-9076-2.
- Mairesse, Fran ois & Marilyn A. Walker. 2011. Controlling user perceptions of linguistic style: Trainable generation of personality traits. *Computational Linguistics* 37(3). 455–488. doi 10.1162/COLI_a_00063.
- Mairesse, Fran ois. 2008. *Learning to adapt in dialogue systems: Data-driven models for personality recognition and generation*: University of Sheffield. Tese de Doutorado.
- Maziero, Erick G., Thiago A. S. Pardo, Ariani di Felippo & Bento C. Dias da Silva. 2008. A base de dados lexical e a interface web do TeP 2.0: Thesaurus eletr nico para o portugu s do Brasil. Em *14th Brazilian Symposium on Multimedia and the Web*, 390–392. doi 10.1145/1809980.1810076.
- McCormick, Christopher. 2012. *Evaluating the perception of personality and naturalness in computer generated utterances*: University of Dublin, Trinity College. Tese de Mestrado.
- Paraboni, Ivandr  & Vera Lucia Strube de Lima. 1998. Possessive pronominal anaphor resolution in Portuguese written texts. Em *17th International Conference on Computational Linguistics*, 1010–1014. doi 10.3115/980691.980735.
- Paraboni, Ivandr , Danielle Sampaio Monteiro & Alex Gwo Jen Lan. 2017. Personality-dependent referring expression generation. Em *Text, Speech and Dialogue (TSD)*, 20–28. doi 10.1007/978-3-319-64206-2_3.
- Plank, Barbara & Dirk Hovy. 2015. Personality traits on Twitter - or - how to get 1,500 personality tests in a week. Em *6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 92–98. doi 10.18653/v1/W15-2913.
- Portet, Fran ois, Ehud Reiter, Albert Gatt, Jim Hunter, Somayajulu Sripada, Yvonne Freer & Cindy Sykes. 2009. Automatic generation of textual summaries from neonatal intensive care data. *Artificial Intelligence* 173(7–8). 789–816. doi 10.1016/j.artint.2008.12.002.
- Ramos, Ricelli Moreira Silva, Georges Basile Stavrakas Neto, Barbara Barbosa Claudino Silva, Danielle Sampaio Monteiro, Ivandr  Paraboni & Rafael Felipe Sandroni Dias. 2018. Building a corpus for personality-dependent natural language understanding and generation. Em *11th International Conference on Language Resources and Evaluation (LREC)*, 1138–1145.
- Reiter, Ehud & Anja Belz. 2009. An investigation into the validity of some metrics for automatically evaluating natural language generation systems. *Computational Linguistics* 35(4). 529–558. doi 10.1162/coli.2009.35.4.35405.
- Reiter, Ehud & Robert Dale. 2000. *Building natural language generation systems*. New York: Cambridge University Press.
- Reiter, Ehud, Roma Robertson & Liesl Osman. 2003. Lessons from a failure: Generating tailored smoking cessation letters. *Artificial Intelligence* 144(1–2). 41–58. doi 10.1016/S0004-3702(02)00370-3.
- Reiter, Ehud, Somayajulu Sripada, Jim Hunter & Jin Yu. 2005. Choosing words in computer-generated weather forecasts. *Artificial Intelligence* 167(1–2). 137–169. doi 10.1016/j.artint.2005.06.006.
- dos Santos, Vitor Garcia, Ivandr  Paraboni & B rbara Barbosa Claudino Silva. 2017. Big five personality recognition from multiple text genres. Em *Text, Speech and Dialogue (TSD)*, 29–37. doi 10.1007/978-3-319-64206-2_4.

- Silva, Bárbara Barbosa Claudino & Ivandré Paraboni. 2018a. Learning personality traits from Facebook text. *IEEE Latin America Transactions* 16(4). 1256–1262. doi 10.1109/TLA.2018.8362165.
- Silva, Bárbara Barbosa Claudino & Ivandré Paraboni. 2018b. Personality recognition from Facebook text. Em *13th International Conference on the Computational Processing of Portuguese (PROPOR)*, 107–114. doi 10.1007/978-3-319-99722-3_11.
- Viethen, Jette & Robert Dale. 2010. Speaker-dependent variation in content selection for referring expression generation. Em *Australasian Language Technology Association Workshop*, 81–89.
- Walker, Marilyn, Grace Lin, Jennifer Sawyer, Ricky Grant, Michael Buell & Noah Wardrip-Fruin. 2011a. Murder in the arboretum: Comparing character models to personality models. Em *Intelligent Narrative Technologies*, Santa Cruz: AAAI.
- Walker, Marilyn A., Ricky Grant, Jennifer Sawyer, Grace I. Lin, Noah Wardrip-Fruin & Michael Buell. 2011b. Perceived or not perceived: Film character models for expressive NLG. Em *International Conference on Interactive Digital Storytelling: Interactive Storytelling (ICIDS)*, 109–121. doi 10.1007/978-3-642-25289-1_12.
- Xu, Kelvin, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron C. Courville, Ruslan Salakhutdinov, Richard S. Zemel & Yoshua Bengio. 2015. Show, attend and tell: Neural image caption generation with visual attention. Em *32nd International Conference on Machine Learning, ICML*, 2048–2057.
- Zhang, Xingxing & Mirella Lapata. 2014. Chinese poetry generation with recurrent neural networks. Em *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 670–680. doi 10.3115/v1/D14-1074.