

ISSN 1646 - 6195

VOLUME 13

ANO 2018

LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS
DA UNIVERSIDADE DO PORTO

CENTRO DE LINGUÍSTICA
DA UNIVERSIDADE DO
PORTO

FACULDADE DE LETRAS
DA UNIVERSIDADE DO
PORTO

LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS
DA UNIVERSIDADE DO PORTO

Diretora:

Ana Maria Brito (Universidade do Porto)

Diretora convidada deste volume:

Fátima Oliveira (Universidade do Porto)

Secretariado Editorial:

Maria da Graça Lisboa Castro Pinto (Universidade do Porto)

Fátima Oliveira (Universidade do Porto)

João Veloso (Universidade do Porto)

António Leal (Universidade do Porto)

Conselho Científico:

Alexandra Guedes Pinto (Universidade do Porto)

Alexandra Soares Rodrigues (Instituto Politécnico de Bragança)

Álvaro Iriarte (Universidade do Minho)

Anabela Gonçalves (Universidade de Lisboa)

Ana Cristina Macário Lopes (Universidade de Coimbra)

Ana Maria Martins (Universidade de Lisboa)

Antónia Coutinho (Universidade Nova de Lisboa)

António Leal (Universidade do Porto)

Augusto Soares da Silva (Universidade Católica)

Belinda Maia (Universidade do Porto)

Carmen Matzenauer (Universidade Católica de Pelotas, Rio Grande do Sul)

Celeste Rodrigues (Universidade de Lisboa)

Clarinda Azevedo Maia (Universidade de Coimbra)

Claudia Brescancini (Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre)

Cristina Martins (Universidade de Coimbra)

Elena Galvão (Universidade do Porto)

Fátima Henriques da Silva (Universidade do Porto)

Fátima Oliveira (Universidade do Porto)

Fernanda Irene Fonseca (Universidade do Porto)

Francisco Lacerda (Universidade de Estocolmo)

Gabriela Matos (Universidade de Lisboa)

Georges Kleiber (Universidade de Strasbourg)

Gerd Wotjak (Universidade de Leipzig)

Gianpaolo Salvi (Universidade Eötvös Loránd, Budapeste)

Graça Rio-Torto (Universidade de Coimbra)

Ignacio Bosque (Universidade Complutense de Madrid)

Ildikó Szijj (Universidade Eötvös Loránd, Budapeste)

Inês Duarte (Universidade de Lisboa)

Isabel Galhano Rodrigues (Universidade do Porto)

Isabel Margarida Duarte (Universidade do Porto)

Isabel Pereira (Universidade de Coimbra)

Jean-Pierre Angoujard (Université de Nantes)

João Costa (Universidade Nova de Lisboa)

João Paulo Silvestre (Universidade de Lisboa)

Joaquim Brandão de Carvalho (Universidade de Paris 8)

Krista Varantola (Universidade de Tampere)

Lluís Payrató (Universidade de Barcelona)

Luís Filipe Cunha (Universidade do Porto)

Maria Clara Barros (Universidade do Porto)

Maria da Graça Lisboa Castro Pinto (Universidade do Porto)

Maria Helena Mateus (Universidade de Lisboa)

Maria Helena Paiva (Universidade do Porto)

Maria João Freitas (Universidade de Lisboa)

Marina Vígário (Universidade de Lisboa)

Norval Smith (Universidade de Amsterdão)

Olívia Figueiredo (Universidade do Porto)

Petra Sleeman (Universidade de Amsterdão)

Purificação Silvano (Universidade do Porto)

Rogélio Ponce de León Romeo (Universidade do Porto)

Rui Marques (Universidade de Lisboa)

Rui Sousa Silva (Universidade do Porto)

Sónia Frota (Universidade de Lisboa)

Telmo Mória (Universidade de Lisboa)

Telmo Verdelho (Universidade de Aveiro)

Thomas Hüsgen (Universidade do Porto)

Tjerk Hagemeier (Universidade de Lisboa)

Tony Berber Sardinha (Pontifícia Universidade Católica de São Paulo)

Violeta Demonte (Universidade Autónoma de Madrid)

LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS
DA UNIVERSIDADE DO PORTO

VOL. 13, 2018

FICHA TÉCNICA

Linguística
Revista de Estudos Linguísticos da Universidade do Porto
Vol. 13, 2018

ISSN: 1646-6195

Periodicidade: Anual

Diretora:
Ana Maria Brito

Diretora convidada deste volume:
Fátima Oliveira

Editores:
Faculdade de Letras da Universidade do Porto
e Centro de Linguística da Universidade do Porto

Revisão:
Manuel Caramelo

Capa:
José Osswald

Impressão e acabamentos:
Invulgar - Artes Gráficas

Depósito Legal:
248653/06

Tiragem:
200 Exemplares

Os artigos publicados estão sujeitos a “peer review”.

Esta edição é integralmente financiada por:
Fundação para a Ciência e a Tecnologia através do Fundo de Reestruturação de Unidades 2016
(UID/LIN/0022/2016) atribuído ao Centro de Linguística da Universidade do Porto

A Revista está registada no DOAJ e Latindex e está indexada na base de dados Fonte Académica.

<http://ler.letras.up.pt/site/default.aspx?qry=id04id191&sum=sim>

Espaço da direção

Este volume da revista *Linguística* é dedicado ao Professor Óscar Lopes, insigne linguista que fundou o Centro de Linguística da Universidade do Porto e trouxe importantes e inovadores contributos para a Linguística, em particular na área da Semântica e também na sua articulação com a Pragmática. Mas a sua investigação sobre a linguagem e a sua complexidade é muito anterior à sua entrada na Universidade do Porto, que só aconteceu depois do 25 de abril de 1974, dado que tal lhe tinha sido vedado anteriormente por razões políticas.

Com efeito, desde pelo menos 1944 que o estudo e o ensino da língua são objeto das suas preocupações. Num texto desse ano intitulado “Preliminares a uma Gramática”, Óscar Lopes, criticando uma certa visão de gramática, diz o seguinte: “É preciso evitar a falsa clareza das ideias feitas”. Essa preocupação, que revela uma postura científica, perpassa todos os seus estudos gramaticais e em particular na *Gramática Simbólica - um esboço* cuja primeira edição data de 1971 e em que apresenta, frequentemente de modo simples, mas não simplista, recursos formais da Lógica e da Matemática que permitem tornar mais claro ou até compreender uma diversidade de temas e problemas linguísticos que trabalha.

A sua conceção de investigação e o seu anseio por respostas cada vez mais satisfatórias levaram-no a considerar que a necessária redução do campo de análise em Linguística, como em qualquer domínio científico, não deve fazer-nos esquecer os dados nem a própria reflexão crítica. É essa uma das razões fundamentais por que defendeu a formalização como um meio ou, como dizia, uma metodologia, para tornar mais claras e passíveis de infirmação a descrição e a explicação em Linguística.

A sua visão extremamente inovadora de Gramática levou-o a trabalhar um número considerável de questões linguísticas que vão desde a distinção semântica entre nomes próprios e nomes comuns, questões de quantificação, definidos e indefinidos, demonstrativos, ‘coordenação’ e ‘subordinação’ de vários tipos, até questões de tempo, aspeto e modo, utilizando instrumentos teóricos do cálculo proposicional, da lógica de predicados, da teoria dos conjuntos, da teoria dos grafos, de topologia, de relações ma-

temáticas e lógicas de vários tipos e suas propriedades. Quem hoje queira estudar alguns destes temas, em particular em Português, não pode deixar de consultar esta gramática, embora de difícil acesso, por não ter sido reeditada.

Depois desta obra fundamental, Óscar Lopes, já na Universidade, continuou a desenvolver investigação em Linguística alargando ou aprofundando vários temas, entre os quais é de salientar em primeiro lugar a semântica dos nomes massivos e problemas de quantificação, usando a formalização de sistemas de lógica intensional. Mas também continuou a fazer estudos sobre contrastivas, condicionais, concessivas e causais, passando por estudos sobre o Presente e a deixis, e interessando-se também por partículas discursivas como '*pois*'. Vários destes textos, dos quais alguns publicados pela primeira vez, estão reunidos no terceiro volume das suas obras completas, publicado em 2005, *Entre a Palavra e o Discurso*.

A sua curiosidade infinita e a vontade de saber mais e compreender melhor articularam-se sempre com a vontade de estabelecer diálogo entre áreas e teorias da Linguística, com outras áreas do saber que também dominava, e também com todos nós. Os seus textos, escritos há dezenas de anos, continuam a interpelar-nos e a tornar claro como foi um investigador inovador, colocando alguns problemas que só recentemente têm sido objeto de estudo. Mostram ainda como soube ouvir o som, a princípio ténue, do novo por apresentar outras perspetivas mais rigorosas, mais profundas e mais claras e, assim, abrir caminhos renovados.

Sobre Óscar Lopes podemos dizer com toda a propriedade que foram não só caminhos novos que construiu e trilhou, mas também foi novo o seu jeito de caminhar.

Por isso, é urgente ler os seus textos e refletir sobre as suas propostas, criando, se possível, novos olhares e formulando novas perguntas, pois é isso que sempre esperou de nós.

O primeiro texto, da autoria de Inês Duarte, reflete sobre o percurso de Óscar Lopes em particular como linguista e sobre o enquadramento intelectual em que se movia, assim como e sobretudo sobre o papel extraordinário

riamente inovador, abordado de um ponto de vista científico, que os seus estudos apresentam.

Este texto corresponde em grande medida à conferência proferida pela autora em junho de 2018 no âmbito da celebração de "Óscar Lopes - figura Eminente da Universidade do Porto"

No segundo texto, Ana Maria Brito e Gabriela Matos apresentam alguns aspetos sintáticos e semânticos das relativas livres e das interrogativas subordinadas parciais próprias e impróprias, correlacionando-os com a forma como Óscar Lopes descreveu estas construções e destacando a ambiguidade estrutural e semântica de algumas delas.

O texto seguinte, da autoria de Fátima Oliveira e António Leal, aborda a especificidade semântica do Pretérito Perfeito Composto em Português dado que, sob certas condições, permite, contrariamente a outras línguas, uma leitura de repetição de situações. Os autores propõem que a leitura de eventualidade única e a leitura iterativa deste tempo pode ser explicada de forma unificada recorrendo à noção de *homogeneidade segmental* e de *homogeneidade incremental*, e defendem, assim, que a propriedade da homogeneidade é uma propriedade aspetual central do PPC em Português Europeu.

O quarto texto, da autoria de Fátima Silva, Purificação Silvano, António Leal, Fátima Oliveira, Pavel Brazdil, João Cordeiro e Débora Oliveira, é um trabalho de interface entre a linguística e as ciências da computação, tendo como objetivo realizar a análise computacional de artigos de opinião na área da economia e finanças, seguindo o quadro teórico da análise de sentimento. O trabalho realizado permite considerar que, para além da elaboração do modelo de anotação linguística adotado, foi possível verificar o papel central dos adjetivos para a determinação da polaridade do sentimento nos segmentos analisados do *corpus*. O estudo permitiu ainda proceder ao desenvolvimento de um novo léxico de sentimento específico português, o EconoLex, e verificar a melhoria do desempenho computacional do EconoLex⊕SentiLex em comparação com o SentiLex isoladamente.

O texto seguinte, da autoria de Isabel Galhano, apresenta um estudo de caso que testemunha a robustez do gesto em que é descrita uma listagem, uma atividade comum de organização do discurso, que consiste na enunciação de uma sequência de unidades caracterizadas por paralelismo

sinfático, prosódico e gestual. O estudo evidencia a forte ligação entre gesto e fala e também a corporização da atividade de listar.

Agradecemos aos autores a sua participação neste volume dedicado a Óscar Lopes e a todos os avaliadores que também colaboraram com os seus comentários. Desejamos que estes artigos possam contribuir para continuar o diálogo que Óscar Lopes sempre desejou, sendo por isso também uma forma de o homenagear.

Fátima Oliveira
(Diretora convidada)
Dezembro de 2018

SUMÁRIO

Espaço da direção.....	3
Artigos	
Óscar Lopes e a Linguagem: o olhar de um cientista <i>Inês Duarte</i>	11
Relativas livres e interrogativas parciais: como Óscar Lopes viu algumas destas construções <i>Ana Brito, Gabriela Matos</i>	29
Sobre a natureza homogénea do Pretérito Perfeito Composto em Português Europeu <i>Fátima Oliveira, António Leal</i>	57
Análise semântica de sentimento em artigos de opinião <i>Fátima Silva, Purificação Silvano, António Leal, Fátima Oliveira, Pavel Brazdil, João Cordeiro, Débora Oliveira</i>	79
Gestos com os pés: descrição de uma listagem <i>Isabel Galhano Rodrigues</i>	115
Bibliografia	
Óscar Lopes - <i>Bibliografia sobre Gramática e Linguística</i>	135

Artigos

Óscar Lopes e a linguagem: um olhar de cientista

Inês Duarte¹
iduarte@letras.ulisboa.pt

Encontrei Óscar Lopes pela primeira vez na viagem de finalistas do meu curso de licenciatura. Estávamos em Abril de 1973, e tínhamos conseguido um pequeno subsídio do então Instituto de Alta Cultura para assistir a uma conferência sobre Semântica Formal em Cambridge, organizada por Ed Keenan, que contava com a participação de investigadores americanos e europeus que partilhavam a convicção de que o conhecimento sobre a linguagem verbal humana não poderia avançar sem o recurso a ferramentas formais matemáticas ou lógicas. Conosco foram também três dos nossos professores e o Óscar.

17/3/1973

Meu caro

Estou em Cambridge de 9 a 12 de Abril, no início das férias da Páscoa. De acordo com os planos actuais, devo ir de comboio por Paris, juntamente com 10 linguistas de Lisboa (Cintra, Mira Mateus, M. Emília Marques, etc.).

[Neves, org., 2004. *AJS & OL: Correspondência*: 356]

Não me lembro se na ida se no regresso do congresso, alguns de nós almoçámos com o Óscar numa esplanada à beira do *Serpentine*, em Hyde Park. É sempre um momento que deixa marcas, encontrar alguém que só se conhece

¹ Centro de Linguística da Universidade de Lisboa, Faculdade de Letras da Universidade de Lisboa.

pelo que se leu dele (a *História da Literatura Portuguesa*, que me acompanhou desde os tempos do liceu) e sobre ele. Sobretudo quando se é muito jovem.



O Óscar impressionou-me pela simplicidade, por aquele ar sério que em público poucas vezes o abandonava e pela intensidade do olhar.

Poucas semanas depois, tive o privilégio de o reencontrar num seminário organizado pela Fundação Calouste Gulbenkian com o trio BCG (assim chamado pelos jovens *habitués* dos seminários da Rue d'Ulm): François

Bresson, psicólogo influenciado por Piaget; Antoine Culioli, linguista inspirado por Benveniste e fundador do primeiro departamento de linguística de uma universidade francesa (Paris-7); e Jean-Blaize Grize, matemático suíço apostado na construção de uma lógica natural. Os três partilhavam a convicção de que a formalização era um instrumento indispensável para que as hipóteses formuladas nas ciências humanas pudessem ser testadas. Desta convivência intelectual nasceria a teoria das operações enunciativas de Culioli, que chegou a contar com alguns discípulos em Portugal.

Porto, 8/3/1973

Meu Caro

[...] Em princípio estarei em Cambridge num seminário de Semântica Formal, precisamente entre 9 e 12 de Abril, logo no início das minhas férias, e ido do congresso de Aveiro. Talvez possa passar por Paris, mas a correr, porque tenho um outro seminário de linguística, com o Grize, o Culioli e o Bresson em Lisboa, logo a seguir.

[Neves, org., 2004. AJS & OL: Correspondência: 355]

Na primavera de 1973, este grupo de investigadores tinha uma grande recepção em França e chegou a Lisboa convencido de que iria encontrar um público pouco mais do que analfabeto.

Previa-se que os conferencistas fizessem exposições durante a manhã, que os participantes se organizassem em grupos de discussão com relatores durante a primeira parte da tarde, e que as conclusões de cada grupo fossem apresentadas em reunião plenária ao fim da tarde.

Tive a sorte de ficar no mesmo grupo do Óscar. Todos os membros do grupo tentaram que fosse ele o relator, mas, ou por timidez, ou porque não se sentia muito à vontade a falar francês, o Óscar não aceitou e acabei por ser eu a desempenhar essa função.

Logo no primeiro dia, o Óscar liderou a discussão no nosso grupo sobre a exposição muito básica, profundamente desinteressante e nada informada sobre o que estava a acontecer de novo do outro lado do Atlântico e na Europa que não falava línguas românicas.

Ora bem, no final do dia, lá fui eu relatar no meu melhor francês as conclusões do nosso grupo, quase todas com origem em observações do Óscar, lamentando o nível demasiado elementar da exposição, muitas das suas assunções de base e o seu fechamento teórico.

Adivinhem o que aconteceu... O trio BCG pediu a interrupção do seminário por um dia, para reformular o programa...

Que personalidade multifacetada era afinal a deste senhor de 56 anos, lutador pela causa da liberdade e da democracia, autoridade na história da literatura e na crítica literária, mas que regressava de uma conferência sobre Semântica Formal e cujas perguntas por interposta voz formuladas tanto perturbaram Bresson, Culioli e Grize?

Para a sua faceta de linguista, a que recordamos aqui hoje, a resposta é muito fácil: em Portugal, **o Óscar foi sem dúvida a primeira voz a anunciar o futuro no campo da Linguística.**

De facto, apesar de uma tradição de grande qualidade nas áreas da Filologia, História da Língua e Dialectologia, a Linguística (chamemos-lhe moderna) só se desenvolveu em Portugal a partir dos anos sessenta. Em Lisboa, só no ano letivo de 1969-1970 foi lecionado o primeiro curso em que o estudo das correntes estruturalistas norte-americanas (pós)-bloomfieldianas e a sua aplicação à análise de fenómenos do português tiveram lugar; e só

em 1970-1971, na disciplina de “Linguística Portuguesa II”, se estudaram muito superficialmente os dois primeiros modelos da Gramática Generativa-Transformacional (os modelos de 57 e de 65). Em Coimbra, dominavam os estudos dialectológicos e uma interpretação neo-tomista de Saussure e Coseriu. E, como afirma José Victor Adragão num texto sobre a pré-história da Linguística em Portugal, no Porto, onde o curso de Filologia Românica só tinha recomeçado a funcionar em 1968, havia ... Óscar Lopes.

Óscar Lopes que, já nos tempos de estudante (licenciou-se na Faculdade de Letras de Lisboa em 1941), tinha como preocupação desenvolver uma teoria matemática que permitisse captar as propriedades da linguagem humana. Como testemunha Vitorino Magalhães Godinho num texto sobre o Óscar intitulado *Definição fundamental: um homem de carácter* e datado de 2007, cito:

Conhecêramo-nos na Faculdade de Letras de Lisboa, nesse casarão tão pouco atraente por baixo da Academia das Ciências [...]. Um corpo docente medíocre [...], onde um ou outro competente não se interessava e procurava passar despercebido, se acaso no seu íntimo albergava ideias “incorrectas”. Mas, incólumes a esse ambiente [...] andávamos à descoberta desse *admirável mundo novo* para lá das fronteiras salazaristas. Arranjávamos livros, emprestávamo-los um ao outro (emprestei-lhe, por exemplo, o Husserl, mais tarde o Tarski), discutíamos linguística e lógica moderna [...], filosofia e política, história e a sua teoria. Eu estava empenhado em investigações de lógica relacional, esforçava-me por conhecer a matemática e a física, porque nelas alicerçava em boa parte a teoria do conhecimento e/ou epistemologia [...]. Óscar Lopes queria desenvencilhar-se dos enredos da linguística, construindo uma teoria matemática relacional que correspondesse às especialidades da linguagem. Será a sua preocupação dominante, até assentar outro pilar na história da literatura.

Assim, na segunda metade dos anos 30, estes brilhantes parceiros intelectuais, ávidos de romper as fronteiras impostas ao conhecimento pelo Estado Novo, estudam e discutem os trabalhos do Círculo de

Viena, fortemente influenciado, como sabemos, pelos avanços da física, da matemática e da lógica simbólica de finais do século XIX e início do século XX. Conseguimos facilmente compreender a atração que estes dois jovens, a viver num ambiente intelectual empoeirado e asfixiante, terão sentido por um programa que poderia vir a instaurar os estudos humanísticos (ou, pelo menos, partes deles) como ramos de uma ciência unificada através de uma linguagem lógico-matemática comum e de um critério de validação também comum: a verificação empírica.

E abro aqui um parêntesis, para destacar o papel da Universidade do Porto na divulgação do empirismo lógico em Portugal. Tanto quanto sei, os historiadores da ciência são unânimes em atribuir a Abel Salazar o papel maior na divulgação entre nós desta escola de pensamento. Para além de artigos esparsos que publicou desde 1934 nas revistas *Seara Nova* (Lisboa), *Sol Nascente* (Porto) e *Síntese* (Coimbra), Abel Salazar foi autor de cinquenta artigos entusiásticos sobre a Escola de Viena publicados n' *O Diabo* entre 1936 e 1940.²

E teve igualmente papel de relevo na disseminação do empirismo lógico Rui Luís Gomes, tanto através dos artigos sobre a Teoria da Relatividade que escreveu na *Seara Nova* como nas quatro conferências que foi convidado a proferir sobre o mesmo tema no Instituto Superior Técnico em 1937. Na lição magistral que proferiu nesse ano, comemorativa do primeiro centenário da Academia Politécnica e da Escola Médico-Cirúrgica do Porto, subordinada ao título "Análise neopositivista das noções

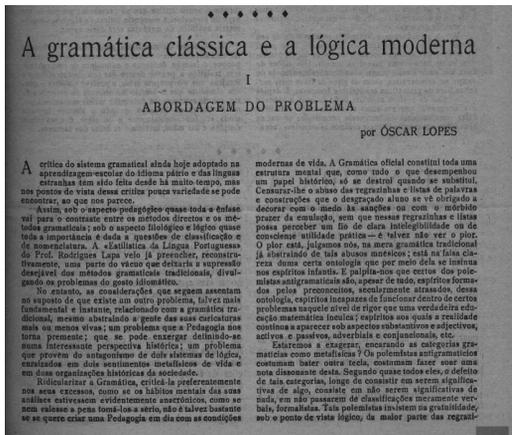


² Ver Fitas (2012).

de espaço e de tempo”, este então jovem professor catedrático afirmava que o empirismo lógico permitia, cito, “enfrentar os problemas de um domínio que dizem ser supra-científico – o filosófico –, por métodos exclusivamente científicos.”³

Regressando a Óscar Lopes, julgo ser possível afirmar que o contacto com a Escola de Viena iniciado nos seus tempos de faculdade e o intenso debate que sobre o assunto manteve com Vitorino Magalhães Godinho marcaram para sempre a sua visão sobre o que poderia ser uma teoria que permitisse não só descrever e compreender mas também fazer previsões sobre a linguagem e as línguas. Se o historiador se demarcou desta escola de pensamento logo na tese de licenciatura, em Óscar Lopes ficou a convicção (que nunca o abandonou) de que as teorias linguísticas, para atingirem uma nível de adequação satisfatório, precisam de importar ou de adaptar formalizações lógico-matemáticas.

Esta convicção foi, aliás, sendo fortalecida pela sua experiência docente: todos os testemunhos que li e ouvi de ex-alunos seus revelam que foi um excelente pedagogo. Não só foi capaz de fazer a transposição pedagógica inerente à atividade de quem ensina, como soube sempre retirar, do diálogo orientado para a aprendizagem que é uma aula, consequências para a sua investigação.



Este duplo aspeto – a necessidade de formalização na investigação sobre a linguagem e as línguas e o exercício pedagógico – estão já presentes no seu texto de duas páginas publicado na Seara Nova em 1947, com o título inesperado *A Gramática Clássica e a Lógica Moderna*.

³ *Apud* Bebiano (2006: 7).

A atitude de indagação característica dos cientistas, essa atividade de inquirir o porquê das coisas oculto pelo sedimento das ideias aceites, está bem presente nos parágrafos finais deste texto:

Porque haverá, precisamente, sete ou oito categorias morfológicas: substantivo, adjetivo, pronome e artigo, verbo, advérbio, preposição, conjunção? Porque será *franqueza* substantiva e não adjectiva? Serão, pois, as qualidades uns *substantivos abstractos*? Mas substância abstracta não será auto-contradição, não será assim como “uma coisa de ser uma qualidade”, isto é: “a coisa de não ser coisa”? Os pronomes indefinidos, porque são “pronomes”? Isto é: em vez de que nome pròpriamente dito estão as palavras “nada”, “tudo”, “alguém”, “cada qual”? E, vendo bem, que têm eles de indefinido, de mais indefinido que muitos substantivos? Que “acção” exprimem os verbos de *ter*, *ser* ou *sentir*? Que há de circunstancial nas palavras “*sim*”, “*eis*”, “*talvez*”, etc., que nos permita classificá-las entre os advérbios, ou até como inclui-las em qualquer das outras categorias morfológicas? Numa oração em que um substantivo é passivo de outro (agente da voz passiva), ou até numa oração em que se exprime uma circunstância de causa - onde reside, afinal de contas, a substância agente ou causal?

Estas perguntas sem resposta, mesmo quando se não definam no espírito do aluno e do mestre, desconfortam a sua inteligência. Ou mal vai ao ensino se não há ao menos o senso crítico que as pressente; porque só um espírito insensível à coerência íntima da álgebra, da geometria analítica, da mecânica será disso capaz.

A educação da metafísica gramatical e a educação matemática coexistem contraditòriamente hoje, só talvez porque a fracção dirigente e educadora da humanidade tem resistido a pensar dentro das relações científicas e técnicas que ela própria criou, e prefere visionar o quadro simplista que o mundo lhe oferecia no princípio da sua grande aventura, há dois mil e tantos centos de anos, no Mediterrâneo helénico, quando o alcance humano era fundamentalmente o do olhar, das máquinas simples, e das mãos, incluindo as mãos dos escravos.

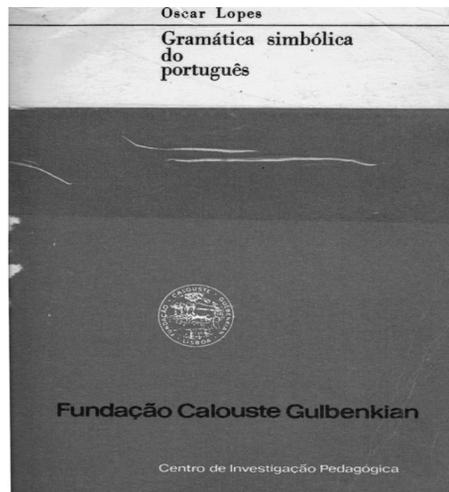
O parágrafo luminoso que fecha este texto é exemplar da capacidade que Óscar Lopes tinha de articular conhecimentos provenientes de vários domínios do saber e de os entrelaçar numa urdidura que tornava claros àqueles que o ouviam ou liam nexos que até então nem suspeitavam que pudessem ser estabelecidos.

A crítica a esta “educação da metafísica gramatical” é novamente tema da comunicação que apresentou no *I Simpósio Luso-Brasileiro sobre a Língua Portuguesa Contemporânea*, que teve lugar em Coimbra em 1967, com o título provocatório *Gramática escolar, reduto de um anacronismo*, em que se pode ler:

A minha tese central é a de que, quer no terreno da investigação, quer no da didática, a linguística não pode dispensar o estojo, ou antes, os estojos da racionalidade hoje mais comum aos diversos terrenos da investigação científica. Refiro-me à lógica simbólica, ou, melhor, a instrumentários de formalização como a teoria dos conjuntos, a das relações interproposicionais [...] e a das relações. Palpita-me, de resto, que a teoria dos grupos de transformação e a topologia seriam também muito úteis aos linguistas e professores de língua, sobretudo da língua materna.

É esta tese central que Óscar Lopes demonstra magistralmente na *Gramática simbólica do português (um esboço)*, publicada em 1971 e rapidamente esgotada, com 2.^a edição corrigida dada à estampa um ano depois.

Graças a uma bolsa da Fundação Calouste Gulbenkian e a uma equiparação a bolseiro concedida pelo Instituto de Alta Cultura durante quatro anos letivos, o Óscar pôde



desenvolver esta tese na abordagem de vários aspetos da gramática da nossa língua e teve a oportunidade de experimentar a sua transposição pedagógica em turmas experimentais de crianças do então denominado ciclo preparatório.

O resultado é impressionante de todos os pontos de vista.

Antes de mais, pelo conjunto de temas trabalhados inovadoramente. Relembro alguns: a distinção entre nomes próprios e comuns, atributos e apostos, coordenação copulativa e disjuntiva; os valores semânticos dos artigos definidos e indefinidos na sua relação com a flexão de número; a distinção entre expressões indefinidas e expressões partitivas; a deixis pessoal (incluindo, claro, pronomes pessoais e demonstrativos); as orações relativas restritivas e apositivas na sua relação com a frase superior; a aproximação entre orações coordenadas adversativas e subordinadas concessivas; o tempo como relação de ordem entre o momento do discurso e o da situação descrita.

O resultado é impressionante também pelos instrumentos lógico-matemáticos a que Óscar Lopes recorre, provenientes da teoria dos conjuntos, da lógica de proposições, do cálculo de predicados de 1.^a ordem, da teoria dos grafos, da topologia.

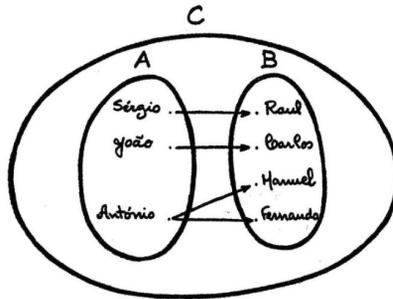
O resultado é impressionante, ainda, do ponto de vista pedagógico. É notável o cuidado nas indicações dadas aos leitores-professores para a transposição pedagógica das análises que propõe, destinada a crianças de 10 a 12 anos, bem como as formas de representação gráfica que usa para as ajudar a compreender as relações lógico-semânticas que são um dos pilares do significado.

O passo seguinte da *Gramática simbólica* ilustra suficientemente esta avaliação. Ao exemplificar relações binárias orientadas através da relação “x pai de y” que, como dizemos hoje, é um antónimo converso da relação “y filho de x”, Óscar Lopes afirma, a pp. 157-158:

Tal relação pode exemplificar-se do modo seguinte, supondo que os nomes próprios se referem cada qual a uma só pessoa num dado conjunto básico C.

E continua:

Isto quer dizer que, dentro do nosso conjunto-base exemplificativo, a relação P se verifica, por exemplo, entre Sérgio e Raul, mas não, por exemplo, entre João e Manuel; e podemos supor que ao conjunto C pertençam outros elementos, os quais, dentro desse mesmo conjunto, se não encontram associados pela mesma relação P. Por outro lado, [...] não basta dizer que a relação P, “pai de”, se verifica entre Sérgio e Raul; é preciso especificar que se estabelece de Sérgio para Raul, pois a relação de Raul para Sérgio é já diferente (Raul é filho de Sérgio).



O resultado é impressionante, além disso, pela bibliografia comentada (em que se incluem textos publicados em 1970 e em 1971) e pelas extensíssimas notas, que revelam a profundidade da sua reflexão sobre os temas de que se ocupa, ancorada na familiaridade com que manuseia a literatura relevante, quer se trate da tradição gramatical ocidental, do edifício lógico-matemático com alicerces em Frege e Russell, ou das teorias estruturalistas e generativistas que vieram a permitir as primeiras descrições fonológicas, morfológicas e sintáticas merecedoras do atributo “científicas”.

Vou exemplificar. A pp. 39, a propósito dos nomes próprios e comuns, Óscar Lopes afirma que, cito, “o uso de um nome postula a existência de um dado objecto (intensional); [...]”. No final do longo período de que apenas citei o início, surge a nota 4. Ora a nota 4, a pp. 287, sobre objetos intensionais, esclarece:

A concepção de objecto intensional, já debatida pelos escolásticos medievais e contra a qual reagem as tendências nominalistas, empenhadas em usar a chamada “navalha de Occam” para “rapar as barbas a Platão”, isto é, empenhados em eliminar meros “entes de razão” ou arquétipos platónicos, foi reabilitada por Franz Brentano (1874) e serviu de ponto de partida, quer para a “teoria dos objectos” (1904) de A. Von Meinong, alvo predilecto da crítica neo-positivista, quer da fenomenologia de Edmund Husserl (1900-1901). Os neo-positivistas, ou positivistas lógicos, que foram os nominalistas modernos, como os adeptos do “Círculo de Viena” do decénio de 1930 (Schlick, Neurath, Carnap), Wittgenstein, B. Russell e Tarski, empenharam-se em eliminar os objectos não-empíricos, quer impondo o princípio da verificabilidade experimental (Schlick), quer postulando dados sensoriais *protocolares* ou absolutos (E. Mach), quer exigindo a redução de todas as linguagens significativas à da física moderna (*fisicalismo* de Neurath e Carnap, noutra sua fase comum), quer, no decénio de 1940, construindo regras de *modelização* ou *interpretação semântica* para os sistemas formais (Tarski). Russell, na fase em que aderiu a uma espécie de *atomismo lógico*, sustentou que os *nomes próprios* (ao contrário das simples *descrições* singularizantes equivalentes ao uso normal do artigo definido singular, como em “o rio”, “o fauno”) postulariam a existência absoluta dos respectivos elementos designados, o que acarretou a sua concepção de *nomes próprios lógicos*, diferentes dos usuais; mas acabou por admitir (e, conforme veremos, paradoxalmente) que os únicos *nomes próprios lógicos* seriam “isto” e “aquilo” (“this” and “that”). Como verificaremos ao longo do nosso estudo, as teses lógico-positivistas radicais não se ajustam à análise da linguagem corrente, embora não escondamos as dificuldades inerentes à ontologia dos *objectos intensionais*.

Óscar Lopes revisita nesta nota episódios da história e da filosofia das ciências, da escolástica medieval à psicologia empírica de Brentano e à fenomenologia de Husserl, referindo as diferentes fases e contributos do Círculo de Viena e dos matemáticos e logicistas Wittgenstein, Tarski e Russell, para se centrar na teoria dos nomes próprios deste último, rematando com

uma avaliação crítica da adequação das teses lógico-positivistas radicais à análise da linguagem humana.

O resultado é impressionante, finalmente, por ter sido obra de um homem só, apenas incentivada por “animadoras críticas [n]a fase inicial dos trabalhos” de Sebastião e Silva, Simões Neto, Joana Lopes e Maria Helena Mateus. E pense-se no que isto significava na época: sem internet, sem bases de dados bibliográficas de fácil consulta, num país sem cultura de investigação científica nas Humanidades, com escasso acesso a livros e revistas estrangeiros. Como Óscar Lopes confessa, pesaroso, a pp. IX, a *Gramática simbólica* é o resultado de

um trabalho quase todo feito isoladamente, sem a colaboração de um seminário de investigação interdisciplinar (como devia ser o caso), sem acesso à enorme quantidade de comunicações mimeografadas, orais, por vezes de simples informação pessoal e directa, que abundam nas referências de obras desta ciência de ponta, que é actualmente a linguística em fase de rápida formalização matemática.

Em 1974, Abril devolveu-nos o sonho feito facto de um país a começar de novo. E Óscar Lopes entra finalmente na universidade: é nomeado professor catedrático e diretor da Faculdade de Letras da Universidade do Porto. Aqui, e até à sua jubilação, concebe e leciona uma unidade curricular inovadora, Linguística Matemática, dá aulas de licenciatura e de mestrado, orienta e co-orienta teses de mestrado e de doutoramento, funda o Centro de Linguística da Universidade do Porto, o CLUP. E estuda, estuda e investiga.

Assim, quatro anos depois da *Gramática simbólica*, a evolução do seu pensamento como linguista é clara. Na comunicação que apresentou ao 1.º Encontro Nacional para o Ensino e a Investigação do Português, que reuniu em Lisboa docentes e investigadores de todo o país, intitulada *Necessidade de um Mínimo de Consciência Formal da Língua na Prática Escolar*, Óscar Lopes traz à colação o que considera serem, cito, “dois sentidos na actual investigação de ponta”⁴: o primeiro é o

⁴ Lopes (1976: 380-381).

[...] das lógicas intensionais do tipo de R. Montague, que procurou formalizar as condições de verdade de textos linguísticos com características determinadas, em função de índices como o tempo e a atitude proposicional do loquente, índices que será possível alargar até abranger o espaço de locução, as relações entre o loquente, o interlocutor imediato, o público indirecto, etc. [...]. (7)

O outro sentido de investigação de ponta [...] assenta na pragmática, isto é, na teoria da linguagem como acto social preciso, como forma específica de comportamento. Estas investigações pragmáticas põem, pelo menos provisoriamente, de parte o tratamento algébrico da semântica linguística, mas tendem para a representação por meio de diagramas que se aproximam dos diagramas do cálculo da programação linear, da teoria dos sistemas, e de outros métodos próximos, se não integráveis, no conjunto das disciplinas das matemáticas. (8)

A nota 7 remete para a colectânea de artigos de Montague organizada por Thomason e publicada em 1974, que inclui o célebre (e complicadíssimo) *PTQ* (*The Proper Treatment of Quantification in Ordinary English*). Como é sabido, o *PTQ* é o artigo em que Montague propõe e aplica a fragmentos do inglês um sistema de lógica intensional temporalizada. Por sua vez, a nota 8 remete para uma coletânea de textos de pragmática, organizada por Dieter Wunderlich e publicada em 1975, intitulada *Linguistische Pragmatik*. Não pode deixar de espantar a capacidade intelectual e de trabalho de Óscar Lopes que, num período simultaneamente empolgante e conturbado da nossa história coletiva, mantém intensa atividade como cidadão, entra no mundo universitário, é diretor da FLUP, dá as primeiras aulas de linguística matemática, funda o Centro de Linguística da Universidade do Porto e se mantém a par das investigações de topo nos domínios da semântica e da pragmática.

É novamente disso exemplo a comunicação que apresentou em 1985 ao 1.º Congresso da Associação Portuguesa de Linguística, intitulada “A noção de definido e a de presente da enunciação”. É um texto espantoso, de grande complexidade, em que Óscar Lopes “descasca” os diferentes valores têmporo-aspetuais do presente do indicativo e do pretérito perfeito

composto, passeia por formas linguísticas de expressão de incerteza e de contrafactualidade, revisita os conceitos de referência definida e indefinida, aproxima deixis e anáfora, e ata este ramallete através do conceito de presente de enunciação, encarado como, cito “um acto discreto de linguagem unilinearmente, ou irreversivelmente, ligado a outros actos de linguagem.” (p. 138). E, noutra passagem, a pp. 136, clarifica o conceito:

Trata-se de um presente de enunciação no sentido em que presentifica um dado, ou conjunto de dados, que na interpretação do destinatário (ou destinatários) pode até vir a funcionar de maneira diferente daquela que está prevista pelo enunciador, o que é capaz de motivar um reacerto dialogal.

A pp. 141-142, para dar conta da dependência do contexto dos tempos e da referência definida e indefinida, Óscar Lopes sugere que se opte por representações semânticas como as propostas nos primeiros esboços da *DRT (Discourse Representation Theory)*, uma teoria semântica dinâmica desenvolvida por Hans Kamp a partir de 1981, uma novidade na época, e uma das mais usadas hoje pelos semanticistas.

Depois dos meus primeiros encontros com o Óscar em 1973, vim para a Faculdade de Letras do Porto em 1976, juntamente com duas colegas de curso, as minhas amigas Ana Brito e Fátima Oliveira, para poder aprender e trabalhar com ele. Posso dizer que foi aqui que dei os primeiros passos como docente universitária e como aprendiz de linguista. Os capítulos que escrevi para a 1.^a edição da *Gramática da Língua Portuguesa*, publicada em 1983 pela Almedina, são quase todos resultado do que aprendi com ele de viva voz e através dos seus escritos. Foi com ele que aprendi que, cito, “as línguas naturais não são absolutamente naturais, são humanamente naturais.” (1976: 389)

Quando voltei para Lisboa, pedi ao Óscar que fosse co-orientador do meu doutoramento. E, embora tenha decidido desenvolver a investigação conducente à tese no campo da sintaxe, a própria escolha do tema foi a consequência dos anos do Porto: os tópicos marcados, um tema que trazia para o centro do trabalho o modo como as configurações sintáticas permitem exprimir relações semântico-discursivas, num tempo em que não se falava

ainda da interface sintaxe-discurso.

Cientista e pedagogo, o Óscar vivia e praticava a emoção da descoberta em diálogo. Ouçamo-lo, numa intervenção na homenagem que lhe fez a revista *O Professor* em 1988⁵:

Há qualquer coisa de mais sagrado que uma pessoa é o encontro entre pessoas. Sente-se isso de um modo exaltante quando se canta em coro ou quando se faz parte de um grupo instrumental. [...]

Tive a sorte de exercer a profissão de professor durante cerca de meio século [...] Ora, a profissão de professor é emocionante precisamente porque a aula, ou o grupo de estudo constitui um espaço de convivência na acepção mais profunda do verbo “conviver”. É um espaço de frequente descoberta, de surpreendente encontro ou sintonização onde em diálogo real ou virtual nos sentimos chegar muitas vezes às mesmas e novas opiniões ou às mesmas e novas razões. (p. 61)

Retiro desta última citação dois aspetos fundamentais, que mostram bem que fazer ciência não é tecer um perpétuo bordado de Penépole. Em primeiro lugar, Penélope sabia bem qual o resultado final do bordado: o que desfazia numa noite era refeito de um modo exatamente igual durante o dia seguinte. Não havia aqui a emoção da descoberta. Mas em ciência, nunca sabemos o que vamos encontrar no fim do caminho a que nos leva cada investigação. Cada hipótese descartada pelos dados empíricos ou por uma teoria com maior poder explicativo ajuda-nos a conhecer um pouco mais, a procurar melhores respostas para os problemas que pretendemos resolver e, sobretudo, a identificar novas perguntas de investigação.

Em segundo lugar, fazer ciência é cada vez mais um trabalho de equipas, ou de orquestras, como o Óscar preferiria certamente dizer. Precisa de um ambiente em que maestro, primeiros solistas, músicos de cada naipe e cordas-sopro-percussão aprendam a ouvir-se, dialoguem, e treinem, treinem, treinem. Privado que foi de um ambiente desses durante os anos de maior potencial criativo da sua vida, Óscar Lopes sabia disso, tinha-o

⁵ Lopes (1988: 61).

aprendido na pele, e o seu magistério universitário e a criação do CLUP foram a semente de um viveiro linguístico que já deu saborosos frutos.

A demanda do conhecimento através da ciência é uma formidável empresa democrática, porventura uma das mais conseguidas da espécie humana. Foi também um crime contra a ciência que o Estado Novo cometeu ao negar a Óscar Lopes o seu lugar de direito na universidade. Com isso privou gerações de estudantes portugueses de se aventurarem pelos caminhos que logo nos anos quarenta ele começava a trilhar. Se o seu legado na Linguística é o que todos nós lhe reconhecemos, o que não teria sido, se o seu magistério tivesse tido o seu início trinta anos antes? É que Óscar Lopes é um desses seres raros que percebem antes dos outros os caminhos que fazem avançar as fronteiras do conhecimento.

Por isso, nada melhor que torná-lo presente, como procurei fazer hoje aqui, dando voz à sua voz e dando testemunho do privilégio que foi conhecê-lo e trabalhar com ele.

REFERÊNCIAS

- Bebiano, N. (2006). Ruy Luís Gomes - Vida e Obra. *Gazeta de Matemática*, 151: 6-16.
- Fitas, A. J. (2012). Abel Salazar e a difusão da teoria da relatividade na Imprensa Cultural Portuguesa. *Vértice*, 162: 41-53.
- Lopes, Ó. (1947). A Gramática Clássica e a Lógica Moderna. *Seara Nova*, 1031: 7-8.
- Lopes, Ó. (1967). Gramática escolar, reduto de um anacronismo. *I Simpósio Luso-Brasileiro sobre a Língua Portuguesa Contemporânea*.
- Lopes, Ó. (1972). *Gramática simbólica do português (um esboço)*. Lisboa: IGC - CIP.
- Lopes, Ó. (1976). Necessidade de um Mínimo de Consciência Formal da Língua na Prática Escolar. In AA, (1977) *1.º Encontro Nacional para o Ensino e a Investigação do Português - 1976 actas: 377-390*. Águeda: Centros de Linguística das Universidades de Lisboa, da Universidade do Porto, da Universidade de Coimbra, Núcleos de Linguística de Aveiro e Braga.
- Lopes, Ó. (1985). A noção de definido e a de presente de enunciação. In Direcção da APL, *1.º Encontro da Associação Portuguesa de Linguística. actas: 129-143*. Lisboa: APL
- Lopes, Ó. (1988). Caros amigos... *O Professor*, 105: 61-64.
- Magalhães Godinho, V. (2007). *Definição fundamental: um homem de carácter*.
<http://www.porto.pcp.pt/index.php/oscar-lopes/nas-palavras-de/2445-vitorino-magalhaes-godinho>
- Neves, L. C., org. (2004). *António José Saraiva e Óscar Lopes: Correspondência*. Lisboa: Gradiva.

Relativas livres e interrogativas parciais: como Óscar Lopes viu algumas destas construções¹

Ana Maria Brito
abrito@letras.up.pt

Faculdade de Letras /Centro de Linguística da Universidade do Porto

Gabriela Matos
gmatos@letras.ulisboa.pt

Faculdade de Letras /Centro de Linguística da Universidade de Lisboa

RESUMO

Neste artigo analisamos alguns aspetos sintáticos e semânticos das relativas livres e das interrogativas subordinadas parciais próprias e impróprias, desenvolvendo Brito (1991), Brito & Duarte (2003), Mória (1996), Veloso (2013), Matos & Brito (2013), Miotto & Lobo (2016), correlacionando-os com a forma como Óscar Lopes descreveu estas construções e destacando a ambiguidade estrutural e semântica de algumas delas. Defendemos a existência de diversos tipos de relativas livres com leitura definida, com leitura universal e com leitura indefinida.

PALAVRAS-CHAVE

interrogativas subordinadas parciais, interrogativas verdadeiras, interrogativas impróprias ou falsas, relativas livres, Óscar Lopes, ambiguidade

ABSTRACT

In this paper we analyze some syntactic and semantic properties of free relatives and *wh* subordinate interrogatives, both proper and improper, developing Brito (1991), Brito & Duarte (2003), Mória (1996), Veloso (2013), Matos & Brito (2013), Miotto & Lobo (2016), correlating them with the way Óscar Lopes described these constructions; we highlight structural and semantic ambiguity of some of *wh* constructions. We argue for the existence of several types of free relatives, mainly ones with definite, universal and indefinite meaning.

KEYWORDS

wh subordinate interrogatives, proper and improper interrogatives, free relatives, Óscar Lopes, ambiguity

¹ Uma versão prévia deste texto foi apresentada no *Colóquio Comemorativo do 100º Aniversário do Nascimento de Óscar Lopes*, ocorrido nos dias 6 e 7 de Junho de 2017, na FLUP. Agradecemos à audiência os comentários feitos e a um revisor anónimo algumas observações que nos permitiram melhorar o texto original.

1. Introdução

No conjunto das chamadas construções *wh*² destacam-se as orações interrogativas e relativas, que têm em comum o emprego das expressões *wh*, tradicionalmente designadas pronomes e advérbios interrogativos e relativos.

As orações interrogativas parciais independentes, como (1a), caracterizam-se pela presença de formas *wh*, em geral colocadas na periferia esquerda da frase, embora na modalidade do oral seja também frequente o uso de *wh in situ*, como no exemplo (1b)³. Pelo contrário, as interrogativas parciais subordinadas não admitem a estratégia *in situ* quando apenas um constituinte é interrogado, como se pode ver no contraste de aceitabilidade entre (1c) e (1d). Esta última propriedade aproxima as interrogativas parciais subordinadas das orações relativas, como mostra (2):

- (1) a. O que comeste?
 b. Comeste o quê?⁴
 c. Perguntei-lhe o que comeu.
 d. *Perguntei-lhe comeu o quê.
- (2) a. Devolvi o que (tu) acabaste de ler.
 b. *Devolvi (tu) acabaste de ler o que/o quê.

No entanto, outras propriedades distinguem as interrogativas das relativas. Em especial, para muitos autores, entre os quais Óscar Lopes, as frases interrogativas caracterizam-se por não terem valor de verdade. Em interrogativas globais, como (3), a frase tem valor ilocutório de pergunta, mas, do ponto de vista semântico, é um molde proposicional, nem verdadeiro nem falso, sendo a resposta *sim* ou *não* que lhe atribui valor de verdade (Lopes (1971: 14, 249)⁵.

² Por estarem muito divulgadas nos estudos sintáticos, usaremos as siglas das categorias em inglês, assim como a própria designação de palavra ou sintagma *wh* para os constituintes interrogativos e relativos.

³ Sobre *wh in situ* ver, e.o., Miotto & Lobo (2016: 275-277).

⁴ Sobre a sintaxe das interrogativas *wh in situ* ver, entre outros, Kato (2013).

⁵ Uma análise alternativa foi adotada por Karttunen (1977) e Adger & Quer (2001), que assumem que as

- (3) a. Estará a chover?
- b. Podes chegar aqui?

No caso das interrogativas parciais, como (1a)-(1c), há a pressuposição de existência de uma entidade (ou circunstância de tempo, lugar, maneira) e o locutor solicita uma informação sobre a identificação da variável, cujo valor desconhece.⁶

As interrogativas parciais em subordinadas são em geral selecionadas por predicados de inquirição (*perguntar(-se)*, *inquirir*, *querer saber*, cf. (4)); de desconhecimento (*ignorar*, *desconhecer*, *não saber*, cf. (5)); de comunicação e percepção (*dizer*, *ver*, *ouvir*), quando estão sob o escopo de um elemento negativo e de um operador de interrogação da frase raiz (cf. (6)), ou com o verbo *dizer*, quando este ocorre em frases imperativas que solicitam uma informação (cf. (7)).⁷

- (4) Perguntaram quem veio.
- (5) Desconheço o que aconteceu.
- (6) a. Não viste o que poderia perturbar a receção?
- b. Disseram-te quem veio?
- (7) a. Diga a quem gostaria de falar.
- b. Diga se precisa de alguma coisa.

Por sua vez, as orações relativas são subordinadas que funcionam como modificadores de um constituinte, tipicamente uma expressão nominal, como em (8), concebida como o antecedente do constituinte relativo:

interrogativas globais devem ser semanticamente analisadas como uma disjunção de proposições.

⁶ Porque o desconhecimento pode recair sobre *mais do que uma* entidade, o locutor pode solicitar o valor de *mais do que* uma variável, produzindo neste caso, interrogativas *wh* múltiplas, como ilustrado em (i). Em línguas como o Português só um dos elementos pode ocupar a posição inicial da frase, ficando os restantes nas respetivas posições básicas:

(i) Quem deu o quê a quem?

Em línguas como o Romeno, o Búlgaro, o Checo, o Polaco ou o Russo, há movimento-*wh* múltiplo para a periferia esquerda das frases (cf., por exemplo, Rudin 1988, Stoyanova, 2008).

⁷ Veja-se o trabalho de Adger & Quer (2001) sobre interrogativas não selecionadas em diferentes línguas.

- (8) a. Eu devolvi o livro que acabaste de ler.⁸
b. Cumprimentei a pessoa que chegou.

Para além das relativas com antecedente⁹, existe um outro tipo de relativa em que a expressão nominal antecedente não é explícita, surgindo então a forma *wh* em posição inicial da oração subordinada, como em (9) e (10):

- (9) Eu devolvi o que acabaste de ler.¹⁰
(10) Cumprimentei quem chegou.

Os exemplos em (9) contêm uma oração subordinada que tem sido designada *relativa livre* (Brito 1991)¹¹. É sobre este tipo de relativa que nos debruçaremos.

Uma das características deste tipo de construção é a sua semelhança aparente com interrogativas parciais subordinadas¹², vejamos os exemplos em (4) e (5). Em línguas como o português, tanto as relativas livres, como (9) e (10), como as interrogativas parciais subordinadas, como (4) e (5), são iniciadas pelas expressões *quem*, *o que*, *onde* e *são*, por isso, confundíveis.

Na sua *Gramática Simbólica do Português*, Óscar Lopes, ao debruçar-se sobre algumas construções *wh* subordinadas do tipo das mencionadas, em particular sobre os exemplos (11), (12) e (13), tece acerca delas considerações semânticas interessantes e inovadoras, sugerindo que a sua classificação é talvez o que menos interessa:

⁸ Sobre a natureza de *que* nas relativas com antecedente ver Brito (1991), Brito & Duarte (2003), Veloso (2013), mais recentemente Rinke & Assmann (2017).

⁹ Designadas em inglês “headed relatives”, na tradição linguística portuguesa estas relativas são geralmente denominadas “relativas com antecedente”. Para uma discussão sobre a sintaxe das orações relativas com antecedente ver Brito (1991), Brito & Duarte (2003), Cardoso (2017), Alexandre (2012), entre outros.

¹⁰ O exemplo (9) tem ainda uma possível leitura de elipse nominal (*Devolvi o [-]_N que acabaste de ler*, em que [-]_N pode ter o valor de *livro*, mencionado no discurso anterior. Trata-se das chamadas falsas relativas livres, que não nos vão interessar neste artigo.

¹¹ Esta construção foi igualmente designada pelos termos de relativas *sem antecedente exposto* (Móia 1996), ou de *antecedente implícito* (Veloso 2013).

¹² Quer as interrogativas parciais quer as interrogativas globais encaixadas / subordinadas e selecionadas por um verbo matriz têm recebido nos estudos gramaticais a designação de ‘interrogativas indiretas’. Como as designações de ‘direta’ e ‘indireta’ podem ter uma incidência pragmática, no texto usaremos preferencialmente a designação de ‘subordinada’ ou ‘encaixada’.

(11) Disseram-me quem veio.

(12) Quem nasceu morrerá.

(13) É meu (tudo) o que vês.

No entanto, importa-nos analisar as propriedades destas construções à luz da discussão sobre interrogativas parciais subordinadas e relativas livres e, em particular, responder às seguintes questões: (i) Serão alguns dos exemplos em causa ambíguos entre a estrutura de relativa livre e a estrutura de interrogativa parcial subordinada? (ii) Por que razão, de diferentes modos, as relativas livres têm, apesar de tudo, mais pontos em comum com interrogativas parciais subordinadas do que com relativas com antecedente? (iii) Por que razão as relativas livres não são todas iguais e a sua interpretação é muito mais variável do que a das interrogativas parciais?

O texto tenta responder às questões apresentadas e está organizado da seguinte maneira. No ponto 2, depois desta introdução, faz-se um estudo acerca das semelhanças e diferenças entre relativas livres e relativas com antecedente, e entre relativas livres e interrogativas parciais subordinadas. No ponto 3, analisamos a forma como Óscar Lopes descreveu estas construções, em particular os casos ambíguos, procurando explicar o que poderá estar na base dessas ambiguidades e desenvolvendo as pistas suscitadas pelos dados considerados. Assim, analisaremos subtipos de relativas livres e abordaremos a leitura definida e universal das relativas livres, assim como a possibilidade de leitura indefinida. Seguem-se as conclusões e a bibliografia.

2. Relativas livres: semelhanças e diferenças com relativas com antecedente e interrogativas subordinadas próprias e impróprias

Na presente secção, recordaremos algumas propriedades que aproximam e distinguem relativas livres e com antecedente, e relativas livres e interrogativas subordinadas parciais próprias e impróprias.

2.1. Relativas livres e relativas com antecedente: semelhanças e diferenças

As relativas livres apresentam contrapartidas em relativas com antecedente realizado, como ilustrado em (14):

- (14) a. Eu cumprimentei quem chegou.
b. Eu cumprimentei a pessoa que chegou.

De facto, as relativas livres podem surgir exatamente nos mesmos contextos que os sintagmas nominais contendo relativas com antecedente: para além de poderem ser selecionadas por verbos da oração matriz que não exijam frases por argumento — é o caso de *cumprimentar* em (14a) —, as relativas livres podem ser selecionadas por preposições (15a) e podem ocorrer em contextos não subcategorizados: veja-se (16a), em que a relativa livre é sujeito da frase matriz, (17a), em que a relativa livre é predicado, e (18a), em que a relativa livre está em Tópico:

- (15) a. As novas tecnologias de comunicação são boas para quem é novo.
b. As novas tecnologias de comunicação são boas para as pessoas que são novas.
- (16) a. Quem chegou é meu amigo.
b. O indivíduo que chegou é meu amigo
- (17) a. A Maria é quem está mais atenta nas aulas.
b. A Maria é a pessoa que está mais atenta nas aulas.
- (18) a. Com quem eu quero casar, esse é mesmo um querido!
b. A pessoa com quem eu quero casar, esse é mesmo um querido!

Há línguas com relativas livres mais complexas, com uma leitura universal, e que não admitem facilmente uma paráfrase com antecedente. É o caso do Inglês, como mostra Riemsdijk (2007: 360), em que *whatever*, diferentemente de *what*, não pode coocorrer com *no matter*:

- (19) a. (*No matter) whatever Bill says, it always comes out wrong.
b. No matter what Bill says, it always comes out wrong.

Mas também há diferenças entre relativas livres e relativas com antecedente. Como mostrado em Brito (1991), há restrições no uso dos morfemas *wh* em relativas livres; assim, *cujo* e *o qual* nunca são usados em relativas livres, sendo exclusivos das relativas com antecedente.¹³

- (20) a. Acabei de ler um jornal cujo editorial trata da crise do *Facebook*.
b. *Acabei de ler cujo editorial trata da crise do *Facebook*.
(21) a. Acabei de ler um jornal com o qual me identifico muito.
b. *Acabei de ler com o qual me identifico muito.¹⁴

Este é um aspeto que aproxima as relativas livres mais das interrogativas parciais subordinadas (**Pergunto cujo leste; *Pergunto o qual leste*) do que das relativas com antecedente. Contudo, como veremos a seguir, há muitas diferenças entre relativas livres e interrogativas parciais subordinadas.

2.2. Relativas livres e interrogativas parciais subordinadas: semelhanças e diferenças¹⁵

Como acabámos de ver, relativas livres e interrogativas subordinadas são aproximáveis quanto ao não emprego de algumas formas *wh* (nomeadamente *cujo* e *o qual*), o que, por sua vez, as afasta das relativas com antecedente. No entanto, há propriedades que distinguem relativas livres e interrogativas

¹³ Quanto ao emprego de *que*, a questão é mais complicada; na verdade, se *que* pode ser usado em relativas infinitivas (Brito 1991, Móia 1996, Mioto & Lobo 2016) ele é, no entanto, impossível numa relativa livre finita, embora seja o elemento mais comum em relativas com antecedente em relativas de sujeito e de objeto:

- (i) Procuo (o) que ler nas férias.
(ii) *Devolvi que acabaste de ler.
(iii) Devolvi o livro que acabaste de ler.

Como sabemos, *que* pode ser usado em interrogativas quer como variante de *o que* (como em (iv) e (v), quer como determinante (vi), o que mostra os seus diferentes valores e revela que certamente o *que* interrogativo não é da mesma natureza do *que* relativo (sobre esta questão, sobre a qual não nos alongaremos aqui, há uma extensa bibliografia para o Português, ver Brito 2001, Veloso 2013, Rinke & Assman 2017, entre outros):

- (iv) Que leste?
(v) O que leste?
(vi) Que livro leste?

¹⁴ Este exemplo seria sempre agramatical como relativa livre por não exibir conformidade categorial.

¹⁵ Vamos neste ponto seguir muito de perto Matos & Brito (no prelo).

parciais subordinadas. Consideremos algumas das principais diferenças entre relativas livres e interrogativas subordinadas parciais, que têm sido enunciadas por vários autores (Brito 1991, Matos & Brito 2013, Veloso 2013, entre outros, para o português europeu).

(i) Uma dessas propriedades é a ocorrência num mesmo domínio oracional de mais do que um morfema *wh*. O fenómeno de interrogativas *wh* múltiplas é uma característica que parece ser universal (cf. (22)). Pelo contrário, as relativas livres múltiplas são tipicamente agramaticais, como mostram os exemplos em (23)-(25), embora com formas do tipo *o que quer que seja*, *quem quer que seja* seja possível detetar alguma melhoria de aceitabilidade:

(22) Eu perguntei *quem* deu *o quê* a *quem*. (interrogativa subordinada)

(23) *Eu devolvi *o que* deu *quem* a *quem*. (relativa livre)

(24) ??*O que quer que seja que faça a quem*, eu apoio-te.

(25) ??Eu admiro *quem quer que seja que faça o quê* nestes momentos de crise.

(ii) Já vimos que as interrogativas parciais subordinadas são essencialmente selecionadas por predicados matriz de comunicação, de inquirição ou de desconhecimento (como *perguntar*, *não saber*, etc.) que selecionam frases por argumento, enquanto as relativas livres podem ser selecionadas por predicados da frase matriz que não obedecem a essa restrição (ver exemplos (14a), (15a) e (16a)).

(iii) Tais verbos podem, em certas circunstâncias, selecionar um sintagma de natureza nominal (cf. (27a)), mas de modo muito mais restrito do que os verbos que selecionam relativas livres (veja-se (27b))¹⁶:

¹⁶ Como tem sido notado, um verbo superior de inquirição, desconhecimento ou comunicação só pode selecionar um sintagma nominal nas chamadas interrogativas escondidas (*concealed questions*) (ii) ou quando o argumento interno pode ser objeto de conhecimento (iii) (Matos & Brito 2013, Veloso 2013):

(i) * Eu perguntei o livro / * Eu ignoro o livro / * Eu disse o livro.

(ii) Eu perguntei as horas.

(iii) Eu perguntei / não sei / ignoro o caminho.

- (27) a. Pergunto {√ *isso* /?*essa questão* / **essa atitude* / **essa causa*}
b. Admiro {√ *isso* /√ *essa atitude* / √ *essa causa* /? *essa questão*}

A propriedade (iii) decorre do facto de as interrogativas parciais subordinadas serem seleccionadas por predicados que preferencialmente aceitam argumentos oracionais, como indicado em (ii), acima.

(iv) Estes mesmos verbos podem seleccionar uma interrogativa parcial subordinada iniciada por *qual* seguido de *ser* com valor de identidade ou por determinante interrogativo *que* seguido de N (Velooso 2013: 2125):

- (28) a. Ele perguntou *quais eram as pessoas que* estavam na sala no momento da votação.
b. Ele perguntou *que pessoas* estavam na sala no momento da votação.

Pelo contrário, as relativas livres não admitem nenhuma destas construções (Velooso 2013: 2125):

- (29) a. *Eu admiro *quais são* as pessoas que vêm jantar.
b. *Eu admiro *que pessoas* vêm jantar.

(v) Decorre de (ii) e (iii) que as interrogativas parciais subordinadas, mas não as relativas livres, admitem a aproximação à construção de *se* + oração, uma vez que essas orações são o correspondente a interrogativas subordinadas globais (Velooso 2013: 2124):

- (30) a. Pergunto *se* eles fizeram isto.
b. *Admiro *se* eles fizeram isto.

(vi) As interrogativas subordinadas, como (31), ao contrário das relativas livres no português europeu padrão (32), podem coocorrer com *é*

*que*¹⁷, sugerindo que nas interrogativas parciais o constituinte interrogado se encontra focalizado, enquanto tal não acontece com o constituinte-*wh* nas relativas livres.

- (31) Pergunto/ignoro o que é que fizeste.
 (32) ??/*Cumprimentei quem é que chegou.

Contudo, as orações argumentais exibindo o morfema *wh* não são todas iguais, devendo distinguir-se entre as interrogativas parciais subordinadas verdadeiras ou próprias, como (4) e (5), repetidos em (33), que relatam uma pergunta ou exprimem um desconhecimento acerca do valor de uma variável, e as falsas ou impróprias, como (34), que têm um valor declarativo, e não solicitam o valor do constituinte instanciado pelo morfema *wh*, como mostra o facto de o valor desse constituinte poder ser explicitado no enunciado pelo próprio locutor/sujeito (cf. (34b)), sendo, por isso, proposições verdadeiras ou falsas (veja-se Suñer 1991, 1993, 1999, Rivero 1994, Lahiri 2002, para o Espanhol, Matos & Brito 2013, para o português)¹⁸:

- (33) a. Perguntaram *quem veio*.
 b. Desconheço *o que aconteceu*.
 (34) a. Ela explicou-me *o que a incomodava*.
 b. Descobri *quem chegou: o Pedro e a Paula*.

Usualmente, a natureza do predicado superior é suficiente para determinar o valor interrogativo (cf. *perguntar, desconhecer*) ou declarativo

¹⁷ Cardoso & Alexandre (2013) encontraram no corpus do CORDIAL-SIN algumas relativas livres com *é que*:

(i) “Portanto, o que é *que* cai é a neve e da neve faz água.”(ii) “Quem é que tinha matado aquele bicho casava com a filha.” Assinale-se que o corpus do CORDIAL-SIN é constituído por produções orais selecionadas de inquéritos dialectais representativos de 42 localidades ou micro-regiões no território português continental e insular. A proveniência é sociolinguisticamente homogénea, sendo as amostras produzidas por falantes naturais e residentes nas localidades rurais inquiridas, idosos e pouco escolarizados ou analfabetos. As autoras consideram que nas frases acima ilustradas há uma leitura de foco e de exaustividade, que pode ter uma paráfrase com advérbios do tipo de *precisamente / exatamente*. Do ponto de vista da análise sintática, este fenómeno justificará a hipótese de que CP ou ForceP pode marginalmente admitir FocP nas relativas livres.

¹⁸ Recorde-se que as interrogativas verdadeiras não são asserções (declarativas) elas solicitam ao interlocutor sobre o valor de verdade ou falsidade de uma dada situação ou estado de coisas (ver a este propósito Lopes 1971: 14, 249), não tendo por isso um estatuto de verdade ou de falsidade.

assertivo (cf. *explicar; descobrir*)¹⁹ da subordinada *wh*; neste segundo caso, trata-se de um predicado que também seleciona um sintagma nominal contendo uma relativa (*Ela explicou-me o problema que a incomodava / descobri a pessoa que chegou*). Porém, há casos em que tal não acontece e temos então uma construção, que, no seu conjunto, é ambígua estrutural e semanticamente, como (35a), em que a frase encaixada pode ter as interpretações explicitadas de (35b) a (35d): (i) interrogativa parcial subordinada própria, como indica a paráfrase em (35b) em que se introduz a subordinada; (ii) interrogativa imprópria, como mostra a possibilidade de ser parafraseada por uma frase declarativa iniciada pelo complementador *que* (cf. (35c)); (iii) e relativa livre, exibindo a paráfrase explicitada em (35d):

- (35) a. Ele ignorava o que o esperava.
b. Ele ignorava se alguma coisa boa/má o esperava.
c. Ele ignorava que alguma coisa boa/má o esperava.
d. Ele ignorava aquilo que o esperava.

Tendo em vista propriedades como as apresentadas acima, tem sido assumido que as interrogativas parciais subordinadas são frases, argumentos internos oracionais de certos predicados superiores, que se distinguem claramente das relativas livres, que estão integradas numa estrutura de natureza nominal (Brito 1991, Mória 1996, Veloso 2013, Matos & Brito 2013, Miotto & Lobo 2016). No entanto, considerando as propriedades das interrogativas impróprias, apercebemo-nos de que as fronteiras entre estas construções *wh* não são completamente nítidas, não obedecendo a uma lógica de bipartição. Emerge, então, uma gradação de construções *wh*, tendo em vista o seu estatuto como orações sem ou com um valor de verdade atribuível e a sua não inclusão ou inclusão num constituinte nominal: interrogativas próprias e interrogativas impróprias e relativas livres. Finalmente, há casos ambíguos, em que a natureza do predicado da oração superior por si só não permite decidir que tipo de construção está presente.

¹⁹ Trata-se de verbos que, na classificação de Hinzen & Sheehan (2011), são fracamente assertivos, cognitivos e definidos.

3. A análise de Óscar Lopes: problemas e pistas levantadas

As considerações apresentadas na secção anterior permitem-nos analisar algumas posições assumidas por Óscar Lopes em relação a construções *wh* do tipo de (11), (12) e (13) e evidenciar em que medida elas se justificam.

3.1. Interrogativas próprias e impróprias e ambiguidade com relativas livres

Ao analisar a frase (11), aqui renumerada como (36),

(36) Disseram-me quem veio.

Óscar Lopes (1971: 218-9) aproxima-a de outras frases em que a seguir a “disseram-me” podemos ter: “qual é o caminho”, “que perda sofreste”, “quem veio”, “quantos faltam”, “quanto custa”, “onde estiveste”, “como procedeste”, “porque fugiste”, “quando ele virá”, propondo que estamos “perante a expressão originariamente interrogativa ou interrogativa-relativa de uma variável referentes a coisa(s), pessoa(s) ou circunstância(s) determinada(s).” E continua: “O modo interrogativo pode quase sempre interpretar-se como expressão de variável *livre* e portanto de molde proposicional. Mas as variáveis normalmente expressas por tais pronomes ou advérbios (*que, quem, quanto, onde, como, porque, quando*) aparecem nestas frases exemplificativas como quantificadas.”

E acrescenta: “Essa quantificação admite duas interpretações que o contexto permite, ou não distinguir. Com efeito, “disseram-me quem veio” pode significar “disseram-me que pessoa veio” ou “disseram-me que pessoas vieram” ou a disjunção lógica inclusiva das duas alternativas” (Lopes 1971: 219).

Assim, para Óscar Lopes, a ambiguidade de (36) reside no valor de *quem*, que pode designar tanto uma entidade singular ou uma entidade plural, ambiguidade essa que pode ser explicitada em paráfrases como (37), que são frases estruturalmente distintas e que patenteiam, ao nível da concordância plural no nome (*pessoa/pessoas*) e no verbo (*veio/vieram*), cada uma das leituras de (36):

- (37) a. Disseram-me que pessoa veio.
b. Disseram-me que pessoas vieram.

Note-se que, em (36), nenhum valor é solicitado pelo locutor para *quem*, não havendo, portanto, uma subordinada interrogativa parcial, e, pelo contrário, infere-se que o locutor detém o conhecimento de *quem são as pessoas que vieram*, ainda que não o explicita; contudo, pode fazê-lo, como atestado em (38). Adicionalmente, a frase tem um valor de verdade atribuído: é assumida como verdadeira, como mostra a estranheza de exemplos como (39):

- (38) Disseram-me quem veio: o Pedro, o Manuel e a Paula.
(39) #Disseram-me quem veio, mas é falso, pois não veio ninguém.

Quer dizer, a oração subordinada em (36) não tem valor interrogativo, não podendo ter como paráfrase uma interrogativa global encabeçada por *se*, mas sim com o complementador *que*, característico das frases subordinadas declarativas (cf. (40)). Comportamento oposto é evidenciado pelas interrogativas subordinadas parciais prototípicas em (41):

- (40) a. *Disseram-me *se veio alguém*.
b. Disseram-me *que veio alguém*.
(41) a. Perguntaram-me *quem veio*.
b. Perguntaram-me *se veio alguém*.
c. *Perguntaram-me *que veio alguém*.

Vimos acima que as chamadas interrogativas falsas ou impróprias diferem ainda das interrogativas verdadeiras pelo facto de ocorrerem obrigatoriamente em contexto de subordinação, i.e., não são interpretáveis na ausência do predicado da frase matriz que as seleciona, como ilustra o contraste de gramaticalidade entre a interrogativa parcial em (42a) e uma declarativa contendo um constituinte *wh*, como em (42b).

- (42) a. Quem veio?
b. *Quem veio.

Quer dizer, ainda que vigente em autores influentes dos anos sessenta a oitenta do século XX (e.g., Kartunen 1977), a classificação da oração encaixada em exemplos como (36) como interrogativa subordinada tem sido repensada a partir dos anos 80 do século XX (Plann 1982, Suñer 1991, 1993, 1999). De acordo com estas últimas abordagens, enquanto as frases selecionadas por verbos de inquirição, como *perguntar* em (41), são classificadas como *interrogativas subordinadas verdadeiras ou próprias*, as subordinadas selecionadas por verbos como *dizer* em (36) são tipicamente designadas como *interrogativas subordinadas falsas ou impróprias*.²⁰ Esta designação visa captar o carácter híbrido destas construções, como destacado por Matos & Brito (2013); trata-se, por um lado, de frases assertivas, com valor ilocutório declarativo; por outro, de frases que partilham propriedades com as interrogativas subordinadas parciais verdadeiras, nomeadamente, funcionando como argumento frásico do predicado do verbo da frase matriz (nos exemplos em (43) como seu objeto direto), apresentam constituintes com o morfema *wh* e admitem a focalização deste constituinte com *é que* (veja-se (43)):

- (43) a. Disseram-me *quem é que veio*.
b. Perguntaram-me *quem é que veio*.

3.2. A ambiguidade estrutural entre as interrogativas (impróprias) e as relativas livres

Embora Óscar Lopes tenha visto a frase subordinada em (36) (*quem veio*) apenas como exemplo de ambiguidade semântica, envolvendo o constituinte *quem*, podemos aprofundar tal ambiguidade, de forma a evidenciar a ambiguidade estrutural entre frases interrogativas subordinadas parciais impróprias e relativas livres.

Como referido na *Introdução* deste artigo, as relativas livres caracterizam-

²⁰ Mas vejam-se os casos referidos em (6b) e (7a)-(7b).

se por não exibirem o antecedente do constituinte relativo expresso, como ilustrado em (44a), em que \emptyset representa o antecedente não expresso do pronome relativo. No entanto, as relativas livres são parafraseáveis por expressões nominais que comportam um antecedente realizado (cf. (44b)). Por isso, as relativas livres podem surgir com um verbo como *cumprimentar*, que seleciona por complemento apenas expressões nominais e não frases (cf. (44c)):

- (44) a. Eu cumprimentei [\emptyset *quem chegou*]
b. O Paulo cumprimentou a *pessoa que chegou*.
c. *O Paulo cumprimentou *que a pessoa chegou*.

Voltando agora a nossa atenção para *dizer*, vemos que este verbo pode selecionar tanto uma frase como um sintagma nominal, como seu argumento interno. Esse sintagma nominal designa tipicamente algo que possa ser objeto de comunicação, como mostra o contraste em (45):

- (45) a. O João disse uma graça / uma anedota / uma adivinha / uma receita / um poema.
b. *O João disse a pessoa / uma pessoa / um livro / uma casa.

Contudo, quando *dizer* assume, no contexto em que ocorre, o significado de *indicar*, *revelar*, pode selecionar um sintagma nominal designando uma entidade, nomeadamente nos casos em que este sintagma nominal tem a sua referência restringida por uma relativa, como em (46a); em (46b) a oração encaixada pode então ser considerada uma relativa livre:

- (46) a. Para eu ficar a par da situação, disseram-me *as pessoas que vieram*.
b. O misterioso visitante disse *quem (ele) era*.
(quem era = a pessoa que (ele) era)

Deste modo, uma frase como *Disseram-me quem veio* é uma frase ambígua. À ambiguidade semântica descrita por Óscar Lopes, pode

adicionar-se a ambiguidade estrutural entre uma frase *interrogativa subordinada parcial imprópria* e uma *relativa livre*.

3.3. Interrogativas impróprias, relativas livres e valor condicional

Ao analisar construções *wh* num contexto não subcategorizado, mais especificamente na posição de sujeito de uma frase como (47), Óscar Lopes faz considerações interessantes.

(47) Quem nasceu morrerá.

Óscar Lopes começa por sugerir que (47), tal como outras frases similares, constitui um exemplo das “chamadas interrogativas indiretas” (*Gramática Simbólica do Português*: 250). Porém, não sendo esta oração selecionada por nenhum verbo interrogativo, de desconhecimento ou de comunicação, nem estando sob um verbo assertivo de conhecimento, ela não parece ser uma interrogativa subordinada própria ou imprópria, mas antes uma relativa livre. Com efeito, dado que *morrer* não é um verbo que seleccione um argumento frásico, *quem nasceu* só pode ser um constituinte nominal complexo, uma relativa livre, que alterna com uma relativa com antecedente, no caso um antecedente quantificado, como em (48).

(48) Toda a pessoa que nasceu morrerá.

Que terá levado Óscar Lopes a usar a expressão “as chamadas interrogativas indiretas” ao referir-se a exemplos como (47)? Uma primeira hipótese de explicação pode prender-se com o facto de a designação de relativa livre só se ter tornado corrente na literatura sobre o português a partir de Brito (1991), apesar de a expressão de pronome relativo indefinido ou sem antecedente já aparecer em Said Ali (1931: 109, 110).²¹ Por outro lado, a não assunção da distinção entre interrogativas próprias e impróprias poderá ter motivado o autor a identificar o que atualmente designamos como relativa livre com uma oração

²¹ A designação de *relativo sem antecedente* surge, de facto, em Said Ali (1931:109), que o designa como pronome *relativo indefinido* e considera que ele “gera orações de carácter substantivo”, enquanto “o relativo propriamente dito (com antecedente) dá origem a orações de carácter adjetivo (cf. Said Ali 1931: 110).

declarativa assertiva patente em interrogativas subordinadas impróprias.

Por sua vez, levanta-se a questão da origem etimológica dos morfemas *wh* numa língua como o português, bem como a da diversidade de valores das subordinadas relativas.

Com efeito, sobre a construção em (47), Óscar Lopes escreve que orações deste tipo “se podem exprimir em latim e em grego clássico como condicionais (“se alguém nasceu morrerá”), portanto, como definidamente proposicionais, o que de resto está de acordo com a nossa interpretação, atrás proposta, segundo a qual esse *quem* (nas línguas clássicas alternativamente relativo, condicional, interrogativo e indefinido) se encontra quantificado, equivalendo ora a *alguém*, ora a *todo aquele que*” (Lopes 1971: 250).

De facto, em Latim, a frase (48) corresponderia a (49a), ou, sem a conjunção condicional, a (49b):

- (49) a. *Si quis natus est morietur.*
se alguém nascido é morrerá
b. *Quis natus est morietur.*
quem nascido é morrerá

Quanto ao Grego, a frase equivalente seria (50a), ou sem a conjunção condicional (50b):

- (50) a. εἰτιςγενετο αποθανειται
[ei tis egeneto apothaneitai]
‘Se alguém nasceu morrerá’
b. οστιςγενετο ποθανε ται
[hostis egeneto apothaneitai]
‘Aquele que nasceu morrerá’

A forma σις (*hostis*) é designada por várias gramáticas de grego antigo como relativo-indefinido, porque junta as formas dos pronomes relativo (ς) e do pronome indefinido (τις), sendo normalmente traduzido como

*todo aquele que; quem quer que*²². Quer dizer, Óscar Lopes apercebe-se claramente da íntima relação entre as formas *wh* / *Qu-* (relativas ou interrogativas), o valor condicional e a quantificação.

A diversidade de usos das orações relativas, nomeadamente em orações temporais e condicionais, tem vindo a ser destacada na literatura do século XXI para várias línguas. Assim, por exemplo, Lecarme (2008: 210) considera que em Somali as orações condicionais são relativas livres modalizadas e Haegeman (2010: 220-221) faz notar que mesmo línguas como o inglês ou o francês apresentam estratégias de relativização para expressar condicionais:

- (51) a. In the event that the proposal is not accepted...
b. In case the proposal is not accepted...
(52) Au cas où tu le verrais, téléphone-moi.

3.4. Relativas livres, relativas com antecedente, morfemas *wh* e quantificadores

Ao apresentar o exemplo (53), Óscar Lopes detém-se mais uma vez sobre o valor quantificacional do constituinte *wh*:

- (53) É meu (*tudo*) o que vês.

Convém, antes de mais, referir que, ao colocar entre parêntesis o quantificador *tudo*, Óscar Lopes sugere, desde logo, que a interpretação deste exemplo é compatível tanto com uma relativa com antecedente quantificado explícito (cf. (54a)), como com uma relativa livre (cf. (54b)), em que o valor quantificacional do antecedente não linguisticamente expresso é recuperado pelo contexto:

- (54) a. É meu tudo o que vês.
b. É meu o que vês.

²² Agradecemos reconhecidamente a Jorge Deserto a tradução para latim e grego dos exemplos de Óscar Lopes, assim como o comentário sobre a natureza das formas.

Óscar Lopes propõe que na frase (53) se exprimem “duas funções coincidentes, uma proposicional e a outra designatória, como explicitado em (55):

- (55) “1 - “a função proposicional “vês x”, cuja variável (expressa pelo pronome relativo) foi presa por um quantificador, o que origina um produto lógico do tipo de $\forall x$ (vês x \wedge x é meu); 2 –a função designatória, cujo elemento variável, ou genérico, percorre o conjunto dos valores que satisfazem a essa função proposicional, ou seja, o conjunto das coisas vistas pelo interlocutor nas circunstâncias supostas.” (Lopes 1971, p. 307).

Óscar Lopes destaca, pois, dois aspetos importantes:

- (i) A aproximação entre morfemas *wh* e quantificadores, aproximação que o Grego Clássico mostra de maneira explícita (ver secção 3.3 do presente estudo);
- (ii) O que designamos atualmente por relativas livres é muitas vezes ambíguo entre uma interpretação definida (*aquele que ...*) e uma interpretação universalmente quantificada (*todo aquele que... / quem quer que...*).

Desenvolveremos estes dois aspetos seguidamente.

3.4.1. Morfemas *wh* e quantificadores

Em várias abordagens, que remontam pelo menos a Chomsky (1973, 1977), as expressões *wh* são consideradas quantificadores; vejam-se as representações em Forma Lógica (FL) atribuídas a (56) e a (57):

- (56) Quem viste?

FL: para x: x uma pessoa, tu viste x.

- (57) Tu viste toda a gente.

FL: para todo o x: x uma pessoa, o João viu x.

Um dos argumentos a favor deste paralelismo entre formas *wh* e quantificadores é o chamado efeito de *Weak Cross Over*:

(58) * Quem_i é que a sua mãe viu t_i?

FL: [para x: x uma pessoa], a sua_i mãe viu t_i

(59) * A sua_i mãe gosta de [todo o estudante]_i.

FL: [para todo o estudante]_i, a sua_i mãe gosta de t_i

A agramaticalidade de (58), paralela à de (59), não se deve a uma violação do Princípio C da Teoria da Ligação, pois em (58) o pronome *sua* não é ligado no domínio restrito da frase a que pertence, mas a um efeito de cruzamento do pronome contido em *sua* por *quem*, conhecido sob a designação de *Weak Cross Over*. Independentemente da explicação formal deste fenómeno, ele mostra o paralelismo de (56) e (57) e favorece a ideia de, nos dois casos, estarmos perante um certo tipo de ligação, uma dependência com um constituinte não-argumental (A-barra), visto que se concebe que tanto o morfema *wh* como o quantificador sobem para uma posição periférica exterior ao domínio frásico propriamente dito, mas arrastando consigo o papel temático que o verbo da oração tem para atribuir. Tal tipo de dependência parece afetar tanto movimentos *wh* em sintaxe visível (como em Português e Inglês) como em sintaxe não visível, i.e., em Forma Lógica (como no Chinês), nas chamadas línguas *wh in situ* (Huang 1982).

Em suma, Óscar Lopes evidencia o paralelismo entre morfemas *wh* e quantificadores, paralelismo tanto mais notório quanto se sabe que em certas línguas, como o Grego Clássico, há similitude nas formas usadas para interrogar e construir relativas livres e para quantificar.

3.4.2. Subtipos de relativas livres e o problema da interpretação

Tem sido notado por vários linguistas que, ao contrário dos morfemas *wh* em interrogativas parciais, em que ao morfema *wh* está associada uma interpretação existencial no escopo de um operador de interrogação, aos

morfemas *wh* nas relativas livres podem estar associados diferentes tipos de interpretação, em particular a interpretação definida e a interpretação universal. Em inglês essa distinção é clara porque nessa língua os morfemas *wh* em relativas livres podem ou não ligar-se ao sufixo *-ever*. Assim, os exemplos em (60a) e (61a) exibem respetivamente, a interpretação definida explicitada em (60b) e a interpretação universal presente em (61b) (exemplos de Riemsdijk 2007).

- (60) a. I ate *what the waiter put on my plate*.
b. I ate the thing that the waiter put on my plate.
- (61) a. I will ate *whatever the waiter will put on my plate*.
b. I will eat anything / everything that the waiter will put on my plate.

Mesmo sem o sufixo *-ever*, o morfema *wh* pode ter interpretação universal, dependendo do tipo de predicção, como em (62):

- (62) I will eat what happens to be on the menu.

Por outras palavras, os morfemas *wh* em relativas livres são à partida ambíguos e é muitas vezes o tipo de predicção (genérica ou específica) contida na oração que nos esclarece sobre se estamos perante a interpretação universal ou a interpretação definida. Jacobson (1995), citada por Riemsdijk (2007: 362-3), mostra que o que acontece com os morfemas *wh* em relativas livres é que eles se relacionam com a noção de “entidade máxima plural”: isto é, “se num dado domínio há só um indivíduo atómico com uma dada propriedade, então a entidade plural máxima com essa propriedade é aquela que consiste nesse indivíduo atómico” (tradução nossa). Assim, “uma relativa livre como *o que a Maria recomendou* denota o conjunto de entidades máximas plurais que a Maria recomendou (e não apenas um conjunto de indivíduos recomendados pela Maria). Uma vez que esse conjunto pode ser limitado a uma única entidade singular, então damos conta do facto de as relativas livres serem umas vezes definidas, outras vezes universais.” (tradução nossa).

Esta propriedade foi notada igualmente por Grosu & Landman (1998) e

Grosu (2000) sob a designação de “efeito de maximalidade” das relativas livres. Ambos os autores referem, por isso, que um dos contextos adequados à ocorrência de relativas livres de interpretação universal são os provérbios; vejam-se, por exemplo, (63) e (64):

(63) Quem vai ao mar perde o lugar.

(64) O que não tem remédio, remediado está.

Mas há ainda um outro tipo de relativa livre, aquele que tem uma interpretação indefinida, que nos parece estar ausente das reflexões de Óscar Lopes. Veja-se exemplos como (65a), e a sua paráfrase de relativa com antecedente, em que, no domínio do verbo *haver*, a interpretação é indefinida:

(65) a. Há quem ache que Donald Trump vai ganhar.

b. Há algumas pessoas que acham que Donald Trump vai ganhar.

O mesmo tipo de interpretação indefinida ocorre em frases como (66), exemplo apresentado em Riemsdijk (2007: 371), que aqui traduzimos:

(66) Ele transportava o que a tripulação pensou ser uma lata de gasolina.

Riemsdijk (2007) designa este tipo de relativas livres como “transparentes”, e apresenta os seus contextos de ocorrência característicos:

(i) Estas relativas livres surgem em contextos que potenciam o aparecimento de expressões indefinidas, como o verbo *haver* (veja-se o contraste entre (67a) e (67b)-(67c):²³

(67) a.* Havia o avião na autoestrada.

²³ Um revisor fez notar que o efeito de definitude que caracteriza *there to be* em Inglês não se verifica com *haver*, em Português, como alguns dados do corpus do CRPC ilustram: (i) “E depois, há os macaenses. Os filhos da terra”; (ii) “Dito de outro modo, há o verso e o reverso, a outra face da moeda.” Concordamos que o efeito de definitude não parece ser tão forte em Português como o é em Inglês. No entanto, a possibilidade de DPs definidos em construções com *there to be* em Inglês é notada por diversos autores se houver uma interpretação de lista. (cf., entre outros, Reuland & ter Meulen (1987: 17).

b. Havia um avião na autoestrada.

c. Havia o que parecia ser um avião na autoestrada.

(ii) Podem surgir com modificadores intensionais, seja adjetivais, como *alegado*, *presumível* ou adverbiais (*alegadamente*, *presumivelmente*):

(68) Serviram-me o que alegadamente se poderia chamar um bife.

(iii) Quando colocada na posição de sujeito, a relativa livre iniciada por *que* / *what* pode desencadear concordância plural:

(69) O que parecia serem aviões privados estavam na autoestrada.

(iv) Neste tipo de relativas *o que* pode referir-se a humanos:

(70) Ela contactou o que eu tomei por um polícia para ficar de noite na entrada da casa.

(v) Em inglês o sufixo *-ever* é impossível neste tipo de relativas livres mostrando que a leitura universal é impossível:

(71) *There was whatever appeared a jet on the free way.

(vi) Pode haver coordenação de SNs predicativos, de interpretação indefinida, e de uma relativa livre deste tipo, como em (72):²⁴

(72) Naquela instituição há incompetentes e quem se faça passar por bom.

(73) * Naquela instituição há incompetentes e os estudantes.

Estamos aqui diante de um tipo distinto de relativas livres, que, para além do carácter indefinido, tem um conteúdo predicativo.

²⁴ Isto não exclui outros casos de coordenação em contextos semelhantes, como os exemplos seguintes evidenciam:

(i) Naquela instituição há incompetentes e estudantes.

(ii) Naquela instituição há os incompetentes e os estudantes.

Apenas queremos realçar aqui a possibilidade de coordenação da relativa livre com um SN indefinido, como em (72).

Em síntese: neste parágrafo referimos alguns tipos distintos de relativas, as relativas livres que podem ter quer uma leitura definida quer uma leitura universal; neste último caso, há línguas que marcam morfologicamente essa leitura com o sufixo *-ever*, como é o caso do Inglês. Em português tal sufixo não existe mas há um tipo de construção *o que quer que seja / quem quer que seja*, que resulta certamente de um processo de gramaticalização muito complexo, que aqui não analisámos, mas que evidencia bem a leitura universal das relativas livres. A este tipo de relativas livres Óscar Lopes faz uma menção especial na *Gramática Simbólica do Português*, anunciando a noção de “efeito de maximalidade” a que vários autores se referiram muitos anos mais tarde. Por fim, fizemos referência a um tipo de relativa livre de leitura indefinida, a que Óscar Lopes não faz referência na sua obra, e cujo estudo se impõe. Deixamos para outra ocasião essa análise.

4. Síntese e conclusões

Neste artigo fez-se um estudo acerca das semelhanças e diferenças entre relativas livres e relativas com antecedente, e entre relativas livres e interrogativas parciais subordinadas: mostrámos, desenvolvendo estudos clássicos sobre o tema, que as relativas livres se distinguem das relativas com antecedente por não poderem conter certos morfemas *wh* que requerem ter um antecedente explícito (*cujo* e *o qual*) e que as relativas livres são diferentes de interrogativas subordinadas parciais por várias propriedades, uma das quais é a sua natureza categorial: as relativas livres são expressões nominais, as interrogativas subordinadas parciais são frases.

Analisámos, em particular, a forma como Óscar Lopes descreveu estas construções e principalmente casos ambíguos como *disseram-me quem veio*; para além de ambiguidade semântica tentámos mostrar que há também aqui ambiguidade estrutural, uma vez que frases deste tipo tanto podem conter uma interrogativa subordinada parcial imprópria (um constituinte oracional) como uma relativa livre (um constituinte nominal). Óscar Lopes apercebeu-se da aproximação entre morfemas *wh* e quantificadores, aproximação que o Grego Clássico mostra de maneira explícita. Por sua vez, notou que aquilo que designamos atualmente relativas livres é muitas vezes ambíguo

entre uma interpretação definida (*aquele que ...*) e uma interpretação universalmente quantificada (*todo aquele que... / quem quer que...*), antecipando aquilo que é hoje designado como o “efeito de maximalidade” das relativas livres. Para além das relativas livres de leitura definida e de leitura universal, referimos igualmente as propriedades fundamentais das relativas livres de leitura indefinida.

A análise permitiu mostrar que há uma “gradação” de construções *wh*: 1 - nas interrogativas subordinadas parciais próprias há uma variável sob o escopo de um operador de interrogação expresso por um certo tipo de verbo (*perguntar, ignorar*); como consequência, a variável tem uma interpretação unicamente existencial; 2 - nas interrogativas subordinadas parciais impróprias ou falsas a variável está sob o escopo de um verbo que poderíamos classificar como fracamente assertivo e cognitivo (*descobrir, explicar*), tendo a variável, quando expressa numa forma *wh*, igualmente valor existencial; 3 – nas relativas livres a variável não está, em geral, sob o escopo de nenhum dos verbos referidos e, encontrando-se a oração integrada numa estrutura nominal, a construção pode ter diferentes valores semânticos, definido, universal e indefinido; 4 - pode haver ambiguidade (semântica e estrutural) se a construção reunir algumas das condições previstas em 2 e 3, nomeadamente quando a variável está sob o escopo de um verbo de comunicação como *dizer*, talvez por este ser o mais neutro dos verbos de comunicação e por poder assumir, consoante o contexto, um valor semântico próximo dos verbos referidos em 2, isto é, quando sinónimo de *descobrir, revelar*.

REFERÊNCIAS

- Alexandre, N. 2012. *The Defective Copy Theory of Movement: Evidence from Wh-Constructions in Cape Verdean Creole*, Amsterdam / Philadelphia: John Benjamins. <http://benjamins.com/#catalog/books/cll.41/main>
- Adger, D. & J. Quer. 2001. The syntax and semantics of unselected embedded questions, *Language*, 77(1): 107-133.
- Brito, A. M. 1991. *A sintaxe das orações relativas em Português. Estrutura, mecanismos interpretativos e condições sobre a distribuição dos morfemas relativos*. Instituto Nacional de Investigação Científica. Centro de Linguística da Universidade do Porto. Porto.
- Brito, A. M. & Duarte, I. 2003. Orações relativas e construções aparentadas. In Mateus, M. H. et al. 2003. *Gramática da Língua Portuguesa*, 6ª ed., Lisboa: Caminho, 675 – 684.
- Cardoso, A. 2017. *Portuguese Relative Clauses in Synchrony and Diachrony*. Oxford: Oxford University Press.
- Cardoso, A. & Alexandre, N. 2013. Relativas clivadas em variedades não standard do português europeu. In F. Silva; I. Falé & I. Pereira (orgs.). *Textos Seleccionados do XVIII ENAPL, 2012*, Porto: APL, 205-227.
- Chomsky, N. 1973. Conditions on Transformations. In S. Anderson & P. Kiparsky (eds.), *A Festschrift for Morris Halle*. New York: Holt Rinehart and Winston, 232-286.
- Chomsky, N. 1977. On Wh-Movement. In P. W. Culicover, T. Wasow & A. Akmajian (eds.), *Formal Syntax*. New York: Academic Press, 71-132.
- Grosu, A. & Landman, F. 1998. Strange Relatives of the Third Kind. *Natural Language Semantics* (6) : 125-170.
- Grosu, A. 2000. Type resolution in relative constructions. Feature marking and dependency encoding. In Alexiadou, A., Law, P., Meinunger, A. & Wilder, C. 2000 (eds.), *The Syntax of Relative Clauses*. Amsterdam: John Benjamins, 83-120.
- Haegeman, L. 2012. *Adverbial Clauses, Main Clause Phenomena, and the Composition of the Left Periphery*. Oxford: Oxford University Press.
- Hinzen, W. & Sheehan, M. 2011 Moving towards the edge: the grammar of reference, *Linguistic Analysis* 37,405-458.
- Huang, C. T. J. 1982. *Logical relations in Chinese and the theory of grammar*. PhD dissertation, MIT.

- Jacobson, P. 1995. On the Quantificational Force of English Free Relatives. In E. Bach, E. Jelinek, A. Kratzer & B. Partee (eds.) *Quantification in Natural Languages*. Dordrecht: Kluwer, 451-486.
- Karttunen, L. 1977. Syntax and Semantics of Questions. *Linguistics and Philosophy* 1: 3-44.
- Kato, M. 2013. Deriving “wh-in-situ” through movement in Brazilian Portuguese. In V. Camacho-Aboada, A. Jimenéz-Fernández, Martín-J. Gonzáles, & M. Reyes-Tejedor (eds.) *Information Structure and Agreement*, Amsterdam: John Benjamins, 175-191.
- Lahiri, Utpal. 2002. *Questions and Answers in Embedded Contexts*, Oxford: Oxford University Press
- Lecarme, J. 2008. Tense and modality in nominals. In Jacqueline Guéron & Jacqueline Lecarme (eds.), *Time and modality*. Dordrecht: Springer, 195–225.
- Lopes, Ó. 1971. *Gramática Simbólica do Português*, Lisboa, Fundação Calouste Gulbenkian, 1ª edição.
- Matos, G. & Brito, A. M. 2013. The alternation between improper indirect questions and DPs containing a restrictive relative. In V. Camacho-Aboada, A. Jimenéz-Fernández, Martín-J. Gonzáles & M. Reyes-Tejedor (eds.) *Information Structure and Agreement*, Amsterdam: John Benjamins, 83-116.
- Matos, G. & Brito, A.M. 2018. Relativas livres e interrogativas parciais: paralelos e diferenças, *Revista da Associação Portuguesa de Linguística*, Nº 4 – 09 / 2018 | 152-167 | <https://doi.org/10.26334/2183-9077/rapln4ano2018a38>
- Mioto, C. & Lobo, M. 2016. Wh movement: interrogatives, relatives and clefts. In W. L. Wetzels, S. Menuzzi & J. Costa (eds.) *The Handbook of Portuguese Linguistics*. Wiley-Blackwell, 275-293.
- Móia, T. 1996. A sintaxe das orações relativas sem antecedente expresso do Português. In Gonçalves, A. & Colaço, M. & Miguel, M. & Móia, T. 1996. *Quatro estudos em sintaxe do Português. Uma abordagem segundo a teoria dos princípios e parâmetros*. Lisboa: Edições Colibri, 149-188.
- Plann, S. 1982. Indirect Questions in Spanish, *Linguistic Inquiry*, 12: 297-312.
- Reuland, E. & ter Meulen, A. 1987 Introduction. In Reuland, E. & ter Meulen, A. (eds.) *The representation of (In)definiteness*, Cambridge, MA: The MIT Press, 1-20.
- Riemsdijk, H. 2007. Free Relatives. In Everaert, M. & van Riemsdijk, H. (eds.) *The Blackwell Companion to Syntax*, Blackwell, 338-382.
- Rinke, E. & Assmann, E. 2017. The Syntax of Relative Clauses in European Portuguese. Extending the Determiner Hypothesis of Relativizers to Relative *que*. *Journal of*

- Portuguese Linguistics*, 16 (4): 1–26. DOI: <https://doi.org/10.5334/jpl.172>
- Rivero, M. L. 1994. On indirect questions, commands and the Spanish quotative *que*. *Linguistic Inquiry* 25: 547-55.
- Rudin, C. 1988. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6: 445-501.
- Said Ali, M. 1931 *Gramática Histórica da língua Portuguesa*. São Paulo: Edições. Melhoramentos, 1964.
- Stoyanova, M. 2008. *Unique focus: languages without multiple wh-questions*. Amsterdam/Philadelphia: John Benjamins.
- Suñer, M. 1991. Indirect Questions and the structure of CP: Some consequences. In H. Campus & F. Martínez-Gil (eds.) *Studies in Romance Linguistics*. Dordrecht, Foris Publications, 183-203.
- Suñer, M. 1993. About indirect questions and semi-questions. *Linguistics and Philosophy*, 16: 45-77.
- Suñer, M. 1999. La subordinación sustantiva: la interrogación indirecta. In I. Bosque & V. Demonte (eds.) *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 2147-2195.
- Veloso, R. 2013. Subordinação relativa. In Raposo, E., F. Nascimento, M.A. Mota, L. Segura & A. Mendes. 2013. *Gramática do Português*. Lisboa: Fundação Calouste Gulbenkian, Vol. II: 2061 – 2134.

Sobre a natureza homogénea do Pretérito Perfeito Composto em Português Europeu

Fátima Oliveira
moliv@letras.up.pt

Faculdade de Letras /Centro de Linguística da Universidade do Porto

António Leal
jleal@letras.up.pt

Faculdade de Letras /Centro de Linguística da Universidade do Porto

ABSTRACT

In European Portuguese, ‘Pretérito Perfeito Composto’ (PPC) is a tense that presents some peculiarities distinguishing it clearly from similar constructions in other languages. Although this construction exhibits a single-eventuality reading in other languages, in EP there is typically a multiple eventuality reading. So, in the first part of this paper, we discuss the quantification over situations underlying the PPC in EP considering the notions of frequency, habituality and iteration. We argue that iteration is the relevant concept, as PPC converts an undetermined number of basic events of the same type into a single event.

However, in some cases, the PPC in EP has a second reading available: a single-eventuality, like in other languages. Therefore, in the second part of this paper, we present an aspectual analysis of the PPC that includes both the single-event reading and the iterative reading, in order to provide a unified explanation for both readings. To do so, we will resort to the notion of homogeneity as defined in Landman & Rothstein 2012a e 2012b). We will argue that homogeneity is a central aspectual property of PPC in EP and that the distinction between segmental homogeneity and incremental homogeneity proposed by these authors for ‘for x time’ adverbials allows us to explain both readings of this EP tense.

KEY-WORDS

pretérito perfeito composto, tense, aspect, homogeneity

RESUMO

O Pretérito Perfeito Composto do Indicativo (PPC) é um tempo que apresenta certas peculiaridades em Português Europeu que o distinguem inequivocamente de construções similares em outras línguas. De facto, enquanto, nas restantes línguas, esta construção veicula apenas uma eventualidade, em PE surge tipicamente uma leitura de repetição de eventualidades. Assim, na primeira parte deste trabalho, é analisada a natureza desta repetição de eventualidades – se se trata de iteração, frequência ou habitualidade. Defenderemos que

a mais adequada é a noção de iteração, na medida em que o PPC converte um número não determinado de eventos básicos do mesmo tipo numa única eventualidade, da qual os eventos básicos são subfases.

Contudo, esta leitura iterativa não surge em todos os casos: sob certas condições, o PPC pode ter uma leitura de eventualidade única. Deste modo, na segunda parte deste trabalho, apresentamos uma análise aspetual do PPC que inclui tanto a leitura de eventualidade única como a leitura iterativa, no sentido de fornecer uma explicação unificada para as leituras do PPC. Para isso, recorreremos à noção de *homogeneidade* tal como é definida em Landman & Rothstein (2012a e 2012b). Defendemos que a propriedade da homogeneidade é uma propriedade aspetual central do PPC em Português Europeu e que a oposição entre homogeneidade segmental e homogeneidade incremental proposta por estes autores para os advérbiais do tipo *'for x time'* permite explicar ambas as leituras deste tempo gramatical.

PALAVRAS-CHAVE

pretérito perfeito composto, tempo, aspeto, homogeneidade

1. Introdução

O Pretérito Perfeito Composto do Indicativo (PPC) do Português Europeu é um tempo que, no quadro das línguas românticas (e não só), apresenta traços peculiares. De facto, as predicções que comparecem com este tempo podem ter um leque de interpretações variado. Em primeiro lugar, em certos casos, podem ter uma leitura – obrigatória – de repetição de eventualidades, como em (1), leitura esta que não surge em construções similares em outras línguas românicas (cf. Paiva Boléo 1937; Castilho 1968; Squartini & Bertinetto 2000, e.o.). Em (1), a única interpretação possível é a de que houve uma sucessão de eventos de “o rapaz tossir”, sucessão essa que teve início no passado e que se prolonga até ao momento da enunciação.

(1) O rapaz tem tossido.

Para além dos casos de leitura de repetição de eventualidades, há exemplos com o PPC que apenas têm leitura de eventualidade única, como em (2). A interpretação deste exemplo requer considerar a existência de uma única situação, denotada por “o rapaz estar doente”, que teve início no passado e que se prolonga até ao momento da enunciação.

(2) O rapaz tem estado doente.

Finalmente, há a apontar casos em que ambas as leituras – de repetição de eventualidades e de eventualidade única – estão disponíveis, como em (3). Esta dupla possibilidade interpretativa pode ser comprovada, por exemplo, pela existência de duas opções de ligação anafórica, que são ilustradas em (3a) e (3b). Assim, o sujeito de (3a), no singular, retoma anaforicamente a predicação expressa em (3) na leitura de eventualidade única; já o sujeito de (3b), no plural, faz a retoma a partir da leitura de repetição de eventualidades.

(3) O rapaz tem perseguido a vizinha.

(3a) Esta perseguição começou hoje de manhã.

(3b) Estas perseguições começaram na semana passada.

Esta dupla possibilidade interpretativa das predicções com o PPC é condicionada por diversos fatores (cf. Oliveira, Leal & Silva, 2014), nomeadamente a localização do ponto de perspetiva temporal da predicação, o modo verbal do verbo auxiliar e certas propriedades semânticas de argumentos verbais e de certos adjuntos. Para além disso, existe também uma relação entre as leituras evidenciadas pelas predicções combinadas com o PPC e certas características aspetuais das próprias predicções básicas. Assim, (i) estados *stage-level* e situações eventivas durativas e atélicas podem ter tipicamente ambas as leituras, conforme ilustrado em (4)-(5); situações eventivas não durativas e situações eventivas durativas e atélicas têm a leitura de repetição das eventualidades, como se mostra em (6)-(8); finalmente, (9) mostra que predicados de indivíduo não ocorrem com o PPC.

(4) O rapaz tem estado no jardim.

(5) O glaciar tem deslizado pela encosta.

(6) O rapaz tem vencido a corrida.

(7) O rapaz tem tossido.

(8) O rapaz tem lido o texto.

(9) * O rapaz tem tido olhos azuis.

Neste trabalho, vamos concentrar-nos, em primeiro lugar, nos casos em que as predicacões com o PPC evidenciam leitura de repetição de eventualidades, no sentido de procurar respostas para as seguintes questões: qual é a natureza da repetição subjacente a estas leituras? Trata-se de habitualidade, pelo que o PPC cria um estado habitual? Ou trata-se de frequência, havendo uma mera quantificação sobre situações, sem a criação de um tipo derivado? Ou será antes um caso de iteração, sendo o PPC um mecanismo de criação de um processo derivado pela repetição de situações do mesmo tipo?¹

Com o intuito de procurar uma resposta para estas questões, analisaremos, na secção 2, as relações que se podem estabelecer entre o PPC e cada uma das categorias anteriormente enunciadas.

Na secção 3, fazemos uma proposta de explicação de ambos os tipos de leitura que o PPC evidencia, tendo por base a noção aspetual de homogeneidade.

2. Relação do PPC com estruturas que envolvem repetição de situações

Nesta secção, analisamos o PPC nos casos em que há leitura de repetição de situações, tendo por base a proposta de Cunha (2006) relativa às noções de iteração, de frequência e de habitualidade.

2.1. PPC e estados habituais

Em termos gerais, a habitualidade está relacionada com uma repetição de situações que decorre em intervalos de tempo necessariamente longos e preferencialmente não delimitados. Esta categoria estabelece propriedades

¹ Veja-se a propósito destas questões, Lopes (1986/2005:108): “[...] em formas compostas como as que encontramos em *tem estado doente*, *tem as cartas já escritas*, *tem escrito muitas cartas*, o auxiliar *ter*, no presente morfológico viabiliza a hipótese de que afinal se trate de variantes aspetuais durativas, resultativas ou durativas-frequentativas do presente”.

caracterizadoras das diversas entidades que predica, criando predicados de indivíduo a partir de uma generalização sobre propriedades episódicas. De acordo com Cunha (2006), os estados habituais em PE comportam-se como perspetivadores aspetuais.

As predicções com PPC apresentam alguns aspetos que podem também ser encontrados em construções que veiculam habitualidade. Em primeiro lugar, tanto o PPC como os estados habituais denotam, tipicamente, a repetição de eventualidades. Vejam-se os exemplos (10), com o presente do indicativo, e (11), com o PPC: em ambos os casos está subjacente uma repetição da eventualidade “o João fumar”.

(10) O João fuma.

(11) O João tem fumado.

Um segundo aspeto que aproxima o PPC de construções habituais é o facto de todas as classes aspetuais poderem ocorrer em construções de habitualidade e com o PPC, com a exceção, em ambos os casos, dos predicados de indivíduo não faseáveis. Vejam-se (12), com exemplos de estados habituais (com o presente do indicativo e o advérbio “habitualmente”) e (13), com o PPC: em ambos os casos, há agramaticalidade apenas quando está em causa o predicado de indivíduo não faseável “o João ter olhos azuis”.

(12) a. Habitualmente, o João almoça na cantina. (processo culminado)
b. Habitualmente, o João fecha a janela do quarto à noite. (culminação)
c. Habitualmente, o João é simpático. (predicado de indivíduo faseável)
d. *Habitualmente, o João tem olhos azuis. (predicado de indivíduo não faseável)

(13) a. O João tem almoçado na cantina. (processo culminado)
b. O João tem fechado a janela do quarto à noite. (culminação)
c. O João tem sido simpático. (predicado de indivíduo faseável)
d. * O João tem tido olhos azuis. (predicado de indivíduo não faseável)

Finalmente, tanto com o PPC como com construções de habitualidade, a coocorrência com a expressão “sempre que” dá origem a exemplos agramaticais (cf. Cunha 2006), como se ilustra em (14). Os exemplos em (15) são gramaticais na medida em que a expressão “sempre que” opera sobre eventos.

- (14) a. * Sempre que o João habitualmente almoça na cantina, encontra os amigos.
b. */??? Sempre que o João tem almoçado na cantina, encontrou os amigos.
- (15) a. Sempre que o João almoça na cantina, encontra os amigos.
b. Sempre que o João almoçou na cantina, encontrou os amigos.

Apesar de existirem estas semelhanças entre o PPC e os estados habituais, existem também diferenças relevantes a apontar. Assim, e ao contrário dos estados habituais, o PPC não requer intervalos de tempo longos. Vejam-se os exemplos em (16): com a predicação “a equipa A comprar jogadores exclusivamente europeus”, o PPC é perfeitamente compatível com o adverbial “na última semana” em (16a), ao contrário do presente do indicativo, em (16b), que cria um estado habitual.

- (16) a. Na última semana, a equipa A tem comprado jogadores exclusivamente europeus.
b. */??? Ultimamente/*na última semana, a equipa A compra jogadores exclusivamente europeus.

Em segundo lugar, ao contrário dos estados habituais, o PPC não parece corresponder a propriedades gerais das entidades que predica. É por este motivo que, em (17a) se pode negar ser um hábito do João chegar atrasado às aulas, mas, em (17b), não se pode negar que fumar é um hábito do João.

- (17) a. O João tem chegado tarde às aulas nesta semana, o que nem é hábito nele.
b. * O João fuma, o que nem é hábito nele.

Em terceiro lugar, os estados habituais não se combinam com adverbiais de localização (cf. Cunha, 2006), o que explica a agramaticalidade de (18a). Contudo, (18b) mostra que o PPC admite a combinação com o advérbio de localização “hoje”.

- (18) a. * O João fuma hoje.
b. O João tem fumado hoje.

Finalmente, o PPC não é verdadeiro em todos os subintervalos do intervalo a que se aplica, enquanto os estados, inclusivamente os estados habituais, são, pelo contrário, verdadeiros em todos os subintervalos do intervalo a que se aplicam. Por este motivo, a partir de (19), se pode inferir que é verdade que o João é fumador ininterruptamente desde 2010. Contudo, a partir de (20), não se pode fazer a mesma inferência.

- (19) O João fuma desde 2010.
(20) O João tem fumado desde 2010.

Apresentamos a seguir um quadro que sistematiza as semelhanças e diferenças apontadas ao PPC e estados habituais.

Quadro I – PPC e estados habituais: semelhanças e diferenças

Semelhanças	Diferenças
<ol style="list-style-type: none"> 1. Repetição de eventualidades 2. Compatibilidade com todas as classes aspetuais, exceto predicados de indivíduo não faseáveis 3. Coocorrência agramatical com “sempre que” 	<ol style="list-style-type: none"> 1. Ao contrário dos estados habituais, o PPC não requer intervalos de tempo longos 2. Ao contrário dos estados habituais, o PPC não corresponde a propriedades gerais das entidades 3. Os estados habituais não se combinam com advérbias de localização; o PPC admite a combinação com o advérbio “hoje” 4. Ao contrário dos estados habituais, o PPC não é verdadeiro em todos os subintervalos do intervalo a que se aplica

Embora haja aspetos em comum entre o PPC e as estruturas de habitualidade, não parece ser possível considerar o PPC como um operador de habitualidade. Salientamos, em particular, as diferenças 2 e 4 no quadro I como cruciais para esta conclusão. De facto, se o PPC não é verdadeiro em todos os subintervalos do intervalo a que se aplica e não corresponde a propriedades gerais das entidades, então não parece ser um tipo de estado. Assim, na secção seguinte, viramos a nossa atenção para a comparação entre o PPC e as estruturas de frequência.

2.2. PPC e estruturas de frequência

Resumidamente, podemos dizer que a frequência expressa padrões de repetição de situações (baixa, média, alta frequência) enquadradas num determinado intervalo temporal cuja extensão pode ser muito variável. A quantidade de situações envolvidas nesta repetição pode ser especificada (e.g. “três vezes”) ou não especificada (e.g. “frequentemente”).

Em Português, a frequência está associada a certos advérbias, como “frequentemente”, “ocasionalmente”, “muitas vezes”, “várias vezes”, “n

vezes”, “de vez em quando”).

De acordo com Cunha (2006), a frequência não altera o perfil aspetual básico das predicções com que coocorre, no sentido em que não cria, como no caso da iteração, um evento de tipo processual derivado a partir da repetição de situações básicas. Para além disso, a frequência requer que haja pausas ou intervalos significativos entre as situações (i.e. descontinuidade entre os intervalos temporais que correspondem aos eventos repetidos).

Tal como no caso da habitualidade, podemos encontrar pontos de convergência e de divergência entre o PPC e estruturas que veiculam frequência. Começemos pelas semelhanças.

Tanto o PPC como as estruturas que expressam frequência têm subjacente a noção de repetição de situações e, em ambos os casos, essa repetição pode corresponder a baixa, média ou alta frequência, como se ilustra em (21).

- (21) a. O João tossiu {ocasionalmente/algumas vezes/frequentemente}.
- b. O João tem tossido. (a frequência com que as situações se repetem não está definida)

Para além disso, em ambos os casos, há pausas de extensão variável entre os eventos que compõem a sequência. Assim, em (22), o intervalo entre cada dois trovões é variável, quer se trate de uma estrutura de frequência (22a), quer se trate do PPC (22b).

- (22) a. Trovejou algumas vezes durante esta semana.
- b. Tem trovejado durante esta semana.

Finalmente, tal como acontece com o PPC, todas as classes aspetuais podem ocorrer em construções de frequência (cf. (23), com eventos e predicado de indivíduo faseável), com a exceção dos predicados de indivíduo não faseáveis (cf.(24)).

- (23) a. O João almoça na cantina frequentemente. (cf. O João tem almoçado na cantina.)
 b. O João corre frequentemente na praia. (cf. O João tem corrido na praia.)
 c. O João fecha a janela do quarto à noite frequentemente. (cf. O João tem fechado a janela do quarto à noite.)
 d. O João tosse frequentemente. (cf. O João tem tossido.)
 e. O João é simpático frequentemente. (cf. O João tem sido simpático.)
- (24) * O João tem olhos azuis frequentemente. (cf. * O João tem tido olhos azuis.)

Passemos agora aos aspetos que afastam o PPC das estruturas de frequência. Em primeiro lugar, com certas predicções (alguns estados e processos), o PPC não acarreta necessariamente uma pluralização de situações; pelo contrário, a ocorrência de adverbiais de frequência força a leitura de repetição de situações. Deste modo, tanto em (25a), como em (26a), o estado “o João estar no jardim” e o processo “o João perseguir a vizinha” podem ter tanto uma leitura de eventualidade única, como uma leitura de repetição de eventualidades. A escolha de uma ou outra leitura pode ser favorecida por certos adverbiais: um adverbial que denote um intervalo curto, como “na última meia hora”, favorece uma leitura de eventualidade única, enquanto um adverbial que denote um intervalo prolongado, como “na última semana”, favorece uma leitura de repetição de situações. Contudo, no caso dos exemplos (25b) e (26b), com o adverbial “frequentemente”, indutor de frequência, apenas a leitura de repetição de situações está disponível.

- (25) a. O João tem estado no jardim {na última meia hora/ na última semana}.
 b. O João esteve no jardim frequentemente.
- (26) a. O João tem perseguido a vizinha {na última meia hora / na última semana}.

b. O João perseguiu a vizinha frequentemente.

Em segundo lugar, há casos com o PPC em que existem alterações aspetuais na predicação básica (nomeadamente com alguns processos culminados), dando lugar a leituras imperfectivas. É o que se passa em (27a): esta frase permite a inferência de que o João ainda não leu o livro, ou seja, permite que seja retirada a culminação do núcleo aspetual da predicação “o João ler o livro”. Pelo contrário, uma construção de frequência nunca permite esta leitura imperfectiva, na medida em que a ocorrência de adverbiais de frequência força a leitura de repetição de situações completas, como se exemplifica em (27b)

(27) a. O João tem lido o livro.

→ O João (ainda) não leu o livro.

b. O João leu o livro frequentemente.

→ O João leu o livro até ao fim diversas vezes.

Um terceiro aspeto que permite afastar o PPC das estruturas de frequência prende-se com o facto de o PPC não se combinar com adverbiais que indiquem a cardinalidade dos eventos repetidos (cf. Oliveira & Leal 2012 e.o.), ao contrário das estruturas de frequência. Vejam-se os exemplos (28): é possível apontar a cardinalidade de tossidas numa estrutura de frequência (28a), mas não com o PPC (28b).

(28) a. O João tossiu 3 vezes.

b. * O João tem tossido 3 vezes.

Apresentamos novamente um quadro que sistematiza as semelhanças e diferenças apontadas ao PPC e estruturas de frequência.

Quadro II – PPC e estruturas de frequência: semelhanças e diferenças

Semelhanças	Diferenças
<ol style="list-style-type: none"> 1. Repetição de eventualidades 2. Essa repetição pode corresponder a baixa, média ou alta frequência 3. Há pausas de extensão variável entre os eventos que compõem a sequência 4. Todas as classes aspetuais podem ocorrer, exceto predicados de indivíduo não faseáveis 	<ol style="list-style-type: none"> 1. Com estados e processos, o PPC pode não implicar uma pluralização de situações 2. Com processos culminados, o PPC pode acarretar alterações aspetuais 3. O PPC não se combina com advérbiais que indiquem a cardinalidade dos eventos repetidos

Tal como no caso anterior, relativamente à habitualidade, também agora podemos apontar aspetos que distinguem o PPC das estruturas de frequência e que, no nosso entendimento, são cruciais. Referimo-nos às diferenças apontadas em 1 e 2. De facto, apesar de haver pontos de contacto, não podemos considerar o PPC como um operador de frequência, na medida em que, quando combinado com situações durativas e atélicas, pode não implicar a repetição de situações e, com situações durativas e télicas, pode proceder a alterações aspetuais no núcleo de uma mesma predicação.

Na secção seguinte, procederemos à comparação entre o PPC e o terceiro tipo de estruturas que envolvem a repetição de situações, neste caso, a iteração.

2.3. PPC e estruturas iterativas

A iteração é considerada um operador aspetual que cria uma repetição de situações idênticas, tipicamente em intervalos de tempo curtos e delimitados e sem que haja pausas significativas entre cada situação (Cunha 2006). Esta repetição de situações é tida, no seu conjunto, como um processo de tipo derivado, pelo que as eventualidades que se repetem são subfases, ou partes, desse processo (*multiple-events activities*, em Smith 1991)

De acordo com Cunha (2006), a iteração cria, portanto, um processo derivado, pelo que, aspetualmente, é dinâmico, durativo e atélico. Podemos acrescentar que este processo derivado é, também, cumulativo e divisível. Relativamente à cumulatividade, veja-se (29). Considerando as frases em

(29), se é verdade que “o João tossir” aconteceu num intervalo de tempo entre as 2h e as 2h05m e se é verdade que “o João tossir” aconteceu num intervalo de tempo entre as 2h05m e as 2h10m, então é verdade que “o João tossir” aconteceu num intervalo de tempo entre as 2h e as 2h10m. Quanto à divisibilidade, veja-se (30): se é verdade que “o João tossir” aconteceu num intervalo de tempo de 10 minutos, então é verdade que “o João tossir” aconteceu num subintervalo desse intervalo de tempo.

(29) a. O João tossiu das 2h às 2h05m.

b. O João tossiu das 2h05m às 2h10.

→ O João tossiu das 2h às 2h10m.

(30) O João tossiu durante 10 minutos

→ O João tossiu durante os primeiros 5 minutos.

Note-se que a divisibilidade das estruturas iterativas não se aplica a qualquer instante do intervalo considerado, mas está limitada pela existência de subfases sucessivas, constituídas pelos eventos básicos repetidos (cf. Cunha 2006).

A iteração pode ser desencadeada por propriedades lexicais ou morfológicas, mas também por propriedades de argumentos internos, assim como por adverbiais temporais durativos, quando combinados com predicções básicas não durativas (cf. Cunha 2006).

Por fim, de salientar que, do ponto de vista aspetual, os estados e processos estão excluídos de estruturas iterativas e que a iteração não provoca alterações no tipo aspetual básico das situações repetidas (Cunha 2006).

Uma análise das construções com PPC põe em evidência alguns aspetos que nos podem levar a pôr em questão que se trate de casos de iteração. Em primeiro lugar, e segundo Cunha (2006), na iteração, os eventos têm de estar organizados de forma contígua, sem pausas significativas entre si (cf. (31)); com o PPC, esse requisito não é relevante (cf. (32)).

(31) O João tossiu durante 10 minutos.

(32) O João tem tossido (desde esta manhã / desde a semana passada).

Para além disso, de acordo com Cunha (2006), a iteração requer um intervalo de ocorrência relativamente curto e delimitado (cf. (31)); pelo contrário, o intervalo a que corresponde um PPC pode ter uma extensão muito variável (cf. (32)).

Finalmente, e ao contrário da iteração, o PPC com leitura de repetição das eventualidades permite a combinação com processos e estados. Vejam-se os exemplos (33) e (34).

(33) O João tem estado no jardim na última semana.

(34) O João tem perseguido a vizinha na última semana.

Contudo, há um conjunto de aspetos que aproximam o PPC das restantes estruturas de iteração. Em primeiro lugar, tal como as estruturas de iteração, o PPC converte um número não determinado de eventualidades numa única eventualidade (cf. (35)) e as eventualidades repetidas constituem as subfases dessa eventualidade processual. Por outras palavras, em ambos os exemplos de (35) se considera uma eventualidade constituída por subfases que correspondem a cada um dos espirros dados pelo João.

(35) a. O João espirrou durante meia hora.

b. O João tem espirrado.

Para além disso, as eventualidades grupais criadas pelo PPC têm as propriedades da divisibilidade e da cumulatividade, tal como as estruturas iterativas. Quanto à divisibilidade, se a proposição em (36a) é verdadeira, então também é verdadeira em qualquer subintervalo contido no intervalo de tempo cuja fronteira inicial é marcada por “as três horas” e cuja fronteira final coincide com o momento da enunciação. No que concerne à cumulatividade, vejam-se (36b-c). Se é verdade (36a) e se é verdade (36b), é necessariamente verdade (36c)

- (36) a. O João tem tossido desde as 3 horas.
b. O João tossiu entre as 2 horas e as 3 horas.
c. → O João tem tossido desde as 2 horas.

É de salientar que a repetição de situações do PPC tem uma natureza cumulativa, apesar de ter uma fronteira temporal à esquerda (implícita ou explícita) e uma fronteira à direita (momento da enunciação).

Note-se, por fim, que, com alguns processos culminados, os subeventos repetidos denotados pelo PPC não correspondem às eventualidades básicas, mas apenas a partes do seu processo preparatório. É o caso de (37), em que o que se repete são subpartes do processo preparatório de “escrever a tese”.

(37) O João tem escrito a tese.

Apresentamos a seguir um quadro que sistematiza as semelhanças e diferenças apontadas ao PPC e estruturas de iteração.

Quadro III – PPC e estruturas de iteração: semelhanças e diferenças

Semelhanças	Diferenças
<ol style="list-style-type: none">1. Repetição de eventualidades.2. Converte-se um número não determinado de eventualidades numa única eventualidade.3. Os eventos básicos repetidos são as subfases da eventualidade derivada.4. As eventualidades derivadas são divisíveis e cumulativas.	<ol style="list-style-type: none">1. Na iteração, os eventos têm de estar organizados de forma contígua, sem pausas significativas entre si (de acordo com Cunha, 2006); o PPC exige que haja pausas significativas entre os eventos que se repetem.2. A iteração requer um intervalo de ocorrência relativamente curto e delimitado (de acordo com Cunha, 2006); o intervalo a que corresponde um PPC pode ter uma extensão muito variável.3. O PPC combina-se com processos e estados (as estruturas iterativas, não, de acordo com Cunha, 2006).

Embora haja diferenças, elas não parecem ser significativas, como nos casos da habitualidade e da frequência. Por exemplo, a diferença

apontada em 3 sugere que o PPC tem um alcance maior do que os restantes desencadeadores da iteração. Os dados apresentados indiciam, portanto, que o PPC corresponde a uma estrutura de iteração com um alcance mais vasto do que o que é proposto em Cunha (2006) e que a definição dada por este autor requererá provavelmente alguma reformulação.²

3. Leituras iterativa e não iterativa do PPC: para uma explicação

A análise que é efetuada na secção anterior leva-nos a defender que o processo criado pelo PPC é um processo de natureza iterativa, e não habitual ou frequentativa. Este processo evidencia uma homogeneidade “relativa”, na medida em que, apesar de o processo não se aplicar ao longo de todo um intervalo (dado que tem pausas), os falantes o interpretam como se a predicação se aplicasse a todos os instantes do intervalo (cf. Kratzer 1989; Landman & Rothstein 2012a, 2012b). Contudo, há a realçar que a criação de um processo de natureza iterativa não acontece em todos os casos, havendo a possibilidade de surgimento de leitura não iterativa com alguns tipos aspetuais de predicações - processos e estados stage-level. Torna-se, portanto, necessário conciliar estas duas interpretações do PPC, ou seja, determinar o que permite que este tempo gramatical permita leituras aparentemente antagónicas.³

Para tal, avançamos com a seguinte hipótese: o PPC é um operador que cria eventualidades que se caracterizam pela sua homogeneidade. As duas leituras possíveis do PPC prendem-se com tipos diferentes de homogeneidade: (i) no caso de haver leitura de eventualidade única, trata-se de homogeneidade segmental (Landman & Rothstein 2012a); (ii) no caso de haver leitura iterativa, trata-se de homogeneidade incremental (Landman & Rothstein 2012b).

A homogeneidade segmental corresponde à propriedade de subintervalos.

² Neste trabalho, não temos como objetivo averiguar se a iteração associada ao PPC tem subjacente um mecanismo que pode ser classificado como “operador” ou como “perspetivador” aspetual (cf. Cunha, 2004).

³ A propósito desta questão, veja-se Ilari (2001:148-149): “Para tratar do PASSADO COMPOSTO nessa perspectiva unificada, a regra semântica precisa aplicar-se tanto a grandezas discretas como a grandezas contínuas, ou seja, ao invés de tentar reduzir a duração à iteratividade de eventos momentâneos, ou vice-versa, precisamos buscar uma matriz capaz de aceitar os dois tipos de eventos”

De acordo com Landman & Rothstein (2012b:100), “um estado s é homogéneo sse, para cada subintervalo i incluído no tempo de duração de s , existir um estado s' que pertence ao mesmo tipo de eventualidade que s tal que s e s' são transtemporalmente idênticos (i.e. para efeitos de contagem de eventualidades, são a mesma eventualidade) e o tempo de duração de s' corresponde a i ”⁴. Para ilustrar esta definição, atente-se no exemplo (38).

(38) O rapaz tem estado no jardim (na última meia hora).

Relativamente a (38), o estado de ‘o rapaz estar no jardim’, que acontece na última meia hora, é segmentalmente homogéneo, pois, em cada subintervalo do tempo de duração do estado ‘o rapaz estar no jardim’, existe um estado de ‘o rapaz estar no jardim’ que conta como “o mesmo estado” que ‘o rapaz estar no jardim’ (o estado global, ou total).

Deste modo, considerando apenas a propriedade da homogeneidade segmental, não existem diferenças entre (38), com o PPC, e (39), com o Pretérito Perfeito Simples: as diferenças entre estes tempos gramaticais são, de certa forma, anuladas.

(39) O rapaz esteve no jardim (na última meia hora).

Por seu lado, quando há leitura iterativa, estamos perante o segundo tipo de homogeneidade, a homogeneidade incremental, que caracteriza os eventos de tipo processual. Para explicar este segundo tipo de homogeneidade, Landman & Rothstein (2012b) partem da noção de *onset*. De acordo com Landman e Rothstein, todos os predicados eventivos têm um evento *onset*. Este evento *onset* é um evento que corresponde a um estádio inicial de um evento e e que não é um ponto. Tomemos como exemplo, para ilustrar esta noção, o predicado ‘comer uma maçã’, o qual, aspetualmente, corresponde a um processo culminado. O *onset* do predicado ‘comer uma maçã’ é um predicado eventivo com um núcleo verbal que corresponde a ‘comer’. Este *onset* (i) corresponde a um estádio inicial de ‘comer uma maçã’ ou seja, faz

⁴ Tradução nossa de: “State s in stative event type α is *homogenous with respect to* α iff if $\tau(s,w)$ is defined then for every subinterval $i \subseteq \tau(s,w)$, there is a state $s' \in \alpha$ such that $s' \sim s$ and $\tau(s',w) = i$ ” Landman & Rothstein (2012b:100)

parte do evento de ‘comer uma maçã’, (ii) está em decurso num tempo de duração que é um subintervalo inicial de ‘comer uma maçã’, (iii) é maior do que um ponto, (iv) é o menor evento que corresponde à denotação de ‘comer’ e (v) está na denotação de ‘comer’, mas não na denotação de ‘comer uma maçã’, pelo que corresponderá à sequência composta por: *trincar, mastigar e engolir o primeiro pedaço de maçã*.

Apresentada a noção de *onset*, passemos à noção de homogeneidade incremental. Um evento *e* é incrementalmente homogéneo sse, para cada subintervalo inicial próprio *i* que faça parte do tempo de duração de *e* e que se encontre incrementalmente entre o tempo de duração do *onset* de *e* e o tempo de duração de *e*, (i) existir um evento do mesmo tipo, que é transtemporalmente idêntico a *e* e (ii) que tem *i* como o seu tempo de duração. Vejamos a aplicação desta noção ao predicado ‘dançar’: se *e* for um evento de ‘dançar’, encontramos, para cada subintervalo próprio do tempo de duração de *e* que estenda incrementalmente o tempo de duração do *onset-dançar* de *e*, um evento transtemporalmente idêntico a *e*, que é ele próprio um evento de ‘dançar’.

Este segundo tipo de homogeneidade, por ser de natureza incremental, permite a existência de pausas. Para ilustrar esta afirmação, tomemos novamente como exemplo o predicado ‘dançar’, em (40).

(40) A Maria dançou na festa.

Em (40), podemos assumir a existência de um evento de ‘dançar’ entre as 21h e as 22h, mesmo que tenha havido pausas (por exemplo, mesmo que a Maria tenha parado de dançar para beber um copo de água duas vezes). Para tal, é necessária uma decisão contextual que determine que um evento de ‘dançar’ é o prolongamento incremental de um evento de ‘dançar’ anterior (ou seja, há identidade transtemporal).

Tendo apresentado as noções de homogeneidade propostas em Landman & Rothstein (2012a, 2012b), vejamos de que forma se pode aplicar a noção de homogeneidade incremental ao PPC com leitura iterativa. Considere-se (41).

(41) O miúdo tem tossido.

Seguindo Landman & Rothstein (2012b), propomos o seguinte tratamento para a leitura iterativa da frase (41). Uma sequência de eventos de ‘o miúdo tossir’ é transformada num processo através da adição de pausas, tal como é ilustrado no Quadro 4.

Quadro 4 – representação da leitura iterativa do PPC

evento1	evento2	evento3	evento4
	pausa1	pausa2	pausa3

A situação começa com o evento1. A fase seguinte será a união da subeventualidade constituída por *evento 1 + pausa1* com *evento2*, tidos como transtemporalmente idênticos. A decisão de *evento1* e *pausa1* contarem como transtemporalmente idênticos é de natureza contextual. Assim, *evento1*, neste contexto, não conta como uma eventualidade por si só, mas é apenas uma fase de um processo maior. O *onset* deste processo é a fase desde o primeiro evento de ‘o miúdo tossir’ até ao segundo evento – i.e. o primeiro momento em que se verifica que há uma iteração.

Note-se que, de acordo com Landman & Rothstein (2012b), a decisão de gerar mais fases transtemporalmente idênticas em processos iterativos é uma decisão contextual. Já de acordo com a proposta que aqui avançamos, o PPC caracteriza-se por forçar a criação de fases transtemporalmente idênticas até ao momento da enunciação. Ou seja, no caso de frases com o PPC, a criação de fases transtemporalmente idênticas não tem a ver com uma decisão contextual, mas é informação marcada gramaticalmente.

4. Observações finais

A análise que efetuámos ao longo deste trabalho permitiu pôr em evidência alguns aspetos do PPC em PE. Assim, apesar de este tempo gramatical partilhar características com construções de habitualidade e de frequência,

apresenta também diferenças que nos parecem irredutíveis. Referimo-nos especificamente ao facto de, ao contrário dos estados habituais, o PPC não corresponder a propriedades gerais das entidades e não ser verdadeiro em todos os subintervalos do intervalo a que se aplica. Para além disso, e ao contrário das estruturas frequentativas, o PPC com estados e processos, não leva a uma pluralização de situações quando se combina com predicções que sejam estados e processos e, com alguns processos culminados, pode acarretar alterações aspetuais.

Assim, propusemos que o processo criado pelo PPC é um processo de natureza iterativa e que esta iteração veiculada pelo PPC, contrariamente ao que é proposto em Cunha (2006) para outras construções iterativas, não exige um intervalo curto, nem a contiguidade dos eventos, assim como não exclui processos e estados.

A propriedade central dos processos que são criados por iteração pelo PPC é, na nossa proposta, a homogeneidade “relativa” ou incremental. Isto significa que, apesar de o processo não se aplicar ao longo de todo um intervalo (por haver pausas), os falantes o interpretam como se a predicação se aplicasse a todos os instantes do intervalo (cf. Kratzer 1989; Landman & Rothstein 2012b). Assim, o PPC denota uma eventualidade derivada que, na leitura iterativa, é obrigatoriamente tida como não se aplicando a todos os instantes do intervalo relevante, i.e. tem necessariamente pausas entre as suas fases.

O PPC é, por conseguinte, um operador que cria eventualidades homogéneas, no sentido que é dado em Landman & Rothstein (2012a, 2012b): no caso de haver leitura de eventualidade única (com estados e processos), a homogeneidade é de tipo segmental; no caso de haver leitura iterativa (com eventos e estados *stage-level*), a homogeneidade é de tipo incremental.

REFERÊNCIAS

- Cunha, L.F. 2006. Frequência vs. Habitualidade: Distinções e Convergências. In: *Actas del XXXV Simposio Internacional de la Sociedad Española de Lingüística*, 333-357. Retirado da Internet, a 25 de maio de 2012: <http://www3.unileon.es/dp/dfh/SEL/actas/Cunha.pdf>.
- Ilari, R. 2001. Notas sobre o Passado Composto em Português. *Letras*, Curitiba, 55: 129-152.
- Kratzer, A. 1989. An Investigation of the Lumps of Thought. *Linguistics and Philosophy*, 12, nº5: 607-653.
- Laca, B. 2006. Indefinites, quantifiers and pluractionals. What scope effects tell us about event pluralities. In: S. Vogeleer; L. Tasmowski (Eds.). *Non-definiteness and plurality*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 191-217.
- Laca, B. 2010. Perfect Semantics: How Universal Are Ibero-American Present Perfects?. In: C. Borgonovo et al. (Eds.). *Selected Proceedings of the 12th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project, 1-16. Retirado, a 25 de maio de 2012, da Internet: <http://www.lingref.com/cpp/hls/12/paper2401.pdf>.
- Landman, F. & Rothstein, S. 2012a. The felicity of aspectual *for*-phrases, part 1: homogeneity. *Language and Linguistic Compass*. Oxford: Blackwell-Wiley, 85-96.
- Landman, F. & Rothstein, S. 2012b. The felicity of aspectual *for*-phrases, part 2: incremental homogeneity. *Language and Linguistic Compass*. Oxford: Blackwell-Wiley, 97-112.
- Link, G. 1983. The logical analysis of plurals and mass terms: a lattice-theoretical approach. In: R. Bäuerle; C. Schwarze; A. von Stechow (Eds.). *Meaning, use and interpretation of language*. Berlin: Walter de Gruyter, 302-323.
- Lopes, O. 1986/2005. Sobre as noções de definido e de presente de enunciação. *Entre a Palavra e o Discurso*. Estudos de Lingüística (1977-1993). Edição coordenada por F. Oliveira e A. M. Brito. Porto: Campo das Letras, 105-120.
- Martinez-Atienza, M. 2008. Dos formas de oposición en el ámbito románico.

- In: Carrasco Gutiérrez, A. *Tiempos compuestos y formas verbales complejas*. Madrid: Lingüística Iberoamericana, 204 -229.
- Oliveira, F. 2003. Tempo e Aspeto. In: Mateus, M. H. et al., *Gramática da Língua Portuguesa*. Lisboa: Caminho, 127-178.
- Oliveira, F. & Leal, A. 2012. "Sobre a iteração do Pretérito Perfeito Composto em Português Europeu". In. *Linguística - Revista de Estudos Linguísticos da UP*, vol.7, nº1: 65-88.
- Oliveira, F., Leal, A. & Silva, F. 2014. Pretérito Perfeito Composto e quantificação em Português Europeu. *Textos Seleccionados do XXIX Encontro Nacional da Associação Portuguesa de Linguística*, 407-418.
- Smith, C. 1991. *The Parameter of Aspect*. Dordrecht: Reidel Pub. Comp.
- Squartini, M. & Bertinetto, P.M. 2000. The Simple and Compound Past in Romance Languages. In: Ö. Dahl (Ed.). *Tense and Aspect in the Languages of Europe*. Berlin: Mouton de Gruyter, 403-440.
- Van Geenhoven, V. 2004. For-adverbials, frequentative aspect, and pluractionality. *Natural Language Semantics*, 12, nº2: 135-190.
- CETEMPúblico (Corpus de Extractos de Textos Electrónicos MCT/Público) - <http://www.linguateca.pt/>

Análise de sentimento em artigos de opinião

Fátima Silva

mhenri@letras.up.pt

Faculdade de Letras /Centro de Linguística da Universidade do Porto

Purificação Silvano

msilvano@letras.up.pt

Faculdade de Letras /Centro de Linguística da Universidade do Porto

António Leal

jleal@letras.up.pt

Faculdade de Letras /Centro de Linguística da Universidade do Porto

Fátima Oliveira

moliv@letras.up.pt

Faculdade de Letras /Centro de Linguística da Universidade do Porto

Pavel Brazdil

pbrazdil@inescporto.pt

*Faculdade de Economia da Universidade do Porto/LIAAD – INESC Tec,
Porto*

João Cordeiro

jpaulo@di.ubi.pt

Universidade da Beira Interior/LIAAD – INESC Tec, Porto

Débora Oliveira

livino.debora.uporto@gmail.com

Faculdade de Economia da Universidade do Porto/LIAAD – INESC Tec, Porto

ABSTRACT: The present study, which is developed in the interface between linguistics and computer science within the framework of sentiment analysis, aims at making a computational analysis of opinion articles in the area of economics and finance. The main objectives of the study are: i) to determine the semantic orientation of text segments that express opinion by annotating the polarity (positive or negative) and the strength (scale from -3 to 3) of nouns and adjectives, and ii) to verify if a specific lexicon for the area of economics and finance has advantages in automatic annotation of sentiment over a general lexicon. To achieve these objectives, a corpus of 45 texts was selected and analyzed in 2 phases, by annotators with different training. First, a sample of 10 texts was annotated by linguists, co-authors of this paper, with the objective of developing a linguistic annotation model to ascertain the polarity and strength of words in opinion articles and extract the relevant words for this area of study. Then, a set of 35 texts was annotated by university students, replicating the annotation model developed during the first phase. Based on the linguistic annotation, the computer science team tried to establish to what extent a general sentiment lexicon for Portuguese - SentiLex - was sufficient to extract the sentiment of a sentence in a satisfactory manner or whether EconoLex, a specific sentiment lexicon, would be more efficient. The specific lexicon includes terms and multiword expressions that are relevant to the area of economics and finance and to Portuguese language, and it was developed by the authors of this study. The data was analyzed according to a blending methodology, qualitative and quantitative. The results of the analysis allow us to consider the following items as contributes of this study: i) the development of a linguistic annotation model for the analysis of the polarity and strength of the lexicon, especially of nouns and adjectives; ii) the key role, though not exclusive, of the adjectives to determine the polarity of opinion segments of the corpus articles; iii) the creation of a new specific sentiment lexicon for Portuguese in the area of economics and finance; iv) the improvement of the computational performance of EconoLex@SentiLex in relation to SentiLex regarding the performance in automatic annotation of sentiment. In spite of these positive results, there are some limitations, which we intend to overcome in the continuity of this interdisciplinary work, namely a more detailed linguistic analysis of the word classes that we studied, the consideration of other elements/ linguistic structures that are essential to ascertain the sentiment in NP/sentence, the extension of the corpus, the expansion of the specific lexicon of the area of economics and finance and the improvement of automatic methods for identifying evaluative words in texts of opinion and for assigning them polarity and strength.

KEYWORDS: sentiment analysis, opinion article, automatic assignment of sentiment, economics and finance, polarity and strength, lexicon, EconoLex

RESUMO: O estudo apresentado realiza-se na interface entre a linguística e as ciências da computação, tendo como objetivo fazer a análise computacional de artigos de opinião na área da economia e finanças, seguindo o quadro teórico da análise de sentimento. Os principais objetivos do trabalho são i) determinar a orientação do sentimento, positivo ou negativo, e a intensidade dessa orientação através da anotação da polaridade do léxico, com incidência nos nomes e adjetivos, nos segmentos em que ocorre a expressão da opinião, e ii) verificar se um léxico específico para a área de economia e finanças tem vantagens na atribuição automática de sentimento sobre um léxico geral. Para atingir esses objetivos, foi

selecionado um corpus de 45 textos, analisado em duas fases por anotadores com formação distinta. Primeiro, uma amostra de 10 textos foi obtida e anotada pelos investigadores da área de linguística, coautores deste artigo, com o objetivo de desenvolver um modelo linguístico para determinar a orientação e intensidade da polaridade de termos em artigos de opinião e extrair termos de léxico relevantes para esta área de estudo. Em seguida, um conjunto de 35 textos foi anotado por estudantes universitários, seguindo o método utilizado na primeira amostra. Com base na anotação linguística, a equipa das ciências da computação procurou determinar até que ponto um léxico de sentimento geral para a língua portuguesa – SentiLex - é suficiente para caracterizar o sentimento de uma frase de maneira satisfatória ou se o EconoLex, um léxico específico de sentimento, seria mais eficaz. O léxico específico inclui termos e expressões multpalavra relevantes para o domínio da economia e finanças e para a língua portuguesa, e foi elaborado pelos autores deste estudo. Os dados foram analisados usando uma metodologia mista, qualitativa e quantitativa. Os resultados obtidos permitem-nos considerar os seguintes itens como contributos desta investigação: i) a elaboração do modelo de anotação linguística adotado para a análise da orientação e da intensidade da polaridade do léxico, em especial dos nomes e adjetivos; ii) o papel central, ainda que não exclusivo, dos adjetivos para a determinação da polaridade do sentimento nos segmentos opinativos dos artigos do corpus; iii) o desenvolvimento de um novo léxico de sentimento específico português para a área da economia e finanças; iv) a melhoria do desempenho computacional do EconoLex⊕SentiLex em relação ao SentiLex no que se refere ao desempenho na caracterização automática de sentimento. Apesar destes resultados positivos, há algumas limitações que constituem os elementos a desenvolver na continuidade deste trabalho interdisciplinar, nomeadamente a análise linguística mais detalhada das classes gramaticais estudadas, a consideração de outros elementos/estruturas linguísticas determinantes para a caracterização do sentimento em SN/frase, o alargamento do corpus, o aumento do léxico específico do domínio e a afinação dos métodos automáticos de identificação de termos de sentimento em textos de opinião e determinação da sua intensidade.

PALAVRAS-CHAVE: análise do sentimento, artigo de opinião, atribuição automática do sentimento, economia e finanças, polaridade e intensidade do sentimento, léxico, EconoLex

1. Introdução¹

Este trabalho centra-se na análise computacional de artigos de opinião na área de economia e finanças, situando-se no quadro da investigação sobre análise de sentimento.

Trata-se de uma investigação realizada na interface entre a linguística e as ciências da computação, sendo objetivos centrais: i) determinar a

¹ Agradecemos aos avaliadores deste artigo os valiosos comentários, muitos dos quais foram integrados no texto. No entanto, a responsabilidade por qualquer falha remanescente é dos autores do artigo.

importância do léxico, em especial dos adjetivos e nomes, para transmitir uma opinião, positiva ou negativa, sobre o(s) tópico(s) principal(ais) do texto e ii) verificar se um léxico específico para a área de economia e finanças tem vantagens na atribuição automática de sentimento.

Através da análise do corpus, que consiste essencialmente em determinar a orientação de sentimento associado ao léxico, em especial aos adjetivos e aos nomes, e a classificação da intensidade dessa orientação², tanto em termos individuais como sintagmáticos, extraindo o léxico relevante do domínio temático em estudo, estabelecemos um modelo de anotação linguística para anotar um conjunto mais vasto de textos, tanto do ponto de vista manual como computacional, assim como para a constituição de um léxico específico do domínio, o EconoLex.

O EconoLex é um léxico de sentimento com termos e expressões multipalavra relevantes para o domínio da economia e finanças e para a língua portuguesa, tendo sido elaborado sob a direção de alguns dos autores deste estudo, investigadores da área da linguística. A primeira versão do léxico é constituída pelos termos resultantes da anotação linguística de 10 artigos, que serviu de base para a anotação de mais 35 textos, realizada por estudantes universitários, tendo, por conseguinte, o estudo empírico relativo à previsão computacional dos valores de sentimento recaído sobre 45 textos.

Com base nos resultados da análise linguística, a análise computacional procura determinar até que ponto um léxico de sentimento geral para a língua portuguesa - SentiLex (cf., e.o., Carvalho & Silva, 2015) é suficiente para caracterizar o sentimento de uma frase de maneira satisfatória ou se o EconoLex, um léxico específico para o domínio em causa (economia e finanças), é mais eficaz.

Assim, os principais contributos desta investigação são: (i) o desenvolvimento de um modelo de anotação linguística para delimitação da

² Neste trabalho, seguimos a terminologia adotada por Liu (2012, 2015), que considera, como parâmetros de análise, a 'orientação do sentimento' (sentiment orientation), positiva, negativa ou neutra, e a 'intensidade do sentimento' (sentiment intensity), o nível de intensidade ou força do sentimento, estreitamente relacionado com o que designa de 'classificação do sentimento' (sentiment rating'), correspondente a uma classificação discreta para exprimir a intensidade do sentimento (no caso do nosso trabalho, os níveis 3 a -3). Note-se, no entanto, que outros autores usam uma terminologia diferente e com aplicação pelo menos ligeiramente diferenciada em contexto similar, facto de que Liu dá também conta: "Sentiment orientation is also called polarity, semantic orientation, or valence in the research literature." (Liu, 2015: 21). 'Orientação semântica' é o termo usado, por exemplo, entre outros, por Taboada et al. (2011: 267-268) para referir "the polarity and strength of words, phrases, or texts".

orientação e da intensidade da polaridade do léxico em artigos de opinião; (ii) o desenvolvimento de um novo léxico (EconoLex) para Português, (iii) a comparação com o Sentilex, e (iv) o desenvolvimento de uma metodologia computacional com vista à classificação automática do sentimento de frases.

O artigo está organizado em 4 secções, a primeira das quais é a introdução. Na secção 2, fazemos um breve enquadramento teórico sobre a análise de sentimento e o género artigo de opinião. Na secção 3, descrevemos o estudo realizado caracterizando o corpus e a metodologia seguida na análise linguística e na análise computacional. A secção 4 apresenta os resultados das duas análises. Terminamos com algumas considerações finais, que nos permitem retomar o percurso efetuado e avaliar do grau de consecução dos objetivos apresentados na introdução.

2. Algumas considerações teóricas

2.1 Sobre análise de sentimento

A análise de sentimento (*sentiment analysis*) é um campo de investigação relativamente recente (Das & Chen, 2001; Tong, 2001; Turney, 2002; Pang, Lee & Vaithyanathan, 2002; Dave, Lawrence & Pennock, 2003; Nasukawa & Yi, 2003; Pang & Lee, 2008), mas que tem conhecido grande desenvolvimento (Benamara, Taboada & Mathieu, 2017). Tem como objeto de estudo as opiniões, avaliações, atitudes e emoções relativamente a certas entidades, como produtos, serviços, organizações, pessoas, eventos ou tópicos (Cambria & Hussain, 2015; Liu, 2012, 2015). Embora o número de estudos neste domínio seja muito elevado, a maior parte do trabalho na área teve como objeto a língua inglesa, havendo muito menos estudos para o Português, em especial para o Português Europeu (e.g. Carvalho, Sarmiento, Silva & Oliveira; 2009; Silva, Carvalho, Costa & Sarmiento, 2010; Silva & Team, 2011; Silva, Carvalho & Sarmiento, 2012; Antunes, 2015; Marques Lucena et al., 2015; Forte & Brazdil, 2016).

De um modo geral, o foco principal da análise de sentimento consiste na capacidade de um sistema automático poder aferir a subjetividade e a manifestação de emoções num texto escrito em relação a certas entidades nomeadas de forma explícita ou implícita. Assim, consideramos o termo

sentimento de forma abrangente, incluindo os *afetos* e as *emoções* que a Psicologia define e caracteriza de forma rigorosa (Fiorin, 2007).

Têm sido estudadas diferentes formas de modelar as emoções, que vão desde a simples deteção de polaridade (positiva vs. negativa) (Turney, 2002), à combinação desta com níveis de subjetividade-objetividade (Baccianella, Esuli & Sebastiani, 2010) e até modelos mais ricos que consideram um conjunto de “emoções base” e depois definem um estado emocional como uma combinação destas emoções (Russell, 1980; Ekman, 1999).

Em termos de análise e abordagem ao problema da análise de sentimento, vários modelos e metodologias têm sido experimentados (Ravi & Ravi, 2015; Liu, 2015), quase todos tomando como base de trabalho um corpus de textos de opinião, parcialmente anotado por humanos, quanto às expressões e intensidades emocionais contidas naqueles. Estes corpora anotados permitem que sistemas de aprendizagem automática (Kodratoff & Michalski, 2014; Witten & Frank, 2016) consigam induzir conceitos gerais de identificação e caracterização de emoções em texto, ou sistemas baseados em léxico de sentimentos, que englobam os termos gerais facilitadores da identificação das emoções (e.g. Hung & Lin, 2013; Forte & Brazdil, 2016) ou combinações de ambos.

Neste contexto, a Linguística tem tido um papel crucial e várias têm sido as propostas teóricas entre as quais se destacam (cf. segundo Taboada, 2016; Benamara, Taboada & Mathieu, 2017) a teoria da avaliação (Martin & White, 2005), postura (Biber & Finegan, 1989), avaliação (Hunston & Thompson, 2000) e contrafactualidade (‘nonveridicality’) (Taboada & Trnavac, 2013). Apesar de não haver nenhum estudo que integre todas as componentes avaliativas disponíveis na língua para a expressão do sentimento nos textos (Benamara, Taboada & Mathieu, 2017), em grande parte pela sua complexidade e abrangência, é possível identificar alguns dos elementos mais relevantes na análise de sentimento na investigação em curso.

Um desses elementos, e um dos que tem recebido maior enfoque, é o léxico, e, em particular, as classes dos adjetivos, dos nomes e, em menor número, os verbos (e.g. Levin, 1993; Mathieu, 2005; Taboada, Anthony & Voll, 2006; Freitas, 2013). Estas classes de palavras têm sido analisadas quanto a polaridade positiva, negativa e neutra, muitas vezes associada a

escalas com diferentes valores, que variam de acordo com as propostas (Goldberg & Zhu, 2006; Pang & Lee, 2008). Uma das tarefas mais complexas relacionada com a atribuição de valores de polaridade a adjetivos, nomes e verbos de sentimento e com a criação de um léxico com estes valores (Levin, 1993; Neviarouskaya, Prendiger & Ishizuka, 2009; Freitas, 2013) surge sempre que o significado base dos itens não é evidente e/ou depende do contexto em que é usado.

2.2. Sobre o género artigo de opinião

O artigo de opinião é um género de discurso que se caracteriza pelo seu traço marcado de comentário (Adam 1997; Charaudeau 2006; Cunha 2012). No domínio jornalístico, consiste essencialmente na discussão, por parte de um autor, mais ou menos especializado, de um assunto atual e considerado de relevância para o leitor, ocupando um espaço que oscila habitualmente entre meia página e uma página no jornal, numa secção dedicada a questões de economia e finanças, no caso dos textos que analisámos. Este autor, é, de acordo com Cunha (2012: 75), “um especialista externo à instância mediática comentando um facto ou mesmo provocando a sua emergência”. Nesse sentido, existe uma assimetria entre a instância que produz o texto e a instância que o recebe, na medida em que a primeira é institucionalmente legitimada na exposição do seu ponto de vista enquanto a segunda o recebe de modo a informar-se e esclarecer-se sobre o tópico em questão (Rodrigues, 2005; Cunha, 2012). Tipicamente, o objeto de análise é um tema que se centra sobre factos recentes, sobre os quais é frequente a geração de diferentes pontos de vista, sendo, por conseguinte, polémico. Assim, consiste frequentemente num facto que “mobilizou a atenção da opinião pública e agora pede que os veículos de comunicação apresentem as análises “esclarecidas” de especialistas da área em que o fato se deu” (Cunha, 2012: 76).

Tendo em consideração que a simples explicitação da opinião do autor do texto não é suficiente para garantir a adesão do leitor ao seu ponto de vista, até pelo facto de existirem habitualmente pontos divergentes sobre o tema, são desenvolvidas estratégias de argumentação que visam convencer o leitor deste ponto de vista, fazendo-o aderir à leitura proposta sobre o

tema objeto de discussão.

A textualização das estratégias argumentativas no artigo de opinião, a partir da qual se pode analisar a orientação do sentimento ou da opinião, em termos da sua polaridade positiva, negativa ou neutra, faz uso de vários recursos linguísticos, entre os quais o léxico. A análise decorrente da descrição do léxico fornece pistas para a determinação da polaridade do documento no seu todo, como expressando sentimento positivo ou negativo (cf. Silva *et al.* 2015, 2018). É verdade que a sua combinação com outros aspetos da análise de sentimento deve ser tida em conta, porque a expressão de sentimento no texto recorre a outros elementos de natureza linguística e discursiva, que não se esgotam nas categorias gramaticais consideradas. No entanto, é indiscutível que a análise semântica lexical e frásica dos nomes, adjetivos e verbos e da sua combinatória em sintagmas fornece indicações relevantes para a análise da expressão do sentimento em textos de opinião do domínio em análise.

3. O estudo

3.1. Corpus

O estudo tem como base um corpus de 45 textos. Numa primeira fase, foi analisada uma amostra de 10 textos do domínio discursivo da economia e finanças do género artigo de opinião, para teste da metodologia de análise linguística e extração dos termos a integrar no EconoLex, o léxico específico para a área de economia e finanças. Esses textos foram extraídos de forma aleatória de três jornais digitais diários, *Público* (1), *Diário de Notícias* (1) e *Jornal de Negócios* (8). Enquanto os dois primeiros jornais referidos são generalistas, o terceiro é um jornal especializado em assuntos de economia e finanças. A recolha foi realizada entre 3 de janeiro e 8 de fevereiro de 2017, tendo sido subordinada a uma unidade temática, a Caixa Geral de Depósitos, para evitar uma dispersão do léxico e aferir os valores do léxico de sentimento encontrado em contextos similares.

Numa segunda fase, foram considerados mais 35 textos aos quais foi aplicada a mesma metodologia de análise da amostra. Estes textos foram igualmente selecionados de forma aleatória, a partir de jornais generalistas e

diários - *Público* (3), *Jornal de Notícias* (2), *Diário de Notícias* (2) e *Observador* (1) -; de jornais especializados, diário e semanário, respetivamente - *Jornal de Negócios* (1) e *Jornal Económico* (12) -; de um jornal generalista semanário - *Expresso* (9) -; e de uma revista semanal - *Visão* (5). Ao contrário da amostra, a recolha destes textos não seguiu um critério temático, na medida em que se pretendia alargar o léxico de sentimento, embora se mantivesse sempre no domínio de economia e finanças.

3.2. Metodologia de Análise Linguística

Feita a seleção e extração do corpus, a análise de sentimento dos 10 artigos de opinião referidos em 3.1. realizou-se em várias etapas:

1. Extração dos segmentos dos textos em que se exprime opinião, isto é, em que está implicado um sentimento positivo ou negativo;
2. Delimitação, nos segmentos considerados, das frases relevantes para análise, isto é, dos contextos em que ocorre uma expressão de sentimento;
3. Determinação das categorias que exprimem sentimento – nomes, adjetivos e multipalavras – e dos sintagmas nominais por eles constituídos;
4. Anotação da orientação do sentimento de cada uma destas categorias – positiva, negativa ou neutra;
5. Anotação da intensidade da polaridade dessa orientação, numa escala de 3 a -3 valores;
6. Avaliação dos resultados da análise;
7. Identificação das palavras relevantes para integração num léxico específico do domínio da economia e finanças, o EconoLex.

A anotação do corpus foi feita manualmente por 3 investigadores especialistas em linguística. Nos casos em que a anotação deu resultados discrepantes em termos da orientação de sentimento e, sobretudo, da intensidade da polaridade dessa orientação, foi realizada uma reanálise e reavaliação conjunta dos dados em questão para aferição de um valor. As palavras e os sintagmas anotados foram sujeitos a inquérito junto de dois informantes especializados em linguística, mas externos à investigação em curso, para aferição da anotação efetuada pelos investigadores, em

geral, e discussão dos casos de anotação não concordante, em particular³. A anotação da orientação e da intensidade de sentimento baseou-se nos seguintes critérios: (i) avaliação por falantes nativos, especialistas e não especialistas; (ii) consulta do significado lexical dos termos em dicionários gerais; (iii) e consulta de dicionários gerais de sentimento de Português e de outras línguas (Sentilex, So-Cal, entre outros).

3.2.1 Identificação dos segmentos relevantes

Como referimos, a análise dos textos constitutivos do corpus consistiu, em primeiro lugar, na identificação dos segmentos em que se veicula de forma explícita (cf. (1)) ou implícita (cf. (2)) a expressão de opinião sobre o tema relevante, tendo sido excluídos os segmentos em que a expressão de sentimento não se centra sobre este tópico de modo direto (cf. (3)). Assim, em (1), a expressão de opinião é marcada, nomeadamente, através do verbo de opinião 'julgar' e do SN 'infeliz marca'. Por sua vez, em (2), a opinião é dada através de uma analogia com um filme descrito no 1.º parágrafo do texto. Finalmente, em (3), o segmento de opinião não se refere diretamente à Caixa Geral de Depósitos, mas às declarações das duas entidades citadas.

- (1) Tem-se falado de prémios de gestão. Prémios de gestão avultados são uma infeliz marca do sistema financeiro. Não julgo apropriada, e até prejudicial, a existência para a CGD de prémios da natureza dos da banca privada, em função dos lucros. Os prémios estão na origem dos males do sistema financeiro, colocando na mira dos gestores bancários a perspectiva de ganhos pessoais elevadíssimos na dependência dos lucros das instituições. [texto 8]
- (2) É um bonito dia de sol que nos acolhe no início do filme "La La Land". Estamos em Los Angeles e o trânsito está engarrafado numa auto-estrada e nada se move. Vamos seguindo então a música diferente que sai de cada carro parado. É uma sinfonia pouco sincronizada, porque cada um escuta uma coisa diferente. Quando paramos, a

3 Agradecemos a Idalina Ferreira e a Luís Filipe Cunha a colaboração na anotação da polaridade do léxico extraído dos 10 textos analisados na primeira fase do trabalho.

música fica mais alta. Então a personagem sai do carro e começa a dançar. Rapidamente todos os outros condutores saltam para as capotas dos carros e dançam sem parar. Todos dançam como se fossem Gene Kelly em “Singing in the Rain”. Os musicais são assim. Serão sempre. A Caixa Geral de Depósitos é uma “La La Land” à nossa dimensão... [texto 4]

- (3) As declarações de Fernando Faria de Oliveira e Carlos Santos Ferreira na comissão parlamentar de inquérito à Caixa Geral de Depósitos (CGD) vieram enfatizar aquilo que há muito é uma evidência: o banco só será pacificado quando se realizar e forem tornadas públicas as conclusões da auditoria forense, aprovada pela Assembleia da República a 20 de Julho de 2016, mas que ainda se encontra numa qualquer gaveta do Banco de Portugal. [texto 5]

3.2.2 Identificação de frases e categorias de palavras a analisar

No âmbito dos segmentos selecionados, foram identificadas as frases relevantes para a análise de sentimento, nas quais estavam contidas as categorias sintáticas analisadas: nomes e adjetivos. Neste contexto, foram considerados não só nomes e adjetivos que exprimem intrinsecamente sentimento (cf. (4) e (5)), mas também nomes e adjetivos que estão relacionados com o domínio específico da economia e finanças e são usados na textualização de sentimento, independentemente de terem, no seu significado lexical nuclear, uma polaridade neutra (cf. (6) e (7)).

- (4) Tal como não é novo que a Caixa vai cortar o número de balcões e de funcionários de forma **agressiva**: são menos 2200 trabalhadores e perto de 200 balcões.
- (5) Os prémios estão na origem dos **males** do sistema financeiro, colocando na mira dos gestores bancários a perspectiva de **ganhos** pessoais elevadíssimos na dependência dos lucros das instituições.
- (6) A **gestão** é pouco qualificada e a mão-de-obra também.
- (7) A escolha de Macedo foi inteligente, tanto do ponto de vista **técnico** como **político**.

Assim, enquanto o adjetivo ‘agressiva’ e os nomes ‘males’ e ‘ganhos’ exprimem no seu significado lexical nuclear a expressão de um sentimento, o nome ‘gestão’ e os adjetivos ‘técnico’ e ‘político’ não contêm esse traço no seu significado lexical básico, sendo necessário atender ao significado lexical atribuído contextualmente para legitimar a sua consideração na delimitação de valores de sentimento a uma dada expressão ou frase. Isto não significa, no entanto, desvalorizar o significado contextualmente atribuído a nomes e adjetivos tipicamente ocorrentes em léxicos de sentimento, dado que a orientação de sentimento e da intensidade dessa orientação é determinada na articulação do significado lexical nuclear com o significado contextualmente adquirido, em particular quando se considera o sintagma nominal (SN) em que ocorrem e o contexto frásico integral.

No que se refere aos adjetivos, foram considerados os seguintes critérios de análise: função - predicativa ou atributiva - e, em relação a esta última, foi definida a posição pré-nominal ou pós-nominal. Os exemplos (8) e (9) atestam a diferença de posição do adjetivo (atributiva e predicativa), enquanto os exemplos (10) – (11) exemplificam a posição pré ou pós-nominal dos adjetivos em posição atributiva.

(8) A dívida ingerível, o crescimento **anémico**, os impostos **altos**, o investimento **nulo**, a regulação **ineficaz**, a separação dos portugueses entre protegidos e excluídos exigem mais do que um simples “virar de página” da austeridade. [texto 2]

(9) A escolha de Macedo foi **inteligente**... [texto 3]

(10) Esse é, aliás, o **grande** alerta da OCDE ao referir que o problema da banca está longe de estar resolvido. [texto 1]

(11) São necessárias taxas de crescimento de 3%, 4% ou 5%, ou seja, taxas **ambiciosas**... [texto 1]

Por outro lado, foram analisadas as subclasses dos adjetivos – *adverbiais*, *numerais*, *qualificativos* e *relacionais*⁴. Os exemplos (12)

⁴ Na análise das classes semânticas, seguimos a proposta de Ferreira (2013). Agradecemos à autora a colaboração na anotação semântica dos adjetivos relevantes para a análise de sentimento proposta.

– (15) ilustram, respetivamente, cada uma das subclasses de adjetivos atualizadas nos adjetivos ‘fortes’, ‘última’, ‘ágil’ e ‘agressiva’, ‘estruturais’. Os *adverbiais* são genericamente aqueles que podem ser transformados em advérbios. No exemplo (12) isso acontece por causa da posição do adjetivo. Se o adjetivo estivesse em posição pós-nominal seria *qualificativo* (“recapitalizações fortes”). Os *numerais* relacionam-se com ordenação (cf. (13)); os *qualificativos* atribuem propriedades e são graduáveis (cf. (14)); os *relacionais* derivam em geral de nomes (‘estrutura’ - ‘estrutural’) (cf. (15)). Note-se que, se usarmos os conceitos da área de análise de sentimento, o adjetivo *adverbial* no exemplo (12) representa um *intensificador*.

(12) ...exigiram aos seus bancos que procedessem a **fortes** recapitalizações dos seus capitais. [texto 9]

(13) Esta é a **última** oportunidade para a Caixa. [texto 3]

(14) Na prática, referia-se ao atraso com que o banco aumentou ‘spreads’ e comissões e cortou remuneração dos depósitos, face à concorrência, muito mais **ágil** e **agressiva**. [texto 5]

(15) Está na altura de Portugal e a UE se capacitarem que os problemas **estruturais** se resolvem com medidas de fundo. [texto 9]

3.2.3. Expressões Multipalavra

Além de nomes e de adjetivos, foi também considerada a ocorrência de multipalavras, de que se dá exemplo nos segmentos (16) e (17), com as multipalavras ‘banco público’ e ‘sistema político’.

(16) Na verdade, a vantagem para a economia que se pode tirar da existência de um **banco público** é ser diferente na sua política de aplicações. [texto 8]

(17) ... o défice está em dia, o diabo não veio, não há muitas greves e experimentamos a genialidade de um **sistema político** “plástico” que impediu o vazio do poder. [texto 2]

Entende-se que uma multipalavra é “uma sequência de palavras que

atuam como uma simples unidade em algum nível da análise linguística” (Calzolari *et al.*, 2002: 1934), pode englobar um variado número de construções, tais como expressões fixas, compostos nominais e construções verbo-partícula (Sag *et al.*, 2002: 191) e ocorre com frequência em domínios técnicos. O conceito de *multipalavra* é complexo tanto do ponto de vista da análise linguística como computacional. Neste trabalho, de natureza exploratória, não foi nosso objetivo rever de forma aprofundada este conceito, mas usamo-lo no seguimento de McCarthy (1990); Nattinger & DeCarrico (2001); Calzolari *et al.* (2002); Sag *et al.* (2002); Ranchhod (2003); Gómez Molina (2004); Thornbury (2007); Abalada *et al.* (2010); entre outros.

No corpus, a delimitação de multipalavras da área de economia e finanças fez-se considerando sobretudo a existência de expressões lexicalizadas. Verificamos que, sempre que ocorre uma multipalavra constituída pela estrutura ‘nome + adjetivo’, o adjetivo é relacional, como é o caso da sua ocorrência nos exemplos (16) e (17), respetivamente ‘banco público’ e ‘sistema político’.

3.2.4. Anotação do Corpus

Com base nestas categorias (nomes, adjetivos e multipalavras), procedemos à análise do corpus, cuja anotação se fez numa grelha considerando oito parâmetros: contexto, frase relevante, polaridade do nome, polaridade do adjetivo, posição do adjetivo, subclasse semântica do adjetivo, polaridade do SN e polaridade da frase. A tabela 1 contém a exemplificação da grelha usada para a anotação.

A tabela 2 mostra apenas exemplos de algumas destas frases com os respetivos valores de sentimento atribuídos (colunas um e dois da tabela). Os valores de sentimento variam de -3 (fortemente negativo) a 3 (fortemente positivo), isto é, podem incluir qualquer valor do conjunto {-3, -2, -1, 0, 1, 2, 3}. Nas colunas três e quatro da tabela podemos ver os valores de sentimento associados aos termos uni-palavra que ocorrem nas frases, divididos pelas classes gramaticais de ‘nome’ e ‘adjetivo’ (ex. ‘crise’, ‘anémico’). As expressões multipalavra são maioritariamente do tipo ‘nome + adjetivo’, como, por exemplo: ‘carga fiscal’, ‘investimento público’,

‘sistema político’.

Note-se que, na determinação da intensidade da polaridade, é necessário ter em conta que, além do conhecimento linguístico e da competência textual dos potenciais anotadores e, de forma mais global, dos leitores, intervém o que poderíamos designar por conhecimento do mundo e um sistema de crenças de natureza diversa, na base do que poderíamos discutir em termos da distinção entre uma polaridade mais subjetiva, dependente da ideologia ou da conceção do mundo do falante/ouvinte, e uma outra, mais universal/objetiva. Este sistema de crenças, socioculturalmente marcado, pode interferir na polaridade atribuída (cf., e.o., Li & Liu 2012: 128). A título de exemplo, consideremos o SN “pensamento conservador”, que ocorre na tabela 2. Embora o valor negativo seja claro no texto, é preciso ler o parágrafo e interpretar a ideologia económica do autor para lhe atribuir esse valor negativo. Noutro contexto ideológico, de índole mais conservadora e neoliberal, essa mesma frase nominal poderia eventualmente ter valor positivo.

As questões colocadas por este tipo de diferenças podem ser limitadas por vários procedimentos como sejam a anotação por um conjunto alargado de informantes e o recurso a léxicos gerais e específicos; a consideração de outros elementos que concorrem para a configuração do sentido dos segmentos de opinião; a análise da percentagem de concordâncias obtidas entre anotadores, do ponto de vista linguístico, e a aplicação de algoritmos computacionais. Estes foram os critérios que nortearam a nossa anotação do corpus já referido.

TABELA 1: Grelha de anotação manual da análise linguística de sentimento

contexto	frase relevante	nome: polaridade	adjetivo: polaridade	posição predica-tiva/atributiva	posição atributiva: pré/ pós-nominal	adjetivo: sub-classe	SN: polaridade	polaridade da frase
Portugal não pode estar a governar só para os mercados, ou seja, para tentar demonstrar que o défice está melhor. Mas a verdade também é que, se não baixar o défice, é penalizado na dívida.	Portugal não pode estar a governar só para os mercados, ou seja, para tentar demonstrar que o défice está melhor.	<i>défice</i> -1	<i>melhor</i> +2	pred.	-	qualif.	pos.	neg.
Mas a verdade também é que, se não baixar o défice, é penalizado na dívida. O país sente-se manietado e não é só por isso, mas também pelo crescimento anémico, pelo investimento reduzido e por um tecido empresarial, no geral, ainda pouco competitivo.	O país sente-se manietado e não é só por isso, mas também pelo crescimento anémico, pelo investimento reduzido e por um tecido empresarial, no geral, ainda pouco competitivo.	<i>país</i> 0	<i>manietado</i> (particípio) -2	pred.	-	qualif.	neg.	neg.
		<i>crecimento</i> +2	<i>anémico</i> -2	atrib.	pos	qualif.	neg.	
		<i>investimento</i> +1	<i>reduzido</i> -1	atrib.	pos	qualif.	neg.	
		<i>tecido</i> 0	<i>empresarial</i> 0	atrib.	pos	relac.	neutro	
		<i>tecido</i> 0	<i>competitivo</i> +1	atrib.	pos	qualif.	neg.	

TABELA 2 - Exemplos de frases, termos relevantes e valores de sentimento

Frase (F)	V(F)	V(nome)	V(adjetivo)	V (express)
Face às eleições anteriores, o Bloco perdeu metade dos votos, metade dos deputados e entrou numa crise profunda .	-3	<i>crise</i> : -2	<i>profunda</i> : -2	<i>crise profunda</i> : -3
Por mais avanços que a tecnocracia europeia se mostre disposta a dar, na sequência da crise do euro, a predominância de um <i>pensamento conservador</i> nas esferas de poder torna previsível que fossem impostas de novo <i>medidas violentas de contenção orçamental</i> .	-2	<i>pensamento</i> : 0 <i>medidas</i> : 0 <i>contenção</i> : -1	<i>conservador</i> : 0 <i>violentas</i> : -1 <i>orçamental</i> : 0	<i>pensamento conservador</i> : -1 <i>medidas violentas</i> : -2 <i>contenção orçamental</i> : -1
O país sente-se manietado e não é só por isso, mas também pelo <u>crescimento anémico</u> , pelo investimento reduzido e por um tecido empresarial, no geral, ainda pouco <u>competitivo</u> .	-3	<i>manietado</i> : -2 <i>crescimento</i> : 2 <i>investimento</i> : 1	<i>anémico</i> : -2 <i>reduzido</i> : -1 <i>competitivo</i> : 1	<i>crescimento anémico</i> : -1 <i>investimento reduzido</i> : -1

A anotação e análise dos 10 artigos iniciais deram origem à primeira versão do EconoLex. Uma outra parte do léxico, bastante significativa, resultou da anotação de mais 35 artigos por estudantes do ensino superior. Assim, no total, foram usados 45 artigos da área da economia e finanças e, após uma pré-análise, foram escolhidas 370 frases.

3.3. Análise computacional

Para este estudo, usamos uma metodologia que segue uma abordagem baseada em métodos computacionais, através do uso de corpora e léxicos eletrónicos, bem como procedimentos que envolvem cálculo de valor previsto de sentimento de frases.

3.3.1. Uso de léxico no cálculo de valor de sentimento

Foi implementado um procedimento automático para caracterizar o sentimento de uma determinada frase. Para cada palavra da frase, é feita uma procura no léxico. Caso a palavra ocorra, tomar-se-á o valor de sentimento associado à mesma. Neste procedimento, o cálculo final do valor de sentimento da frase é processado em duas fases. Na primeira, o procedimento elabora uma soma dos valores de todos os termos encontrados no léxico. Na segunda fase, estes valores são reescalados, dividindo-os por um fator que tem um efeito equalizador, tendo em conta certas gamas de valores de sentimento, para os diferentes léxicos considerados. No final, cada frase fica associada a um valor numérico que traduz o sentimento geral dessa frase e que poderá ser negativo, neutro (zero), ou positivo, em diferentes níveis de intensidade. Este processo está descrito com mais pormenor na secção 3.3.2.

Léxico geral SentiLex

O SentiLex (Silva *et al.*, 2012), mais precisamente a versão SentiLex-PT02 que foi usada neste estudo, foi desenvolvido para uso geral (i.e. é *general-purpose lexicon*). O léxico contém uma lista de cerca 82.347 entradas, que são palavras flexionadas ou expressões idiomáticas e multipalavras em Português. Este conjunto corresponde a 7.014 termos lematizados. Cada entrada está associada a um valor de sentimento do conjunto $\{-1, 0, 1\}$, significando respetivamente valores de sentimento *negativo*, *neutro* e *positivo*. Assim, por exemplo, para a entrada ‘abafada’ e ‘combaterá o desemprego’ temos respetivamente os valores -1 e 1 associados, sendo a primeira um adjetivo, cujo lema é ‘abafado’, e a segunda uma expressão multipalavra, que na área de lexicologia é referida como *colocação*. Na

tabela 3 encontramos algumas entradas do SentiLex, com os respetivos lemas, classe gramatical e valor de sentimento associado.

TABELA 3 - Entradas típicas no SentiLex

Termo	Lema	Classe	Valor
abafada	abafado	Adj	-1
condenou	condenar	Vrb	-1
votou	votar	Vrb	0
encorajando	encorajar	Vrb	1
irregularidade	irregularidade	Nom	-1
combaterá o de- semprego	combater o de- semprego	Multi	1

Ao trabalhar com o SentiLex, notou-se que este não inclui alguns termos importantes da área da economia e finanças, como, por exemplo, 'inflação', 'despedir', e 'incumprimento', entre outros. Na versão lematizada, a maioria dos termos do SentiLex são adjetivos. No entanto, quando se considera a versão flexionada, como seria previsível, há mais verbos do que adjetivos (ver tabela 4).

TABELA 4 - Distribuição de ocorrências de lemas por tipo (classe)

Classe	Lema	Forma flexionada
Nome	1,080	1,280
Adjetivo	4,779	16,863
Verbo	489	29,504
Multipalavra	666	34,700
Total	7,014	82,347

A maioria das entradas (66%) tem o valor negativo (-1), enquanto só um quarto (25%) tem o valor positivo (1) e as restantes entradas (9%) têm o valor neutro (0).

Léxico específico EconoLex

O EconoLex é, tal como foi referido anteriormente, um léxico de sentimento com termos e expressões multipalavra relevantes para o domínio da economia e finanças em Português. Este léxico contém menos entradas do que o SentiLex, pois inclui 2811 termos não flexionados, que correspondem a 1246 termos lematizados. Neste léxico os valores positivos e negativos estão mais equitativamente distribuídos, relativamente ao SentiLex, pois há 46% de termos positivos, 35% de negativos e 19% de neutros.

3.3.2. O cálculo final do valor de sentimento

O léxico específico EconoLex foi usado como extensão do léxico geral SentiLex. A junção dos dois léxicos é aqui designada de EconoLex \oplus SentiLex.

Como um dado termo pode ocorrer em ambos os léxicos, torna-se necessário decidir qual o valor de sentimento a aplicar na caracterização do sentimento de uma frase. Assim, para cada termo $t \in \{\text{EconoLex}, \text{SentiLex}\}$ escolhe-se o *valor(t)* no primeiro ou no segundo léxico? Decidimos dar preferência ao léxico específico EconoLex, atendendo a que este caracteriza de forma mais adequada o domínio do texto em análise. Em termos procedimentais, esta preferência é executada dando prioridade aos termos do léxico EconoLex sobre o léxico SentiLex.

Cada léxico é ordenado de acordo com o tamanho da expressão que corresponde ao número de palavras envolvidas. As expressões de tamanho maior são colocadas no topo da lista de prioridades relativamente às expressões de tamanho inferior e aos termos uni-palavra que aparecem mais para o final da lista. Esta ordenação tem como objetivo aplicar o valor da expressão multipalavra, caso ocorra, em vez dos valores das uni-palavras que a constituem. Considere-se, por exemplo, que se pretende obter o valor do seguinte fragmento, usando o léxico na Tabela 5 e concentrando especial atenção no segmento ‘crescimento anémico’:

- (18) O país sente-se manietado e não é só por isso, mas também pelo crescimento anémico, pelo investimento reduzido e por um tecido empresarial, no geral, ainda pouco competitivo.

TABELA 5 - Esboço de entradas no léxico

Termo	Tamanho	Valor
crescimento anémico	2	-1
anémico	1	-2
crescimento	1	2

Tendo em conta a ordem das entradas na tabela, o valor do segmento corresponde à primeira entrada: 'crescimento anémico' com o valor -1. Caso não fosse aplicada esta ordem de prioridade, aplicar-se-iam os valores dos termos que constituem esta expressão, isto é, os valores atribuídos individualmente a 'anémico' e a 'crescimento'. Repare-se que, nesse caso, o valor deste segmento seria obtido através da soma dos seus valores individuais, resultando numa neutralização (zero) do valor final do segmento, o que seria claramente inapropriado.

3.3.3. Reescalonamento de valores de sentimento

Como o objetivo é prever os valores de sentimento usando um método computacional baseado no léxico, é necessário garantir que os valores gerados são da mesma grandeza e, se não o forem, ajustá-los e integrá-los da forma mais conveniente. Para atingir este fim, usámos um reescalonamento de valores, tendo por base os desvios padrão de pares de populações, sendo cada população caracterizada pela aplicação de um determinado léxico, no cálculo do sentimento das frases.

Em primeiro lugar, usámos o nosso procedimento para gerar as classificações de sentimento das frases. Os valores gerados foram usados para construir um histograma que pode ser caracterizado pela sua *média* e *desvio de padrão* σ_{prev} . Os valores de σ_{prev} para SentiLex, EconoLex e EconoLex \oplus SentiLex foram respetivamente: 1,02, 3,65 e 3,85. Repare-se no baixo valor de σ_{prev} para o SentiLex e que se deve fundamentalmente à gama de valores de sentimento mais reduzida neste léxico ($\{-1,0,1\}$). Um processo semelhante pode ser usado para caracterizar os valores atribuídos pelos especialistas humanos que classificaram manualmente o valor de

sentimento para cada frase. Assim, vamos obter um histograma semelhante, que pode ser caracterizado por σ_{true} . Neste caso, o valor obtido foi de 1,45.

Subsequentemente, decidiu-se tomar como fator de reescalonamento o rácio $\sigma_{\text{prev}}/\sigma_{\text{true}}$. No caso de EconoLex, por exemplo, o fator é de $3,65/1,45 = 2,51$. Após o cálculo do fator, todos os valores previstos são divididos pelo respetivo fator, sendo transformados para uma escala normalizada.

3.3.4. Avaliação de previsões de polaridade

Em termos de avaliação da eficácia de uma abordagem, o método convencional utilizado nas áreas de *Aprendizagem Automática (Machine Learning)*, *Prospecção de Dados (Data Mining)* e *Prospecção de Texto (Text Mining)* consiste na comparação das previsões da abordagem com os valores corretos (*golden standard*). Assim, neste estudo também foram comparados os valores de sentimento calculados para uma frase f , $\text{ValS}_{\text{prev}}(f)$, com os valores de sentimento atribuídos pelos especialistas para essa mesma frase, $\text{ValS}_{\text{true}}(f)$, que se considera como sendo o *correto*. Deste modo, calculou-se a diferença absoluta (erro absoluto) entre os dois valores, nomeadamente o erro absoluto (EA) para cada frase f :

$$\text{EA}(f) = | \text{ValS}_{\text{prev}}(f) - \text{ValS}_{\text{true}}(f) |.$$

Por exemplo, se $\text{ValS}_{\text{prev}}(f) = 1$ e $\text{ValS}_{\text{true}}(f) = -1$, então o valor de EA é igual ao valor 2 ($1 - (-1)$). Pretende-se que a diferença dos dois valores seja a mais baixa possível. Esta medição é aplicada a todas as frases e depois é calculada a média dos erros absolutos, ou o chamado *Erro Absoluto Médio (EAM)*, como medida de desempenho geral.

3.3.5. Metodologia de validação cruzada com ‘deixa-um-fora’

Na área da *Aprendizagem Automática*, os procedimentos propostos (e.g. um classificador) são normalmente avaliados seguindo a metodologia designada de *validação cruzada*. Trata-se de um conjunto de dados que normalmente é dividido em dados de treino e de teste, sendo o primeiro usado para construir um modelo para gerar os valores previstos.

Seguidamente, mede-se o desempenho do procedimento nos dados de teste.

Uma versão mais elaborada deste processo consiste em dividir o conjunto de dados em N blocos de tamanho aproximadamente igual (normalmente $N = 10$) e executar N ciclos treino/teste. Para o i -ésimo ciclo, o procedimento é treinado com todos os blocos, exceto o bloco i que é usado para teste. O desempenho final é obtido a partir da média do desempenho nos N blocos.

No contexto deste trabalho, um bloco de dados é simplesmente um bloco de textos. Aqui o “treino” envolve a elaboração do léxico a partir de $N-1$ blocos (textos), bem como o reescalonamento descrito na subsecção 3.3.3.

A metodologia *Validação Cruzada com Deixa-Um-Fora (Leave-One-Out Cross-Validation, LOOCV)* é semelhante à descrita anteriormente, mas o número de blocos N é igual ao número de casos (textos) nos dados. Neste estudo, foi esta a metodologia de avaliação adotada. Assim, usámos $N-1$ textos (no nosso caso 44 textos) para elaborar o léxico específico e o texto que ficou de fora foi usado para avaliar as previsões. O processo é repetido N vezes (i.e., 45 vezes) e toma-se a média dos resultados de desempenho como desempenho final, medido em termos de EAM.

4. Discussão dos resultados

4.1. Análise linguística

Tendo explicitado na secção 3.2. os parâmetros usados na análise do corpus, apresentamos nesta secção os resultados dessa análise, subdividindo-os em duas componentes. Na primeira, apresentamos a análise dos dados; na segunda, descrevemos as regularidades encontradas no corpus no que se refere à análise de sentimento e à sua orientação.

4.1.1. Dados quantitativos

O corpus dos 10 artigos de opinião perfaz no seu total 5407 palavras, o que equivale a uma média de 540,7 palavras por texto. Tendo sido contabilizadas as ocorrências de nomes (N) e de adjetivos (A) e de multipalavras (MP)

no corpus, verificamos que elas correspondem, respetivamente, a 1,96%, 2,75% e 0,61%, o que constitui uma percentagem globalmente pequena de ocorrências lexicais na totalidade do corpus, embora forneça um contributo fundamental para a argumentação veiculada na explicitação da opinião do autor. A tabela 6 especifica estes resultados, assim como a sua distribuição por texto analisado.

TABELA 6: Quantificação das ocorrências relevantes em função do número total de palavras

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	totais
n.º total de ocorrências de palavras	411	1144	425	345	496	344	455	716	585	486	5407 (100%)
n.º de ocorrências N	15	19	10	5	9	3	6	11	11	17	106 (1,96%)
n.º de ocorrências ADJ	18	21	13	7	14	3	7	19	24	23	149 (2,755%)
n.º de ocorrências de multipalavras	1	3	1	0	2	0	0	3	2	2	14 (0,61%)

Na tabela 7, é explicitada a quantificação dos adjetivos em função da sua subclasse.

TABELA 7: Quantificação dos adjetivos ocorrentes no corpus por subclasse semântica

	tipos	n.º de ocorrências	%
adjetivos 149	adverbiais	11	7,39
	numerais	1	0,67
	qualificativos	110	73,82
	relacionais	27	18,12

Neste âmbito, verificamos que há uma dominância de adjetivos qualificativos no contexto da expressão de valores de sentimento (73,82%),

logo seguida pelos adjetivos relacionais (18,12%). Com um valor menos expressivo, 7,39%, ocorrem os adjetivos adverbiais, sendo residual o número de adjetivos numerais nas ocorrências relevantes. Esta tendência exprime, de forma evidente, o papel dos adjetivos qualificativos no contexto da expressão de valores de sentimento, decorrendo das suas propriedades semânticas de expressão de uma qualidade, o que, no contexto do presente estudo, se coaduna com a manifestação da opinião do autor do texto relativamente a determinadas entidades (essencialmente denotadas pelos nomes). A ocorrência de adjetivos relacionais em número relativamente elevado está, em certa medida, ligada à sua associação a nomes com os quais, em contextos vários, produzem expressões lexicalizadas do tipo multipalavras.

Quanto ao adjetivo, verificamos que há uma diferença percentual significativa em relação à sua posição, que é sobretudo atributiva, com 81,8% das ocorrências, o que vai ao encontro da sua função central na expressão da avaliação da polaridade das entidades textuais, frequentemente denotadas por nomes da área da economia e finanças. Relativamente à posição dos adjetivos com função atributiva, observamos que 61,44% das suas ocorrências nesta posição é pós-nominal, o que constitui mais de metade das ocorrências. Estes dados podem ser observados de forma circunstanciada na tabela 8.

TABELA 8: Quantificação da ocorrência de adjetivos quanto à posição

		T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	totais
n.º de ocorrências ADJ	total	18	21	13	7	14	3	7	19	24	23	149 (100%)
	predicativa	6	1	3	1	3	0	1	6	3	3	27 (18,12%)
	atributiva pré-nominal	1	4	1	1	6	0	2	3	4	8	30 (20,134)
	atributiva pós-nominal	11	16	9	5	5	3	4	10	17	12	92 (61,744)

4.1.2. Regularidades encontradas na expressão dos valores de sentimento

De seguida, apresentamos algumas regularidades identificadas através da análise do corpus.

Nos casos em que a polaridade dos nomes e dos adjetivos é distinta entre si, é o adjetivo que é responsável pela polaridade do SN, como se verifica nos exemplos (19) e (20), em que é o adjetivo a determinar a polaridade, respetivamente, positiva e negativa do SN, quer em posição atributiva, quer em posição predicativa.

(19) Portugal não pode estar a governar só para os mercados, ou seja, para tentar demonstrar que o **défi**ce está **melhor**. [texto 1]

N -1 Adj 2 SN Pos

(20) Que encare **verdades inconvenientes**, como o fracasso a prazo do sistema de pensões. [texto 2]

N 3 Adj -2 SN Neg

Há casos em que não é possível determinar o contributo do adjetivo para a valoração do SN, pois ambos têm a mesma polaridade, como é o caso dos exemplos (21) e (22).

(21) A **dívida ingerível**, o crescimento anémico, os impostos altos, o investimento nulo, a regulação ineficaz, a separação dos portugueses entre protegidos e excluídos exigem mais do que um simples «virar de página» da austeridade. [texto 2]

N -3 Adj -3 SN Neg

(22) O que a OCDE contrapõe é uma dúvida a esse estado de **pura felicidade**. [texto 2]

N 3 Adj 3 SN Pos

Por outro lado, os adjetivos qualificativos têm, geralmente, valoração positiva ou negativa, e os relacionais, valoração neutra. Os adjetivos qualificativos ‘excelentes’ (cf. (23)) e ‘chocante’ (cf. (24)) são adjetivos qualificativos com especificação da orientação de sentimento, positivo e negativo, respetivamente, mas ao adjetivo ‘políticos’ (cf. (25)), de natureza relacional, atribui-se valoração neutra.

- (23) Esta partidarização da CGD foi visível em administrações sucessivas, onde no meio de **excelentes gestores**, existiam prateleiras de luxo para quem saía da órbita governamental. [texto 4]
- (24) Este **jogo de sombras** é **chocante**. [texto 5]
- (25) O que passa por se proteger dos **lóbis políticos** e alterar de forma substantiva a sua actividade creditícia. [texto 3]

Apesar de, como se disse os adjetivos serem maioritariamente responsáveis pela polaridade do SN, há casos em que são subespecificados quanto à valoração em virtude de o conteúdo semântico do nome ser decisivo para atribuição da polaridade, como acontece nos exemplos (26) e (27). Neste contexto, o adjetivo 'sério' pode assumir uma polaridade positiva ou negativa conforme o nome ao qual se associa. Assim, em (26), ligado ao nome 'forma' e dado o contexto, assume um valor positivo, sendo a intensidade da polaridade que lhe foi atribuída de 2 (na escala de 3 a -3). Já no caso da sua associação a 'avisos' (cf. (27)), por força do significado lexical do nome, assume uma polaridade negativa, com intensidade -2.

- (26) Os problemas existem e têm de ser resolvidos, haja a coragem política de os reconhecer, mas também de os expor de **forma séria** e direta à UE, procurando zelar pelos interesses dos cidadãos portugueses e não pelos índices de popularidade junto destes. [texto 9]
- (27) Desta vez, a OCDE poupou-nos a raspanetes. Mas deixou **avisos sérios**. Manter a santa paz das rendas e privilégios é parar no tempo. E permanecer no atraso que faz de Portugal o país mais injusto da Europa. [texto 2]

O estudo exploratório de natureza linguística descrito consistiu, portanto, na anotação dos valores de sentimento de um conjunto de artigos de opinião da área da economia e finanças, permitindo não só a análise do contributo dos adjetivos e do nome para a expressão da orientação do sentimento ou polaridade (positiva, negativa ou neutra), mas também da sua intensidade (3 a -3), e ainda a elaboração de um léxico de adjetivos, nomes e multipalavras relevantes para a criação de um léxico computacional

específico de economia e finanças, o EconoLex. Além disso, esta análise usou uma metodologia replicável a um conjunto mais vasto de textos (45 no total) ao qual se aplicou uma análise de tipo computacional, com os objetivos, a metodologia e os resultados apresentados na secção 3.3.

4.2. Previsão computacional de valores de sentimento

Nesta secção, pretende-se explorar e conhecer a possibilidade efetiva de um sistema automático poder prever corretamente a polaridade de uma frase, bem como a intensidade do sentimento presente. Seguimos uma abordagem baseada em léxico, dirigindo especial atenção a textos de opinião no domínio da economia e finanças.

4.2.1 Resultados quantitativos

Os resultados deste estudo são apresentados na Tabela 10. Assim, o melhor resultado é obtido com a combinação de léxicos EconoLex \oplus SentiLex, pois o valor de EAM é o mais baixo de todos.

TABELA 9 - Valores de EAM resultantes da aplicação de três léxicos de sentimento.

Léxico	EAM
SentiLex	1,53
EconoLex	1,41
EconoLex \oplus SentiLex	1,38

O uso exclusivo do EconoLex apresenta também um valor de erro inferior ao do uso exclusivo do léxico SentiLex, mas, na realidade, a junção dos dois atinge o menor erro. Todavia, importa verificar, em termos estatísticos, se esta diferença é ou não significativa.

De modo a verificar a significância estatística dos resultados obtidos, realizou-se o teste não-paramétrico de Wilcoxon com nível de confiança de 95%. Este teste não-paramétrico classifica os valores absolutos das diferenças entre as observações emparelhadas para cada frase e o respetivo léxico

utilizado, calculando assim uma estatística sobre o número de diferenças negativas e positivas.

O *p-value* resultante foi de 0,027 para o EconoLex \oplus SentiLex versus SentiLex. Como o *valor* foi inferior a 0,05, a diferença entre as observações pode ser considerada estatisticamente significativa. Um teste semelhante foi realizado também para o EconoLex versus SentiLex, mas, neste caso, o *p-value* era superior (0.057), isto é, o resultado não é estatisticamente significativo.

4.2.2. Discussão dos resultados

Neste trabalho, comparámos o efeito da introdução de um léxico específico de sentimento (EconoLex) do domínio da economia e finanças na caracterização automática de sentimento, partindo de um léxico de sentimento mais geral (SentiLex). O EconoLex foi especificamente criado a partir de um número modesto de artigos de opinião do domínio da economia e finanças. Seguiu-se, portanto, uma abordagem de análise de sentimento baseada em léxico extraído de um corpus específico de um domínio. Procedeu-se a uma medição do *erro absoluto médio* (EAM), numa avaliação cruzada dos dados e para três configurações de léxicos de sentimento (SentiLex, EconoLex, EconoLex \oplus SentiLex).

Os resultados mostraram uma melhoria, embora não muito expressiva, do EconoLex \oplus SentiLex em relação ao SentiLex. Na comparação dos resultados EAM com o teste Wilcoxon, constata-se um *p-value* de 0,027, o que dá evidência estatística suficiente para afirmar que o acréscimo de léxico específico (EconoLex) ao léxico geral melhorou o desempenho.

O estudo empírico aqui apresentado poderá ser melhorado com a inclusão de um maior número de textos do domínio. Além disso, poderá também explorar-se a possibilidade de expansão automática do léxico de sentimento deste domínio seguindo/adaptando alguns trabalhos recentes nesta área (Almatarneh & Gamallo, 2018).

5. Considerações finais

Este trabalho teve como objetivos construir um léxico de sentimento específico do domínio da economia e finanças, partindo da avaliação da relevância de nomes, adjetivos e multipalavras na orientação e intensidade dos segmentos de opinião ocorrentes em artigos de opinião, e aferir a eficácia desse léxico na anotação automática do sentimento.

Os resultados da análise do corpus evidenciam a necessidade de melhorar não só o EconoLex, mas também a análise linguística. Em termos de dados, é fundamental alargar o número de textos, idealmente para algumas centenas, bem como proceder a um aumento do léxico específico do domínio, que continua a ser reduzido, relativamente ao SentiLex. Embora não se pretenda que tenha a mesma dimensão, uma vez que se trata de um domínio específico, há, no entanto, um maior leque de termos, expressões multipalavra e construções linguísticas específicas, inerentes ao domínio que nos interessa considerar.

Relativamente ao aumento do léxico de sentimento, no domínio económico, existem diversas possibilidades promissoras para a sua execução de forma automática. Uma abordagem mais clássica consiste no uso de um *thesaurus* (e.g. WordNet-PT; cf. Marrafa 2004), de modo a identificar sinónimos de termos já conhecidos e expandir em vários incrementos sucessivos os termos originais, resultando assim num léxico mais rico.

Uma outra abordagem, mais atual, é usar os chamados *word embeddings*, tais como *word2vec* (Mikolov et al. 2013) e *GloVe* (Pennington et al. 2014), para fazer o mesmo. A grande diferença neste contexto é que as relações semânticas entre os termos são induzidas automaticamente a partir de corpora, envolvendo o treino de uma rede neuronal multicamada. Uma vantagem evidente desta abordagem é a possibilidade de treinar os modelos com textos específicos de um domínio. No nosso caso, com um bom volume de textos do domínio da economia e finanças, poder-se-ão induzir relações semânticas específicas para estes domínios, menos comuns nos modelos gerais.

Uma outra possibilidade seria introduzir regras semelhantes às que são usadas no tratamento de expressões polares, intensificadores e atenuadores no trabalho de Polanyi & Zaenen (2006) ou de Forte & Brazdil (2016).

No que se refere à análise linguística, estudos como este são prova de

que uma investigação linguística mais detalhada de classes gramaticais como os adjetivos e nomes é necessária, não só para analisar as suas propriedades lexicais na base, mas também as propriedades composicionais na constituição de SN, SADJ. Para além disso, é importante considerar outros elementos/estruturas linguísticas que são determinantes para a caracterização do sentimento em SN/frase, nomeadamente verbos, advérbios e a negação, porque a expressão de sentimento no texto recorre a outros elementos de natureza linguística e discursiva, que não se esgotam nas categorias gramaticais consideradas. Benamara, Taboada & Mathieu (2017) sintetizam esses elementos, dando especial relevo à ocorrência de determinados advérbios, por exemplo, intensificadores e de negação, que podem fazer mudar a polaridade contextual decorrente da consideração de dados lexicais apenas, ou ainda a modalidade, quantificação, relações retóricas, etc.

Tendo em consideração os aspetos que, de acordo com esta investigação, necessitam de melhoramento, é nosso propósito, na sua continuidade, contribuir para o avanço do conhecimento na área da análise de sentimento na interface entre linguística e ciências da computação, com o alargamento do EconoLex e a melhoria dos métodos automáticos de identificação de sentimento em textos de opinião da área da Economia em Português Europeu, em articulação com os dados resultantes da investigação linguística.

REFERÊNCIAS

- Abalada, S., Cabarrão, V. & Cardoso, A. 2010. Proposta de Classificação Semântica de Unidades Lexicais Multipalavra Nominais. In: *Textos Seleccionados do XXV Encontro Nacional da APL*. Porto, 81-94.
- Adam, J-M. 1997. Unités rédactionnelles et genres discursifs : cadre général pour une approche de la presse écrite. *Pratiques*, 94: 3-18.
- Almatarneh, S. & Gamallo, P. 2018. A Comparative Study of Polarity Lexicons to Identify Extreme Opinions. In: *Proceedings of SNAMS 2018, Fifth International Conference on Social Networks Analysis, the Second International Workshop on Advances in*

- Natural Language Processing (ANLP 2018) Management and Security*. Valencia, Spain, 296-301.
- Antunes, P. 2015. *Sentiment Analysis in Financial News*. Dissertação de Mestrado. Porto: FEP.
- Baccianella, S., Esuli, A. & Sebastiani, F. 2010. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In: *Proceedings of the Seventh conference on International Language Resources and Evaluation*. Valtetta, Malta, 2200-2204.
- Benamara, F., Taboada, M. & Mathieu, Y. 2017. Evaluative Language Beyond Bags of Words: Linguistic Insights and Computational Applications. *Computational Linguistics*. 43(1): 201-264.
- Biber D, Finegan E. 1989. Styles of stance in English: lexical and grammatical marking of evidentiality and affect. *Text*. 9: 93–124.
- Calzolari, N., Fillmore, C. J., Grishman, R., Ide, N., Lenci, A., MacLeod, C. & Zampolli, A. 2002. Towards Best Practice for Multiword Expressions in Computational Lexicons. In: *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC 2002)*. Las Palmas, Canary Islands, 1934-1940.
- Cambria, E. & Hussain, A. 2015. *Sentic computing: a common-sense-based framework for concept-level sentiment analysis*. Vol. 1. Dordrecht: Springer.
- Carvalho, P. & Silva, M.J. 2015. Sentilex-pt: principais características e potencialidades. In: A. Simões, A. Barreiro, D. Santos, R. Sousa-Silva & S.E.O. Tagnin (Eds.). *Linguística, Informática e Tradução: Mundos que se Cruzam. Oslo Studies in Language*. 7(1): 425–438.
- Carvalho, P., Sarmiento, L., Silva, M. J., Oliveira, E. 2009. Clues for Detecting Irony in User-Generated Contents: Oh...!! It's "so easy" ,-. In: *Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion*. Hong Kong, China, 53-56.
- Charaudeau, P. 2006. Discours journalistique et positionnements énonciatifs. *Frontières et derives*. *Semen*, 22. 1-9. Retirado, a 20 de maio de 2017, da Internet: <https://journals.openedition.org/semen/2793>.
- Cunha, G. X. 2012. A articulação discursiva do gênero artigo de opinião à luz de um modelo modular de análise do discurso. *Filologia Linguística Portuguesa*. 14(1): 73-97.
- Das, S. & Chen, M. 2001. Yahoo! for Amazon: Extracting market sentiment from stock message boards. In: *Proceedings of the 8th Asia Pacific Finance Association Annual*

- Conference (APFA 2001)*, Bangkok, Thailand.
- Dave, K., Lawrence, S. & Pennock, D.M. 2003. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In: *Proceedings of International Conference on World Wide Web (WWW-2003)*.
- Ekman, P. 1999. Basic emotions. In: T. Dalgleish & M. Power (Eds.). *Handbook of cognition and emotion*. Chichester: John Wiley & Sons, 45-60.
- Ferreira, I. 2013. *Para o estudo semântico dos adjetivos adverbiais temporais e aspetuais do Português Europeu*. Tese de Doutoramento. Porto: FLUP.
- Fiorin, J.L. 2007. Paixões, afetos, emoções e sentimentos. *CASA: Cadernos de Semiótica Aplicada*. 5 (2), 1-15. Retirado, a 20 de setembro de 2017, da Internet: file:///C:/Users/Fatima%20Silva/Downloads/541-1486-1-PB%20(1).pdf.
- Forte, A.C. & Brazdil, P. 2016. Determining the Level of Clients' Dissatisfaction from Their Commentaries. In: J. Silva, R. Ribeiro, P. Quaresma, A. Adami & A. Branco (Eds.). *Computational Processing of the Portuguese Language. PROPOR 2016. Lecture Notes in Computer Science*. Vol. 9727. New York: Springer, 74-85.
- Freitas, C. (2013). Sobre a construção de um léxico da afetividade para o processamento computacional do português. *Revista Brasileira de Linguística*, 13 (4), 1031-1059.
- Goldberg, A. & Zhu, J. 2006. Seeing stars when there aren't many stars: Graph-based semisupervised learning for sentiment categorization. In *TextGraphs: HLT/NAACL Workshop on Graph-based Algorithms for Natural Language Processing*, 45-52.
- Gómez Molina, J. R. 2004. Las unidades léxicas en español. *Carabela*, 56: 27-50.
- Hung, C., & Lin, H-K. 2013. Using objective words in SentiWordNet to improve sentiment classification for word of mouth. *IEEE Intelligent Systems*. 28(2): 47-54.
- Hunston S, Thompson G. 2000. Evaluation: an introduction. In: S. Hunston & G. Thompson (Eds.). *Evaluation in Text: Authorial Distance and the Construction of Discourse*, Oxford: OUP. 1-27.
- Kodratoff, Y. & Michalski, R.S. 2014. *Machine learning: an artificial intelligence approach*. Vol. 3. Massachusetts: Morgan Kaufmann.
- Levin, B. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
- Li, G. & Liu, F. 2012. Application of a clustering method on sentiment analysis. *Journal of Information Science*. 38(2), 127-139.
- Liu, B. 2012. *Sentiment Analysis and Opinion Mining: Synthesis Lectures on Human Language Technologies*. California: Morgan & Claypool Publishers.
- Liu, B. 2015. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge:

- Cambridge University Press.
- Marques-Lucena, M., Sarraipa, J., Fonseca, J., Grilo, A., Jardim-Gonçalves, R. 2015. Framework for customers' sentiment analysis. In: P. Angelov, K.T. Atanassov, L. Doukowska, M. Hadjiski, V. Jotsov, J. Kacprzyk, N. Kasabov, S. Sotirov, E. Szmidt, S. Zadrozny (Eds.). *Intelligent systems'2014. Advances in Intelligent Systems and Computing*. Vol 322. Cham: Springer, 849-860.
- Marrafa, P. 2004. Extending WordNets to Implicit Information. In: M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa & R. Silva (Eds.). *Proceedings of LREC 2004 - International Conference on Language Resources and Evaluation*. Paris: ELRA - European Language Resources Association, 1135-1138 (CD-ROM).
- Martin, J.R. & White, P.R.R. (2005). *The Language of Evaluation*. New York: Palgrave.
- Mathieu, Y. 2005. Annotation of Emotions and Feelings in Texts. In: J. Tao, T. Tan & R. W. Picard (Eds.). *Affective Computing and Intelligent Interaction. First International Conference, AII 2005, Beijing, China, October 22-24, 2005, Proceedings*, 350-357.
- McCarthy, M. 1990. *Vocabulary*. Oxford: Oxford University Press.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. & Dean, J. 2013. Distributed representations of words and phrases and their compositionality. In: C. Burges, L. Bottou, M. Welling, Z. Ghahramani & K.Q. Weinberger (Eds.). *Advances in Neural Information Processing Systems 26 (NIPS 2013)*, 3111-3119.
- Nasukawa, T. & Yi, J. 2003. Sentiment analysis: Capturing favorability using natural language processing. In: *Proceedings of the 2nd international conference on Knowledge capture*. New York: ACM, 70-77.
- Nattinger, J. R., & DeCarrico, J. S. 1992. *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Neviarouskaya, A., Prendinger, H. & Ishizuka, M. (2009). Semantically distinct verb classes involved in sentiment analysis. In: Weghorn, H. & Isaias, P. (Eds.). *Proceedings of the IADIS International Conference on Applied Computing*, 27-34. Retirado, a 28 junho de 2017, da Internet: <http://www.iadisportal.org/applied-computing-2009-proceedings>.
- Pang, B. & Lee, L. 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*. 2(1-2): 1-135. Retirado, a 28 junho de 2017, da Internet: <http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>.
- Pang, B., Lee, L., & Vaithyanathan, S. 2002. Thumbs up?: Sentiment classification using machine learning techniques. *EMNLP '02 - Proceedings of the ACL-02 conference on Empirical methods in natural language processing*. Vol.10, 79-86.

- Pennington, J., Socher, R. & Christopher, M. 2014. Glove: Global vectors for word representation. In: *Proceedings of the 2014 conference on empirical methods in natural language processing*. Stroudsburg: The Association for Computational Linguistics, 1532-1543.
- Polanyi, L. & Zaenen, A. 2006. Contextual valence shifters. In: J.G. Shanahan, Y. Qu, Yan & J. Wiebe (Eds.). *Computing attitude and affect in text: Theory and applications*. Dordrecht: Springer, 1-10.
- Ranchhod, E. M. M. 2003. O Lugar das Expressões Fixas na Gramática do Português. In: I. Castro & I. Duarte (Eds.). *Razões e Emoção. Miscelânea de estudos oferecida a Maria Helena Mira Mateus*. Lisboa: Imprensa Nacional - Casa da Moeda, 239-254.
- Ravi, K. & Ravi, V. 2015. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems*. 89: 14-46.
- Rodrigues, R. H. 2005. Os gêneros do discurso na perspectiva dialógica da linguagem: a abordagem de Bakhtin. In: J. L. Meurer, A. Bonini & D. M. Roth (Eds.). *Gêneros: teorias, métodos, debates*. São Paulo: Parábola Editorial, 154-183.
- Russell, J.A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology*. 39: 1161-1178.
- Sag, I.A., Baldwin, T., Bond, F., Copestake, A. & Flickinger D. 2002. Multiword Expressions: A Pain in the Neck for NLP. In: Gelbukh A. (Ed.) *Computational Linguistics and Intelligent Text Processing. CICLing 2002*. Lecture Notes in Computer Science. Vol. 2276. Berlin: Springer, 189-206.
- Silva, F, Leal, A., Silvano, P., Ferreira, I. & Oliveira, F. 2018. Crítica cinematográfica: análise linguístico-textual. In: J. Veloso, P. Silvano, J. Guimarães & R. Sousa e Silva (Eds.). *A linguística em diálogo: volume comemorativo dos 40 anos do Centro de Linguística da Universidade do Porto*. Porto: FLUP / CLUP, 431-458.
- Silva, F., Leal, A., Ferreira, I., Oliveira, F. & Silvano, P. 2015. Marcas linguísticas no texto de apreciação crítica. In *Literatura e Gramática: um diálogo infinito*. Lisboa: Associação de Professores de Português.
- Silva, M. J. & Team, R. 2011. *Notas sobre a realização e qualidade do twitómetro. Technical report*. Lisboa: FCUL/LASIGE.
- Silva, M. J., Carvalho, P. & Sarmiento, L. 2012. Building a Sentiment Lexicon for Social Judgement Mining. In: H. Caseli, A. Villavicencio, A. Teixeira & F. Perdigão. (Eds.). *International Conference on Computational Processing of the Portuguese Language (PROPOR)*. Berlin: Springer, 218-228.
- Silva, M. J., Carvalho, P., Costa, C. & Sarmiento, L. 2010. *Automatic Expansion of a Social*

- Judgment Lexicon for Sentiment Analysis*. Relatório Técnico: DI-FCUL-TR-2010-08. Lisboa: FCUL.
- Taboada, M. & Trnava, R. 2013. *Nonveridicality and Evaluation Theoretical, Computational and Corpus Approaches*. Brill Academic.
- Taboada, M. 2016. Sentiment Analysis: An Overview from Linguistics. *Annual Review of Linguistics* 2016. 2(1): 325-347.
- Taboada, M., Anthony, & Voll, K. 2006. Methods for creating semantic orientation dictionaries. In: N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odiijk & D. Tapias (Eds.). *Proceedings of Fifth International Conference on Language Resources and Evaluation (LREC 2006)*. Genoa, Italy, 427-432.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K. & Stede, M. 2011. Lexicon-Based Methods for Sentiment Analysis. *Association for Computational Linguistics*. 37(2): 267-307.
- Thornbury, S. 2007. *How to teach vocabulary*. Malaysia: Pearson, Longman.
- Tong, R. M. 2001. An operational system for detecting and tracking opinions in on-line discussion. In: *Working Notes of the SIGIR Workshop on Operational Text Classification*. New Orleans, 1-6.
- Turney, P. D. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. Philadelphia, 417-424.
- Witten, I. H. & Frank, E. 2016. *Data Mining: Practical machine learning tools and techniques*. San Francisco: Morgan Kaufmann.

Gestos com os pés: descrição de uma listagem*

Isabel Galhano

irodrig@letras.up.pt

Faculdade de Letras / Centro de Linguística da Universidade do Porto

ABSTRACT:

Following a linguistic approach of gesture studies, this paper presents a case that bears witness to the robustness of gesture. It describes a listing procedure, recurrent in discourse organization, which consists in the production of a sequence of units characterised by syntactic, prosodic and gestural parallelism. In this example the speaker (a victim of Thalidomide) has no upper limbs, but he gesticulates with his feet. The analysis of the modalities involved in the listing activity will show that feet gestures can assume the functions performed by manual gestures. The example points out not only the tight link between gesture and speech, but also the embodiment of the listing activity, resulting from the necessity to index ideas or objects to the real world.

KEY-WORDS: feet-gestures; listing gestures; gesture studies; multimodality in face-to-face interaction

RESUMO

Numa abordagem linguística dos estudos do gesto, apresenta-se um estudo de caso que testemunha a robustez do gesto. Será descrita uma listagem, uma atividade comum de organização do discurso, que consiste na enunciação de uma sequência de unidades caracterizadas por paralelismo sintático, prosódico e gestual. Neste exemplo, o falante (vítima da Talidomida) não tem membros superiores, mas gesticula com os pés. Através da análise das modalidades envolvidas na execução da listagem, mostra-se como os gestos dos pés assumem as funções desempenhadas pelos gestos manuais. O exemplo evidencia não só a forte ligação entre gesto e fala, mas também a corporização da atividade de listar, resultante da necessidade de indexar ideias ou objetos ao mundo real.

PALAVRAS-CHAVE: gestos dos pés; gestos de listar; estudos do gesto; multimodalidade na interação face a face

*Para o Álvaro.

Introdução

Quando falamos, gesticulamos. A informação que provém da boca é captada pelos ouvidos, a que parte das mãos é percebida pelos olhos. Ou seja, também ouvimos os gestos. Só não os ouvimos se fecharmos os olhos. Assim explica Godwin-Meadow (2003: 241) o entrelaçamento entre gesto e fala. Esta ligação intrínseca entre fala e gestos tem vindo a ser explorada em diferentes abordagens dentro da área dos estudos do gesto, em análises de registos de eventos comunicativos feitos em contextos naturais (Kendon, 2004; Streeck, 2009; de Ruyter, 2007) e em contextos experimentais (McNeill, 1992; Goldin-Meadow, 2003; Kita; Özyürek, 2003). De forma resumida, esta ligação justifica-se pelas seguintes razões: o gesto coordena-se com os elementos da fala semanticamente correlacionados (Kendon, 2004); o gesto reflete os padrões linguísticos (Özyürek; Kita, 1999; Kita; Özyürek, 2003); a sincronização entre gesto e fala depende dos parâmetros temporais da língua falada (McNeill; Duncan, 2000); o gesto aligeira a carga cognitiva do falante, poupando esforços que podem ser usados para outras tarefas, e fomenta a mudança a nível cognitivo (Goldin-Meadow, 2003: 184 segs.); apoia o falante na estruturação da sua mensagem e fornece uma visão das representações mentais, assim como a informação essencial sobre eventos concretos (Goldin-Meadow, 2003: 15-16); por conseguinte, favorece a atividade de percepção da mensagem pelo ouvinte/observador (Goldin-Meadow, 2003: 97 segs.); o gesto desempenha um papel essencial na aquisição da linguagem, porque a partir do momento em que as crianças começam a pronunciar as primeiras palavras isoladas, se une à fala, formando um só sistema caracterizado por coerência semântica e temporal (Goldin-Meadow, 2003: 208 segs.; 2016). A integração do gesto na produção verbal e a sua robustez estão também patentes na gesticulação dos cegos de nascença, que apresenta características semióticas idênticas às da gesticulação dos normovisuais (Goldin-Meadow, 2003: 141 segs.).

O que acontece, então, quando um indivíduo não tem membros superiores? Melhor, quando nunca teve membros superiores? Será que a modalidade alternativa aos gestos manuais – neste caso, os gestos executados com os pés – se coordena, de idêntico modo, com a fala e se encontra ligada ao sistema conceptual subjacente? Foram estas questões que coloquei ao

observar um indivíduo sem membros superiores que gesticulava com os pés. Os seus gestos espontâneos eram ações visíveis coordenadas com a fala, em que os pés substituíam de modo tão eficaz os membros superiores, que não causavam estranheza. Pedi a AC (o falante em questão) o seu consentimento para analisar os seus gestos e, tendo-o obtido, realizamos algumas gravações. Estou-lhe muito grata por ter proporcionado os conteúdos para a concretização deste estudo, que foi apresentado à academia pela primeira vez em 2007¹. Numa primeira análise dos registos feitos, pude verificar que os gestos executados com os membros inferiores se coordenam de facto com a fala e com outras modalidades cinésicas (como as expressões faciais e os movimentos do torso e da cabeça), comportam iconicidade e desempenham funções pragmáticas de forma semelhante aos gestos coverbais manuais, sobretudo aos mais recorrentes (Ladewig, 2014). Para a elaboração deste primeiro artigo, escolhi a sequência de uma listagem, uma estratégia discursiva comum, da qual sobressai a função pragmática do gesto na organização do discurso.

Os fundamentos teóricos da análise que aqui se apresenta provêm da área dos estudos do gesto (Mülleret al, 2013, 2014); a identificação e descrição dos gestos baseia-se em Kendon (2004; 2013), que sugere grupos de gestos de acordo com as suas características semióticas e funções pragmáticas que, por sua vez, implicam diferentes formas de relacionamento com os elementos da fala. Este assunto será brevemente apresentado na secção 1. Segue-se, na secção 2., a descrição do *corpus* e, na secção 3., a micro-análise das modalidades envolvidas na execução de uma listagem. Termina-se o trabalho com uma conclusão e informações breves relacionadas com aspetos neurológicos desta forma de gesticular.

1. Gestos coverbais

Já é do senso comum que gesto² e fala partilham um mesmo sistema

¹ Este estudo foi apresentado em 2007 na terceira conferência da ISGS (International Society for Gesture Studies) *Integrating Gestures*, Northwestern University, Chicago, USA, June 18–21, 2007) e, dez anos mais tarde, no Colóquio Internacional de Homenagem ao Professor Óscar Lopes, realizado a 6 e 7 de junho de 2017, na Faculdade de Letras da Universidade do Porto.

² Na linguagem comum, a palavra *gesto* é geralmente usada como sinónimo de gesto emblemático, um tipo específico de gesto espontâneo cujo significado e características formais estão convencionalizadas (Galvão, 2015: 552). Como termo científico da área dos estudos do gesto, a palavra refere a ação visível (sobretudo a que é executada

conceptual subjacente (Müller, 2013; Streeck, 2009) e que certos aspetos de uma mensagem não são exclusivamente determinados pelo falante/ator, mas também condicionados pelas respetivas experiências corporais no contexto envolvente (Enfield, 2011; Streeck, 2009). Experiências corporais, hábitos motores, manipulação de objetos e interação social são fatores determinantes para a conceptualização de dimensões abstratas, para a sua corporização e para o desenvolvimento da linguagem (Alibali, 2014). É através das ações sensório motoras que se desenvolvem capacidades cognitivas, como a perceção e categorização, sendo os gestos um tipo específico de ações físicas num contexto social estreitamente ligadas ao desenvolvimento cognitivo da linguagem. Outras aptidões que se desenvolvem na interação social e da experiência de comunicar através das modalidades corporais disponíveis (como a fala, gestos manuais, movimentos da cabeça, orientação do olhar, expressões faciais e postura) são as que possibilitam a compreensão mútua ou o desenvolvimento de sensibilidade interpessoal. Ou seja, as aptidões de um indivíduo não só para produzir pistas de contextualização através das modalidades cinésicas e de variações prosódicas correlacionadas com as palavras dos enunciados de acordo com as suas intenções comunicativas, mas também para, simultaneamente, percecionar as intenções comunicativas do seu interlocutor através das pistas de contextualização que este fornece (Gumperz, 1992).

1.1. Como o gesto se relaciona com a fala

Nem todos os gestos manuais são executados com a intenção de transmitir uma mensagem (com ou sem fala). Alguns resultam da expressão espontânea de emoções e atitudes, não fazendo parte da intenção comunicativa de um falante – como é o caso dos *Butterworths*, gestos que emergem em momentos de falhas na produção verbal, expressando o esforço cognitivo na estruturação do enunciado ou na procura de palavras (McNeill, 1992: 77), podendo ser designados por pausas cheias, mesmo não sendo acompanhados por vocalizações (Galhano-Rodrigues, 2007: 185). Outros gestos estão diretamente envolvidos na produção do enunciado e nele integrados,

com os membros superiores) que participa na produção do enunciado e é parte dele, ou seja, a gesticulação que acompanha a produção verbal e se relaciona com o conteúdo do enunciado (Kendon 2004; McNeill, 1992).

relacionando-se com o conteúdo ou com a sua estrutura. Por este motivo, são designados por co-verbais. Segundo a proposta de Kendon (2013) de agrupar os gestos conforme a sua relação semântica e/ou pragmática com a fala, os gestos distinguem-se por terem prioritariamente ou propriedades referenciais, ou pragmáticas: os primeiros estão semanticamente relacionados com o significado referencial ou proposicional do enunciado, ou com características semânticas do enunciado que não são parte do significado referencial; os segundos não expressam significados correlacionados com objetos, ações, ou referências a lugares. O grupo dos gestos referenciais abrange os emblemáticos (de significado mais convencionalizado, podendo substituir a fala), os deícticos (gestos de apontar) e os descritivos (que representam, desenham ou modelam características semânticas do referente) (Kendon, 2004: 158, 2013). No desempenho de funções pragmáticas³ destacam-se os gestos estruturantes ou batuta (*beats*, McNeill, 1992; *batonic movements* ou *parsing gestures*, Kendon, 2013), que marcam o ritmo e enfatizam partes do enunciado, os gestos operadores, que acrescentam informação sobre a atitude do falante relativamente ao conteúdo do enunciado, os gestos performativos, ligados a um valor ilocutivo do enunciado (de pedir, perguntar, interrogar, censurar ou de dar uma ordem), os gestos modais (ou metacomunicativos), que dão informações sobre o conteúdo ou sobre o referente do enunciado, e os gestos interativos, que organizam o discurso e a alternância de vez (Kendon, 2013). Atendendo às propriedades polifuncionais dos gestos, qualquer gesto pode ser combinado (*blend*) quando as suas propriedades (por exemplo, a configuração da mão e o percurso do golpe) permitem que seja inserido em mais do que um só grupo (Kendon, 2004; Galhano-Rodrigues, 2012; Galvão, 2015). Muitos gestos com funções pragmáticas também têm propriedades icónicas, mais ou menos acentuadas, que parecem derivar de experiências corporais no mundo físico. A recorrência de alguns destes gestos faz com que estejam agrupados de acordo com características formais dominantes. Um exemplo são os gestos que Streeck (2009) descreve como *Open Hand Supine* (gesto de mão aberta, palma orientada para cima) um resquício de “*ceivings*”, formas de pren-

³ Note-se que os gestos com funções pragmáticas (que não são o mesmo que gestos pragmáticos) podem relacionar-se de vários modos com as características do significado do enunciado que não são parte do seu significado referencial ou conteúdo proposicional (Kendon, 2004: 158; Payrató; Teßendorf, 2014: 1532).

der, apreender e de conceber. São gestos em que as mãos ajudam o falante a pensar e a apreender conceitos, com base no seu conhecimento do mundo (Streeck, 2009: 151). Kendon refere-se a famílias de gestos, como os *G-gestures*, gestos em forma de pinça (*grappolo*) (Kendon, 1995, 2004: 231), ou de mão aberta palma orientada para cima (*Open Hand Supine gestures*), ou de mão aberta palma para baixo (*Open Hand Prone gestures*) (Kendon, 2004: 248 segs.; Müller, 2004), os primeiros com o significado de apresentar alguma coisa, partilhando-a convidando à atenção conjunta (Müller, 2004: 252), os segundos com o significado nuclear de cortar uma linha de ação (Kendon, 2014: 282). Estes gestos são entendidos como resultando de experiências físicas no mundo envolvente e encontram-se corporizados em representações mentais (Zatlev, 2013). Apresentam um elevado grau de convencionalidade (Kendon 2014: 282), como o caso do gesto de mão aberta palma para cima, que, acompanhado por um encolher dos ombros e por uma expressão facial específica, já tem uma representação iconográfica “globalizada” com o significado de “não sei”: $\backslash _ (\Psi) _ /$. Também os gestos de listar, recorrentes na atividade de enumerar vários elementos ou ideias, refletem experiências corporais com objetos concretos. São descritos como corporizando a necessidade natural e espontânea de ligar as ideias ao mundo real (Goldin-Meadow, 2003: 164), reduzindo a carga cognitiva.

No *corpus* em análise, detetou-se um momento em que o falante realiza uma listagem, expressa não só por elementos verbais, mas também por ações dos membros inferiores (sobretudo do pé direito). Com o objetivo de averiguar se, na atividade de listar, as ações dos pés se comportam de forma idêntica às ações das mãos, resolvi analisar detalhadamente este fenómeno, que será descrito no capítulo 2. A secção seguinte fornece uma descrição mais específica sobre a forma como gesto e fala estão envolvidos nesta atividade organizadora do discurso.

1.2. Gestos de listar

A enunciação de uma lista de itens (objetos, ideias, argumentos, etc.) caracteriza-se por paralelismo sintático, prosódico e cinésico (i.e., relativo aos movimentos do corpo, gestos, olhar, movimentos da cabeça) (Galhano-Rodrigues, 2015: 255 segs). Cada parte de uma lista é composta por ele-

mentos linguísticos (palavras, sequências de palavras ou frases mais longas) de estrutura sintática idêntica; constitui uma unidade entoacional, que se caracteriza por um padrão prosódico específico, de altura de tom ascendente no contorno final. A proeminência prosódica de cada parte da lista (de maior altura de tom ou maior intensidade da voz) coincide com a verbalização do item listado, que, tendencialmente, é o referente mais importante. Os gestos envolvidos nesta atividade, designados por “gestos de listar” (Galhano-Rodrigues, 2015: 255) fornecem as pistas visíveis sobre o enunciado. Os seus golpes acompanham o ritmo do enunciado verbalizado e apresentam características específicas que seguem um sistema próprio. No caso do português europeu, de outras variedades da língua portuguesa, assim como de outras culturas do sul da Europa, acontece que, em cada golpe, os dedos de uma mão são sucessivamente tocados, ou pressionados pelo dedo indicador da outra mão, começando pelo dedo mínimo e terminando com o polegar. A partir do momento em que o falante executa o primeiro gesto de listar, o ouvinte infere o tipo de atividade que vai ser realizada e prepara-se para ouvir uma sequência limitada de elementos. Além de fornecer pistas de interpretação ao ouvinte/observador, estes gestos apoiam o falante na organização do discurso, aligeirando o esforço cognitivo de ter mentalmente presentes os elementos da lista que quer comunicar ao ouvinte/observador. Muitas vezes, cada gesto de listar é acompanhado por uma inclinação da cabeça. Ambas as modalidades coincidem temporalmente com os momentos de proeminência prosódica (Galhano-Rodrigues, 2015: 272). Deste modo, cada elemento constitutivo (ou parte) da listagem está marcado, de forma idêntica, através de modalidades de diversa natureza. Numa perspetiva do contexto interacional, considerando ainda a perceção do enunciado pelo ouvinte/observador, a repetição de cada composto multimodal da lista pelo falante/ator confere ao enunciado um ritmo que, por sua vez, cria expectativas no seu interlocutor, orientando a sua atenção para o que vai ser dito a seguir (Auer; Couper-Kuhlen, 1994: 82 segs.; Galhano-Rodrigues, 2007: 175). Além disso, esta marcação tem um efeito coesivo sobre a totalidade dos elementos constitutivos da lista, mostrando ao ouvinte/observador as partes que a compõem. Por outras palavras, estrutura frásica, características prosódicas e movimentos do corpo co-verbais funcionam como dispositivos de contextualização, através dos quais o falante informa o(s) parceiro(s) da interação sobre

o tipo de atividade que realiza.

2. Sujeito analisado, corpus e método de análise

O corpus em análise consiste numa passagem com a duração de 00:00:52, selecionada do seguinte registo vídeo:

ref.	AC/2/2007	formato/ Mb	Mpeg / 1440x1080 / 48.000Mz // 120 MB
data	4-04-2007	câmara	Sony handycam DCR-PC110 mini DV
local	Porto; res. AC	duração	34'57"
part.	AC; IG		

O sujeito em questão, AC, é vítima da Talidomida, um produto da indústria farmacêutica, que, nos anos 50, era prescrito às mulheres grávidas para combater enjoos matinais, causando a síndrome de focomelia, um problema congénito que impede a formação dos membros superiores e inferiores. AC tem uma escoliose forte, uma formação deficiente dos membros inferiores, sendo estes de diferentes dimensões, e não tem braços nem mãos. Como alternativa aos membros superiores, usa os inferiores na execução de diversas atividades, como escrever no teclado do computador, manipular objetos e “gesticular”. Na produção do enunciado executa ações visíveis com os pés, envolvendo ainda movimentos da cabeça, sobancelhas, ombros e a orientação do olhar. À semelhança dos gestos manuais de indivíduos com braços, a modalidade mais diretamente ligada às ideias expressas na fala e com elas coordenada é a movimentação dos pés e dos dedos dos pés, com os quais executa movimentos amplos, em diferentes planos espaciais.

A análise foi feita com o software de anotação de vídeo e áudio ELAN (Slotjes; Wittenburg, 2008). Para a análise do sinal acústico, recorreu-se ao software PRAAT (2008) e para a transcrição prosódica da fala, a uma versão simplificada do sistema de transcrição prosódica GAT (Selting et al., 1998) adaptado ao português (Galhano-Rodrigues, 2007). As unidades cinésicas analisadas foram os movimentos dos membros inferiores, da cabeça e dos ombros, assim como a orientação do olhar correlacionados com a fala; na

descrição dos movimentos dos pés considerou-se o pé direito, o pé mais ativo, visto o pé esquerdo não ter a mesma mobilidade. As unidades de análise foram classificadas de acordo com a sua pertença aos grupos de gestos descritos sob 2.

3. Microanálise da listagem

Como referido, o segmento escolhido para a presente análise foi uma breve atividade de listagem (linhas 14–27). Para dar maior transparência a esta análise, apresenta-se a respetiva transcrição prosódica, correspondendo cada linha a uma unidade entoacional (Galhano-Rodrigues, 2007: 222). Na transcrição inclui-se a descrição dos movimentos coverbais, os mais relevantes ilustrados através dos fotogramas 1-9. A descrição da sequência divide-se em dois subcapítulos: o primeiro correspondente a uma fase preparatória de estruturação do discurso, o segundo à própria listagem.

3.1 - Estruturação do discurso: hesitação

14. -isto `é - eh : : : -

|_____||

|
pé apoiado no chão, ergue e volta a baixar os dedos, contraídos, um por um, começando com o dedo grande e depois no sentido inverso – dois movimentos semi-circulares; desvia olhar para cima e move ligeiramente a cabeça para a esquerda

15. (inspiração profunda) - eh : : `eh `eh

|_____||_____||

|
Mantém mantém a posição do pé, assim como a orientação do olhar e da cabeça

|
mantendo configuração do pé, repete movimentos precedentes, de um lado para o outro, elevando e baixando os dedos dos pés, um por um; orienta a cabeça e o olhar para a frente

Preparação para a listagem (fotograma 1): à execução de alguns movimentos de sobe-e-desce alternado com os dedos dos pés contraídos segue-se uma paragem (congelamento do gesto), em que o pé se mantém elevado, com o dedo maior retraído; simultaneamente, o falante orienta o olhar para cima/lado. As características destas modalidades (congelamento do gesto e orientação do olhar⁴) são indicadores recorrentes de esforço cognitivo, mostram que o falante se encontra num processo de estruturação do discurso (Galhano-Rodrigues, 2007: 255). A forma como movimenta todos os dedos dos pés antes da paragem tem certa iconicidade, porque inspira a ideia de pluralidade, certamente correlacionada com as ideias que precisam de ser organizadas.

3.2 - Listagem

16. - aslimita`ção :: s :: - `eh :: :

|_____||
|

orienta olhar para cima/frente, depois para cima/esq.; ergue ligeiramente sobrançelas vira ligeiramente cabeça para dir, apoia pé dir. no chão, dedo grande mais estirado que os outros (acompanhando a verbalização de "limitações") – 1º elemento da listagem; seguem-se leves acenos com cabeça

17. de `de ´de -

|_____||
|

vira cabeça para cima, ligeiramente voltada para dir.; orienta olhar para cima; sobrançelas erguidas; mantém o pé apoiado no chão, dedo grande ligeiramente elevado

⁴ A orientação do olhar para cima-lado verifica-se também num segundo momento de hesitação, após a verbalização da primeira unidade da listagem (ver fotograma 3, linhas 17 e 18 da transcrição), revelando um momento idêntico de estruturação do discurso.

18. ((inspiração)) - a FAL`ta - de=experiÊN : : : cia -

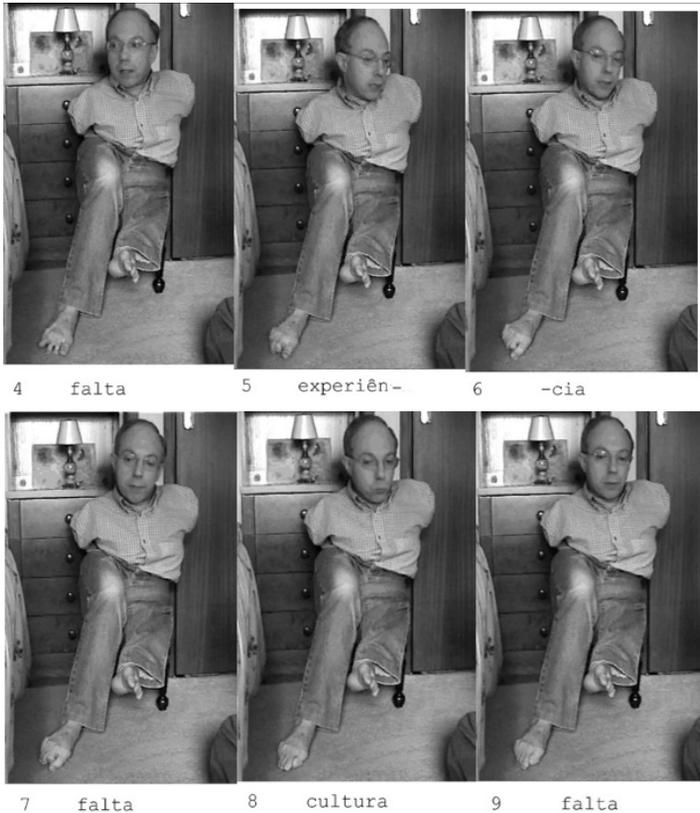
_____	_____	_____
vira cabeça para frente; mantém orientação do olhar; ergue pé dir. e estira dedo grande	começa a virar cabeça para dir/baixo; toca chão com 2º dedo (2º elemento da listagem)	vira cabeça para esquerda e encolhe ombro esquerdo; eleva o pé e começa a começa a baixá-lo de novo; orientação do olhar para dir/baixo toca chão com 2º dedo (2º elemento da listagem)

19. - afALta de – cultU : : : `ra = <<a> a falta’

_____	_____	_____
cabeça para direita, inclina-a para esq/baixo; toca o chão com dedo grande e desliza o 3º dedo sobre o chão, da frente para trás (3º item da listagem)	ergue de novo a cabeça; ergue ligeiramente o pé; orienta olhar para baixo	toca levemente o chão com os 4º e 5º dedos – estirando-os para a frente e fletindo-os a seguir para trás (esboço de um 4º elemento de listagem)



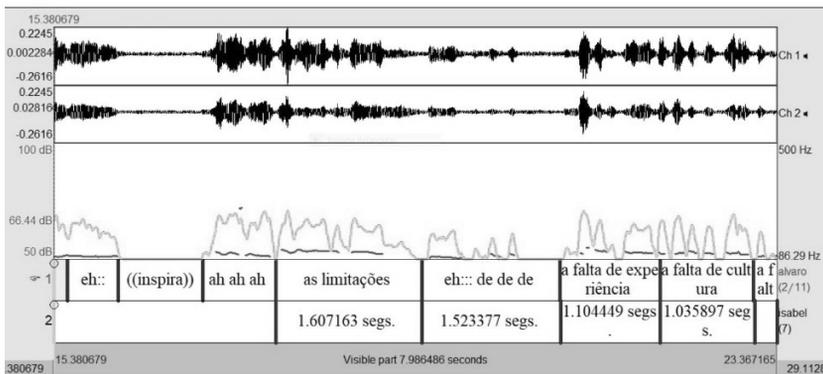
1 eh:: 2 limitações 3 ((inspira))



A sequência transcrita nas linhas 16-19 corresponde à listagem de três argumentos, que vão justificar a opinião do falante, anteriormente expressa, sobre atitudes de determinados indivíduos. Os argumentos são “as limitações”, “a falta de experiência” e “a falta de cultura”; o falante verbaliza ainda o início de outro argumento - “a falta” – que acaba por não explicitar, porque interrompe o discurso. Esta interrupção auto-iniciada pode-se considerar a primeira parte de uma reparação (Galhano-Rodrigues, 2007: 263 segs.). A listagem está contextualizada através de a) paralelismo frásico (repetição da expressão “a falta de” no segundo e terceiro ato de listar), b) paralelismo prosódico e c) paralelismo cinésico (através dos gestos de listar e de movimentos da cabeça). Na representação do sinal acústico (FIGURA 1), pode-se constatar que todas as unidades da listagem apresentam parâmetros

idênticos de altura de tom (globalmente descendente) e de intensidade. A primeira unidade da lista, assim como a fase de hesitação subsequente apresentam uma duração quase idêntica, respetivamente de 1.607 e de 1.523 segundos. Os dois atos de listar que se seguem têm uma menor duração, a saber, de 1.104449 e de 1.035897 segundos. É provável que o retardo causado pela hesitação tivesse implicado uma necessidade de recuperar o ritmo iniciado. A menor duração destes dois atos pode ser interpretada como uma forma de compensação da fase de hesitação transcrita nas linhas 17 e 18 (Galhano-Rodrigues, 2007: 688). No que diz respeito aos gestos de listar com os dedos dos pés, constata-se que, à semelhança dos gestos manuais de listagem, cada elemento é marcado por um toque com um dedo diferente (Galhano-Rodrigues, 2015). Não havendo a possibilidade de tocar os dedos com um dedo do outro membro, o falante usa a superfície do chão.

FIGURA 1: Imagem do sinal acústico da sequência (com PRAAT)



Assim sendo, o toque é concretizado pelo contacto do dedo em extensão com o chão, seguido de uma retração, em que o dedo desliza sobre a superfície do solo, da frente para trás, e se eleva ligeiramente a seguir. Devido à diferente mobilidade de cada dedo, este movimento nem sempre é evidente, mas, numa observação cuidada, verifica-se o seguinte: no primeiro elemento da listagem, o dedo ativo é o grande (FOTOGRAMA 2); no segundo elemento, é o segundo dedo que se movimenta para trás e para a frente (FOTOGRAMAS 4-6); na verbalização do terceiro elemento, o terceiro e

quarto dedos são pressionados e ligeiramente elevados (parte-se do princípio de que por uma questão de mobilidade, não é possível que estes dedos se movimentem independentemente) (FOTOGRAMAS 7 e 8). Constata-se ainda que a produção de cada um destes gestos está coordenada com uma ligeira inclinação da cabeça para o lado esquerdo e para baixo. Ao fim do terceiro elemento da listagem, o falante interrompe a atividade de listar para fazer um comentário sobre o último tópico referido, o conceito de cultura (um tipo de interrupção também muito frequente nas listagens). Recorre, para isso, a uma estratégia de autorreparação, constituída pela interrupção da frase iniciada e pela introdução de um aparte, que passa a ser o tema principal, visto o falante não dar continuidade ao tema principal que antecede a interrupção, ou seja, à atividade de listagem.

4. Conclusão

A análise desta sequência mostra que os gestos dos pés podem assumir inteiramente uma função pragmática recorrente dos gestos manuais. À semelhança do que acontece com os indivíduos com braços, a atividade de listagem envolve um paralelismo de características sintáticas, prosódicas e gestuais.

Verifica-se ainda que, apesar dos condicionamentos físicos do falante no que diz respeito à capacidade (mais limitada) de representar gestualmente formas, ele consegue transmitir alguma iconicidade através de combinações de posições do pé, movimentos ágeis dos dedos do pé e percurso do golpe do gesto⁵. Desde já, e apesar da apresentação de apenas um exemplo, parece possível afirmar que a) os gestos dos pés são suscetíveis de desempenhar as funções pragmáticas dos gestos dos membros superiores e b) a atividade de listar parece estar fortemente corporizada, de tal forma que não ficou omitida nesta modalidade “gestual” alternativa.

⁵ A iconicidade dos gestos dos pés não foi tratada neste estudo, mas merece ser aprofundada num trabalho posterior.

5. Da robustez do gesto fantasma

Goldin-Meadow(2003; 2007⁶) refere-se aos gestos dos pés, assim como a gestos fantasma dos membros superiores como sendo testemunhos da robustez do gesto, quer pela sua interligação forte com a fala desde o processo da aquisição da linguagem, quer pela sua presença constante na oralidade. Sobre o caso da gesticulação de Mirabelle, uma jovem que nasceu sem braços, mas que podia sentir os chamados “membros fantasma” (*phantom limbs*) a autora escreve: “[Gesture] comes with talk (...) You don’t have to be taught to gesture. Strangely you don’t even have to have arms to feel yourself gesturing” (Goldin-Meadow, 2003: 241).

Ramachandran e Balkeslee (1998) transcrevem a descrição que Mirabelle faz dos seus gestos fantasma:

Well, because as I’m talking to you, they are gesticulating. They point to objects when I point to things, just like your arms and hands... When I walk, doctor, my phantom arms don’t swing like normal arms, like your arms. They stay frozen on the side, like this. “She stood up, letting her stumps drop straight down on both sides. “But when I talk, “she said, “my phantoms gesticulate. In fact, they’re moving now as I speak (Ramachandran & Balkeslee, 1998: 41).

Depois de contactar Vilayanur Ramachandran sobre este assunto, recebi uma lista de questões que deveria colocar a AC: *Does he wave good-bye with his feet? Does he balance his arms when he moves? How long are his phantom arms?*⁷.

Como Mirabelle, AC confirmou sentir impulsos motores em membros fantasma, confirmou sentir membros fantasma a balançar quando andava.

Juntamente com o que se demonstra nesta análise, estas informações não só evidenciam a capacidade adaptativa e a flexibilidade do cérebro humano, mas também comprovam a robustez do gesto na corporização da linguagem. Trata-se de fenómenos que merecem ser mais explorados em abordagens interdisciplinares.

⁶ Comunicação pessoal, 18 de junho de 2007

⁷ Comunicação pessoal de V. Ramachandran por correio eletrónico, 14.01.2009

REFERÊNCIAS

- Alibali, M. 2014. Embodiment – the body and its role for cognition, emotion, and communication. In: C. Müller; A. Cienki; E. Fricke; S. H. Ladewig; D. McNeill; S. Teßendorf (Eds.). *Body–language–communication: An international handbook on multimodality in human interaction* Berlin: De Gruyter Mouton, 1833-1840.
- Auer, P.; Couper-Kuhlen, E. 1994. Rhythmus und Tempo konversationeller Alltagsprache. *Zeitschrift für Literaturwissenschaft und Linguistik*, 96, 78-106.
- Boersma, P.; Weenink, D. 2008. *Praat: doing phonetics by computer*. [Computer Program] Version 6.0.37, retrieved 14 March 2018 from <http://www.praat.org/>
- de Ruiter, J.P. 2007. Postcards of the mind: the relationship between speech, imagistic gesture and thought. *Gesture*. 7 (1), 21-38.
- ELAN, the software annotation tool. Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands. Open source software. <https://tla.mpi.nl/tools/tla-tools/elan/> (acedido a 5 de abril de 2018).
- Enfield, N.J. 2011. Elements of Formulation. In: J. Streeck; C. Goodwin; C. LeBaron (Eds.). *Embodied Interaction. Language and Body in the Material World*. Cambridge: Cambridge University Press, 59-66.
- Enfield, N.J.; Kita, S.; de Ruiter, J.P. 2007. Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics*. 40, 1722–1741.
- Galhano-Rodrigues, I. 2007. *O corpo e a fala*. Lisboa: FCG/FCT.
- Galhano-Rodrigues, I. 2012. Vou buscar ali, ali acima! A multimodalidade da deixis no português europeu. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto*. 7, 129-164.
- Galhano-Rodrigues, I. 2015. A tool at hand - gestures and rhythm in listing events: case studies of European and African Portuguese speakers. *OSLa: Oslo Studies in Language*. 7 (1), 253-281.
- Galvão, E.Z. 2015. Gestures. In: F Pöschhacker (Ed.) *Routledge encyclopedia of interpreting studies*. London: Routledge, 552.
- Goldin-Meadow, S. 2003. *Hearing Gesture. How hands help us think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S. 2013. How our gestures help us learn? In: C. Müller; A. Cienki; E. Fricke; S.H. Ladewig; D. McNeill; S. Teßendorf (Eds.). *Body–language–communication: An international handbook on multimodality in human interaction*. Berlin: De

- Gruyter Mouton, 792-803.
- Goldin-Meadow, S. 2017. What the hands can tell us about language emergence. *Psychonomic Bulletin and Review*. 24 (1), 213–218. DOI: 10.3758/s13423-016-1074-x.
- Gumperz, J. 1992 Contextualization and understanding. In: A. Duranti A.; C. Goodwin (Eds.). *Rethinking context: Language as an interactive phenomenon*. Cambridge: Cambridge University Press, 229-252.
- Kendon, A. 2004. *Gesture. Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, A. 2013. Exploring the utterance roles of visible bodily action: A personal account. In: C. Müller; A. Cienki; E. Fricke; S. Ladewig; D. McNeill; S. Teßendorf (Eds.). *Body, Language, Communication: An international handbook on multimodality in human interaction*. Berlin: Walter de Gruyter, 7-28.
- Kita, S.; Özyürek, A. 2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*. 48, 16–32.
- Ladewig, S. 2014. Recurrent gestures. In: C. Müller; A. Cienki; E. Fricke; S. Ladewig; D. McNeill; J. Bressems (Eds.) *Body – Language – Communication: An international handbook on multimodality in human interaction*. Berlin: de Gruyter, 1558–1574.
- Levinson, S.C.; Holler, J. 2014. The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*. DOI:10.1098/rstb.2013.0302.
- McNeill, D. 1992. *Hand and Mind*. Cambridge: Cambridge University Press.
- McNeill, D. & Duncan, S.D. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and Gesture*. Cambridge: Cambridge University Press, 141-161.
- Müller, C. 2013. *Introduction*. In: C. Müller; A. Cienki; E. Fricke; S. Ladewig; D. McNeill; S. Teßendorf (Eds.). *Body, Language, Communication: An international handbook on multimodality in human interaction*. Berlin: Walter de Gruyter, 1-6.
- Müller, C. 2004. Forms and uses of the Palm Up Open Hand: A case of a gesture family? In: C. Müller; R. Posner (Eds.) *The semantics and pragmatics of everyday gestures. Proceedings of the Berlin conference, April 1998*. Berlin: Weidler Buchverlag, 233-256.
- Müller, C; Cienki, A.; Fricke, E.; Ladewig, S.; McNeill, D.; Teßendorf, S. (Eds.) 2013. *Body - Language - Communication: An international handbook on multimodality in human interaction*. (HSK 38.2). Vol. 1. Berlin: Mouton de Gruyter.
- Müller, C.; Ladewig, S.; Cienki, A.; Fricke, E.; McNeill, D.; Bressems, J. (Eds.) 2014. *Body -*

- Language – Communication: An international handbook on multimodality in human interaction.* (HSK 38.2). Vol. 2. Berlin: Mouton de Gruyter.
- Özyürek, A.; Kita, S. 1999. Expressing manner and path in English and Turkish: Differences in speech, gesture, and conceptualization. In: M. Hahn; S. C. Stoness (Eds.). *Proceedings of the Twenty-first Annual Conference of the Cognitive Science Society.* London: Erlbaum, 507-512.
- Payrató, L.; Tessendorf, S. 2014. Pragmatic gestures. In: C.Müller; A. Cienki; E. Fricke; S. Ladewig; D. McNeill; J. Bressemer (Eds.) *Body – Language – Communication. Berlin: de Gruyter: An international handbook on multimodality in human interaction.* Berlin: De Gruyter Mouton, 1531-1539.
- Ramachandran, V.; Blakeslee, S. 1998. *Phantoms in the Brain.* New York: Quill.
- Selting, M.; Auer, P.; Barden, B.; Bergman, J.; Couper-Kuhlen, E.; Günthner, S.; Meier, C.; Quasthoff, U.; Schlobinski, P.; Uhmann, S. 1998. Gesprächsanalytisches Transkriptionssystem (GAT). *Linguistische Berichte.* 173, 91-122.
- Sloetjes, H.; Wittenburg, P. 2008. Annotation by category – ELAN and ISO DCR. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008).*
- Streeck, J. 2009. *Gesturecraft. The manufacture of meaning.* Amsterdam: John Benjamins.
- Streeck, J. 2013. Praxeology of gesture. In: C. Müller; A.Cienki; E. Fricke; S. Ladewig; D. McNeill; S. Teßendorf (Eds.). *Body, Language, Communication: An international handbook on multimodality in human interaction.* Berlin: Walter de Gruyter, 674-688.
- Zlatev, J. 2013. Levels of embodiment and communication. In: C. Müller; A. Cienki; E. Fricke; S. Ladewig; D. McNeill; S. Teßendorf (Eds.). *Body, Language, Communication: An international handbook on multimodality in human interaction.* Berlin: Walter de Gruyter, 533-550.

Bibliografia
de Óscar Lopes

Bibliografia sobre Gramática e Linguística

- Lopes, Óscar. 1943. Equívocos do Velho Humanismo. *Revista da Faculdade de Letras de Lisboa*, X.
- Lopes, Óscar. 1944. Preliminares a uma Gramática. *Gazeta de Filosofia*. 3: 12-23.
- Lopes, Óscar. 1947. A Gramática Clássica e a Lógica Moderna. *Seara Nova*. 1031: 7-8.
- Lopes, Óscar. 1950. Humanismo e Linguística. *Vértice – Revista de Cultura e Arte*. Coimbra: Sociedade Editora Vértice, 81-84.
- Lopes, Óscar. 1958. Lógica Gramatical e Lógica Simbólica. *Labor*. 179: 69-82.
- Lopes, Óscar. 1967. Gramática escolar, reduto de um anacronismo. *I Simpósio Luso-Brasileiro sobre a Língua Portuguesa Contemporânea*.
- Lopes, Óscar. 1967. Possibilidades de Coordenação entre o Ensino do Português e o da Matemática. *Encontro Luso-Brasileiro de Linguística Aplicada*. Coimbra.
- Lopes, Óscar. 1968. Para a modernização da análise gramatical. *Boletim Bibliográfico e Informativo*, nº 7. Lisboa: Fundação Calouste Gulbenkian.
- Lopes, Óscar. 1970. Para a coordenação necessária entre o Português e a Matemática. *Cadernos do Centro de Investigação Pedagógica*. Lisboa: Fundação Calouste Gulbenkian.
- Lopes, Óscar. 1971. Gramática e Lógica Simbólica. *I Simpósio Luso-Brasileiro sobre a Língua Portuguesa Contemporânea*. Coimbra: Ministério da Educação Nacional, Instituto de Alta Cultura.
- Lopes, Óscar. 1971. Perspetivas de uma Gramática Simbólica. *VI Congresso do Ensino Liceal*, Aveiro. Separata: 7 páginas.
- Lopes, Óscar. 1971. *Gramática Simbólica do Português – um esboço, 1ª edição*. Lisboa: Fundação Calouste Gulbenkian.
- Lopes, Óscar. 1972. *Gramática Simbólica do Português – um esboço, 2ª edição*. Lisboa: Fundação Calouste Gulbenkian.
- Lopes, Óscar. 1976. Necessidade de um Mínimo de Consciência Formal da Língua na Prática Escolar. In: *Actas do 1.º Encontro Nacional para o Ensino e a Investigação do Português*. Águeda, 377-390.

- Lopes, Óscar. 1977. Topologias da Deixis em Português. *XV Congresso Internacional de Língua e Filologia Românica*. Rio de Janeiro. In: Lopes, Ó. 2005, 121-150.
- Lopes, Óscar. 1979. A Materialidade da Linguagem. In: "Actas" do *Seminário sobre Ciência e Filosofia*. Porto: Faculdade de Letras da Universidade do Porto.
- Lopes, Óscar. 1980. Sobre a Semântica dos Nomes Massivos. *XVI Congresso Internacional de Linguística e Filologia*. Palma de Maiorca. In: Lopes, Ó. 2005, 41-54.
- Lopes, Óscar. 1983. Algumas Peculiaridades do Português, e Especialmente do Português Europeu, que importam à Teoria da Semântica Linguística Universal. *Actas do Congresso sobre "A Situação Actual da Língua Portuguesa no Mundo"*. Lisboa: Instituto de Cultura e Língua Portuguesa, vol I, 85-104. Também publicado em: Lopes, Ó. 1986. *Uma Arte da Música*. Porto: Oficina Musical, 107-134 e Lopes, Ó. 2005, 15-40.
- Lopes, Óscar. 1983. Sobre as Contrastivas em Português. *Actes du XVII Congrès International de Linguistique et Philologie Romanes*. Aix-en-Provence. Vol. IV: 544-554. Também em: Lopes, Ó. 2005, 177-192.
- Lopes, Óscar. 1984. Perspectivas de Formalização Semântico-Linguística. *Actas do I Encontro de Linguistas Portugueses*. Lisboa, 216-219.
- Lopes, Óscar. 1985. A noção de definido e a de presente de enunciação. *Actas do 1.º Encontro da Associação Portuguesa de Linguística*. Actas. Lisboa: APL, 129-143. Também em Lopes, Ó. 2005, 105-119.
- Lopes, Óscar. 1985. Para um Conceito Dístico de Presente e de Presença. *Encontro sobre a Teoria do Texto*. Évora: Universidade de Évora. In: Lopes, Ó. 2005, 87-104.
- Lopes, Óscar. 1987. Os Modos nas Construções Condicionais e Concessivas em Português. *Conferência no Instituto de Filologia Românica*. Berlim: Universidade Livre de Berlim.
- Lopes, Óscar. 1987/1991. Observações sobre os Actualizadores em Português. *Actas do Encontro em Homenagem a Óscar Lopes*. APL. Porto, 17-41.
- Lopes, Óscar. 1987. O Acordo Ortográfico. In: Castro, I.; Duarte I.; Leiria, I. (Orgs.). *A Demanda da Ortografia Portuguesa*. Lisboa: João Sá da Costa, 129-133.
- Lopes, Óscar. 1989/ 1997. Construções Concessivas – Reflexões Semântico-Pragmáticas. In Lorenzo, R. (org.) *Actas do XIX Congreso Internacional de Lingüística e Filoloxia Románicas*. Tomo I, Sección I - Lingüística Teórica e Lingüística Sincrónica. A Co-ruña: Fundación Pedro Barrié de la Maza, Conde de Fenosa, 895-904.
- Lopes, Óscar. 1989. Construções Concessivas. Algumas reflexões formais logico-pragmáticas. Versão alargada publicada em Lopes, Ó. 2005, 194-209.

- Lopes, Óscar. 1991. Sobre a Semântica da Maneira e da Conformidade. *Actas do VI Encontro da Associação Portuguesa de Linguística*. Lisboa, 3-21. Também publicado em: Lopes, Ó. 2005, 245-263.
- Lopes, Óscar. 1992. Da Partícula «Pois» ao Conceito de «Apodeixis». *Actas do VII Encontro da Associação Portuguesa de Linguística*. Lisboa, 179-192. Também publicado em: Lopes, Ó. 2005, 211-224.
- Lopes, Óscar. 1994/1996. Partículas de Agulhagem do Discurso n' *O Crime do Padre Amaro*. Duarte I.; Leiria, I. (Orgs.) *Actas do Congresso Internacional sobre o Português*. Lisboa: Colibri, vol I, 189-202. Também publicado em: Lopes, Ó. 2005, 225-244.
- Lopes, Óscar. 2005. *Entre a palavra e o discurso (estudos de Linguística 1977-1993)*. Coordenação da edição de F. Oliveira e A. M. Brito. Porto: Campo das Letras.

LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS
DA UNIVERSIDADE DO PORTO

Espaço da direção

Apresentação
Fátima Oliviera

Artigos
Óscar Lopes e a Linguagem: o olhar de um cientista
Inês Duarte

Relativas livres e interrogativas parciais: como Óscar Lopes viu algumas destas construções
Ana Brito e Gabriela Matos

Sobre a natureza homogénea do Pretérito Perfeito Composto em Português Europeu
Fátima Oliveira e António Leal

Análise semântica de sentimento em artigos de opinião
*Fátima Silva, Purificação Silvano, António Leal, Fátima Oliveira,
Pavel Brazdil, João Cordeiro e Débora Oliveira*

Gestos com os pés: descrição de uma listagem
Isabel Galhano Rodrigues

Bibliografia em linguística de Óscar Lopes

VOLUME 13

ANO 2018