

LINGUÍSTICA

REVISTA DE ESTUDOS LINGUÍSTICOS
DA UNIVERSIDADE DO PORTO

VOL. 17, 2022

FICHA TÉCNICA

Linguística
Revista de Estudos Linguísticos da Universidade do Porto
Vol. 17, 2022

ISSN: 1646-6195

ISSN digital: 2182-9713

Periodicidade: Anual

Diretor:
António Leal

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

Capa:
José Osswald

Impressão e acabamentos:
Invulgar - Artes Gráficas

Depósito Legal:
248653/06

Tiragem:
120 Exemplares

Os artigos publicados estão sujeitos a “peer review”.

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>

Esta publicação é financiada por fundos nacionais através da FCT – Fundação para a Ciência e a Tecnologia, I. P., no âmbito do projeto UIDB/00022/2020.

Espaço da direção

O número de 2022 da *Linguística* inclui cinco contribuições em forma de artigo.

O volume começa com o texto “Para a construção de um *corpus* de interações orais em Português Língua Não Materna (PLNM) – algumas reflexões”, de Conceição Carapinha, da Universidade de Coimbra. A autora tece um conjunto de considerações acerca da pertinência da constituição de um *corpus* de interações orais tanto para o ensino como para a investigação em PLNM e em áreas conexas, das dificuldades que uma tal tarefa coloca e da forma como essas dificuldades podem ser ultrapassadas.

O segundo artigo, “Algumas notas sobre a quantificação de adjetivos de cor por *muito* e *pouco*”, é da autoria de Inês Cantante. Trata-se de uma primeira abordagem à semântica de cinco adjetivos de cor em Português Europeu, que parte de dados de *corpus* (CETEMPúblico) para a análise da possível graduabilidade destes adjetivos e do tipo de escala (aberta ou fechada) a que estarão associados.

Leticia Quesada Vázquez, da Universidade de Nebrija (Madrid), é a autora de “Rhythm instruction to improve ESP students’ fluency: A study of pausing”. Esse trabalho descreve uma experiência no âmbito do ensino de aspetos específicos da fonologia do Inglês a falantes nativos do Espanhol/Catalão.

O quarto artigo, “Sobre as formas de tratamento no Português em Angola: o caso de *você*”, de Queneth Pires, descreve também uma experiência, neste caso no âmbito do uso do pronome “*você*” como forma de tratamento no Português de Angola. A abordagem, de índole pragmática, procura pôr em evidência os fatores que favorecem o uso desta forma pronominal numa variedade do Português que está atualmente em reestruturação.

Finalmente, o artigo “Particularidades semânticas das expressões temporais com *até* e desafios de tradução entre português e inglês”, de Telmo Mória, da Universidade de Lisboa, apresenta uma descrição, baseada também em dados de *corpus* (CETEMPúblico), dos diversos valores temporais e discursivos que o conector “*até*” pode ter em Português. O autor reflete ainda sobre os conectores que veiculam os mesmos valores em

Inglês e sobre os problemas que esses valores colocam a alguns sistemas de tradução automática.

Este número não seria possível sem a inestimável colaboração de dezassete avaliadores, a quem agradecemos o tempo despendido e as valiosas sugestões.

Porto, setembro de 2022

António Leal

SUMÁRIO

Espaço da direção	3
Artigos	
Para a construção de um <i>corpus</i> de interações orais em Português Língua Não Materna (PLNM) – algumas reflexões.....	9
<i>Conceição Carapinha</i>	
Algumas notas sobre a quantificação de adjetivos de cor por <i> muito e pouco</i> ...	37
<i>Inês Cantante</i>	
Rhythm instruction to improve ESP students' fluency: A study of pausing.....	63
<i>Leticia Quesada Vázquez</i>	
Sobre as formas de tratamento no Português em Angola: o caso de você.....	95
<i>Queneth Pires</i>	
Particularidades semânticas das expressões temporais com <i> até</i> e desafios de tradução entre português e inglês.....	117
<i>Telmo Mória</i>	

Artigos

Para a construção de um *corpus* de interações orais em Português Língua Não Materna (PLNM) – algumas reflexões

Conceição Carapinha
mccarapinha@fl.uc.pt

Centro de Estudos de Linguística Geral e Aplicada (CELGA-ILTEC)
Universidade de Coimbra

ABSTRACT.

As authentic data sets, different types of corpora increasingly support theoretical and applied linguistic studies. In the same vein, corpora may also constitute an appropriate working basis for second language acquisition and foreign language teaching, especially with regard to the uses of language. However, both Pragmatics, as a discipline focused on the analysis of meaning in context and the area of second language acquisition (SLA) have been quite reticent when it comes to using corpora, whether written or oral. In this text, we present a characterization of spoken learner corpora and advocate their relevance and usefulness to the teaching and learning of Portuguese as a Foreign Language, namely in what concerns Interlanguage Pragmatics. In the same line of thinking, the PL2 oral interaction corpus project presented in this text, with the challenges faced by its implementation, and the possibilities for future research that it will allow, aims at contributing to a better understanding of the scientific and pedagogical advantages of a database containing texts of a more interactive nature.

KEYWORDS.

Corpora; spoken learner corpora; acquisition; oral interaction; Portuguese as a Foreign Language.

RESUMO.

Enquanto acervos de dados autênticos, os diversos tipos de *corpora* sustentam, cada vez mais, estudos de natureza teórica e aplicada, podendo também constituir uma boa base de trabalho no âmbito do ensino/aprendizagem de uma L2, sobretudo no que tange ao domínio dos usos. No entanto, quer a Pragmática, enquanto disciplina focada na análise dos usos linguísticos contextualizados, quer a área de aquisição de línguas segundas (SLA) se têm mostrado bastante reticentes no que toca ao recurso a *corpora*, sejam eles escritos ou orais. Neste texto, apresenta-se uma caracterização dos *spoken learner corpora* e advoga-se a sua pertinência e a sua utilidade no processo de ensino/aprendizagem do PLNM, nomeadamente no que à área da Pragmática da Interlíngua diz respeito. Nessa linha, o projeto de constituição de um *corpus* de interações orais, em PLNM, que aqui se apresenta – descrevendo não apenas os desafios que se colocam à sua implementação, mas também as possibilidades de pesquisa futura que, uma vez constituído e disponibilizado, ele permitirá – visa contribuir com uma

reflexão sobre as valências científicas e pedagógicas de uma base de dados contendo textos de natureza mais interativa.

PALAVRAS-CHAVE.

Corpora; *spoken learner corpora*; aquisição; interação oral; Português Língua Não Materna (PLNM).

0. Introdução

Num momento em que se exige ao investigador a pesquisa a partir de dados autênticos, um *corpus*, enquanto “conjunto de textos escritos (ou excertos de textos) ou de transcrições de registos orais, tipicamente em formato electrónico” (Nascimento 2002), pode ser explorado em áreas tão diferentes como o desenho de materiais instrucionais, a construção de um dicionário ou a investigação fundamental, possibilitando, portanto, investigação quer de cunho mais empírico quer de carácter mais teórico.

É verdade que são muito diferentes os acervos de dados a que este rótulo é atribuído, quer relativamente ao número de *tokens*, quer ao tipo de dados, quer ainda à sua organização interna, por isso, e para que um conjunto de dados seja considerado um *corpus*, é necessário que esses dados sejam representativos do objeto que se pretende analisar – seja ele uma língua, uma variedade linguística, um género textual, um tipo de falante ou qualquer outro fenómeno –, que haja critérios consistentes no que toca à sua recolha e organização e ainda que esses dados respondam a um determinado objetivo de investigação.

Independentemente das questões relacionadas com os diferentes tipos de *corpora*, com a sua constituição, dimensão, objetivos e acessibilidade, são inúmeras as vantagens da sua utilização. Estes acervos traduzem usos autênticos da língua em apreço, quer na sua modalidade escrita quer na sua modalidade oral e o facto de serem informatizados aumenta exponencialmente as suas valências, permitindo pesquisas muito rápidas e confiáveis em largas quantidades de dados, facilitando o reconhecimento de frequências de uso, a identificação de padrões (que escapariam a uma análise mais casuística) e fenómenos de coocorrência em determinados contextos ou géneros.

Flowerdew (2009) chama até a atenção para uma outra mais-valia dos

corpora, quiçá mais invisível; com efeito, as pesquisas baseadas nestas bases de dados permitem evidenciar muitas das incongruências que persistem nas atuais descrições da língua, pois muitos dos fenómenos atestados em gramáticas e outro tipo de textos não encontram suporte nos *corpora*, ao mesmo tempo que estes revelam usos que não estão ainda descritos.

Não negligenciando a relevância teórica da discussão acerca do estatuto da Linguística de Corpus (LC), entendida ora como domínio de pesquisa, ora como metodologia de análise, podemos afirmar que esta área permite estudar muitos aspetos da pesquisa linguística a partir de uma abordagem empírica (McEnery & Wilson 2001). Neste sentido, os *corpora* assumem indubitável relevância, não só porque eles possibilitam uma descrição da língua centrada em usos atestados, como também porque permitem trabalhar variados fenómenos de diferentes planos de descrição linguística a partir de quadros teóricos diversos. Uma terceira e incontornável vantagem diz respeito à aplicabilidade dos *corpora* a áreas de investigação conexas, como a do ensino de línguas estrangeiras, domínio em que os *corpora* constituem também um recurso precioso.

Neste estudo, apresentar-se-á precisamente uma reflexão sobre a forma como a constituição de um *corpus* de interações orais de aprendentes de português língua não materna (PLNM), que se integra no âmbito, mais lato, dos chamados *corpora* orais de aprendentes (*spoken learner corpora*), pode vir a coadjuvar a investigação no âmbito da pragmática do PLNM, elencando, por um lado, os desafios que se colocam à sua constituição e, por outro, as possibilidades de análise pragmática abertas por um acervo deste tipo. Na secção 1., serão abordadas, ainda que de forma necessariamente breve, as relações entre a LC e a Pragmática, por um lado, e a área de investigação em Aquisição de Línguas Segundas (doravante *SLA*¹), por outro. Na secção 2., caracterizar-se-ão os *corpora* orais de aprendentes (*spoken learner corpora*). Partindo da experiência obtida na constituição de um *corpus*-piloto, na secção 3., analisar-se-ão alguns problemas que podem dificultar a constituição de um *corpus* deste tipo, bem como algumas das vantagens deste inovador recurso metodológico para a investigação e para a pedagogia, sobretudo na área dos usos do PLNM.

1 Adota-se a sigla em inglês por ser de mais fácil identificação.

1. Linguística de Corpus e outras áreas de investigação

1.1. Linguística de Corpus e Pragmática

Centrada na análise de usos linguísticos contextualizados, a Pragmática parece ser a área de investigação preferencial para trabalhar com *corpora*. Na verdade, “[a] pragmatic perspective on language implies that the focus is on the use of lexical elements or grammatical structures in their linguistic, social and cultural context” (Aijmer 2020: 29). Ora, ao permitir identificar estruturas e usos no seu contexto de ocorrência, um *corpus* parece estar, de facto, ao serviço das pesquisas pragmáticas.

Porém, esta óbvia convergência nem sempre foi explorada e, durante muito tempo, o trabalho da LC correu em paralelo às pesquisas em Pragmática, sobretudo devido às diferentes abordagens propostas pelas duas áreas. A natureza da investigação pragmática é de índole claramente qualitativa, individual e local, interessada que está na análise da função de uma determinada expressão ou estrutura inserida num contexto específico (Aijmer & Rühlemann 2015; Jucker 2018; Aijmer 2020), ao passo que a LC, ao analisar grandes massas de dados, tem preferido uma metodologia mais quantitativa e vertical.

Estabelecer uma articulação entre os dois domínios tem sido, aliás, um desafio que enfrenta algumas resistências. Muitos investigadores ligados à Pragmática alertam para o facto de os *corpora* servirem ‘apenas’ como porta de entrada para os dados e necessariamente terem de ser sujeitos a um rigoroso escrutínio em função do quadro teórico (e das perguntas de investigação) adotado, sob pena de o estudo se tornar apenas estatístico e não incluir nenhum tipo de explicação causal sobre os factos, ou seja, não testar uma teoria. Outros assinalam que a escassez de dados relativos ao perfil sociológico dos informantes e às condições (contextuais) que envolveram a recolha de dados pode enviesar ou pelo menos não auxiliar as análises pragmáticas (McEnery & Wilson 2001). Por outro lado, sempre que houver uma discrepância entre o que se diz (e que consta do *corpus*, oral ou escrito) e o que se pretende dizer, ou seja, sempre que estiver em jogo um significado mais implícito, a análise pragmática não encontrará respostas no *corpus*. E esta é uma lacuna relevante, à qual vários investigadores são

sensíveis: “the relationship between pragmatics and corpus linguistics is not unproblematic. The reason is simple: corpora record text, not meaning, and they record context only crudely” (Rühlemann 2010: 289). Avulta ainda o problema das ferramentas de anotação pragmática. O desenvolvimento dessas metodologias ainda se debate com alguns desafios de difícil resolução, pois não havendo muitas vezes, no domínio dos usos, uma correlação fixa entre forma e função, nem sempre é fácil detetar determinados fenómenos pragmáticos.

Apesar destes obstáculos, o entrosamento das duas áreas tem vindo, segundo Romero-Trillo (2008), a dar-se paulatinamente e os contributos da LC trouxeram claros benefícios às análises qualitativas realizadas no domínio pragmático, uma vez que é possível combiná-las com metodologias mais quantitativas. Parece ser este, justamente, o grande contributo da LC para a Pragmática: oferecer-lhe uma metodologia de análise de base empírica (Rühlemann & Clancy 2018: 242).² Entretanto, neste recente trilhaço conjunto, as duas áreas têm redefinido os respetivos objetivos. Por um lado, a LC tem vindo a admitir a pertinência científica dos pequenos *corpora*; por seu turno, a metodologia investigativa da Pragmática acomoda-se bastante melhor com esses pequenos e *domain-specific corpora*, que lhe permitem manter a base empírica, enquanto busca padrões pragmáticos numa base de dados mais manuseável (Vaughan & Clancy 2013: 70).

Nesta sequência, a Pragmática de Corpus, jovem disciplina que interseta as valências das duas áreas, vem-se afirmando como a “science that describes language use in real contexts through corpora” (Romero-Trillo 2017: 1)

1.2. Linguística de Corpus e *Second Language Acquisition (SLA)*

A pesquisa realizada no âmbito da área de investigação em Aquisição de Línguas Segundas (*SLA*) visa descrever os conhecimentos linguísticos do aprendente, a progressão desses conhecimentos ao longo das diversas fases de construção da interlíngua, ou seja, a forma como esses conhecimentos

² Note-se que os procedimentos controlados de eliciação de dados, habitualmente usados na área, têm sido criticados precisamente pela sua falta de autenticidade, o que tem levado os investigadores a preferir dados mais autênticos ou, pelo menos, a cruzar os resultados obtidos a partir dessas duas metodologias.

evoluem (ou estagnam, no processo conhecido por ‘fossilização’) e os fatores que afetam, limitando ou incentivando, o desenvolvimento dessa competência. Neste sentido, as dimensões cognitivas do processo de aquisição têm vindo a ser muito exploradas, sobretudo no que respeita ao processamento da informação linguística, aos recursos atencionais e ao papel da motivação na aprendizagem de línguas estrangeiras, análises muitas vezes domiciliadas num quadro teórico de base generativa.³ Ora, o desenho de um quadro teórico deste tipo tende a afastar o interesse dos investigadores da análise de dados autênticos, e muitos destacam que a investigação realizada no âmbito da *SLA* não tem favorecido o recurso a *corpora* (Lozano & Mendikoetxea 2013: 65).

Entretanto, outros autores (Leech 1997; Mukherjee 2006; Johansson 2007) salientam o potencial contributo dos *corpora* no processo de ensino-aprendizagem de uma língua estrangeira, especificando que esse contributo pode concretizar-se sob três formatos distintos: os dados da língua alvo podem ser explorados na sala de aula de língua estrangeira e constituir a base de novas aprendizagens; podem também ser usados, de forma indireta, servindo como base para a construção de material instrucional; uma terceira hipótese prevê o desenvolvimento de *corpora* especificamente orientados para o ensino.

É precisamente nesta última área que avultam os *learner corpora*, acervos de dados produzidos por aprendentes de uma língua estrangeira. Inicialmente, os *learner corpora* englobavam pequenas bases de dados decorrentes da recolha de professores ou investigadores, em aula; no entanto, a investigação efetuada a partir desses dados não permitia a obtenção de resultados com elevado poder explanatório e foram razões desta índole que levaram à constituição de *learner corpora* de maiores dimensões, como o ICLE (International Corpus of English Learner)⁴, o primeiro grande *learner corpus*, resultante da escrita de textos argumentativos produzidos por aprendentes de inglês como L2⁵ e organizado em diferentes *subcorpora* em

3 Hondo (2014: 2) assinala, no âmbito desta linha cognitivista, a emergência de outras orientações. Outras abordagens, de cunho mais funcionalista, têm vindo, entretanto, a surgir neste domínio. Veja-se: Atkinson (2011) e VanPatten et al. (2020).

4 Sobre este corpus, ver Granger et al. 2002.

5 O termo L2 refere, aqui, qualquer língua adquirida após a nativa, evitando assim os problemas definitórios inerentes às expressões ‘língua segunda’ e ‘língua estrangeira’.

função da língua materna dos falantes.⁶

Ainda assim, e nem mesmo com a crescente entrada das abordagens socioculturais na área da *SLA*, ou seja, com a tónica nos fatores sociais que influem na aquisição da L2, o interesse pelos *learner corpora* aumentou. Em sentido inverso ao que seria previsível, o impacto dos *learner corpora* na pesquisa em *SLA* tem sido bastante lento (McEnery et al. 2019).

A preferência pelos dados experimentais (elicitados) – em detrimento dos dados reais produzidos pelos aprendentes – revelada pela área da *SLA* (Granger 2002) deve-se a algumas limitações destes últimos, apontadas pelos investigadores. Em primeiro lugar, os *learner corpora* são constituídos por dados, orais ou escritos, que apenas refletem as competências produtivas dos aprendentes, não dando conta das suas competências recetivas, nem espelhando os processos cognitivos necessários à construção da interlíngua do aprendente, razão por que a sua utilidade é, sob este prisma, reduzida. Um segundo problema relaciona-se com a escassez dos dados; por muito vasto que seja um *learner corpus*, ele pode não representar, de forma adequada, as diversas fases de construção da interlíngua, pode não refletir os verdadeiros conhecimentos do aprendente (*o intake*) e pode não oferecer suficientes evidências de um determinado fenómeno que se pretende estudar. As ferramentas de anotação de *corpora* constituem um terceiro problema, uma vez que ainda não estão suficientemente desenvolvidas para identificar fenómenos numa base de dados que se afasta, em múltiplos aspetos, da variedade padrão, exigindo, portanto, uma metodologia de anotação manual para tratamento dos erros e desvios dos aprendentes.

No global, o investigador que trabalha na área de investigação da *SLA* crê que é difícil, por um lado, controlar as numerosas variáveis que operam em contextos mais autênticos (Granger 2002; Mackey & Gass 2005) e, por outro, considera que os metadados que acompanham estes acervos não explanam exaustivamente todas as diferentes variáveis relevantes para a pesquisa.

Apesar destas fragilidades, a investigação realizada no domínio da *SLA* pode beneficiar do aporte trazido pela análise de um grande *learner corpus*.

⁶ O ICLE foi posteriormente valorizado com o advento de outros *corpora*, como o LOCNES (Louvain Corpus of Native English Essays) que inclui um grande acervo de textos argumentativos produzidos por estudantes nativos ingleses e norte-americanos, possibilitando, assim, análises contrastivas (nativo vs. não nativo).

A vastidão dos dados disponíveis pode permitir ao investigador encontrar tendências, estruturas, expressões, usos e desvios que, de outra forma, lhe escapariam. Com as atuais valências tecnológicas, será certamente possível criar um instrumental de anotação adequado, que permita uma fácil identificação de padrões de coocorrência e de listas de frequência. Estes resultados poderiam alimentar muitos estudos nesta área, nomeadamente no que toca aos aspetos gramaticais mais prontamente adquiridos, mais vulneráveis ou mais sujeitos a fossilização e, por outro lado, ao papel das interfaces na aquisição de uma língua segunda (Lozano & Mendikoetxea 2013).

2. Os *corpora* orais de aprendentes de L2 (*spoken learner corpora*)

Embora o campo de investigação conhecido como *Learner Corpus Research* (LCR)⁷, tenha direcionado a sua atenção sobretudo para os *learner corpora* contendo dados escritos, os *spoken learner corpora* têm vindo a ganhar cada vez mais protagonismo. Ainda raros e principalmente pouco explorados, estes *corpora* resultam da gravação e posterior compilação de discursos orais, produzidos na língua alvo por falantes não nativos.⁸ Estes acervos podem apresentar-se sob três formatos distintos: apenas na versão transcrita, habitualmente apelidados de *mute spoken corpus* (Ballier & Martin 2015); na versão transcrita acompanhada dos respetivos áudios; e, na versão mais moderna, sob formato multimodal, combinando dados provenientes de áudio e de vídeo, na tentativa de apreender gestos, expressões faciais, olhares, comportamento próxémico e háptico e todo o tipo de sinais comportamentais que enriquecem e dão sentido à interação verbal. Muito recentes, estes *corpora* multimodais são, todavia, ainda mais raros, tendo em conta as dificuldades de natureza ética e legal que envolvem a recolha e a divulgação pública da imagem e dos dados pessoais.

A carência de *spoken learner corpora* deve-se a razões várias, relacionadas com a difícil tarefa de coletar os dados, uma vez que cada informante ou

7 Nesta área, pontuam os trabalhos de Sylviane Granger.

8 Assinale-se que, ao mencionarmos os *spoken learner corpora*, tal não significa, como muito bem assinalam Lozano et al. (2021), que estamos a reportar-nos a dados necessariamente autênticos, uma vez que, não raro, estes dados são produzidos em contexto de aula e em condições controladas.

grupo de informantes tem de ser gravado individualmente, com o tempo consumido na respetiva transcrição e com a relativa ausência de ferramentas que deem conta das especificidades do oral na escrita (Luzón et al. 2007: 3). A este respeito, é indubitável que a análise de *spoken learner corpora* se torna ainda mais difícil por requerer uma atenção particular aos “elements beyond the text, such as intonation, gesture and discourse structure, which cannot easily be explored with the use of the kinds of frequency-based techniques used in the analysis of written corpora” (Adolphs & Knight 2010: 38). Com efeito, e tal como foi assinalado por diversos investigadores, um *corpus* oral é sempre multimodal e, portanto, engloba também elementos paraverbais e não verbais que complexificam a tarefa da transcrição. Por este motivo, é sempre necessário tomar decisões quanto ao grau de granularidade da transcrição de um *spoken learner corpus*, a qual pode envolver diferentes níveis de complexidade.

Não por acaso, muitos *spoken learner corpora* foram criados para possibilitar, sobretudo, pesquisas de natureza fonética e fonológica, com exercícios orientados de leitura e com um elevado controlo do *output*, de modo a permitir a comparabilidade dos dados e análises quantitativas rigorosas (Díaz-Negrillo & Thompson 2013), e não tanto para permitir, por exemplo, a descrição de uma interação verbal em língua segunda, mais imprevisível, mais complexa e em que o contexto (conceito difícil de definir, nas suas múltiplas vertentes – epistémica, linguística, interacional e social) é determinante.

De qualquer modo, os *spoken learner corpora* podem fornecer uma valiosa base empírica para analisar o desenvolvimento das competências orais do aprendente, nos diversos planos de descrição linguística. Se forem concebidos com objetivos mais ambiciosos, nomeadamente o de incluírem diferentes amostras de discurso espontâneo e finalístico, monogerado e poligerado, então as suas valências aumentarão consideravelmente, pois permitirão obter materiais adequados à descrição das interlínguas dos aprendentes no plano pragmático-discursivo, possibilitando o acesso a informações diversas que um *corpus* escrito tipicamente não traduz: silêncios, hesitações, interrupções, sobreposições e todos os traços vocais e prosódicos que refletem as dimensões emocionais e atitudinais que habitualmente acompanham a fala (MacWhinney 2021).

3. A pertinência de um *learner corpus* de interações orais em PLN

No que tange à língua portuguesa na sua variante europeia, estão disponíveis três *corpora orais* de aprendentes de Português como L2 que envolvem, aliás, dados de natureza bastante díspar. Os dados constantes do COPLE2⁹ resultam da gravação de conversas entre dois candidatos, ocorridas durante um exame de natureza oral, do CAPLE, e mediadas por um avaliador (Mendes *et al.* 2016). No caso do projeto CAL2¹⁰, os dados orais correspondem a entrevistas realizadas entre um aprendente e um falante nativo do português sobre temas diversos (Ferreira *et alii*, no prelo). O *corpus* COral-Co¹¹ apresenta um conjunto bastante variado de dados de produção oral, que decorrem da aplicação de um inquérito, envolvendo distintas tarefas, orientadas para permitir análises em diferentes domínios de funcionamento da língua portuguesa (fonético-fonológico; morfológico; sintático; semântico; lexical; pragmático; textual).¹²

Estes *corpora* resultam, assim, de atividades de avaliação, da interação entre nativo e não nativo, e de tarefas de elicitación de dados (respetivamente). Não escamoteando a sua inequívoca relevância, é notória a ausência de um *corpus* que dê conta da interação verbal entre aprendentes de PLN.

Apesar de muitas abordagens ao ensino/aprendizagem de L2 estarem, em si mesmas, enformadas pela centralidade atribuída ao conceito de interação verbal, como é o caso das teorias sociointeracionistas (Mondada & Doehler 2004) e, particularmente, da *Communicative Language Teaching* (Richards 2006) e apesar de um bom domínio da competência interacional ser o objetivo último de qualquer processo de ensino/aprendizagem de uma L2, esta é uma área em que os aprendentes exibem, normalmente, grandes dificuldades comunicativas, quer a interação verbal em que participam seja mais espontânea ou mais finalística.

A competência interacional (Celce-Murcia 2007) integra a capacidade

9 *Learner Corpus Português L2 – COPLE2*, do Centro de Linguística da Universidade de Lisboa (CLUL), com coordenação de Amália Mendes.

10 *Projeto CAL2 – Corpus de Aquisição L2 - subcorpus Produção Oral*, do Centro de Linguística da Universidade Nova de Lisboa (CLUNL), com coordenação de Ana Madeira. Para mais informações, consultar: <http://cal2.clunl.fcsh.unl.pt/index.html>.

11 *Corpus Oral de Português L2 – Coimbra (COral-Co)* – do CELGA-ILTEC, da Universidade de Coimbra, com coordenação de Isabel Almeida Santos.

12 Sobre o *corpus* COral-Co, ver Santos *et al.* (2016).

de construir uma conversa, desde a secção de abertura até à fase de encerramento, gerindo todas as vertentes próprias de uma interação verbal que é, também, sempre social. Neste âmbito, o aprendente terá de fazer uso da competência acional, que prevê a construção de atos ilocutórios adequados aos vários momentos da interação, mas também terá de dominar a competência conversacional, gerindo, de forma apropriada, as regularidades inerentes à organização sequencial do diálogo, nas fases mais relacionais e mais transacionais da interação, e ver-se-á ainda instado a organizar adequadamente todo o comportamento paraverbal, cinésico e háptico (gestos, olhares, distância do interlocutor e sinais de retroação, por exemplo). Todas estas imposições representam, assim, um ónus considerável para um aprendente de uma L2, na medida em que lhe exigem a gestão adequada e simultânea dos planos transacional e relacional, ainda por cima enquadrados por normas reguladoras que são, não raro, culturalmente dependentes.

Razões desta natureza exigem, por um lado, um conhecimento profundo e relativamente sistematizado das características e do funcionamento das interações orais, em PLN, produzidas por aprendentes de português europeu e impõem, por outro, um trabalho intensivo em torno do desenvolvimento da competência interacional dos aprendentes. Com efeito, é necessário ter uma descrição fundamentada dessas produções orais, com o objetivo de identificar as áreas de maior dificuldade e fragilidade, as subcompetências mais carenciadas de intervenção, em suma, para realizar um levantamento exaustivo dos aspetos que terão de ser alvo de instrução explícita e de ser contemplados em contexto instrucional.¹³

A constituição de um *corpus* de interações orais produzidas por aprendentes de PLN constitui um primeiro passo para dar consecução a este desiderato.

¹³ Reconhece-se também, naturalmente, a necessidade de obter dados provenientes de falantes nativos. Numa fase posterior deste projeto, poderá equacionar-se a hipótese de vir ainda a incluir subcorpora de controlo de falantes nativos das LM dos informantes.

3.1. Desafios na compilação de um *corpus* de interações orais em PLN – resultados preliminares

O objetivo primordial da constituição deste *corpus* oral é a obtenção de uma base de dados constituída por interações verbais orais, protagonizadas por uma díade ou tríade de participantes – aprendentes adultos de PLN – que discutem, num determinado intervalo temporal, pontos de vista acerca de um tópico sensível, e foi com este propósito que se iniciou, em maio de 2022, a fase-piloto da recolha de dados. As opções metodológicas adotadas nesta fase fizeram emergir algumas questões que serão, agora, objeto de análise.

Ao ambicionar aceder a discurso oral espontâneo, foi necessário refletir sobre a metodologia de recolha de dados. A exigência, epistemológica, de trabalhar com dados autênticos deveria obrigar o investigador a coletar discurso espontâneo, sem o conhecimento dos envolvidos na recolha e, obviamente, sem qualquer intervenção da sua parte, evitando, assim, o recurso a procedimentos de recolha de dados que permitissem ao informante preparar (ou refletir sobre) o seu próprio desempenho. Todavia, de um ponto de vista legal, tal metodologia é impraticável, razão por que a eliciação de dados, enquanto metodologia orientada, tem sido a mais utilizada no domínio da *SLA*.

Nesta linha de raciocínio, e considerados todos os potenciais problemas, a metodologia utilizada na fase-piloto foi a da *elicited conversation*, método que, segundo Taguchi & Roever (2017), constitui um procedimento pouco utilizado, mas promissor. Neste caso, e uma vez lançado, pelo investigador, o tópico de discussão, toda a interação fica a cargo dos intervenientes; a ausência do investigador do contexto conversacional e a não existência de um *script* rígido deixam aos participantes o ónus de coconstruir a sua interação. Esta opção possibilita uma produção oral (quase-) espontânea, a decorrer em situação (quase-) informal e com a participação mínima do moderador (o investigador), razão por que pode aproximar-se bastante do perfil de um verdadeiro debate na língua alvo. Não sendo dados autênticos, estes poderão, seguramente, ser caracterizados como *near-natural*.

A metodologia a utilizar envolveu também a decisão relativa ao assunto que constituiria o móbil do debate. A escolha recaiu sobre um tópico

acessível, mas suficientemente instigante para motivar a discussão, suscitado por um título do jornal Observador: “O Parlamento criou um grupo de trabalho para discutir o fim dos animais nos circos e nos jardins zoológicos. Manter os animais nos jardins zoológicos e nos circos deve ser proibido?”.

Aos intervenientes foram dadas indicações genéricas sobre a atividade, de modo a permitir que esta decorresse da forma mais natural possível e sem a subsequente intervenção do investigador. Para enquadrar a interação, foi criado um cenário verosímil – apresentado aos participantes, pelo investigador, no início da atividade – que contemplava um contexto específico: um encontro entre três colegas, no bar da faculdade, durante um intervalo de 10 minutos entre duas aulas, no âmbito do qual, e a propósito do título do Observador que um deles deveria ler em voz alta, decorreria o debate.

Não tendo previsto fornecer aos aprendentes, previamente, quaisquer informações acerca da atividade em que iam participar, para promover a autenticidade da interação oral, o primeiro grupo de três informantes cuja produção oral foi recolhida na fase-piloto pediu, no entanto, algum tempo de preparação antes do início da tarefa. A imprevisibilidade da situação causou manifesta ansiedade em todos os membros do grupo, evidenciando a complexidade da tarefa e o risco de o grupo poder vir a ter sérias dificuldades em produzir dados orais, e consequentemente, em cumprir os objetivos pretendidos.¹⁴

Estas dificuldades obrigaram a repensar toda a planificação inicial e conduziram à sua reformulação. Num segundo momento da fase-piloto, já com outros informantes, optou-se pela diferenciação de contextos de recolha, exigindo a cada um dos grupos participantes um grau de preparação distinto. Na segunda fase, participaram nove aprendentes, oito de nível C1+ e um de nível B2, constituindo três grupos de debate, e a cada um desses grupos foi apresentado um cenário diferente. Ao primeiro grupo foram dadas instruções prévias (com uma semana de antecedência) sobre os objetivos da atividade, o tema e a forma como iria processar-se o debate, instruções que incluíam, também, o papel interacional a desempenhar por cada um (um dos participantes teria de defender um determinado ponto

¹⁴ Esta dificuldade foi também reiterada na resposta à pergunta – sobre propostas de melhoria da atividade – que constava da ficha de dados individuais solicitada a cada um dos participantes no fim da atividade.

de vista; outro deveria contra-argumentar; o terceiro elementa funcionaria como voz do consenso, tentando equilibrar e harmonizar a discussão). O segundo grupo diferiu do anterior pelo facto de ter conhecimento do tema e das instruções apenas no momento do debate. Também ao terceiro grupo foi apresentado o tema-estímulo pelo investigador, imediatamente antes da interação, não tendo sido fornecida qualquer outra instrução, podendo cada um dos participantes defender o ponto de vista que entendesse.¹⁵

Sendo certo que, segundo Biber (1993), o processo de compilação de um *corpus* deve ser recetivo aos problemas surgidos aquando da coleta de dados, com esta diversidade de opções pretendeu-se calcular em que medida a pré-preparação e a pré-determinação de papéis interacionais do debate 'contaminariam' a naturalidade da interação, por um lado, e avaliar qual o risco de o debate fracassar face à inexistência de instruções rígidas, por outro, ou seja, avaliar qual a melhor metodologia, tendo em conta os objetivos do *corpus*.

Para além desta avaliação, há ainda algumas questões mencionadas pelos participantes e sobre as quais urge também refletir, antes de tomar uma decisão definitiva, nomeadamente as que respeitam: à necessidade de introduzir um momento inicial, que permita contextualizar o assunto e familiarizar os aprendentes com a temática, contrariando a impositividade da tarefa; à possibilidade de formar grupos mais alargados, para que cada participante se sinta mais apoiado na defesa da sua opinião; à eventual rotatividade dos aprendentes nos vários grupos, permitindo a cada um participar no debate mais do que uma vez, mas com diferentes interactantes.¹⁶

A seleção dos informantes constituiu um outro aspeto importante na fase-piloto da recolha de dados.

Contrariamente ao desenho inicial do projeto, que previa a recolha de dados envolvendo informantes com diferentes línguas maternas e de distintos níveis de proficiência – a iniciar no nível B1, dada a complexidade da tarefa –, foi apenas possível obter, na fase-piloto, os contributos de

15 Este terceiro cenário correspondia, como é óbvio, ao inicialmente previsto no protocolo.

16 Se a primeira sugestão parece ser de fácil execução, as duas seguintes podem ser problemáticas. Um grupo mais alargado implicará mais dificuldades na identificação das vozes e só resultará se os dados forem recolhidos em vídeo, o que nos reconduz às questões legais atrás identificadas. Por outro lado, e embora a dinâmica das interações orais, sempre diferente, obrigue os aprendentes a reagir verbalmente de forma diferenciada, a possibilidade de um mesmo aprendente poder participar no debate, em grupos distintos, pode vir a diminuir o grau de naturalidade da sua participação.

quatro grupos, ou seja, de doze aprendentes de PLNM, onze de nível C1+ e um de nível B2+,¹⁷ que se encontravam em contexto instrucional em Portugal e, portanto, em regime de imersão, a frequentar cursos ou unidades curriculares de língua, literatura ou cultura portuguesas.¹⁸ De igual modo, as línguas maternas destes doze informantes são também pouco variadas (nove falantes de chinês mandarim; um falante de árabe; um falante de italiano e um falante de japonês). Tal assimetria pode, por um lado, vir a revelar alguns dados interessantes, no caso dos grupos constituídos por falantes que partilham a mesma língua materna,¹⁹ mas, por outro, pode vir a obstaculizar o equilíbrio do *corpus* relativamente às LM nele representadas, pode vir a inviabilizar a constituição de tríades que partilhem outras LM e pode até dificultar a constituição de grupos com grande diversidade de LM. Esta é uma clara desvantagem que resulta do facto de a base de recrutamento de informantes ser escassa e pouco diversificada.

Um outro potencial problema relacionado ainda com os informantes é o que decorre do número, desigual, de alunos inscritos nos diferentes níveis de aprendizagem, havendo bastante menos alunos a frequentar os níveis mais altos, o que pode dificultar, de novo, a constituição de um *corpus* equilibrado, desta feita no que concerne aos níveis de proficiência.

Finalmente, ao assumir que os contributos dos intervenientes são voluntários, e foi desse pressuposto que partiu o convite que lhes foi endereçado para, de forma espontânea, participarem numa atividade extracurricular, é ainda de equacionar a possibilidade de só os alunos mais proficientes se terem apresentado, o que pode enviesar os dados obtidos e retirar representatividade ao *corpus* (Gilquin 2015).

Um outro aspeto importante no que concerne à recolha de dados orais diz respeito ao próprio processo de gravação, aqui incluindo o espaço e o equipamento. Tentou evitar-se a captação de ruídos de fundo, cuidado fundamental para preservar a qualidade da gravação e, mais ainda, para permitir uma transcrição do oral que pode vir a revelar-se dilatada no tempo

17 Idealmente, os grupos deveriam ser constituídos por informantes com o mesmo nível de proficiência.

18 As produções orais dos quatro grupos da fase-piloto perfizeram um total de cinquenta minutos de gravação, com uma média de doze minutos e meio por grupo.

19 Num dos grupos cuja produção oral foi recolhida na fase-piloto do projeto, constituído por estudantes que partilhavam precisamente o chinês mandarim, como LM, é audível um fenómeno de *code-switching*, em que um dos participantes sussurra algo em mandarim, à espera de obter, dos colegas de grupo, a tradução portuguesa da expressão chinesa.

e exigente quanto à discriminação vocal dos intervenientes, sobretudo se estes forem mais do que dois. Com efeito, e considerada a participação de três intervenientes do mesmo sexo, como aconteceu em dois dos grupos da fase-piloto, ficou patente que o equipamento de gravação a utilizar deveria permitir identificar as diferentes vozes. Para coadjuvar o processo de reconhecimento vocal, foi solicitado às participantes nestas interações que se identificassem no início da gravação (embora essa informação não venha a ser divulgada aquando da futura disponibilização dos ficheiros áudio).²⁰ Por outro lado, e embora tal não seja o objetivo primeiro da constituição deste *corpus* oral, uma boa qualidade sonora permitirá também outras possibilidades de pesquisa. No que concerne ainda ao equipamento, é importante procurar um adequado equilíbrio entre a qualidade e a discrição. Segundo Golato (2017: 23), “the very presence of recording equipment may alter subjects’ speech production”, por isso foi necessário ponderar qual dos mecanismos de captação e de gravação de voz – gravador, computador ou telemóvel – seria o mais fidedigno (em termos de registo sonoro) e o menos invasivo. Durante a fase-piloto, a recolha dos dados foi feita por uma dupla via (telemóvel e *tablet*), precisamente para permitir analisar a qualidade sonora dos dois dispositivos e fundamentar uma decisão posterior.

Apesar de todos estes procedimentos cautelares, no sentido de melhorar o desenho do protocolo e de otimizar o processo de recolha, não descurando os objetivos iniciais do projeto, será ainda necessário avaliar o peso da própria situação de gravação, uma vez que também esta pode influenciar a interação oral que vai ser gravada. A marcação de um dia e de uma hora para reunir os participantes e dar consecução à tarefa retira-lhe também alguma genuinidade. A imposição de um tópico de debate, pelo investigador, e a obrigatoriedade de participação de cada um dos intervenientes, associadas ao facto de os participantes estarem cientes de que o seu desempenho será avaliado por um investigador, ainda que essa entidade se ausente do contexto enquanto a interação decorre, constituem outras restrições que, previsivelmente, terão consequências na produção oral dos falantes.

Tentar minimizar o impacto de todos estes fatores é um objetivo difícil

²⁰ Prevê-se a disponibilização do material áudio acompanhado da respetiva transcrição.

de atingir, mas tal desiderato deve permanecer sempre no horizonte de preocupações do investigador e, mais do que isso, levá-lo a documentar bem todo o contexto em que a gravação tem lugar, com informação acerca dos participantes, do espaço, do tempo, do equipamento e de todos os dados que possam vir a ser relevantes numa audição e análise posteriores (Adolphs & Knight 2010).

Esta atenção aos metadados orientou já todo o processo de recolha de dados na fase preliminar e a elaboração do documento intitulado *Perfil do Informante* foi pensada nesse sentido. Para além de um bem detalhado consentimento informado – na medida em que os dados orais vão tornar-se públicos – que todos os envolvidos tiveram de ler e assinar, e que já visava obter informação pormenorizada sobre aspetos biológicos, geográficos, sociológicos e linguísticos dos envolvidos, foi explicitado o quadro, tão aprofundado quanto possível, das diversas vertentes envolvidas no processo de gravação dos dados orais. De facto, “[t]o increase the rigor, transparency, and usability of learner corpora, collecting and publishing substantial metadata (including how the metadata were collected) is essential (Bell & Payant 2021: 54).²¹

Importa, por fim, considerar ainda um aspeto fundamental: a representatividade de um *corpus* deste tipo. Embora esta questão possa levantar dúvidas, nomeadamente quanto à quantidade de textos/discursos que representam, de forma fidedigna e completa – se é que é possível falar de completude, neste caso –, um determinado género textual, considerando os onerosos procedimentos a que obriga (em termos de tempo, de espaço e de equipamento), um *corpus* oral deste tipo tende necessariamente a ser mais pequeno (Adolphs & Knight 2010), sendo que esta menor dimensão não invalida, necessariamente, a sua validade. De facto, e dependendo dos objetivos visados, um *corpus* mais pequeno pode ser suficientemente representativo de uma determinada população; por outro lado, justifica-se

21 Estas autoras sugerem que é necessário disponibilizar informação sobre: a instituição em que decorre a gravação; o tipo de conhecimento prévio dos informantes acerca do evento em que vão participar; o seu grau de conhecimento do género textual em causa; eventuais pesquisas previamente realizadas acerca do tópico a discutir; o contexto sociointeracional desenhado pelo investigador no início de cada uma das sessões de gravação; o hipotético recurso dos participantes a auxiliares escritos, durante a interação; o tempo que demora a contextualização inicial e a atividade interacional em si mesma, entre outras informações possíveis (Bell & Payant 2021). Somente desta forma se avaliará o grau de autenticidade dos dados recolhidos e o peso detido pelo enquadramento, simulado, que os enquadra.

também o seu menor tamanho se considerarmos que se trata de um *corpus* mais especializado; por fim, havendo um rico conjunto de metadados, como acima se referiu, os investigadores terão à sua disposição todas as informações de que necessitam para avaliar as instâncias discursivas presentes.²²

3.2. Possibilidades de investigação

Conquanto alguns investigadores considerem que um *corpus* deste tipo é pouco interessante para a investigação no domínio da *SLA*, por conter mais informações sobre os usos efetivos que os aprendentes fazem da língua alvo do que sobre o processo de aprendizagem dessa língua, as valências de um *corpus* de interações orais em PLN^M são inúmeras, devem ser ressaltadas e podem mesmo revelar-se aliciantes para o tratamento das questões que, tipicamente, têm ocupado a agenda investigativa na área da *SLA*.

Em primeiro lugar, e tratando-se de um tipo de interação oral relativamente mais espontânea, exterior ao contexto letivo, os dados obtidos serão mais naturais do que os disponibilizados pelos diferentes *corpora* orais existentes, uma vez que grande parte destes últimos foi recolhida em contexto de avaliação e em contexto de sala de aula. É também crível que, neste tipo de trocas verbais, de natureza dialógica e em contexto *near-natural*, seja visível o manifesto e genuíno interesse dos falantes em defender uma causa, o que concorrerá para a maior autenticidade dos dados obtidos. No mesmo sentido, o conforto psicológico advindo de uma interação entre pares, sem a presença e a intervenção do professor/investigador, será, previsivelmente, um adjuvante não apenas de maior produção linguística, como de menor inibição em participar (Philp *et al.* 2014).

O facto de se tratar de uma *elicited conversation* concorre, de igual forma, para a riqueza e para a originalidade dos dados obtidos, pois, neste caso, a produção oral dos participantes não se reduz a palavras ou sequer a frases, mas traduz-se em textos/discursos, ou seja, aquilo a que acedemos é ao produto de uma atividade comunicativa completa, a um conjunto

²² Por limitações de espaço, não serão aqui considerados os constrangimentos impostos pela transcrição destes dados.

de intervenções poligeradas, devidamente articuladas entre si e inseridas num contexto que, não sendo efetivamente real, constitui, nas palavras de Kramsch (1986: 367), “a shared internal context or ‘sphere of inter-subjectivity’ that is built through the collaborative efforts of the interactional partners”.

Por outro lado, não estando dirigido para a elicitación de determinadas estruturas ou para a verbalização de funções linguísticas específicas, este *corpus* oferece uma base de dados mais ampla, relativa ao uso oral do português como língua não materna. Neste mesmo sentido, e dado estar centrado num género específico, que envolve sequências textuais de natureza argumentativa, será um *corpus* inovador, dotado de amplas potencialidades no que diz respeito à investigação em PLNLM.

Como facilmente se depreende, o protocolo que orientará a recolha de dados foi pensado, em primeiro lugar, para permitir análises de natureza pragmático-discursiva que possibilitem a descrição das interlínguas dos participantes neste domínio. De facto, o objetivo primeiro deste *corpus* é o de permitir a análise da competência interacional dos aprendentes de PLNLM, área ampla que contempla: a capacidade de realizar atos de fala e sequências de atos de fala em interações na língua alvo, que visem a troca de informações, a expressão de opiniões, a explicitação de problemas, a criação de cenários hipotéticos, entre outras possibilidades (âmbito da competência acional); a capacidade de gerir a interação verbal, no plano da organização dos turnos de fala, da abertura e do fecho das trocas verbais, da introdução, fecho, alteração ou retoma de tópicos (âmbito da competência conversacional); o domínio dos mecanismos de coesão lexical e gramatical, sobretudo no que concerne à adequada utilização de cadeias anafóricas e de marcadores discursivos; o domínio dos deícticos enquanto marcadores de ancoragem situacional; o domínio dos esquemas formais e das estratégias retóricas que permitem identificar uma determinada estrutura genológica (âmbito da competência discursiva); e o domínio da competência paraverbal (silêncios, pausas e sinais de retroação).

Neste sentido, são evidentes as potencialidades de um *corpus* oral deste tipo. As interações verbais a gravar, mais ou menos extensas, conterão um manancial de dados que refletirá a forma como estes falantes lidam com as exigências de uma interação oral em curso, em que têm de dar consecução

aos seus objetivos comunicativos, integrá-los na sequência discursiva a decorrer, negociar e gerir tópicos, construir argumentos e contra-argumentos, manter a vertente relacional, adotar ora o papel de locutores, ora o papel de ouvintes. No fundo, o acervo exibirá a forma como os participantes interagem, defendem pontos de vista e contra-argumentam numa L2, bem como a forma como coconstroem, momento a momento, uma interação verbal. Neste âmbito particular, há, aliás, uma outra dimensão que deve ser salientada. Na verdade, este *corpus* oral possibilitará a compreensão da forma como as competências de cada um se conformam e se ajustam reciprocamente à medida que a própria interação vai decorrendo. Subscrevemos, assim, as palavras de Young (2011: 428), quando afirma que a competência interacional (IC) “is not the knowledge or the possession of an individual person, but is co-constructed by all participants in a discursive practice, and IC varies with the practice and with the participants.”

Outra das valências deste *corpus* oral – e seguramente uma das mais importantes – diz respeito aos processos de negociação de sentido, isto é, aos momentos de incompreensão, de ocorrência de erros pragmáticos e, até, de quebras comunicativas, que seguramente sucederão, e aos efeitos dessas dificuldades na própria interação (Walsh 2010), bem como à forma como os participantes ultrapassarão (ou não) esses incidentes. As estratégias de evitação e de paráfrase, os mecanismos de reformulação, as autocorreções, os pedidos de clarificação, os comentários de natureza metadiscursiva, o recurso ao *code-switching* ou os fenómenos de transferência a que recorrerão para remediar ou resolver esses momentos de impasse poderão constituir, para o investigador, informação relevante para compreender qual a relação entre as estratégias escolhidas e o nível de proficiência, por exemplo, e até para avaliar quais as características desta competência estratégica em cada uma das fases de desenvolvimento da interlíngua destes aprendentes.

Concomitantemente, um *corpus* deste tipo possibilitará também a análise das competências de natureza sociolinguística (exigidas pela contextualização inerente ao uso da língua) exibidas pelos aprendentes. Considerado o contexto de debate, enquanto confronto de pontos de vista (pelo menos parcialmente) divergentes sobre um determinado tópico, é previsível que surja a necessidade de discordar, de contra-argumentar, de contradizer, e, nessa tentativa de deslegitimar o discurso do(s) outro(s),

emergem as questões ligadas à cortesia, à potencial agressão à face do(s) outro(s), à necessidade de mitigar algum desse conflito ou, pelo contrário, de o extremar para vincar posições. Num contexto como este que aqui se desenha, será pertinente analisar a ocorrência de *face-threatening acts*, entre outro tipo de possíveis expressões de descortesia, mas também o recurso a estratégias atenuadoras e intensificadoras. A definição da relação interpessoal e, no fundo, o sistema de cortesia que vigora no início da interação, enquanto conjunto de direitos e deveres tacitamente aceite por todos os participantes, constituirão, de igual forma, um apelativo objeto de análise, sobretudo se os termos em que o debate decorre se forem alterando e negociando ao longo da interação.

Em rigor, o interesse pelas características e pelo funcionamento da interação, assim como pelos fatores que nela repercutem, já tinha atraído a atenção dos investigadores no âmbito das orientações mais interacionistas da área da SLA – veja-se o trabalho de Long (1981), por exemplo – mas o foco recaía, sobretudo, na interação verbal entre nativo e não nativo e a interação verbal entre dois não nativos, enquanto contexto de aquisição, ainda não tinha sido plenamente explorada. Por esta razão, um *corpus* de interações orais em PLNM pode revelar-se fundamental na investigação em L2, na medida em que possibilita aos investigadores a oportunidade de estudar o papel da própria interação na aprendizagem da língua alvo.

Quanto aos benefícios dos *corpora* orais para a área de investigação na área da SLA, é inegável que a análise destes textos/discursos autênticos permite dar mais consistência à investigação sobre o desenvolvimento das competências linguísticas do aprendente nas diferentes fases de construção da sua interlíngua. Com efeito, algumas das questões que têm estado sob o foco da SLA, pelo menos na sua linha mais cognitivista, podem ser reavaliadas se observadas à luz dos dados empíricos produzidos pelos aprendentes. O contraste entre textos/discursos de aprendentes de diferentes línguas maternas pode auxiliar a perceber a existência de tendências universais no processo de aprendizagem. A comparação das produções orais de aprendentes com a mesma língua materna, de diferentes níveis, pode permitir captar a influência dessa mesma língua no processo de aprendizagem, bem como avaliar o maior grau de complexidade e de fluência que essas produções exibem, à medida que o grau de proficiência

umenta, o que pode ajudar a definir parâmetros avaliativos e a perceber quais as etapas e qual a natureza do processamento da L2. De igual forma, a identificação do erro constitui um precioso auxiliar no acesso às áreas mais críticas da aprendizagem da língua alvo (Myles 2015), ao mesmo tempo que permitirá, se analisado em diferentes níveis de proficiência, retirar ilações acerca dos processos de fossilização. Um bom conjunto de metadados pode também constituir um manancial de informações para estudar variáveis que são relevantes nesta área de investigação, quer se trate de traços de natureza sociolinguística, como a idade, o percurso académico, o número e a ordem de aprendizagem de outras línguas para além da materna, ou contextual, como o modo de comunicação (McEnery *et al.* 2019).

Mas outras questões relacionadas com as teorias de aprendizagem podem também beneficiar do aporte trazido pelos *corpora* dos aprendentes e, sobretudo, dos *corpora* orais. A forma como estes falantes acedem à L2 e a usam em tempo real (Myles 2015) pode ser clarificada com a análise de dados orais não planeados, na medida em que o investigador tem, neste caso, acesso privilegiado ao conhecimento linguístico implícito do aprendente. De facto, a dinâmica e a pressão de uma interação oral não permitem ao aprendente nem a reflexão prévia nem a correção subsequente, razão pela qual MacWhinney (2021: 159), citando Myles (2015: 314) considera até que “[f]rom the viewpoint of SLA theory, oral production is ‘a better window into implicit knowledge’”.

3. Palavras finais

A constituição de um *corpus* de interações orais em PLN M assume grande relevância em vários planos. Em primeiro lugar, na área do ensino/aprendizagem do português como língua não materna. Observar e analisar o comportamento linguístico dos não nativos neste tipo de contexto interacional permitirá obter uma compreensão mais apurada e rigorosa de uma área de investigação em que há evidentes lacunas descritivas e que não é trabalhada, de forma sistemática, em contexto de instrução formal. Analisar a interação oral (informal) em L2 revela-se, assim, uma preciosa pista de trabalho tendente a evidenciar em que medida o processo de

ensino/aprendizagem de uma gramática da L2 serve (ou não) os propósitos e as exigências de uma verdadeira interação oral. Os dados resultantes dessa análise podem auxiliar na melhoria das práticas pedagógicas, bem como ajudar a repensar muitos materiais instrucionais.

Naturalmente, e para que os objetivos delineados neste estudo sejam plenamente atingidos é necessário haver um conhecimento sistematizado das características e do funcionamento do mesmo tipo de interações orais entre nativos. Embora alguns autores questionem a existência de uma 'norma nativa', a possibilidade de detetar desvios e problemas comunicativos no *corpus* de interações orais em PLNM só é possível através do contraste com um *corpus* proveniente de falantes nativos.

Em segundo lugar, é também possível abordar algumas das questões de investigação da área da *SLA* através da análise de um *corpus* de interações orais em PLNM, uma vez que os problemas e os erros nele detetados permitem aceder às estratégias usadas pelos aprendentes no processamento da L2.

Os benefícios que um *corpus* oral deste tipo pode trazer à investigação em PLNM são evidentes, razão por que é necessário continuar a recolher *corpora* orais de diferentes géneros e a enriquecer as bases de dados existentes. Ao avançar neste sentido, um *corpus* de natureza multimodal seria de toda a utilidade, para dar mais consistência às análises do oral, mas também um *corpus* longitudinal, na medida em que seria possível averiguar, a longo prazo, a evolução da competência interacional dos aprendentes. Uma outra via que se afigura essencial diz respeito à necessidade e à urgência de desenvolver novas ferramentas digitais que permitam a análise destes *corpora* orais de aprendentes.

REFERÊNCIAS

- Adolphs, S., & Knight, D. (2010). Building a spoken corpus: what are the basics? In A. O’Keeffe, & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 38-52). Routledge.
- Aijmer, K. (2020). Contrastive pragmatics and corpora. *Contrastive Pragmatics*, 1(1), 28-57.
- Aijmer, K., & Rühlemann, C. (Eds.). (2015). *Corpus Pragmatics: A Handbook*. Cambridge University Press.
- Atkinson, D. (Ed.). (2011). *Alternative Approaches to Second Language Acquisition*. Routledge.
- Ballier N., & Martin, P. (2015). Speech annotation of learner corpora. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 107-134). Cambridge University Press.
- Bardovi-Harlig, K. (2013). Developing L2 pragmatics. *Language Learning*, 63(s1), 68-86.
- Bell, P., & Payant, C. (2021). Designing Learner Corpora: Collection, Transcription, and Annotation. In N. Tracy-Ventura, & M. Paquot (Eds.), *The Routledge Handbook of Second Language Acquisition and Corpora* (pp. 53-67). Routledge.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243-257.
- Celce-Murcia, M. (2007). Rethinking the role of communicative competence in language teaching. In E. Alcón Soler, & M. P. Safont-Jordà (Eds.), *Intercultural Language Use and Language Learning* (pp. 41-58). Springer.
- Díaz-Negrillo, A., & Thompson, P. (2013). Learner corpora: Looking towards the future. In A. Díaz-Negrillo, N. Ballier, & P. Thompson (Eds.), *Automatic Treatment and Analysis of Learner Corpus Data* (pp. 9-30). John Benjamins.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford University Press.
- Ferreira, T.; Santos, I.; Carapinha, C.; Martins, C.; Pereira, I.; Rio-Torto, G.; Pereira, R.; Inverno, L.; Ferreira, C.; Sousa, S. & Chapouto, S. (no prelo). *Building a spoken learner corpus of Portuguese L2: objectives and difficulties*.
- Flowerdew, J. (2009). Corpora in Language Teaching. In M. H. Long, & C. J. Doughty (Eds.), *The Handbook of Language Teaching* (pp. 327-350). Blackwell.
- Gilquin, G. (2015). From design to collection of learner corpora. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 9-34). Cambridge University Press.

- Golato, A. (2017). Naturally Occurring Data. In A. Barron, Y. Gu, & G. Steen (Eds.), *The Routledge Handbook of Pragmatics* (pp. 21-26). Routledge.
- Granger, S. (2002). A bird's eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 3-33). John Benjamins.
- Granger, S. (2008). Learner corpora in foreign language education. In N. Van Deusen-Scholl, & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education* (Vol. 4, pp. 337-351). Springer.
- Granger, S., Dagneux, E., & Meunier, F. (2002). *The International Corpus of Learner English. Version 1.1*. Handbook and CD-ROM. Presses Universitaires de Louvain.
- Hondo J. (2014). Synthesizing Social and Cognitive Approaches in SLA. *Working Papers in Educational Linguistics*, 29(2), 1-23.
- Johansson, S. (2007). Using Corpora: from learning to research. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the Foreign Language Classroom: Selected Papers from the Sixth International Conference on Teaching and Language Corpora (TaLC 6)* (pp. 17-30). Rodopi.
- Jucker, A. H. (2018). Data in pragmatic research. In A. H. Jucker, K. P. Schneider, & W. Bublitz (Eds.), *Methods in Pragmatics* (pp. 3-36). De Gruyter Mouton.
- Kramsch, C. (1986). From language proficiency to interactional competence. *The Modern Language Journal*, 70(4), 366-372.
- Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McNery, & G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 1-23). Longman.
- Long, M. H. (1981). Input, interaction, and second language acquisition. *Annals of the New York Academy of Sciences*, 379(1), 259-278.
- Lozano, C., & Mendikoetxea, A. (2013). Learner corpora and Second Language Acquisition: The design and collection of CEDEL2. In A. Díaz-Negrillo, N. Ballier, & P. Thompson (Eds.), *Automatic Treatment and Analysis of Learner Corpus Data* (pp. 65-100). John Benjamins.
- Lozano, C., Teixeira, J., & Madeira, A. (2021). Corpora and L2 acquisition: the L1 Portuguese – L2 Spanish subcorpus of CEDEL2. *Revista da Associação Portuguesa de Linguística*, (8-10), 137-154.
- Luzón, M. J., Campoy, M. C., Sánchez, M. M., & Salazar, P. (2007). Spoken Corpora: New Perspectives in Oral Language Use and Teaching. In M. C. Campoy, & M. J. Luzón (Eds.), *Spoken Corpora in Applied Linguistics* (pp. 3-26). Peter Lang.

- Mackey, A., & Gass, S. (2005). *Second Language Research: Methodology and Design*. Lawrence Erlbaum.
- MacWhinney, B. (2021). TalkBank for SLA. In N. Tracy-Ventura, & M. Paquot, (Eds.), *The Routledge Handbook of SLA and Corpora* (pp. 158-172). Routledge.
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics. An Introduction*. Edinburgh University Press.
- McEnery, T., Brezina, V., Gablasova, D., & Banerjee, J. (2019). Corpus Linguistics, Learner Corpora, and SLA: Employing Technology to Analyze Language Use. *Annual Review of Applied Linguistics*, 39, 74-92.
- Mendes, A., Antunes S., Janssen, M., & Gonçalves, A. (2016). The COPLE2 Corpus: A Learner Corpus for Portuguese. In *Proceedings of the Tenth Language Resources and Evaluation Conference – LREC'16* (pp. 3207-3214). European Language Resources Association.
- Mondada, L., & Doehler, S. P. (2004). Second Language Acquisition as Situated Practice: Task Accomplishment in the French Second Language Classroom. *Modern Language Journal*, 88(4), 501-518.
- Mukherjee, J. (2006). Corpus linguistics and language pedagogy: the state of the art – and beyond. In S. Braun, K. Kohn, & J. Mukherjee (Eds.), *Corpus Technology and Language Pedagogy* (pp. 5-24). Peter Lang.
- Myles, F. (2015). Second language acquisition theory and learner corpus research. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *Cambridge Handbook of Learner Corpus Research* (pp. 309-331). Cambridge University Press.
- Nascimento, M. F. B. (2002). O lugar do corpus na investigação linguística. In A. Mendes, & T. Freitas (Orgs.), *Actas do XVIII Encontro da Associação Portuguesa de Linguística* (pp. 601-605). APL.
- Orletti, F. (1984). Some methodological problems in data gathering for discourse analysis. *Journal of Pragmatics*, 8(4), 559-567.
- Philp, J., Adams, R., & Iwashita, N. (2014). *Peer interaction and second language learning*. Routledge.
- Richards, J. C. (2006). *Communicative Language Teaching Today*. Cambridge University Press.
- Romero-Trillo, J. (2017). Editorial. *Corpus Pragmatics*, 1(1), 1-2.
- Romero-Trillo, J. (Ed.). (2008). *Pragmatics and corpus linguistics. A mutualistic entente*. Mouton de Gruyter.
- Rühlemann, C. (2010). What can a corpus tell us about pragmatics? In A. O'Keeffe, &

- M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 288-301). Routledge.
- Rühlemann, C., & Clancy, B. (2018). Corpus linguistics and pragmatics. In C. Ilie, & N. Norrick (Eds.), *Pragmatics and its Interfaces* (pp. 241-266). John Benjamins.
- Santos, I. A., Pereira, I., Martins, C., Lopes, A. C. M., Carapinha, C., & Silva, A. (2016). *Corpus* oral de PL2: um novo recurso para a investigação e ensino. *Revista da Associação Portuguesa de Linguística*, (1), 745-760.
- Taguchi, N., & Roever, C. (2017). *Second language Pragmatics*. Oxford University Press.
- VanPatten, B., Keating, G. D., & Wulff, S. (2020). *Theories in Second Language Acquisition: An Introduction* (3.^a ed.). Routledge.
- Vaughan, E., & Clancy, B. (2013). Small Corpora and Pragmatics. In J. Romero-Trillo (Eds.), *The Yearbook of Corpus Linguistics and Pragmatics* (Vol. 1, pp. 53-73). Springer.
- Walsh, S. (2010). What features of spoken and written corpora can be exploited in creating language teaching materials and syllabuses? In A. O’Keeffe, & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 333-344). Routledge.
- Young, R. F. (2011). Interactional Competence in Language Learning, Teaching and Testing. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (Vol. 2, pp. 426-443). Routledge.

Algumas notas sobre a quantificação de adjetivos de cor por *muito* e *pouco*¹²

Inês Cantante
cantante.ines@gmail.com
FLUP/CLUP (Portugal)

ABSTRACT.

The present investigation aimed to assess whether color adjectives, considering that they are qualificatives, could accept gradation and, if so, what type of scale they would be associated with. The paper begins by focusing on some of the existing literature on gradable adjectives, presenting, afterwards, some works specifically focused on color adjectives. The quantification of color adjectives was analyzed through the combination of the selected adjectives (primary colors: *yellow*, *blue* and *red* plus *black* and *white*) with the quantifiers *muito* and *pouco*. The results show that, in EP, the semantic investigation of this type of adjective is complex, since the readings obtained were not linear. In fact, there is, on the one hand, a typically gradable reading, which promotes a movement of the adjectives to a higher or lower point of the scale. It is possible, on the other hand, to obtain metaphorical readings, whose non-literal uses can have diverse meanings, which can either refer to physical or emotional states, skin tones, or even meanings that are very specific for EP, such as, for example, *to make a yellow smile*. It is thus difficult to evaluate the type of scale associated with these adjectives, since, in the case of intensity readings, there may indeed be an approximation to a representative prototype of the color, but the same does not apply to metaphorical readings.

KEYWORDS.

Color adjectives; quantification; gradability; quality reading; metaphoric reading.

RESUMO.

A presente investigação teve como objetivo avaliar se os adjetivos de cor, sendo qualificativos, poderiam aceitar graduação e, se sim, qual o tipo de escala a que estariam associados. O trabalho começa por focar alguma da literatura existente sobre adjetivos graduáveis, focando-se, depois, particularmente, nos adjetivos de cor. O estudo da quantificação deste tipo de adjetivos é feito através da verificação da compatibilidade com os quantificadores *muito* e *pouco*. Os resultados mostram que, em PE, a investigação semântica deste tipo de adjetivo é complexa, já que as leituras obtidas não são lineares. De facto, se existe, por um lado, uma leitura tipicamente graduável, que promove uma movimentação dos adjetivos para um ponto

1 Gostaria de deixar o meu especial agradecimento aos avaliadores anónimos pelos seus comentários e sugestões, cujas contribuições foram fundamentais para melhorar a presente investigação.

2 O presente trabalho foi financiado pela Fundação para a Ciência e a Tecnologia, I.P. (FCT) (referência 2021.04998.BD).

mais elevado ou mais baixo da escala, é possível, por outro, obter leituras metafóricas, cujos usos não literais podem ter significados diversos, que tanto podem referir-se a estados, físicos ou psicológicos, como a tons de pele, passando por sentidos muito específicos do PE, como, por exemplo, *fazer um sorriso amarelo*. Assim, torna-se difícil avaliar o tipo de escala associado a estes adjetivos, já que, no caso das leituras de intensidade pode, sim, haver uma aproximação a um protótipo representativo da cor, mas o mesmo não se aplica às leituras metafóricas.

PALAVRAS-CHAVE.

Adjetivos de cor; quantificação; graduabilidade; leitura de qualidade; leitura metafórica.

1. Introdução

Os adjetivos de cor são tradicionalmente classificados como qualificativos, já que “denotam qualidades básicas, na sua maioria simples, que se aplicam a dimensões constitutivas dos vários tipos de entidades que os nomes podem denotar - p.e., forma, cor e tamanho relativamente aos objetos, características fisiológicas e disposição mental relativamente às pessoas etc.” (Velo 2013: 1362).

Na Gramática do Português, Velo (2013) considera quatro classes para os adjetivos: denotativos, avaliativos, modais e intensionais. É na primeira classe, a dos denotativos, que a autora inclui os adjetivos qualificativos e os relacionais, classes tradicionalmente aceites na literatura, embora, na maioria das vezes, tratadas como classes autónomas. Quanto aos adjetivos de cor, por sua vez, a autora inclui-os na classe dos qualificativos por considerar que estes adjetivos, em particular, atuam sobre um nome no sentido de denotar “propriedades que incidem sobre aspetos variados da entidade representada por esse nome” (p. 1368).

Considerando que podem referir-se a “propriedades associadas a dimensões cujos valores podem variar numa escala quantitativa, como a altura e a temperatura” (Velo 2013: 1385)³, os adjetivos qualificativos aceitam graduação (cf. (1)), ao contrário dos adjetivos relacionais (cf. (2)). A graduabilidade dos adjetivos tem sido um tema bastante trabalhado por vários autores, entre os quais se destacam Kennedy & McNally (2005, 2010),

³ Uma propriedade que caracteriza os adjetivos de cor, especificamente, é o facto de que estes “exprimem propriedades “primitivas” das entidades representadas pelo nome modificado, ou seja, propriedades que se aplicam a diversas dimensões pré-determinadas que estruturam os vários tipos de entidades” (Velo 2013: 1373).

Kennedy (1999, 2007), Kennedy & Levin (2008), entre muitos outros.

- (1) Este vestido é **muito** bonito.
- (2) *Esta guerra é **muito** europeia.

Assim, considerando que os adjetivos de cor pertencem à classe dos adjetivos qualificativos, exemplos como (3) mostram que parece possível quantificar este tipo de adjetivos, em Português Europeu (doravante PE).

- (3) Este vestido é **muito** vermelho.

Por outro lado, não existem, para o PE, muitos trabalhos que estudem a graduabilidade dos adjetivos de cor⁴. Por essa razão, a presente investigação pretende verificar se estes adjetivos são graduáveis ou não, bem como analisar sob que circunstâncias, i.e., em que contextos aparecem e, conseqüentemente, que leituras originam.

Numa primeira parte, far-se-á uma breve introdução acerca dos adjetivos, em particular os graduáveis, seguindo-se uma breve apresentação de alguns trabalhos e autores que focaram, particularmente, os adjetivos de cor. Em terceiro lugar, na secção seguinte, serão analisados exemplos do PE, que apresentam adjetivos de cor quantificados com *muito e pouco*, retirados da base de dados CetemPúblico. Estes exemplos serão analisados em simultâneo com uma breve reflexão acerca das leituras que originam e quais os fatores que as determinam, o que possibilitará, numa secção final, uma breve discussão sobre os dados e os resultados obtidos, juntamente com algumas observações finais.

⁴ Na verdade, para o Português, tanto Europeu como Brasileiro, os trabalhos sobre adjetivos de cor focam, sobretudo, questões morfológicas, relacionadas com a designação da cor, essencialmente de um ponto de vista do léxico, não focando os adjetivos de cor de uma perspetiva semântica, que se preocupe com os contextos de ocorrência e com as interpretações geradas.

Para o PE, Correia & Barbosa (2013) apresentam uma investigação com base nos nomes de cor do português, sob o ponto de vista da sua estrutura interna. O estudo permitiu-lhes concluir que, se, por um lado, “os nomes de cor básicos denominam seja a cor de forma genérica, seja a cor focal”, por outro, “as palavras construídas a partir desses nomes permitem a nomeação de variantes da cor básica em função das suas três principais dimensões: o tom, a luminosidade e a saturação” (Correia & Barbosa 2013: 402). Apesar de ser uma investigação pertinente, que foca algumas questões morfológicas e semânticas destes constituintes, o enfoque é maioritariamente lexicológico, procurando contribuir para a investigação acerca da denominação da cor, pelo que não corresponde, exatamente, ao objetivo do presente estudo. Para o Português do Brasil (PB), Zavaglia (1998, 2006, 2008) contribuiu bastante para o mesmo tipo de investigação, relacionada com os adjetivos de cor, estando a autora associada à apresentação de um modelo de dicionário temático trilingue (português-italiano-francês), precisamente relativo aos adjetivos de cor.

2. Reflexões sobre a graduabilidade dos adjetivos de cor

De um modo geral, pode dizer-se que os adjetivos graduáveis denotam propriedades que podem ser organizadas em escalas, que, segundo Kennedy & McNally (2005: 349), “map their arguments onto abstract representations of measurement, or degrees”. Assim, as escalas organizam os valores em graus, definidos como “points or intervals partially ordered along some dimension (e.g., height, cost, weight and so forth)” (Kennedy & McNally 2005: 349).

Sobre os adjetivos graduáveis, Kennedy & McNally (2005) notam que a sua interpretação depende, muitas vezes, do contexto, já que “what counts as *tall* or *expensive* may vary from context to context” (*ibidem*: 348). Considerando que as escalas associadas a este tipo de adjetivos podem estar relacionadas com três tipos de parâmetros – graus, dimensão de medição e relação de ordenação (cf. Kennedy & McNally 2005) –, as escalas deverão distinguir-se, precisamente, nesses três parâmetros, o que leva os autores a propor uma distinção entre adjetivos absolutos e relativos.

De facto, em ambos os casos, os adjetivos são sensíveis ao contexto, embora de forma diferente. Com efeito, considerando-se os adjetivos relativos, “different contexts pick out different sets of individuals of which the adjective can be truthfully predicated” (Aparício *et al.* 2015: 414). No entanto, a variação contextual associada aos adjetivos absolutos está relacionada com a distância em relação ao topo (máximo ou mínimo) da escala. Por outras palavras, esta variação contextual deverá ser menos restritiva com adjetivos relativos, como *alto*, do que com adjetivos como *cheio*, i.e., absolutos.

Tal facto poderá estar relacionado com o tipo de escala⁵ a que estes adjetivos estão associados: os adjetivos relativos representam escalas abertas e os adjetivos absolutos, escalas fechadas. Significa isto que, no primeiro caso, a escala não tem limites mínimo e máximo, pelo que não tem capacidade para, por si só, fornecer um *standard* de comparação, o que explica que este tenha de ser dado pelo contexto (cf. Aparício *et al.*

5 Note-se que, conforme notado por Kennedy & McNally (2005), os adjetivos podem estar associados a escalas abertas, i.e., sem limites, ou fechadas, que são limitadas em, pelo menos, uma das pontas. Na verdade, os autores reconhecem a possibilidade de quatro tipos de escalas: completamente fechadas (contêm limite máximo e mínimo), fechadas com limite mínimo, fechadas com limite máximo ou abertas (sem limite mínimo nem máximo).

2015: 415). Pelo contrário, os adjetivos absolutos, tipicamente associados a escalas fechadas, podem ter valor máximo e mínimo ou, pelo menos, um dos dois. Assim, nestes casos, a variabilidade contextual parece ser mais reduzida, uma vez que apenas pode dizer respeito a valores fornecidos pela própria escala.⁶ De uma forma geral, parece, de facto, possível afirmar que “gradable adjectives associated with totally open scales have relative standards; gradable adjectives that use totally or partially closed scales have absolute standards (Kennedy & McNally 2005: 361).

2.1. O caso dos adjetivos de cor

Kennedy & McNally (2010), ao olhar para um exemplo como (4), abaixo, refletem sobre o facto de os adjetivos de cor poderem estar associados a diferentes valores de verdade. De facto, segundo os autores, a frase pode ser verdadeira para um pintor/artista, para quem apenas a cor visível das *folhas* importa, mas não para um botânico, que deverá dar prioridade à constituição interna das *folhas* no momento de analisar a veracidade de (4), considerando-a, por isso, falsa.

(4) As folhas são **verdes**. (retirado e adaptado de Kennedy & McNally 2010)

Nesta linha encontram-se vários autores que têm tentado explicar que os adjetivos de cor apresentam variação contextual (cf. Travis 1985, 1994, 1997; Szabó 2001; Rothschild & Segal 2009; Clapp 2012; Hansen 2011, Hansen & Chemla 2017; Kennedy & McNally 2010; Zavaglia 2006, e.o.).

Szabó (2001), por exemplo, considera que o princípio da composicionalidade⁷ poderá ajudar a resolver esta questão. Para explicar a possibilidade de associação de diferentes valores de verdade ao mesmo enunciado, em diferentes contextos, o autor propõe a existência de uma

⁶ Note-se, ainda assim, que os adjetivos associados a escalas fechadas parecem apresentar alguma variação quanto ao valor do *standard* de comparação, já que “the standards for close-scale adjectives default to an endpoint of the scale: the minimum in some cases (e.g. awake and open), the maximum in others (e.g. asleep and straight)” (Kennedy & McNally 2005: 361)”.

⁷ O princípio da composicionalidade, atribuído a Frege (1892), propõe que o significado de uma expressão complexa é definido através dos significados das suas partes constituintes. Neste princípio, é também dada relevância à estrutura, responsável pela forma como os significados se combinam.

variável presente na denotação dos adjetivos de cor, que deverá funcionar como parâmetro, cujo valor será fornecido pelo contexto (cf. Szabó 2001).

Rothschild & Segal (2009), por sua vez, consideram os adjetivos de cor como expressões linguísticas cujo significado muda de contexto para contexto, tal como acontece com “the classic indexicals such as ‘I’, ‘here’, ‘now’, ‘that’, ‘he’, and ‘this’” (Rothschild & Segal 2009: 5). Por essa razão, os autores defendem que devem incluir-se, nesta classe, os adjetivos de cor. No entanto, Kennedy & McNally (2010) consideram que tal proposta não deve ser a resposta para o problema da ambiguidade contextual associada aos adjetivos de cor, já que não dá conta das restrições que condicionam os diferentes significados nos vários contextos (cf. Kennedy & McNally 2010: 6).⁸

Os mesmos autores mostram que o estudo dos adjetivos de cor é, de facto, complexo, quando afirmam que “the interpretation of particular color adjectives can vary systematically in ways that cannot be captured merely by reference to the part structure of their arguments (or comparison classes)” (Kennedy & McNally 2010: 5). Os autores propõem, assim, a distinção entre uma leitura graduável, que relaciona os adjetivos de cor com graus (em que a cor é manifestada num determinado objeto), e uma leitura não graduável, na qual se foca a presença/ausência da cor em questão num determinado objeto. Interessa, assim, saber se um determinado objeto apresenta ou não a cor, não importando, por isso, a medida em que tal acontece. Não se trata, portanto, de um conceito escalar, mas apenas de saber se a cor se manifesta ou não (cf. Kennedy & McNally 2010).

Sob a perspetiva de uma leitura graduável – em que os adjetivos representam uma função que associa objetos a graus, numa escala –, esta escala pode ser variável, representando quantidade de cor, por um lado, ou qualidade da cor, por outro. A leitura de quantidade relaciona-se com a atribuição da cor a um objeto numa determinada quantidade (i.e., *quanto* do objeto é daquela cor) e está associada a uma escala fechada, já que “in principle there exist both minimal and maximal degrees to which the property can be ascribed” (Kennedy & McNally 2010: 14). Já a leitura de qualidade representa a aproximação da

8 Outros autores apresentam diferentes propostas, como Clapp (2012), que olha para os adjetivos de cor como representando adjetivos absolutos, i.e., com escala fechada, mas apenas de *standard* mínimo, o que significa que, para que um objeto seja considerado, por exemplo, *verde*, apenas tem de apresentar a cor num determinado grau (que corresponderá ao grau mínimo da escala). Também McNally (2011) considera os adjetivos de cor como absolutos, mas não de *standard* mínimo, antes representando um valor mais próximo do meio da escala.

cor de um objeto a um determinado ‘protótipo’ representativo da cor atribuída pelo adjetivo. Também Rosch (1975: 193) nota que “color categories appear to be represented in cognition not as a set of criterial features with clear-cut boundaries but rather in terms of a prototype (the clearest cases, best examples) of the category surrounded by other colors of decreasing similarity to the prototype and of decreasing degree of membership” (Rosch 1975: 193). Assim, parece haver uma tendência para considerar que um protótipo deverá definir o grau de correspondência com o valor-padrão, já que, segundo Taylor (2003: 43), a aceitação de uma entidade como pertencente a uma dada categoria, está dependente da distância ao protótipo: quanto mais afastada estiver deste protótipo, menor deverá ser o grau de aceitação. Ainda assim, parece possível afirmar que os protótipos podem ser de dois tipos: focos de cor (‘focal colors’, segundo Kay & McDaniel (1978)), que correspondem, em termos cromáticos, às cores puras, i.e., ao que conceptualizamos acerca das cores, e que permite explicar que, por exemplo, *azul*, fisicamente (i.e., em termos cromáticos), corresponda sempre à mesma cor, independentemente da língua utilizada, como nota Taylor (2003: 53), ao afirmar que “even though different languages might carve up the color space differently, speakers of different languages tend to agree on the focal reference of their respective color terms”; ou, por outro lado, a informação de protótipo de uma cor pode estar relacionada com determinados pontos de referência, como, por exemplo, neve, para *branco*, sangue, para *vermelho*, ou ouro, para *amarelo*.

A leitura de qualidade deverá, assim, estar associada a uma escala fechada, em que o grau máximo corresponderá a um “perfect match with the prototype”⁹ (Kennedy & McNally 2010: 15). Se, no primeiro caso (leitura de quantidade), os adjetivos se podem combinar com os designados *proportional modifiers* (cf. Kennedy & McNally 2005)¹⁰, i.e.,

9 Esta ideia de correspondência com um protótipo tinha já sido notada por Sapir (1944), que refere que as dimensões da graduabilidade dos adjetivos de cor podem estar relacionadas, não apenas com a intensidade da cor, mas também com o “size of surface or volume characterized as red” e, talvez até mais importante, com o grau de conformidade em relação a “some accepted standard of redness” (Sapir 1944: 94).

10 Kennedy & McNally (2005) utilizam os designados *proportional modifiers* (*half, most, most of the way*), como uma forma de identificar o tipo de escala associada aos adjetivos graduáveis. De facto, como afirmam, “proportional modifiers fix the value of an adjective’s degree argument relative to maximum/minimum degrees on an adjective’s scale, and so are limited in their distributions to closed-scale adjectives” (Kennedy & McNally 2005: 369). Assim, adjetivos compatíveis com *proportional modifiers*, deverão representar uma escala fechada (cf. (i)), enquanto adjetivos incompatíveis, deverão estar ligados a escalas abertas (cf. (ii)).

(i) A caixa está **meio** aberta.
(ii) *A estrada está **meio** longa.

advérbios como *totalmente*, *completamente* ou *meio*, no segundo (leitura de qualidade), os adjetivos deverão aceitar a combinação com advérbios como *perfeitamente*. Vejam-se, a propósito desta diferença, os exemplos abaixo, retirados e adaptados de Kennedy & McNally (2010).

(5) A camisola do bebé era **completamente** azul, mas poderia ser de um tom de azul mais forte. = leitura de quantidade

(6) Os olhos do bebé eram **perfeitamente** azuis; não poderiam ser mais azuis. = leitura de qualidade

(Kennedy & McNally 2010: 16)

Kennedy & McNally (2010) defendem, assim, que os adjetivos de cor devem ser tratados, em termos semânticos, como os restantes adjetivos graduáveis, associados a uma escala, caso em que podem denotar uma leitura de quantidade ou de qualidade.

No entanto, deve ser considerada, também, a possibilidade de uma leitura não graduável, que funciona, de certa forma, como uma leitura classificatória, já que representa objetos ou indivíduos que, ou têm a propriedade denotada pelo indivíduo, ou não têm. Os autores defendem que esta leitura reflete a existência de uma correlação entre propriedades. Retomando o exemplo (1), relativo às folhas verdes, os autores defendem que a diferença entre a leitura graduável e não graduável permite distinguir as folhas “on the basis of why they are green (e.g. chlorophyll vs. paint)” ou “how they are green (depth of hue, proximity to a prototype, extent of the object, etc.)” (Kennedy & McNally 2010: 89).

Também Hansen (2011) e Hansen & Chemla (2017) aceitam a distinção entre adjetivos relativos, cuja avaliação é feita segundo “context sensitive standards (of, e.g., duration or price)”, e absolutos, cujos pontos de referência têm um valor fixo (Hansen & Chemla 2017: 240). No entanto, Hansen (2011) começa por rejeitar que os adjetivos de cor devam ser tratados como absolutos, acedendo, no entanto, à ideia de que estes últimos são menos sensíveis ao contexto do que os relativos (cf. Hansen & Chemla 2017: 240). De facto, num primeiro momento, considera que os adjetivos de cor se comportam como adjetivos relativos, porque, segundo este autor, as cores são, elas próprias, uma forma de propriedade relacional, por representarem

“relations between physical properties of objects, observation conditions, and types of observers” (Hansen 2011: 214).¹¹ Posteriormente, no entanto, reconhece que os adjetivos de cor não devem ser tratados como relativos (cf. Hansen & Chemla 2017), acabando por confirmar as intuições de Kennedy & McNally (2010).

Hansen & Chemla (2017) consideram, assim, a possibilidade de três grupos, cuja diferença reside na localização do valor-padrão na escala: *standard* mínimo, meio da escala ou *standard* máximo, o mesmo não acontecendo para a leitura qualitativa, que não apresenta qualquer tipo de padrão comportamental (cf. Hansen & Chemla 2017). Aceitando a existência de um morfema (nulo) para as formas positivas dos adjetivos – *pos* –, os autores consideram que os valores dos *standards* de comparação estão já no significado do adjetivo, uma vez que fazem parte do seu significado lexical, podendo combinar-se, podendo combinar-se com *pos* para “generate different truth conditions” (cf. Hansen & Chemla 2017: 243), sendo estes valores fixos, no caso dos adjetivos absolutos, e variáveis, consoante o contexto, nos adjetivos relativos.

Em relação à leitura quantitativa, acabam por concluir que os adjetivos de cor devem, efetivamente, ser analisados como adjetivos absolutos, não havendo, no entanto, uma resposta conclusiva em relação à localização do *standard* de comparação (ponto mínimo, máximo ou intermédio da escala). Os autores explicam esta variação recorrendo a um fator que denominam de *intolerância à imprecisão*, considerando que, ainda que possa haver um valor fixo para o *standard*, “there is variation among the adult population in how much imprecision is tolerated” (Hansen & Chemla 2017: 264). No entanto, salientam que a própria escala associada aos adjetivos de cor pode ser sensível ao contexto, mesmo que o *standard* se mantenha fixo num determinado ponto da escala (Hansen & Chemla 2017: 368).

De uma forma geral, para Hansen (2011), o estudo dos adjetivos de cor envolve vários fenómenos, nomeadamente a ambiguidade entre leituras graduáveis e não graduáveis (Kennedy & McNally 2010), a variação contextual que pode estar associada ao parâmetro da dimensão de medição

¹¹ O autor acaba por aproximar os adjetivos de cor aos predicados de gosto pessoal (*personal taste*), já que apresentam um elemento em comum, aquilo que o autor designou de ‘*experiencer parameter*’, cujo valor é dado pelo próprio contexto, em função do ‘*experienciador*’ (cf. Hansen 2011).

e, por fim, a natureza relacional dos adjetivos de cor, que representam relações entre as condições de observação e o observador (cf. Hansen 2011).

3. A graduabilidade dos adjetivos de cor em Português Europeu

Tendo em conta todas as considerações anteriores sobre adjetivos de cor, é objetivo deste trabalho analisar este tipo de adjetivos de um ponto de vista da sua graduabilidade. Assim, importa saber em que medida é que, sendo qualificativos, os adjetivos de cor, em PE, são graduáveis ou não. Importa, ainda, tentar perceber a que tipo de escala estão associados estes adjetivos: aberta (adjetivos relativos) ou fechada (adjetivos absolutos).

Neste estudo, a pesquisa foi restringida a cinco adjetivos. Para além de *branco* e de *preto*, foram selecionados os adjetivos que representam as cores primárias: *amarelo*, *vermelho* e *azul*. Todos os exemplos foram retirados do corpus CETEMPúblico.

A análise tentará responder às seguintes questões: qual a posição que os adjetivos de cor podem ocupar na frase (atributiva e predicativa), que tipo de objeto podem estes adjetivos qualificar e, por fim, quais as leituras que originam?

3.1. Graduabilidade das cores primárias em PE

Em PE, os adjetivos de cor podem ocorrer, na frase, em posição atributiva (cf. (7)) e predicativa (cf. (8)), como se confirma nos exemplos seguintes:

(7) A luz **muito amarela** dos candeeiros alumia os edifícios, todos muito iguais, dos arrabaldes de San Sebastian. (*par = ext601893-des-94a-2*)

(8) Ao secar, o lenço **ficou azul** porque o composto químico perdeu quatro das seis moléculas de água. (*par = ext230145-clt-soc-95b-1*)

No primeiro caso, a *luz muito amarela* refere-se, parece-me, ao tom da luz, que pode ter um tom mais ou menos *amarelado*. Já no segundo caso, a leitura é de cor literal, i.e., houve uma reação química que modificou a cor

do objeto *lenço*.

Apesar de não ser comum, foram, ainda, encontrados alguns casos em que os adjetivos de cor se combinam com o nome em posição atributiva pré-nominal, mantendo-se a interpretação de intensidade da cor (em (9), a interpretação é de *um vermelho muito forte/vivo*):

- (9) Chegam à caleche e o lado esquerdo do enquadramento é invadido pelo **muito vermelho** cabedal da carruagem. (*par=ext1209135-nd-95a-1*)

Este comportamento confirma que os adjetivos de cor devem ser tratados como adjetivos qualificativos e, por isso, graduáveis.

3.2. Adjetivos de cor em construções de grau: combinação com os quantificadores *muito/pouco*

Por serem muito utilizados em PE, os quantificadores *muito* e *pouco* poderão fornecer algumas pistas acerca do comportamento dos adjetivos de cor enquanto adjetivos graduáveis. Para as cores em análise, i.e., as cores primárias, os adjetivos que as representam parecem aceitar a combinação com o quantificador *muito*, como se pode verificar de seguida. Em todos os casos, a interpretação parece ser a de cor literal, mas quantificada, o que deverá refletir-se numa leitura (que será designada de intensidade), parafraseável por *um tom de (cor) muito forte*.

- (10) Foi sepultado a 15 e, porque perdera o sangue todo, estava seco como palha, **muito amarelo**, parecia uma figura de cera. (*par=ext1327878-soc-95a-2*)
- (11) Dos cestos das castanhas subia um fuminho muito simpático e **muito azul**, atribuído, enganosamente, por quem observava o espectáculo de longe, a granadas lacrimogéneas. (*par=ext322125-nd-94b-2*)
- (12) E simples: uma loura, pouco vestida, enlaçada por um gentleman, trincando uma maçã **muito vermelha**. (*par=ext712380-eco-93a-1*)

À partida, poder-se-ia pensar que os casos de *branco* e *preto* deveriam apresentar algumas dificuldades de aceitação com este quantificador, já que, representando o primeiro a totalidade das cores e o segundo a ausência de qualquer cor, deveriam corresponder, eles próprios, a topos de escala (i.e., graus máximos). Assim, a movimentação, promovida por *muito*, para pontos mais elevados na escala, deveria estar vedada a estes dois adjetivos, em particular, o que, de facto, não se verifica (cf. (13) e (14)).

- (13) «Um, um», diz, abana a cabeça e ri, os dentes **muito brancos**.
(par = ext353044-soc-98b-2)
- (14) Vestidos **muito pretos**, fita vermelha no chapéu, cravo vermelho bem firme na mão, dançam como a idade deixa, mas com mais convicção que a maioria dos jovens que as rodeiam.
(par = ext378429-pol-94b-2)

Em todos os exemplos apresentados acima, a interpretação é, também, a de graduação, propriamente dita, i.e., os quantificadores atuam, neste caso, no sentido de aumentar o grau de intensidade da cor mencionada. Dito por outras palavras, o que parece estar a ser avaliado, em termos gradativos, nos exemplos acima, é a intensidade da cor, “based on the physical and perceptual characteristics of color” (Kennedy & McNally 2010: 93). Esta avaliação, no entanto, nem sempre se aplica de forma literal, i.e., existem certos usos em PE, associados aos adjetivos de cor e às construções em que estes aparecem, que têm sentidos ligeiramente diferentes – metafóricos. Note-se o caso de *amarelo*, abaixo.

- (15) Um sorriso **muito amarelo**. (par = ext1020861-clt-93a-1)

Aplicado a *sorriso*, *amarelo* descreve, em PE, *um sorriso* que não tem verdadeiramente razão de o ser e que, por isso, é falso, normalmente associado a situações de desconforto ou embaraço. No caso particular do exemplo (15), a utilização do quantificador reforça, precisamente, esta situação.

Azul apresenta também alguns usos não literais, que podem fazer referência, por exemplo, aos tons apresentados pelo *mar* (cf. (16) e pelo *céu*

(cf. (17)), muitas vezes difíceis de definir e associados a noções de calma, profundidade, entre outras. Tratam-se, neste caso, de usos metonímicos do adjetivo.

- (16) Luís Pavão traz as recordações **muito azuis** dos «Mergulhos na Praia Grande» que fez e fotografou, com uma máquina especial, à prova de água, entre Julho e Agosto e 1988. (*par* = *ext400440-nd-91a-2*)
- (17) O sábado passava **muito azul** e muito quente. (*par* = *ext174293-pol-95a-2*)

No caso de (16), a expressão *recordações muito azuis* é uma construção metafórica que faz, precisamente, referência à cor do mar. O mesmo acontece com (17), em que a atribuição da característica *muito azul* à forma de passar de um *Sábado* é, claramente, metafórica, fazendo-se referência ao conceito de um dia quente (efetivamente referido logo de seguida), com céu limpo e, por isso, azul (i.e., sem nuvens).

Em casos muito específicos, em PE, *azul* pode ainda fazer referência a um estado de raiva, como o demonstra o exemplo abaixo.

- (18) Só com a ideia, Pinto da Costa **ficou azul**. (*par* = *ext591817-pol-94b-1*)

Já *vermelho*, por sua vez, tem, em termos metafóricos e metonímicos, um vínculo bastante marcado com questões políticas, sendo, normalmente, utilizado para referir partidos ou ideologias políticas mais ligadas à comumente designada 'esquerda' (cf. (19)). Pode, ainda, ser associado a estados emocionais (cf. (20))¹²:

- (19) (...) Só que naquela altura estávamos numa época vermelha, **muito vermelha** – Politicamente falando, é? (...). (*par* = *ext143689-pol-92b-3*)
- (20) Fulo, **vermelho de raiva**, gritou o seu inconformismo, agora libertador (*par* = *ext237195-pol-95b-2*)

¹² Note-se que, apesar de não se analisarem todos os adjetivos de cor no presente trabalho, adjetivos relativos a outras cores podem apresentar o mesmo tipo de comportamento, como é o caso de verde, abaixo:

(i) Um aspecto de relevância suficiente para fazer qualquer lisboeta ficar **verde de inveja**. (*par* = *ext157087-eco-95b-1*)

Contudo, quando modificado por *muito*, este adjetivo não parece apresentar, de forma tão visível, este vínculo.¹³ Na verdade, esta construção quantificada parece ser bastante utilizada para fazer referência a uma pessoa que se encontre corada, seja por questões de embaraço (cf. (21)), de raiva (cf. (22)), por ter feito um esforço superior ao normal (cf. (23)) ou, ainda, por ter a pele queimada (cf. (24)), casos em que a leitura deixa de ser apenas metafórica, passando a representar, também, usos metonímicos do adjetivo.

- (21) «Ah, não parlas?, pois devias», desdenhou um, indiferente ao espanto de um casal muito louro e **muito vermelho**, que se esforçava por compreender porque é que um bando de jovens pulava à sua volta e os bombardeava com perguntas. (*par = ext868982-nd-91a-1*)
- (22) Decreta 24 contos de multa ou 40 dias de prisão, mas o advogado levanta-se (**muito vermelho**) e pede recurso. (*par = ext1026157-soc-94a-2*)
- (23) A transpirar, **muito vermelho**, que o esforço não fora pouco, Luís Silva põe um ar compenetrado: (*par = ext1478822-clt-soc-95a-1*)
- (24) Sam Greenfield, um americano enorme, com quase dois metros de altura e **muito vermelho** pelos dois dias de sol na piscina do hotel Riviera, veio na primeira leva de «invasores», (...). (*par = ext1499684-des-91b-1*)

Os adjetivos *preto* e *branco*, mantendo as interpretações tipicamente associadas ao aumento do grau de intensidade da cor (cf. (25)), acrescentam a esta lista de interpretações a possibilidade de considerar variações para o tom de pele (cf. (25a)), por vezes associadas a questões étnicas (25b.), às quais, em alguns casos, se associam opiniões pessoais se associam opiniões pessoais, como a ironia representada em (26), através do uso do diminutivo (*familiazinha*). Neste último caso, a interpretação parece ser semelhante a

13 Se, no entanto, se tiver em atenção construções comparativas, com *mais* e/ou *menos*, a tendência para a associação do *vermelho* com a política volta a emergir, como se comprova nos exemplos abaixo. No entanto, por não ser o foco da presente investigação, este tipo de construção terá de ser analisado em trabalhos posteriores.

- (ii) Apenas a remota esperança de conservar algum poder político nos distritos alentejanos, progressivamente **menos vermelhos**, leva os comunistas a acreditar nas regiões-distritos. (*par = ext631579-soc-93a-2*)
- (iii) Há mesmo quem garanta que a influência de Carvalhas sobre o presidente da Câmara do município **mais vermelho** do país não foi alheia à sua recandidatura como cabeça-de-lista da CDU. (*par = ext1407363-pol-93a-1*)

temos muitas características/comportamentos de família caucasiana (i.e., branca).

- (25) Sabes que a tua árvore, a magnólia, já esta cheia de flores **muito brancas**? (*par = ext228570-soc-92b-2*)
- Até porque alguns de nós somos **muito brancos**» de pele. (*par = ext588806-soc -98a-1*)
 - Sem tirar o chapéu, mostrando umas pernas **muito brancas** a sair das calças arregaçadas, o homem lá avança contra as vagas purificadoras. (*par = ext1072634-clt-92b-2*)
- (26) Sem esquecer que a diversidade existe também entre os «portugueses»: ao contrário do que os governos pensam quando dizem que devem governar para o «consenso», não somos todos familiazinhas **muito brancas**, mui católicas e mui assustadas com as mudanças deste mundo. (*par = ext69868-opi-96b-2*)

Além disso, a cor branca é, muitas vezes, associada ao tom pálido ou descolorido da pele, principalmente do rosto, quando alguém está fisicamente debilitado (cf. (27)).

- (27) Foi então que notei que estava **muito branco** e deitava espuma pelo nariz.” (*par = ext429740-soc-92a-2*)

Por sua vez, *preto*, além da interpretação associada à graduação (cf. (28)), e ao tom de pele (cf. (29))¹⁴, tem, também, em PE, uma interpretação essencialmente associada a um acontecimento negativo, ou uma situação menos boa, como fica patente em (30), através do uso da expressão *a coisa está preta (muito preta, neste caso)*, de uso mais ou menos comum em PE.

14 No caso do adjetivo *preto*, contrariamente a *branco*, a sua utilização para fazer referência ao tom de pele, em particular quando quantificado por *muito*, envolve uma recategorização do adjetivo em nome (cf. (i) abaixo), e não existem muitos exemplos no CETEMPúblico. A explicação poderá estar relacionada com a conotação negativa (e racista) a que se associam estes usos de *preto*.

(i) Souberam testemunhar da vida miserável que **muitos pretos** levam em Portugal, da exploração de que são alvo por parte de empresários sem escrúpulos, do racismo que suportam, da saudade que sentem da terra natal. (*par = ext926967-clt-93a-1*)

- (28) Sou filha de nobres, bonita, elegante, de pele cheirosa e cabelos macios, **muito pretos** e lisos, tão brilhantes que até dá gosto! (par=ext590650-nd-91b-1)
- (29) Eu quando estou sóbrio não sou nada assim, disse Aurélio a estimular a circulação nos braços **muito pretos**. (par=ext709627-soc-96b-1)
- (30) «Oh, diabo... a coisa está **muito preta**. (par=ext816749-nd-95b-2)

Em alguns casos, a leitura pode ser ambígua. Note-se que, em (31), abaixo, a tonalidade preta associada aos *braços*, poderá dever-se, não ao tom de pele da *mulher* identificada, mas sim ao facto de estes conterem muitos *pelos*, como enfatizado através do adjetivo *peludos*. Já em (32), a ambiguidade prende-se com o facto de, ao referir-se uma *figura muito branca*, poder estar a fazer-se referência a uma de duas coisas: ou o tom de pele de *Santana Lopes*, ou o facto de este estar, por exemplo, vestido de branco.

- (31) Era «uma mulher alta, normal, com muito cabelo e os braços **muito pretos** e peludos, fazendo lembrar um homem». (par=ext298897-soc-93a-1)
- (32) Efusivamente saudado por mais de duas mil pessoas que encheram o espaço onde se realizou a festa-convívio, Santana Lopes não se cansou de distribuir beijos a novas e velhas senhoras que se abeiravam da sua figura «muito **branca** e charmosa», um pouco diferente daquela que a televisão e as revistas retratam. (par=ext17452-nd-98a-1)

Também em (33) parece haver a possibilidade de duas leituras. As interpretações possíveis podem dizer respeito à intensidade da cor do *hotel*, i.e., o grau com que o *branco* se manifesta, ou, por outro lado, estar mais relacionado com a quantidade de *branco*, i.e., pode ter várias áreas interiores ou exteriores brancas, bem como outros componentes, como as janelas e as paredes, que, quando combinadas, contribuem para uma grande ‘mancha’ de *branco*, para quem observa. Esta última interpretação aproxima-se de uma leitura do tipo mereológico – relativa às várias partes do *hotel*.

- (33) Aquele hotel **muito branco**, de paredes rendadas, surgindo subitamente por detrás das dunas, pareceu-me um prodígio.
(par = ext1298664-nd-95a-1)

Também em (34), embora possa haver a possibilidade de uma indicação de tonalidade, ou seja, intensidade da cor, a utilização do adjetivo pode, também, indicar que *muitas partes* da camisola (ou *uma grande parte* da camisola) têm a cor azul. Note-se a possibilidade da substituição de *muito* por um advérbio do tipo *parcialmente* (equivalente aos *proportional modifiers*) (cf. (34a.)), que comprova o caráter absoluto destes adjetivos (e, por isso, a sua associação a uma escala fechada).

- (34) No meio do diálogo, aproxima-se Milivoje Cukic, um sérvio de 61 anos com o nome preso a uma camisola **muito azul** que imita uma marca de renome. (par = ext717836-pol-96b-2)

- a. No meio do diálogo, aproxima-se Milivoje Cukic, um sérvio de 61 anos com o nome preso a uma camisola **parcialmente/totalmente azul** que imita uma marca de renome.¹⁵

Existem, ainda, situações em que a utilização do quantificador pode referir-se a uma grande quantidade de objetos da referida cor, como o demonstram os exemplos abaixo (cf. (35)-(38)). Nestes casos, portanto, a leitura adequada passa pela interpretação da quantidade, e não deverá ser confundida com a intensidade, observada acima.

- (35) Esplanada, cores vivas (**muito amarelo** e vermelho), música acelerada e empregados vestidos de preto com aventais coloridos.
(par = ext1306023-soc-98a-2)

¹⁵ Note-se que, em PE, a utilização de *meio*, neste contexto, equivalente ao inglês *half*, não se aplica, exatamente da mesma forma que os restantes *proportional modifiers*. No exemplo abaixo, *meio* não está a indicar que metade da camisola é azul (veja-se, a este propósito, (ia.)), servindo, antes, como uma forma de o autor indiciar que não está certo quanto ao facto de a camisola ser, de facto, da cor mencionada:

(i) No meio do diálogo, aproxima-se Milivoje Cukic, um sérvio de 61 anos com o nome preso a uma camisola **meio azul** que imita uma marca de renome.
a. No meio do diálogo, aproxima-se Milivoje Cukic, um sérvio de 61 anos com o nome preso a uma camisola **meia azul** que imita uma marca de renome.

- (36) Para os homens, **muito preto**, lisos, xadrez e canelado em blazer, blusão, polo, calças com cortes anatómicos. (*par=ext856181-soc-93a-1*)
- (37) Obras de Vieira da Silva e Arpad Szenes, azulejos de Manuel Cargaleiro e **muito branco** preenchem o interior. (*par=ext1359278-soc-97b-1*)
- (38) Do meu poiso, pois, além de um polvilhar breve de branco (os afloramentos das casas alfortinas), além de muito verde, vê-se **muito azul**, barrando ao fundo o quadro: é sob o céu, o mar. (*par=ext617541-opi-96b-1*)

Ainda sobre esta leitura de quantidade, notem-se os casos abaixo, em que os nomes aos quais os adjetivos se aplicam estão elididos. No primeiro caso (cf. (39)), provavelmente, *amarelos* e *vermelhos* deverão referir-se, em linguagem tipicamente associada ao mundo futebolístico, a cartões mostrados a um *jogador marroquino*. No segundo caso (cf. (40)), é possível recuperar o nome *votos*, referente à expressão *votos brancos*.

- (39) Este marroquino tem o defeito de ver **muitos amarelos e vermelhos**. (*par=ext1380184-des-96a-2*)
- (40) cinismo», de «desresponsabilização estatal que já desde há vários anos vem incidindo» sobre o sector. **Muitos brancos**, alguns nulos. Os mais de 60 mil votos brancos registados constituem um resultado fora do comum...

No entanto, a combinação com o quantificador *pouco* não parece tão comum, já que alguns destes adjetivos apenas a aceitam como parte da construção *um pouco* (cf. (41)), e outros não parecem aceitá-la, de todo, já que não foram encontrados exemplos representativos desta construção.

- (41) «Fair-play» acima de tudo, os visados acolheram com sorrisos (às vezes **um pouco amarelos**) as graças do cantador. (*par=ext1298137-soc-96a-1*)

Ainda assim, parece-me que poderá ser útil proceder à manipulação de alguns exemplos, para que se possa verificar quais as leituras obtidas e saber se há algum tipo de correspondência – de sentido inverso – com os resultados tidos para *muito* . Assim, (42), abaixo (manipulação do exemplo (12)), apesar da duvidosa aceitabilidade, pode ser interpretado, apenas no sentido em que se refere à graduação dos adjetivos, i.e., à intensidade da cor. Também em (43)¹⁶, a inserção de *pouco* torna a frase difícil de aceitar.

- (42) ? E simples: uma loura, pouco vestida, enlaçada por um gentleman, trincando uma maçã **pouco vermelha**.
- (43) ? É tudo pouco azul piscina, **pouco amarelo** sol e vermelho de automóvel.

No entanto, esta leitura não se mantém nos exemplos abaixo, em que, no caso da combinação com *muito* , se obtinham leituras de intensidade e, nestes casos, parecem mais relativas à quantidade observável da cor. Assim, em (44) (manipulação de (11)), a interpretação passa a ser a de que o *fumo* tem várias cores, entre as quais o *azul* , que não predomina.¹⁷

- (44) Dos cestos das castanhas subia um fuminho muito simpático e **pouco azul**, atribuído, enganosamente, (...) a granadas lacrimogéneas.

No exemplo seguinte, relativo ao nome *barba* , a utilização de *muito* induz uma leitura de intensidade da tonalidade da *barba* , i.e., é de um branco *muito intenso* (em oposição, por exemplo, a um branco sujo). Pode, ainda, haver a possibilidade de uma leitura mereológica, semelhante a

16 Manipulação do seguinte exemplo, retirado do CETEMPúblico:

(i) É tudo muito **azul piscina**, muito amarelo sol e vermelho de automóvel. (*par=ext166950-clt-93b-1*)

17 Note-se que a utilização de *um pouco* , ao invés de *pouco* , parece transmitir leituras relativas à intensidade da cor. De facto, tanto em (44) como em (45), as leituras obtidas não parecem referir-se à quantidade de partes do *fumo* ou da *barba* , respetivamente, que são da cor denotada pelo adjetivo, mas antes que a intensidade dessa cor não é muito forte. O mesmo acontece com (42), em que o sentido de *maçã um pouco vermelha* não deverá estar relacionado com a quantidade ou as partes da *maçã* que são, efetivamente, vermelhas, mas sim com a intensidade do vermelho atribuído à *maçã* (pouco forte/intenso). Vejam-se os exemplos (manipulados) abaixo:

(i) E simples: uma loura, pouco vestida, enlaçada por um gentleman, trincando uma maçã **um pouco vermelha**.

(ii) Dos cestos das castanhas subia um fuminho muito simpático e **um pouco azul**, atribuído, enganosamente, (...) a granadas lacrimogéneas.

(iii) Como Herbert, velho muito velho, negro muito negro, de barba branca **um pouco branca**, figura gravada na memória.

barba totalmente branca, com o sentido de *todas as partes da barba são brancas*. No entanto, *pouco*, no mesmo contexto, parece promover uma interpretação mereológica: refere-se, assim, que poucas partes da *barba* (i.e., *poucos pelos*) são brancas, o que torna, até, a leitura um pouco contraditória.

- (45) Como Herbert, velho muito velho, negro muito negro, de barba branca **muito branca**, figura gravada na memória (...).
(*par*=*ext384158-clt-97a-2*)
- a. Como Herbert, velho muito velho, negro muito negro, de barba branca **pouco branca**, figura gravada na memória.

A utilização de *pouco*, em combinação com adjetivos de cor, deveria transmitir um afastamento em relação ao protótipo, i.e., a baixa intensidade da cor ou a diminuição da superfície/volume que a cor deveria afetar (relativamente ao referente designado pelo nome). Ainda assim, a difícil aceitação destes adjetivos com a quantificação por *pouco*, e a forte restrição em relação à sua distribuição, parece revelar que esta utilização não é comum, talvez por não ser produtiva, na medida em que não traz informação relevante.

Note-se que, em relação aos usos literais, como mencionado anteriormente, o afastamento em relação ao valor-padrão torna difícil a aceitação de uma entidade como pertencendo a uma determinada categoria (cf. Taylor 2003). Neste caso, tal comportamento traduz-se pela não aceitação da cor, efetivamente: note-se que, em (44), *um fuminho pouco azul* não é, realmente, um fumo com a cor azul. Assim, para que a interpretação se torne possível, parece haver uma transição para a leitura mereológica, em que se define que poucas partes do fumo têm a cor azul.

Já relativamente aos usos não literais, as suas interpretações estão dependentes dos usos de cada língua, o que torna a interpretação relativa, em cada caso. Assim, a não ocorrência de casos de adjetivos de cor com *pouco* parece mostrar que os falantes desta língua (PE) não têm usos prototípicos com este quantificador.¹⁸

18 De facto, note-se a estranheza de (i), abaixo, em oposição a (ii):
(i) ?Ele ficou calmo e **pouco vermelho**.

4. Algumas conclusões possíveis e considerações finais

Os exemplos acima comprovam que o estudo dos adjetivos de cor é, realmente, complexo, pois são vários os fatores que se torna necessário analisar, para corretamente interpretar o tipo de função associada aos quantificadores.

Na presente investigação, apenas foi analisada a possibilidade de quantificação dos adjetivos de cor (relativos às cores primárias e, ainda, a *preto e branco*) por *muito e pouco*. Por essa razão, apenas faz sentido tecer considerações acerca da leitura graduável, associada a estes adjetivos, sendo, por isso, pouco produtiva uma discussão que se foque na leitura não graduável (i.e., classificatória), conforme proposto por Kennedy & McNally (2010).

Ainda assim, e sabendo que os referidos autores propuseram dois tipos de leituras, relativas à graduação deste tipo de adjetivos, parece importante compreender se se tratam, no caso do PE, de leituras relativas à quantidade ou à qualidade da cor.

Com efeito, nos casos em que a leitura está associada a uma graduação da cor, que reflete uma movimentação para pontos mais elevados (*muito*) ou mais baixos (*pouco*) da escala, a interpretação mais natural parece ser a de qualidade, i.e., de aproximação àquilo que, segundo Kennedy & McNally (2010), representa o protótipo da cor. Neste caso, a correspondência ao protótipo parece referir-se aos focos de cor, i.e., à correspondência cromática com o que é, cognitivamente, conceptualizado como uma determinada cor. Note-se (46), abaixo, em que esta leitura parece ficar ainda mais evidente, devido ao uso de *muito pretos*, para se referir a cor dos *pêlos*, reforçado, ainda, pelo uso de *mesmo* (*mesmo muito pretos*).

- (46) No tribunal, o casaco mantivera-se repuxado nos braços cheios de pêlos mesmo **muito pretos**. (*par = ext602455-soc-98a-1*)

Não me parece, neste caso, possível obter uma leitura de quantidade, já que não é feita referência a que quantidade de *pelos* é *preta*, o que segundo Kennedy & McNally (2010), seria a interpretação tipicamente

(ii) Ele ficou furioso e **muito vermelho**.

associada a esta leitura. Note-se que, de facto, para fazer uma interpretação desse tipo teria de procurar-se uma resposta à pergunta *quantos pelos são pretos?*, resposta essa que, numa escala fechada, apenas poderia ser dada se tivéssemos uma forma de contabilizar os referidos *pelos*, o que não é possível (nem relevante).

Porém, como mencionado na secção anterior, existem casos em que esta leitura de quantidade parece evidente, relativa a um grande número de objetos de uma determinada cor, que, no seu conjunto, contribuem para que exista uma grande ‘mancha’ da cor representada pelo adjetivo. Por outro lado, é possível, embora mais raro, obter, em certos exemplos, uma leitura mereológica, que diz respeito, não ao número de objetos, mas às partes de um mesmo objeto que contêm a referida cor.

Todavia, os exemplos recolhidos e analisados revestem-se de uma complexidade acrescida, já que, para todos os adjetivos estudados, foram obtidas leituras não literais, i.e., não exatamente relacionadas com a cor, propriamente dita. De facto, ao associar-se à tonalidade da pele, *amarelo* e *branco* podem ter um significado semelhante a *pálido*, e *vermelho a corado*, *enraivecido* ou, ainda, *escaldado/queimado pelo sol*. O *branco* e o *preto* estão muitas vezes associados, não só aos tons de pele mais ou menos brancos/pretos – dentro da etnia caucasiana e/ou africana – mas também são muito utilizados, nos exemplos em causa, para fazer referência às próprias etnias, estando, nestes casos, ligados a uma interpretação de quantidade de pessoas.

Por outro lado, existem usos muito específicos do PE, como *a coisa está preta*, ou *fazer um sorriso amarelo*, que são expressões utilizadas pelos falantes, no seu quotidiano, e que, de certa forma, são já do conhecimento partilhado da população, pelo que são automaticamente interpretados de forma não literal, não correspondendo, por isso, a uma graduação da cor.

Em suma, pode dizer-se que os adjetivos de cor aceitam quantificação, em PE, embora sejam, de um modo geral, mais compatíveis com *muito* do que com *pouco*, como mostra a tabela abaixo, baseada nos dados recolhidos no *corpus* CETEMPúblico.

TABELA 1 – Quantificação dos adjetivos de cor por *muito* e *pouco*

	Muito	Pouco
Branco	Sim	Sim
Preto	Sim	Não
Amarelo	Sim	Não
Azul	Sim	Sim (apenas 1 ocorrência)
Vermelho	Sim	Não

Quanto ao tipo de leituras obtidas, parece possível afirmar, por um lado, que os adjetivos de cor quantificados podem induzir, maioritariamente, leituras de qualidade, que medem uma aproximação ao protótipo representativo da cor em questão, e que pode ser parafraseada por *um tom de (cor) muito intenso/forte/vivo*. Desse ponto de vista, poder-se-ia considerar que a escala associada a estes adjetivos deveria ser fechada, com o protótipo representativo da cor a corresponder ao seu topo máximo. No entanto, o presente estudo confirma, por outro lado, que, em PE, estes adjetivos não têm um comportamento uniforme quanto ao tipo de leitura que induzem, pois podem ter significados não literais, que variam consoante o adjetivo estudado. Estas diferenças são apresentadas na tabela abaixo, também ela baseada nos dados recolhidos no âmbito da presente investigação.¹⁹ Este comportamento torna difícil uma avaliação do tipo de escala, porque não parece possível tecer considerações acerca do limite máximo ou mínimo para, por exemplo, *um sorriso amarelo*.

¹⁹ Note-se que, em textos de medicina, por exemplo, a aplicação do adjetivo *azul* ao tom de pele poderá ser mais frequente, o que mostra que o tipo de texto em análise pode, também, ter importância quanto ao contexto sobre o qual o adjetivo atua e, conseqüentemente, quanto às leituras que induz.

TABELA 2 – Leituras obtidas através da quantificação de adjetivos de cor por *muito e pouco*

Tipos de leitura		Branco	Preto	Ama- relo	Azul	Verme- lho
Literal	Intensidade	Sim	Sim	Sim	Sim	Sim
Não Literal	Tom de pele	Sim	Sim	Sim	Não	Sim
	Conotação política	Não	Não	Não	Não	Sim
	Emoções	Sim (estado físico)	Não	Não	Sim (uso metafóri- co)	Sim
	Expressões de uso co- mum	Não	Sim	Sim	Não	Não

Assim, apesar de constituir um estudo inicial, que permite levantar algumas pistas acerca do comportamento dos adjetivos de cor em PE, fica, contudo, muito por dizer. Futuramente, poder-se-á fazer um estudo do mesmo tipo – i.e., um levantamento das interpretações possíveis – para casos em que os adjetivos de cor aparecem em estruturas comparativas. Por outro lado, poderia também ser interessante averiguar com que advérbios podem estes adjetivos combinar-se, já que poderiam fornecer uma boa pista acerca do tipo de leitura que induzem e, conseqüentemente, a que tipo de escala deverão estar associados (cf. Kennedy & McNally 2010). Além disso, seria importante aumentar o número de adjetivos em estudo, bem como focar, particularmente, a distinção entre *vermelho* e *encarnado*, para verificar se existem diferenças no uso destas duas designações, i.e., se ambas se referem à mesma cor ou se, pelo contrário, são utilizadas com o intuito de distinguir duas cores ou dois tons da mesma cor.

REFERÊNCIAS

- Aparício, H., Xiang, M., & Kennedy, C. (2015). Processing gradable adjectives in context: a visual world study. In S. D'Antonio, M. Moroney, & C. R. Little (Eds.), *Proceedings of SALT 25* (pp. 413-432). Linguistic Society of America & Cornell Linguistics Circle.
- Clapp, L. (2012). Indexical color-predicates: Truth-conditional semantics vs. truth-conditional pragmatics. *Canadian Journal of Philosophy*, 42(2), 71-100.
- Correia, M., & Barbosa, S. (2013). Para o estudo da denominação da cor em português: estrutura e significado dos nomes e adjetivos construídos. *Filologia linguística portuguesa*, 15(2), 385-406.
- Frege, G. (1892). On sense and reference. In P. Geach, & M. Black (Eds.), *Translations from the Philosophical Writings of Gottlob Frege* (2.^a ed., 1960, pp. 56-85). Basil Blackwell.
- Hansen, N. (2011). Color adjectives and radical contextualism. *Linguistic and Philosophy*, 34(3), 201-221.
- Hansen, N., & Chemla, E. (2017). Color adjectives, standards, and thresholds: an experimental investigation. *Linguistic and Philosophy*, 40(3), 239-278.
- Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language*, 54(3), 610-646.
- Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. Garland. (1997 UCSC Tese de Doutoramento).
- Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30(1), 1-45.
- Kennedy, C., & Levin, B. (2008). Measure of change: The adjectival core of degree achievements. In L. McNally, & C. Kennedy (Eds.), *Adjectives and adverbs: Syntax, semantics and discourse* (pp. 156-182). Oxford University Press.
- Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, 81(2), 345-381.
- Kennedy, C., & McNally, L. (2010). Color, context, and compositionality. *Synthese*, 174(1), 79-98.
- McNally, L. (2011). The relative role of property type and scale structure in explaining the behavior of gradable adjectives. *VIC 2009: Papers from the ESSLI 2009 workshop on vagueness in communication* (Vol. 6517, pp. 151-168). Springer.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104(3), 192-233.
- Rothschild, D., & Segal, G. (2009). Indexical predicates. *Mind and language*, 24(4), 467-493.

- Sapir, E. (1944). Grading, a study in semantics. *Philosophy of Science*, 11(2), 93-116.
- Szabó, Z. (2001). Adjectives in context. In I. Kenesei, & R. M. Hamish (Eds.), *Perspectives on semantics, pragmatics, and discourse: A Festschrift for Ferenc Kiefer* (pp. 119-146). John Benjamins.
- Taylor, J. R. (2003). Prototypes in Cognitive Linguistics. In P. Robinson, & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 39-65). Routledge.
- Travis, C. (1985). On what is strictly speaking true. *Canadian Journal of Philosophy*, 15(2), 187-229.
- Travis, C. (1994). On constraints of generality. *Proceedings of the Aristotelian Society* (Vol. 94, pp. 165-188). Aristotelian Society.
- Travis, C. (1997). Pragmatics. In B. Hale, & C. Wright (Eds.), *A companion to the philosophy of language* (pp. 87-107). Blackwell.
- Veloso, R. (2013). Adjetivo e Sintagma Adjetival. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, & A. Mendes (Eds.), *Gramática do Português* (Vol. 2, Cap. 31, pp. 1359-1493). Fundação Calouste Gulbenkian.
- Zavaglia, C. (1998). Aspectos semânticos dos cromônimos entre as línguas italiana e portuguesa do Brasil. *Estudos Lingüísticos*, 27, 912-917.
- Zavaglia, C. (2006). Dicionário e cores. *Alfa*, 50(2), 25-41.
- Zavaglia, C. (2008). Quem tem boca vai a Roma: as cores dos provérbios. In A. N. Isquierdo, & M. J. B. Finatto (Orgs.), *As ciências do léxico. Lexicologia, Lexicografia e Terminologia* (Vol. 4, pp. 113-131). Editora da UFMS.

Rhythm instruction to improve ESP students' fluency: A study of pausing

Leticia Quesada Vázquez
lquesada@nebrija.es
Universidad Nebrija

ABSTRACT

The present work is aimed at examining how efficient it is to teach language rhythm to enhance English for Specific Purposes (ESP) learners' fluency. 298 Spanish/Catalan speakers participated in the experiment, among which 42 students fulfilled treatment and were examined as subjects of the study. For ten weeks, participants took pronunciation instruction within their regular lessons. The pronunciation sessions lasted thirty minutes and followed a communicative framework. Students were divided into an experimental group, with explicit rhythm training, and a control group, without it. The total number of pauses and unfilled pauses were counted and compared to assess their frequency performing ANOVAs and t-tests. Descriptive analysis shows that the experimental group tends to pause less and make fewer unfilled pauses after training. However, statistical significance is only reached for the independent-samples t-test of the effect sizes of the number of unfilled pauses. Therefore, these findings suggest a greater command over pausing when rhythm is explicitly taught, and a consequent improvement of fluency in the second language, despite statistical results not always being significant.

KEYWORDS

Rhythm instruction; pronunciation teaching; pausing; fluency; pronunciation assessment.

1. Introduction

English for Specific Purposes (ESP) students' proficiency level when speaking varies considerably, and some of them experience rejections when trying to enroll at masters' programs or applying for international companies due to their low communicative competence in the target language (Räsänen & Fortanet-Gómez 2008); despite having the knowledge to address a given situation, these students do not manage to communicate effectively (Douglas 2000). Although ESP courses may contain much listening and speaking practice often implemented as group work and task-based activities, it does not mean that spoken interaction is taught effectively, as aspects such as

input and feedback do not tend to be closely examined (Dudley-Evans & St John 1998). For instance, many times ESP teachers do not know enough about the scientific content of a task so, on the one hand, sometimes the material brought to class is not updated and, on the other hand, the feedback provided might not respond to the final communicative purpose of the task, but rather to the grammatical correctness of the students' interactions.

Pronunciation does not tend to be explicitly taught in ESP context either, although it is essential for successful oral communication to happen. In fact, second language pronunciation teaching remains a hot topic nowadays. After long periods of neglect, a shift in focus towards communication and intelligibility has fostered new research approaches to pronunciation instruction, but these proposals do not usually filter down to the second language classroom (see Derwing & Munro 2015 for a review). Therefore, more class-based studies are needed in order to examine the applicability of the research advances on pronunciation teaching and guide practitioners on how to deal with pronunciation effectively in their classes to enhance students' oral competence.

1.1. Teaching rhythm to improve fluency

Introducing rhythm within the L2 classroom has been revealed to be beneficial to enhance the students' communicative skills (Anderson-Hsieh 1990; Chela-Flores 1997, 2003; Derwing & Munro 2015; Frost & Picavet 2014; Gilbert 2008; Goodwin 2013; Ordin & Polyanskaya 2015; Quené & Van Delf, 2010; Tajima *et al.* 1997; Tuan & An 2010; Wong 1987). However, whether rhythm should be at the forefront of the features to teach has been questioned. Jenkins (2000) did not encourage teaching how to produce an English-like rhythm for intelligibility purposes if English is working as a lingua franca, as nonnative speakers might adopt a different rhythm without jeopardizing communication (as cited in Levis 2018: 143). Levis (2018) argued, however, that in this globalized world L2 speakers are in contact with native Englishes, so it is important that they are aware of how native speakers produce rhythm. He also maintained that even when rhythm does not interfere with the correct interpretation of the message, it is

likely to ease its understanding.

Rhythm anticipates and delimitates crucial information for the successful comprehension of a message. Indeed, a misuse of the right rhythm in the target language can lead to serious communication issues that impede mutual understanding (Gilbert 2008). This is particularly true when the L1 and the L2 clearly differ in rhythmic patterns, such as the case of Spanish speakers learning English. Stress in English is characterized by considerably increasing the length, loudness and pitch of a syllable within words, or words within longer utterances. Knowing that is essential to understand and speak English because not only does it clarify lexical and grammatical misunderstandings, but it also points at the meaningful information of speech. By contrast, in spite of the fact that there are some well-defined rules on how to use syllable stress in Spanish, stressed syllables are not as long, loud and high in pitch as the English ones (Dauer 1987). Besides, sentence stress is not used to highlight important content; in Spanish a more flexible syntax allows pushing this information towards the end of a sentence so as to make it stand out (Chela de Rodríguez 1976).

These rhythmic cues inevitably affect timing: while the English rhythm is based on similar time intervals between stresses, Spanish rhythm is based on similar time intervals between syllables. Consequently, when Spanish speakers use English applying the Spanish rhythm, their speech sounds disjointed. Patel (2007) goes one step beyond and cognitively relates the organization of music with the organization of the discourse. As a piece of music can be arranged in beats, note values and bars, the speech is divided into different elements which are assembled according to perceptual boundaries that go from simple (i.e., the syllable) to more complex (i.e., the utterance) connections. In both cases, the boundaries are mainly established by the use of pausing. The hierarchical structure of the discourse helps us to understand the links among syllables within words, phrases, and full sentences, which constitute the complete meaning of speech. Although all these connections play a role, those that create thought groups (also known as intonational units) are considered key to build the rhythm of a language. Thought groups are semantically and syntactically coherent and meaningful segments into which an utterance is divided. Even though there are no strict rules on how to segment the discourse into intonational units

and the division can vary depending on the speed, people have an intuitive notion of when it is right to pause (Celce-Murcia *et al.* 1996). For example, if in a sentence like “the teacher told me to close the door”, the speaker pauses like “the teacher told/ me to/close the/door”, the sense, intention and meaning of the sentence are lost.

Although some researchers have studied the teachability of rhythm reporting encouraging results in terms of perception, error rate or intelligibility (Chela-Flores 1997; Couper 2006; Hahn 2004; Tsiartioni 2011), no such research has been conducted on how rhythm instruction might affect fluency. Pausing, phrasing or speed of delivery are features that alter the rhythm of the speech and play an important role in how smooth-spoken a speech is. Dudley-Evans and St John (1998) highlighted the importance of silence: controlling when and how to pause can make an impact on the audience, who uses that silent time to process the information given. If the pauses are either too long, wrongly placed, or made too often in a short period of time, speakers sound choppy and ineloquent. Cauldwell (2008) also supported the importance of pausing in his study on air traffic control, where he showed how long pauses can determine the level of fluency of the members of the International Civil Aviation Organization (ICAO). Hence, the transfer of the mother tongue's rhythm when speaking in a second language affects the frequency, location and general use of pauses, which may contribute to creating dysfluent speech. Practicing how and when to pause to adopt the proper language rhythm can enhance students' fluency in the target language.

1.2. Pausing: an indicator of fluency

In Wood's words, “the most complex and one of the most informative elements of fluency studied so far in empirical research involves pause phenomena” (2010: 23). After examining different research studies exploring second language speech, he resolves that, together with location, pausing frequency has been revealed as a crucial indicator of fluency. Lennon (1990) investigated the pause time of four German students doing an international stay in Reading (England). He concluded that the total unfilled pause time

decreased in 25% for three of the four participants while the total filled pause time remained the same. Riggenbach (1991) conducted a study with six Chinese speakers of English who had previously been classified as either fluent or dysfluent to see which linguistic features were fluency indicators. Among the features studied, the pausing analysis revealed that the number of unfilled pauses was a determining indicator of fluency. Besides, the difference in frequency of the total number of pauses between fluent and dysfluent speakers was significant. Derwing *et al.* (2004)'s research on the influence of different perceptual tasks on fluency ratings revealed that pausing stood for more than 60% of the variance in the mean fluency ratings for both the monologue and the picture description tasks. Therefore, analyzing the pauses second language speakers made shed light on how fluent they were in the target language.

The current study belongs to a larger project which investigates the effectiveness of explicit rhythm training to improve engineering undergraduates' comprehensibility and fluency in English. For this paper, only fluency will be examined. To this end, the following hypotheses are formulated:

H1. Students will make fewer pauses when language rhythm is explicitly taught and trained.

H2. Students will make fewer unfilled pauses when language rhythm is explicitly taught and trained.

2. Method

A classroom-based longitudinal study was conducted according to Derwing and Munro's (2015) framework. As explained in Quesada Vázquez & Romero (2020), a ten-week pronunciation module based on the instruction of rhythmic features was designed as part of a six-month technical English course taken at Rovira i Virgili University in 2017. This was a compulsory course for first-year undergraduates of all the engineering disciplines, so initially 298 students participated in the experiment, distributed into six class groups. These class groups were divided into two major treatment groups: rhythm was explicitly taught in three of them (experimental group), while

the other three were trained with conventional pronunciation explanations (control group). Students were recorded before and after instruction to assess their progress in comprehensibility and fluency depending on the instruction received.

2.1. Participants

Students were asked to fill in a consent form to participate in the experiment on the first day of class. Forty-two students (twenty-one experimental; twenty-one control) completed treatment and, hence, were eligible as final participants of the study out of the 298 students enrolled. As shown in Table 1, they were mainly 18-year-old Spanish/Catalan speakers with different initial levels of English.

TABLE 1 – Final participants' profiles (adapted from Quesada Vázquez & Romero, 2020)

Group	Age		First language			English level		
	18	18 +	Sp/Cat bilinguals	Sp	Other	A2-B1	B1-B2	B2-C1
Experimental	21	0	20	1	0	10	6	5
Control	16	5	16	1	4	4	11	6

Note. Sp = Spanish. Cat = Catalan. English level is classified according to the Common European Framework of Reference for Languages (CEFR): A2-B1 = basic user. B1-B2 = independent user (intermediate). B2-C1 = independent/proficient user (upper-intermediate/advanced).

Students took a placement test as their first task submission to determine their language level, based on listening and grammar activities. The experimental group was made up of ten beginners, six intermediates, and five upper-intermediate/advanced students, while the control group consisted of four beginners, eleven intermediates, and six upper-intermediate/advanced. Regarding their mother tongue, students who were not Spanish/Catalan bilinguals reported a high command of both Spanish and Catalan, as all of them pursued primary and secondary education in Catalonia. None of them reported hearing problems.

2.2. The module and instruction

The pronunciation module was designed according to the course syllabus, using vocabulary and grammatical items previously seen and practiced in class, so as to guarantee students' motivation (Anderson-Hsieh, 1990) and ensure the applicability of the knowledge acquired to the students' direct reality, meeting the students' needs (Derwing & Munro 2015; Dudley-Evans & St John 1998). The course content was divided into seven units plus the pronunciation module. As displayed in Figure 1, each of the pronunciation sessions was related to one of the units:

FIGURE 1 – Distribution of the pronunciation module according to the course syllabus

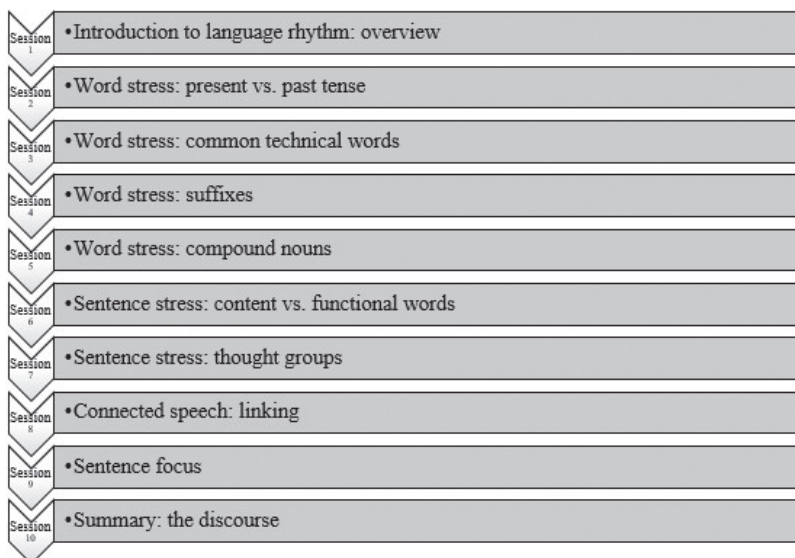
PRONUNCIATION MODULE	SYLLABUS	COURSE
Introduction/Pronunciation of the regular past tense	Chronological description/ Review of verb tenses	UNIT 1
Pronunciation of technical words according to their derivational morphemes	Physical and function description	UNIT 2
Pronunciation of compound nouns vs. adj + noun phrase	Making definitions/ Compound nouns	UNIT 3
Process and graph description	Process description/Visual information	UNIT 4
Talk and debate	Vocabulary on smart technology	UNIT 5
Oral presentation	Vocabulary on robotics and automation/Linking words	UNIT 6
Oral presentation	Oral presentation	UNIT 7

Colors represent the units of speech that were dealt with when working on each pronunciation item (i.e., white represents the word; light gray, the sentence; and dark gray, the discourse). Therefore, instruction was scaffolded.

The main points described at the pronunciation program are related to one or more of the course units. However, the module was divided into ten sessions, so some of the units, which tended to be the ones that took longer to teach, occupied more than one pronunciation session. For example, although sentence instruction was based just on Unit 4, there were two separate pronunciation sessions related to that unit: one worked on process description and the other one on graph description, concepts included within the same regular unit.

Each individual session followed Celce-Murcia *et al.* (1996) communicative framework, starting with an explanation, continuing with listening discrimination, and finishing with oral practice, from controlled to communicative exercises. The only difference between groups was rhythm instruction. Experimental training consisted of practicing word stress, sentence stress, linking and sentence focus.

FIGURE 2 – Distribution of rhythm instruction per session



Rhythm instruction was mainly based on vowel lengthening: as not language-oriented learners, it was considered that length will be a simple concept to assimilate. Besides, vowel lengthening makes a notable difference between English and Spanish rhythms. The introduction of the rhythmic items in class was done progressively, from the simplest to the most complex ones. Its distribution within the sessions is summarized in Figure 2.

Explicit rhythm activities were mainly adapted from textbooks and research papers to suit the technical background of the course. An example of an activity that implied rhythm training is the following: When teaching the pronunciation of common technical words in session 4, the syllables of the word were replaced by *LA* when they carried the main stress and *la* for the rest of them (e.g. *electrician* became *la la LA la*, and *chemical* became *LA la la*), so students got familiarized and practiced the pronunciation of the words through their rhythmic patterns. Secondary stress was excluded in order to focus on the acquisition and practice of primary stress. This technique of using rhythmic patterns to explain word stress, adapted from Chela-Flores (1997: 127), was further used in other sessions to explain other items, such as the difference in pronunciation between compound nouns and adjective + noun combinations in session 5. Those other activities not focusing on rhythm explicitly were created from online resources, such as videos, website texts and images, or adapted from activities for academic purposes and pronunciation exercises.

Control groups did not take rhythm training, so pronunciation was explained and practiced with traditional methods, such as the pronunciation of technical words according to the suffix they carry (i.e., words ending in *-ian* stress the second from last syllable). For a full example of how sessions worked and differed depending on treatment, you can see Figure 2 in Quesada Vázquez & Romero (2020).

2.3. Recordings

Both sessions (before and after training) took place in the library of the School of Engineering. Three semi-sound-proof isolated rooms were booked to carry the recordings out with Two Sony PCM-M10 and a Zoom H4nsp

recorders. Students were recorded individually. The pretest and the posttest were the same in order to guarantee comparability. They consisted of four exercises, including both controlled and extemporaneous activities. The test started with a reading aloud exercise inspired on Couper's (2006) speaking test (p. 61). It was made up of ten sentences, both simple and complex and of an uncontrolled nature (rhythmic patterns were not considered for its usage) that contained technical vocabulary. The second exercise was a diagnostic passage proposed by Celce-Murcia *et al.*'s (1996: 398) to check learners' pronunciation, which students also had to read aloud. The third exercise consisted in introducing themselves, also inspired by Couper's (2006) test. Finally, the fourth exercise was about giving their opinions on social media. The present study only focuses on the assessment of the sentences (i.e., the first exercise).

2.4. Data analysis

The rhythm of the sentences was acoustically measured before and after the instruction. Two aspects were taken into account: VarcoV values (White & Mattys 2007, 2007b), based on vowel duration variability, and pauses. A total of 840 sentences were segmented and acoustically analyzed using PRAAT. An individual sound file and its corresponding text grid were created for each of the sentences. The segmentation was carried out manually following Ordin and Polyanskaya's criteria (2014: 5).

The text grid contained four different tiers: From top to bottom, the first tier showed the transcription of the sounds pronounced. Sound identification followed Olive *et al.* (1993) guidelines; the second tier displayed the syllable structure of each word; the third tier presented the different vocalic and consonant clusters; finally, the fourth tier showed the sentence to pronounce. Blanks represented both filled, i.e., "voiced fillers which do not normally contribute to additional lexical information" (Riggenbach, 1991:426), such as hesitations and repetitions, and unfilled pauses. The latter were mainly silent micropauses of less than .2 seconds as the study was carried out with short sentences, not full speeches (For an example of the text grid used, see Figure 3 in Quesada Vázquez & Romero

2020). For this study, the frequency of pauses was analyzed. Pauses were manually counted, and two analyses were performed: one took into account only the unfilled pauses while the other took into consideration the total number of pauses (both filled and unfilled).

2.5. Statistical analysis

Two different mixed repeated-measures ANOVAs were performed. First, the impact on the overall groups' performances depending on the instruction received was investigated by means of a mixed repeated-measures ANOVA with time and sentence as within-subjects factors, and group as a between-subjects factor. Then, in order to study the relevance of the difference in production before and after training, another mixed repeated-measures ANOVA was run with the difference between pretest-posttest pauses (either all of them or just unfilled ones) as the dependent variable, sentence as the within-subjects factor and group as the between-subjects factor. To further analyze the relevance of this difference and reinforce the ANOVAs' results, t-tests were also performed: two paired-samples t-tests comparing groups' learning progress were completed for each sentence. Finally, one more independent-samples t-test with the two groups' sentence effect sizes was also performed.

3. Results

As shown in Table 2, the means and standard deviations of the pauses for each sentence were estimated and distributed according to time (T1 = pretest; T2 = posttest) and group (experimental; control). The interest fell in examining whether students paused less after treatment. When that happened, values were highlighted in light gray. On average, both groups made approximately the same number of pauses before the instruction. However, despite the fact that both groups made fewer pauses after training, the experimental group's total mean decreased more than the one of the control group. In addition, when examining the sentence means, it was

observed that the experimental group decreased its means after instruction in 9 out of 10 sentences, whereas the control group's means fell on 7 out of 10. Therefore, a preliminary descriptive analysis shows that the experimental group reduced more the total amount of pauses after instruction than the control group.

TABLE 2 – Total number of pauses means (*M*) and standard deviations (*SD*) per sentence

Sentence	Control group				Experimental group			
	Time 1		Time 2		Time 1		Time 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	0.57	0.51	0.62	0.74	0.52	0.6	0.33	0.48
2	1.86	1.39	1.86	1.15	1.86	1.49	1.52	1.36
3	1.76	1.48	1.14	1.01	0.9	0.77	1.1	1.09
4	3.29	1.38	2.62	1.69	3	1.7	2.48	1.57
5	3.76	2.33	3	1.51	4.12	2.66	3.86	1.55
6	3.26	1.38	2.9	1.3	3.1	1.87	2.71	1.38
7	2.95	1.69	2.52	1.78	2.81	1.6	2.29	1.45
8	4.33	2.73	4.24	2.14	3.86	2.08	3.05	2.22
9	2.43	1.4	2.26	1.18	3.21	1.33	2.56	1.07
10	2.53	0.8	2.9	1.33	2.48	1.21	2.14	1.11
Total	2.7		2.5		2.6		2.2	

Appendix A contains the number of pauses made by each of the participants while reading aloud each of the sentences. Generally speaking, intermediate students from the control group and high-intermediate/advanced students from the experimental group were the ones to show more improvement. Besides, although no clear pattern could be established because individual results varied considerably depending on the student and the sentence performed, participants' means showed that 16 out of 21 students from the experimental group paused less after instruction whereas in the control group they were 12 out of 21. Therefore, individual performances support a greater improvement for the experimental group.

In order to investigate the effect of instruction on group performance,

a mixed repeated-measures ANOVA was run. Time and sentence were the within-subjects factors and group was the between-subjects factor. The total number of pauses was the dependent variable. Results did not show significance for group, $F(1,40) = .427, p = .517$, but they did for time, $F(1,40) = 8.995, p < .005$, and sentence $F(9,32) = 46.743, p < .01$. However, no significance was found for the time*sentence interaction $F(9,32) = .578, p = .739$ (see Table 3).

TABLE 3 – Mixed repeated-measures ANOVA with the total number of pauses as the dependent variable

	Effect	df	Error df	F	p-value
Within-subjects	time	1	40	8.995	.005**
	sentence	9	32	46.743	.01**
Interaction	time*sentence	9	32	.578	.739
Between-subjects	group	1	40	.427	.517

Note. ** $p < .01$.

Nevertheless, Figure 3 shows that the number of pauses made dropped more after treatment for the experimental group, agreeing with the tendency observed when comparing group means.

Another ANOVA was run so as to examine the magnitude of the difference between pretest and posttest performances. The dependent variable was the difference in the number of pauses for each sentence before and after treatment, sentence was the within-subjects factor and group was the between-subjects factor. Neither sentence, $F(9,32) = .827, p = .546$, nor group, $F(1,40) = 1.567, p = .218$, showed significance (see Table 4).

FIGURE 3 – Group progress on the total number of pauses according to treatment

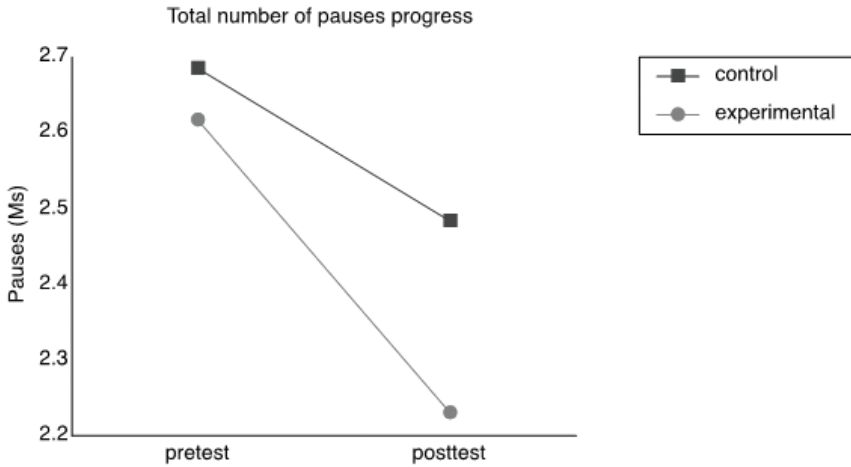


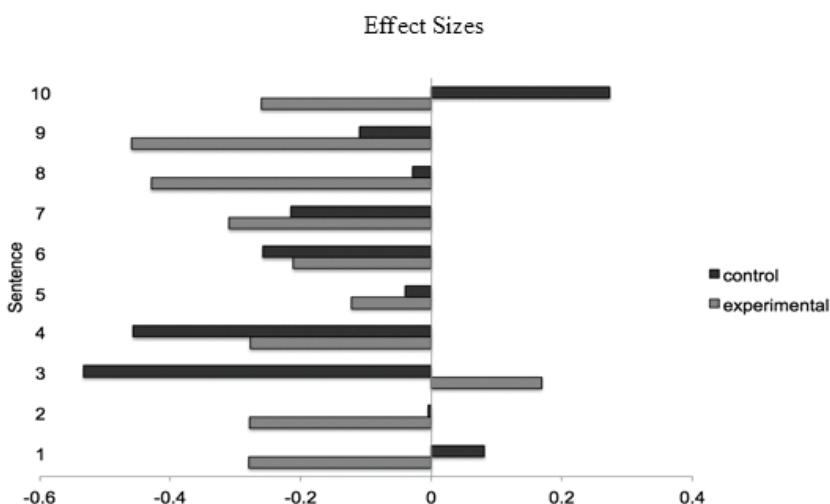
TABLE 4 – Mixed repeated-measures ANOVA with the total number of pauses difference over time as the dependent variable

	Effect	df	Error df	F	p-value
Within-subjects	sentence	9	32	.827	.546
Between-subjects	group	1	40	1.567	.218

To further examine the relevance of the difference, two paired-samples (control T2 vs. control T1; experimental T2 vs. experimental T1) and two independent-samples t-tests (control T1 vs. experimental group T1; control T2 vs. experimental group T2) per sentence were calculated, which reached significance only in a few specific cases (see Appendix B). Independent-samples t-tests were significant when analyzing both groups before treatment in sentence 3, $T(40) = 2.355, p < .023$, and after treatment in sentence 10, $T(40) = 2.044, p < .048$. As for paired-samples t-tests, they showed significance for the control group performance in sentences 3, $T(20) = -2.444, p < .024$, and 4, $T(20) = -2.092, p < .049$, and for the experimental group performance in sentence 9, $T(20) = -2.103, p < .048$. A

final independent t-test comparing the effect sizes of all the paired-samples t-tests was also carried out to examine the relevance of the improvement. A difference between the two treatment groups' progress was not obvious, since both showed mostly negative effect sizes, which corresponded to a reduction of the number of pauses made (see Figure 4). Besides, results did not reveal significance $T(18) = 1.229, p = .235$.

FIGURE 4 – Effect sizes for the total number of pauses paired-sample t-tests



The same analytical process was followed to examine unfilled pauses. As shown in Table 5, the experimental group reduced the number of unfilled pauses in eight out of ten sentences, made the same amount in sentence 3, and made more in sentence 5. The control group, on the other hand, made fewer pauses in six out of ten sentences, the same amount in sentence 8, and more in sentences 1, 9 and 10. In addition, the total mean of pauses decreased more for the experimental group. Hence, the experimental group seemed to perform better after instruction.

TABLE 5 – Unfilled pauses means (*M*) and standard deviations (*SD*) per sentence

Sen- tence	Control group				Experimental group			
	Time 1		Time 2		Time 1		Time 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	0.52	0.51	0.57	0.75	0.48	0.51	0.33	0.48
2	1.95	1.47	1.81	1.03	1.86	1.49	1.52	1.36
3	1.43	1.03	1.1	0.89	0.9	0.77	0.9	0.83
4	2.76	1.04	2.52	1.63	2.81	1.66	2.29	1.62
5	3.43	1.8	3.33	1.35	3.52	1.97	3.79	0.89
6	2.72	0.79	2.71	1.1	2.81	1.54	2.43	1.16
7	2.57	1.36	2.19	1.5	2.29	1.19	1.86	1.06
8	4	2.41	4	2.05	3.71	2.03	2.86	1.93
9	2.19	1.08	2.24	1.37	3.05	1.16	2.33	1.02
10	2.33	1.06	2.71	1.1	2.38	1.07	2.1	1.04
Total	2.4		2.3		2.4		2	

Again, individual performance differs depending on the sentence examined (see Appendix C). However, none of the control beginners improved after instruction this time. As for the participants' individual means, 15 experimental and 10 control students decreased their number of unfilled pauses after training, which reinforces a better performance of the experimental group.

The results of first mixed repeated-measures ANOVA that was run (time and sentence = within-subjects factors; group = between-subjects factor; number of unfilled pauses = dependent variable) are displayed in Table 6. Results were not significant for group, $F(1,40) = .349, p = .558$, although they were for sentence, $F(9,32) = 56.725, p < .01$, and for time, $F(1,40) = 5.478, p < .024$. Nevertheless, the time*sentence interaction was, again, not significant, $F(9,32) = .773, p = .589$. Therefore, results seemed to correlate with the ones obtained when analyzing the total number of pauses.

TABLE 6 – Mixed repeated-measures ANOVA with unfilled pauses as the dependent variable

	Effect	df	Error df	F	p-value
Within-subjects	time	1	40	5.478	.024*
	sentence	9	32	56.725	.01**
Interaction	time*sentence	9	32	.773	.589
Between-subjects	group	1	40	.349	.558

Note. * $p = .05$. ** $p = .01$.

However, Figure 5, which displays the progress of the different groups in regard of the number of unfilled pauses, shows a slightly different scenario from the one presented regarding the total number of pauses in Figure 4. Firstly, while the control group made more pauses than the experimental group on average at the beginning of the course, the number of unfilled pauses was almost the same for both groups. Secondly, when analyzing all the pauses, the progress of both groups was similar, since the control group made 0.2 pauses and the experimental group 0.4 pauses less than the first time they took the test. Nonetheless, when taking into account just the unfilled pauses, the impact of progress changed. The control group remained almost the same at 2.4 pauses, while the experimental group made 0.4 pauses less after instruction. Thus, the improvement of the control group seemed to result from making fewer hesitations and repetitions rather than silent pauses.

The second ANOVA was carried out with the difference in times as the dependent variable, sentence as the within-subjects factor and group as the between-subjects factor. As seen when analyzing the total number of pauses, results did not reach significance for any of the factors under study: sentence, $F(9,32) = .708, p = .643$; group, $F(1,40) = 2.470, p = .124$.

FIGURE 5 – Group progress on the number of unfilled pauses according to treatment

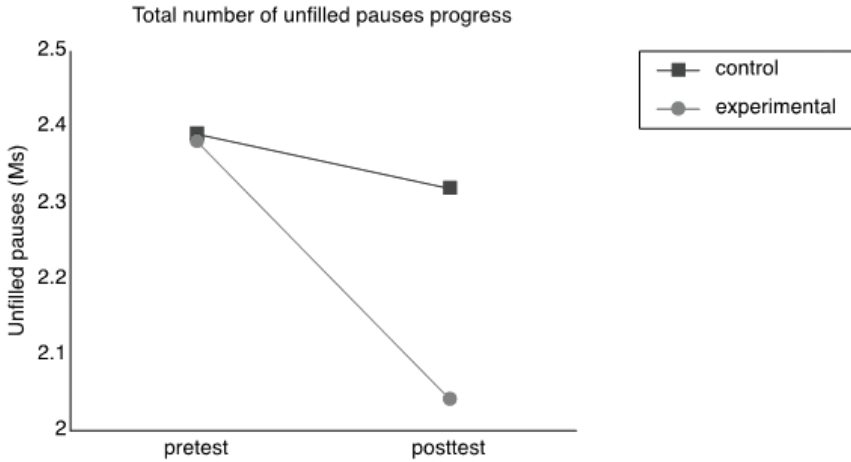
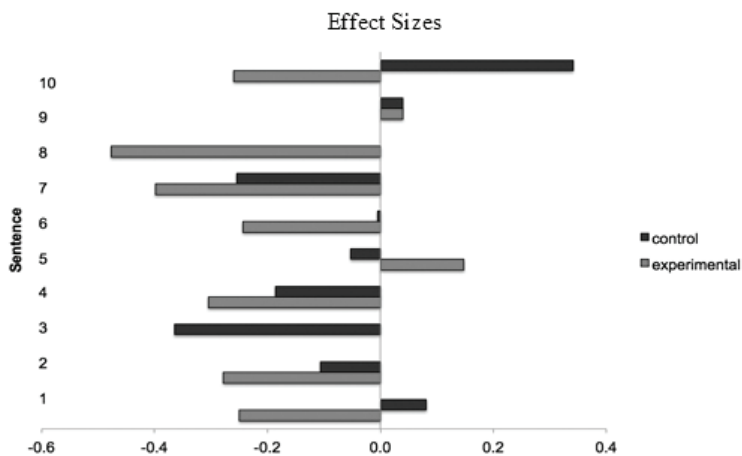


TABLE 7 – Mixed repeated-measures ANOVA with the unfilled pauses difference over time as the dependent variable

	Effect	df	Error df	F	p-value
Within-subjects	sentence	9	32	.708	.643
Between-subjects	group	1	40	2.470	.124

Regarding t-test analyses, these, again, reached significance in only some of the cases (see Appendix D). This time, only one independent-samples t-test showed significance: groups before treatment in sentence 9, $T(40) = -2.480, p < .017$. As for the paired-samples t-tests, results for the experimental group were found significant in sentences 8, $T(20) = -2.186, p < .041$, and 9, $T(20) = -2.855, p < .010$. However, this time the effect sizes of all the paired-samples t-tests showed a greater improvement for the experimental group performances in most of the cases (see Figure 6). What is more, the independent-samples t-test with these effect sizes did show significance $T(18) = 2.343, p < .031$.

FIGURE 6 – Effect sizes for the total number of unfilled pauses paired-sample t-tests



To sum up, the different analyses carried out for the two items under analysis (i.e., total number of pauses and unfilled pauses) suggest favorable effects on students' prosody when pronunciation was introduced in class, especially when rhythm was explicitly taught, even though results were not always statistically significant.

4. Discussion

For both unfilled and the total number of pauses, the descriptive analysis did not reveal a clear difference in the treatment received although the experimental group showed a slightly greater improvement in both cases (see Tables 2 and 5). On the one hand, both the control and the experimental group reduced their sentence means after training and, on the other hand, the overall mean per group was a bit lower for the experimental group after instruction. As for individual values, there were more students from the experimental group whose means decreased after training (see Appendixes A and C). However, no students improved in all the sentences and there is no sentence in which all of them improved. In fact, the sentence standard deviations oscillated between one and two in most of the cases, so a wide

variability among speakers seems to be confirmed.

As for the ANOVAs carried out with the total number of pauses and unfilled pauses as the dependent variable respectively, both the sentence and the time variables were found significant, but the time*sentence interaction was not, so no statistical connection could be established between the two variables. The statistical results are a bit more encouraging when examining the data for just unfilled pauses, though. As suggested in the previous section, the control group progress seems to be thanks to the fact that they made fewer hesitations and repetitions after instruction, which could be the consequence of the in-class practice of the pronunciation of certain words as part of the vocabulary of both the pronunciation module and the regular lectures. However, they made approximately the same number of silent pauses after instruction, which suggests that they kept pausing within thought groups. On the other hand, the experimental group reduced the number of both filled and unfilled pauses to the same extent. Besides, the analysis of the effect sizes of the paired-samples t-tests conducted for the unfilled pauses was significant ($p < .031$), while the one for the total number of pauses was not. These results suggest, hence, that the experimental group recognized thought groups better and avoided pausing in between them more than the control group, producing a more English-like rhythm.

Overall, the statistical analysis of the total number of pauses extracted from the sentences does not confirm hypothesis 1 while the analysis of the number of unfilled pauses provides more evidence of an improvement when rhythm instruction takes place, supporting hypothesis 2. Although the descriptive analysis of the total number of pauses shows that both groups improve after instruction to a certain extent and that the experimental group performed better after treatment, the statistical analyses conducted do not fully support these assumptions. As previously mentioned, individual variability could partially explain the results obtained, but there are other factors that could have also played a role, such as the limited number of sessions, or the dramatical decrease of students who fulfilled treatment and, hence, were taken as the final participants of the study. Nonetheless, the analysis of unfilled pauses seems to provide clearer results on fluency improvement despite a lack of robust statistical significance. The results of this study go along the lines of previous research such as Lennon (1990)

and Riggensbach (1991), who found unfilled pauses to be a good indicator of fluency. It is also interesting to point out that Wood claimed in his compilation of research studies of fluency that “the importance of pause times and frequencies tells us a great deal about speech fluency, particularly that related to the value of unfilled pauses. It appears that analyzing filled pauses yields mixed and inconclusive results (2010: 24)”.

5. Conclusion

Although further research should be conducted to obtain more conclusive results, the findings of this study are enlightening. On the one hand, pronunciation instruction, particularly explicit rhythm training, has been revealed to be helpful for ESP students to manage pausing when speaking in the target language and, thus, enhance their fluency in controlled oral production, such as reading sentences aloud. It is true, though, that this study only focuses on sentences that are read aloud and more research should be done on students' performances in the text reading and opinion tasks to examine longer speeches. In addition, it would be interesting to examine the exact location of the pauses made within the sentences to support the assumption that these were not made within thought groups. On the other hand, this study can become a model for those researchers and practitioners who are looking for approaches to introduce pronunciation effectively within ESP courses. By means of this study, it has been revealed that pronunciation instruction is possible within the ESP field adapting the lessons to the students' professional needs and, besides, can bring benefits to the students' communicative skills. However, more classroom-based studies are still necessary to fully support these hypotheses and help to bridge the gap between practice and research.

REFERENCES

- Anderson-Hsieh, J. (1990). Teaching suprasegmentals to international teaching assistants using field-specific materials. *English for Specific Purposes*, 9(3), 195-214. [https://doi.org/10.1016/0889-4906\(90\)90013-3](https://doi.org/10.1016/0889-4906(90)90013-3)
- Caudwell, R. (2008). Defining fluency for air traffic control. *Speech in Action*, 3, 1-5. <https://www.speechinaction.org/wp-content/uploads/2012/03/Defining-Fluency.pdf>
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.
- Chela-Flores, B. (1997). Rhythmic patterns as basic units in pronunciation teaching. *Onomázein*, (2), 111-134.
- Chela-Flores, B. (2003). Optimizing the teaching of English suprasegmentals to Spanish speakers. *Lenguas Modernas*, (28–29), 255-273.
- Chela de Rodríguez, B. (1976). Teaching English suprasegmentals to Spanish speakers. *Kansas Working Papers in Linguistics*, 1, 63-82.
- Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect*, 21(1), 46-66.
- Dauer, R. M. (1987). *Phonetic and phonological components of language rhythm*. In T. Gamkrelidze (Ed.), Paper presented at the 11th International Congress of Phonetic Sciences (ICPhS) (pp. 447-450). Academy of Sciences of the Estonian S.S.R.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54(4), 655-679.
- Douglas, D. (2000). *Assessing language for specific purposes*. Cambridge University Press.
- Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes. A multi-disciplinary approach*. Cambridge University Press.
- Frost, D., & Picavet, F. (2014). Putting prosody first – Some practical solutions to a perennial problem: The Innovalangues project. *Research in Language*, 12(3), 1-11. <https://doi.org/10.2478/rela-2014-0002>
- Gilbert, J. B. (2008). *Teaching pronunciation: using the prosody pyramid*. Cambridge University Press.
- Goodwin, J. (2013). Pronunciation teaching methods and techniques. In C. A. Chapelle

- (Ed.), *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell. <https://doi.org/10.1002/9781405198431.wbeal0970>
- Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), 201-223.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387-417.
- Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.
- Olive, J. P., Greenwood, A., & Coleman, J. (1993). *Acoustics of American English speech: A dynamic approach*. Springer-Verlag.
- Ordin, M., & Polyanskaya, L. (2014). Development of timing patterns in first and second languages. *System*, 42, 244-257. <https://doi.org/10.1016/j.system.2013.12.004>
- Ordin, M., & Polyanskaya, L. (2015). Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. *The Journal of the Acoustical Society of America*, 138(2), 533-544. <https://doi.org/10.1121/1.4923359>
- Patel, A. D. (2007). *Music, language, and the brain*. Oxford University Press USA.
- Quené, H., & Van Delft, L. E. (2010). Non-native durational patterns decrease speech intelligibility. *Speech Communication*, 52 (11-12), 911-918. <https://doi.org/10.1016/j.specom.2010.03.005>
- Quesada Vázquez, L., & Romero, J. (2020). ESP within ESP: the design and implementation of a pronunciation module in a technical English course. *Onomázein*, (NE VI), 209–228. <https://doi.org/10.7764/onomazein.ne6.11>
- Räisänen, C., & Fortanet-Gómez, I. (2008). The state of ESP teaching and learning in Western European higher education after Bologna. In I. Fortanet-Gómez, & C. Räisänen (Eds.), *ESP in European higher education. Integrating language and content* (pp. 11-51). John Benjamins. <https://doi.org/10.1075/aals.4.03rai>
- Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, 14(4), 423-441. <https://doi.org/10.1080/01638539109544795>
- Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, 25(1), 1-24. <https://doi.org/10.1006/jpho.1996.0031>
- Tsiartsioni, E. (2011). *Can pronunciation be taught? Teaching English speech rhythm to Greek students*. In E. Kitis, N. Lavidas, N. Tpointzi, & T. Tsangalidis (Eds.), *Selected Papers from the 19th International Symposium on Theoretical and Applied Linguistics*,

- Thessaloniki 3-5 April 2009* (pp. 447-458). Aristotle University of Thessaloniki.
<https://doi.org/10.26262/istal.v19i0.5522>
- Tuan, L. T., & An, V. T. P. (2010). Teaching English rhythm by using songs. *Studies in Literature and Language*, 1(2), 13-29.
- White, L., & Mattys, S. L. (2007a). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501-522. <https://doi.org/10.1016/j.wocn.2007.02.003>
- White, L., & Mattys, S. L. (2007b). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró, & M.-J. Sole (Eds.), *Segmental and prosodic issues in romance phonology (Current issues in linguistic theory series)* (pp. 237-257). John Benjamins.
- Wong, R. (1987). *Teaching pronunciation: Focus on English rhythm and intonation*. Regents Prentice Hall.
- Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. Continuum.

Appendix A. Total number of pauses per student

St.	T1										T2										M		
	L	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S2	S3	S4	S5	S6	S7	S8	S9		S10	M
	Total number of pauses – individual values and means for the control group (cleaned data)																						
1	1	1	2	2	5	5	4	4	5	1	4	3.3	1	4	2	5	6	4	5	8	4	4	4.3
2	1	1	2	5	4	3	3.52	3	3	3	3	3.1	1	3	4	1	3	5	1	5	4	2	2.9
3	1	1	1	2	4	1	5	3	1	4	3	2.5	2	2	2	4	4	6	3	5	3	3	3.4
4	1	1	1	3	3	5	4	4	6	2	2	3.1	0	3	1	2	4	3	4	5	2	5	2.9
5	2	0	3	0	3	2	3	2	3	5	3	2.4	0	3	0	4	6	4	6	3	2	3	3.1
6	2	1	1	2	3	3	3	4	2	2	3	2.4	1	1	2	2	4	3	1	4	1	4	2.3
7	2	1	2.14	3	2	9	2	3	6	2	4	3.4	1	2	1	1	4	3	1	5	2	3	2.3
8	2	1	2	1	4	4	5	4	3	2	3	2.9	0	0	1	2	4	4	1	2	1	4	1.9
9	2	0	4	5	5	5	2	5	4	4	1	3.5	0	1	1	2	6	2	5	4	3	3	2.7
10	2	1	5	2	4	5	7	6	4	6	2.57	4.3	2	3	2	6	3	5	8	2.52	4	4.2	
11	2	0	0	1	2	2	2	0	2	1	1	1.1	0	0	0	2	1	1	0	2	0	3	0.9
12	2	1	2	2	4	4	2	4	1.2	3	3	3.7	1	2	2	3	5	3	1	5	3	1	2.6
13	2	1	2	3	4	10	3	3	6	2	2	3.6	1	2	1	4	4	3	3	3	2	3	2.6
14	2	0	3	2	7	5	4	3	5	3	3	3.5	0	3	2	6	3	3	2	8	2	6	3.5
15	2	1	1	1	2	5	2	2	2	1	3	2.0	0	1	1	2	3	2	5	6	2	3	2.5
16	3	0	2	0	2	2	2	2	0	2	2	1.4	1	0	1	1	3	2	1	2	2	2	1.5
17	3	0	0	2	3	2	2	5	7	0	2.57	2.4	0	1	0	1	1	0	2	0	0	1	0.6
18	3	0	0	2	3	2	0	8	2	2	2	1.9	0	2	0	1	4	3	1	5	2	2	2.0
19	3	0	0	0	3	1	3	4	6	2	2	2.1	0	1	0	0	4	2	2	2	4	2	1.7
20	3	1	4	1	2	3	5	1	2	2	2	2.3	2	3	1	3	2	2	2	3	4	2	2.4
21	3	0	2	0	1	2	3	0	4	2	2	1.6	0	2	0	3	2	3	2	4	2	1	1.9
Total												2.7											2.5

Note. T1 = Time 1. T2 = Time 2. St = Student. L = Level. S = Sentence. M = Mean. Light gray shadow = Improvement after treatment in the individual scores. Dark gray shadow = Improvement after treatment in means.

Total number of pauses – individual values and means for the experimental group (cleaned data)

St.	T 1										T 2										M
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	
1	1	0	0	3	4	4	2	2	2	0	2.1	0	0	4	1	2	3	2	0	2	1.6
2	1	1	1	4	10	2	6	4	1	1	3.3	1	1	2	2	4.19	3	6	2	4	2.8
3	1	0	0	0	0	5	0	1	2	0	0.9	0	0	0	1	4.19	2	2	1	0	1.1
4	1	2	2	3	4	4	1	5	3	1	2.7	0	3	1	0	2	4	2	2	4	2.1
5	1	0	0	3	0	2	2	2	1	0	1.2	0	0	1	3	3	4	4	1	1	1.8
6	1	0	3	0	1	8	3	4	6	3	3.1	1	4	0	3	5	3	2	3	3	2.7
7	1	0	3	2	3	4	5	2	4	3	2.9	0	2	3	3	2	2	5	3	2	2.5
8	1	1	2	2	1	4	8	2	7	4	3.2	0	2	3	3	7	4	1	8	4	3.6
9	1	0	1	0	1	2	4	4	3	3	2.1	0	1	0	0	4.19	2	2	2	3	1.5
10	1	0	0	1	3	4	3	4	4	3	2.4	0	1	1	3	3	3	0	1	1	1.4
11	2	0	0	1	5	4	1	4	2	6	2.4	1	2	1	3	4	3	2	2	4	2.5
12	2	0	1	1	6	5	2	2	3	3	2.5	0	0	1	2	6	2	2	2	4	2.1
13	2	1	4	2	3	6	3	3	7	3	3.5	1	1	1	3	6	4	3	6	2	3.0
14	2	1	3	1	5	7	3	4	8	3.48	3.8	0	2	1	5	5	6	3	7	2.76	3.5
15	2	0	1	0	0	1	0	3	2	0	0.8	0	0	0	1	2	1	1	5	2	1.3
16	2	1	4	1	6	4	1	3	4	3	3.2	1	4	1	4	4	2	3	4	2	2.9
17	3	1	2	0	3	5	2	6	3	6	3.1	0	2	1	1	3	1	3	4	3	2.1
18	3	0	1	1	4	7	0	2	1	4	2.3	0	0	0	1	5	0	1	1	1	1.0
19	3	1	4	1	3	6	3	7	4	1	3.4	1	1	1	4	6	4	5	2	2	2.8
20	3	1	4	2	4	8	3	2	5	1	3.4	1	4	1	6	6	3	2	5	3	3.3
21	3	2	3	1	3	4	3	3	3	2	2.6	0	2	0	3	2	1	0	1	2	1.2
Total											2.6										2.2

Note. T1 = Time 1. T2 = Time 2. St = Student. L = Level. S = Sentence. M = Mean. Light gray shadow = Improvement after treatment in the individual scores. Dark gray shadow = Improvement after treatment in means.

Appendix B: Comparison of total number of pauses – t-tests results

Comparison of total number of pauses across Time (T1 vs T2) and Group (Control vs Experimental) – t-tests results

Sentence	T-test		df	T	p-value	Cohen's d
1	Paired	C T2-T1	20	.370	.715	.0814
		E T2-T1	20	-1.284	.214	-.279
		C T1-E T1	40	.277	.783	.090
2	Independent	C T2-E T2	40	1.482	.146	.464
		C T2-T1	20	-.022	.983	-.005
		E T2-T1	20	-1.276	.217	-.278
3	Paired	C T1-E T1	40	.015	.988	.005
		C T2-E T2	40	.855	.398	.269
		C T2-T1	20	-2.444	.024*	-.533
4	Independent	E T2-T1	20	.777	.446	.169
		C T1-E T1	40	2.355	.023*	.729
		C T2-E T2	40	.146	.884	.038
5	Paired	C T2-T1	20	-2.092	.049*	-.457
		E T2-T1	20	-1.272	.218	-.278
		C T1-E T1	40	.597	.554	.187
6	Independent	C T2-E T2	40	.284	.778	.086
		C T2-T1	20	-.183	.857	-.040
		E T2-T1	20	-.561	.581	-.122
7	Paired	C T1-E T1	40	-.741	.463	.228
		C T2-E T2	40	-.764	.450	.236
		C T2-T1	20	-1.183	.251	-.258
8	Independent	E T2-T1	20	-.969	.344	-.211
		C T1-E T1	40	.331	.742	.102
		C T2-E T2	40	.460	.648	.141
9	Paired	C T2-T1	20	-.987	.335	-.216
		E T2-T1	20	-1.419	.171	-.310
		C T1-E T1	40	.281	.780	.085
10	Independent	C T2-E T2	40	.475	.637	.142
		C T2-T1	20	-.133	.895	-.029
		E T2-T1	20	-1.966	.063	-.429
11	Paired	C T1-E T1	40	.636	.528	.194
		C T2-E T2	40	1.766	.085	.544
		C T2-T1	20	-.503	.620	-.110
12	Independent	E T2-T1	20	-2.103	.048*	-.459
		C T1-E T1	40	-1.864	.070	.575
		C T2-E T2	40	-.854	.398	.264
13	Paired	C T2-T1	20	1.251	.225	.273
		E T2-T1	20	-1.195	.246	-.261
		C T1-E T1	40	.172	.865	.053
14	Independent	C T2-E T2	40	2.044	.048*	.629

Note. C = control group. E = experimental group. T = time. Paired = paired-samples t-test. Independent = Independent-samples t-test. Cohen's d = Effect size.
 * p < .05.

Appendix C: Unfilled pauses per student

Unfilled pauses – individual values and means for the control group (cleaned data)		T 2																							
		St.	L	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	M	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	M
1	1	1	2	2	2	5	4	4	4	4	5	1	4	3.2	1	4	2	5	6	4	5	8	4	4	4.3
2	1	1	2	2	2	2	3	3.05	3	3	3	3	2	2.4	1	2	3	1	3	4	1	5	3	2	2.5
3	1	1	1	2	4	1	4	3	1	4	3	1	4	2.3	2	2	2	4	4	5	3	5	3	3	3.3
4	1	0	1	2	3	4	4	3	3	6	2	2	2	2.6	0	2	1	2	4	3	3	5	2	4	2.6
5	2	0	3	0	3	2	3	2	3	4	2	2	2	2.2	0	3	0	4	5	4	5	3	2	3	2.9
6	2	1	1	2	3	3	3	3	3	3	2	2	3	2.3	0	1	2	2	3	3	1	4	1	4	2.1
7	2	1	5	2	2	6	2	2	3	6	2	3	3	3.2	1	2	1	1	3	3	1	5	2	3	2.2
8	2	1	2	1	3	3	4	3	2	2	2	3	3	2.4	0	2	1	2	3	3	1	1	1	4	1.8
9	2	0	4	4	3	5	2	4	4	4	3	1	3	3.0	0	1	1	2	5	2	4	4	3	3	2.5
10	2	1	4	2	4	5	3.05	5	4	4	4	5	3	3.7	2	3	2	5	3	3	5	7	6	3	4.1
11	2	0	0	1	1	2	2	0	2	1	1	1	1	1.0	0	0	0	1	1	1	1	0	2	2	0.7
12	2	1	2	2	4	4	2	2	3	10	3	3	3	3.4	1	2	2	3	5	3	1	5	3	1	2.6
13	2	1	2	2	4	8	3	3	6	2	2	2	2	3.3	1	2	1	4	4	3	2	3	2	3	2.5
14	2	0	3	2	2.95	5	3	3	5	3	3	3	3	3.0	0	3	2	6	3	3	2	7	2	5	3.3
15	2	1	1	1	2	5	2	2	2	2	1	3	3	2.0	0	1	1	2	3	2	3	6	2	3	2.3
16	3	0	2	0	2	2	2	2	2	0	2	2	2	1.4	1	0	1	1	3	2	1	2	2	2	1.5
17	3	0	0	2	3	2	2	4	7	0	0	0	0	2.0	0	1	0	1	0	2	0	0	0	1	0.6
18	3	0	0	0	2	3	2	0	6	2	2	2	2	1.7	0	2	0	1	2	2	1	3	1	2	1.4
19	3	0	0	0	2	1	2	3	5	2	2	2	2	1.7	0	1	0	0	3	2	1	2	2	2	1.3
20	3	1	4	1	2	3	4	1	1	2	2	2	2	2.1	2	3	1	3	2	2	2	3	4	2	2.4
21	3	0	2	0	1	1	2	0	4	1	2	2	2	1.3	0	1	0	3	2	3	2	4	2	1	1.8
Total														2.4											2.3

Note. T1 = Time 1. T2 = Time 2. St = Student. L = Level. S = Sentence. M = Mean. Light gray shadow = Improvement after treatment in the individual scores. Dark gray shadow = Improvement after treatment in means.

Unfilled pauses – individual values and means for the experimental group (cleaned data)

St.	T 1										T 2													
	L	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	M	L	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	M
1	1	0	0	0	3	3	4	1	2	2	3	1.8	0	0	0	3	1	3.38	2	1	0	2	1.4	
2	1	1	1	1	4	7	2	4	4	1	3	2.8	1	1	2	2	3.38	3	4	2	3	3	2.4	
3	1	0	0	0	1	0	4	0	1	2	0	0.8	0	0	0	0	1	3.38	2	1	1	0	0.9	
4	1	1	2	2	3	4	3	1	5	3	2	2.6	0	3	1	0	3.38	3	2	2	3	3	2.0	
5	1	0	0	0	2	0	1	2	1	1	2	0.9	0	0	0	2	3.38	3	2	1	1	1	1.3	
6	1	0	3	0	1	6	3	3	6	3	3	2.8	1	4	0	3	5	2	2	3	3	3	2.6	
7	1	0	3	2	3	4	5	2	4	3	3	2.9	0	2	2	3	3	2	2	5	3	2	2.4	
8	1	1	2	2	1	3	6	2	5	3	1	2.6	0	2	2	2	5	3	1	5	3	3	2.6	
9	1	0	1	0	1	2	4	3	3	3	3	2.0	0	1	0	0	3.38	2	1	2	3	1	1.3	
10	1	0	0	1	3	3	3	3	4	3	2	2.2	0	1	1	1	3	3	3	0	1	1	1.4	
11	2	0	0	1	4	4	1	2	2	5	1	2.0	1	2	1	3	4	3	2	2	3	3	2.4	
12	2	0	1	1	6	3	2	2	3	3	2	2.3	0	0	1	2	3	2	2	2	3	2	1.7	
13	2	1	4	2	3	6	3	3	7	3	3	3.5	1	1	1	3	5	4	3	6	2	3	2.9	
14	2	1	3	1	5	5	3	3	8	6	3	3.8	0	2	1	5	5	5	3	7	5	3	3.6	
15	2	0	1	0	0	1	0	0	3	2	1	0.8	0	0	0	0	3.38	1	1	4	2	1	1.2	
16	2	1	4	1	6	4	1	3	4	3	4	3.1	1	4	1	4	4	2	3	4	2	4	2.9	
17	3	1	2	0	3	5	2	5	3	4	3	2.8	0	2	1	1	3	1	3	4	2	3	2.0	
18	3	0	1	1	3	3	0	2	1	4	3	1.8	0	0	0	0	4	0	1	1	1	1	0.9	
19	3	1	4	1	1	3	5	3	7	4	4	3.3	1	1	1	4	5	4	3	2	2	2	2.5	
20	3	1	4	2	4	6	3	2	2	3	3	3.0	1	4	1	6	5	3	2	5	3	2	3.2	
21	3	1	3	1	2	3	3	2	3	3	1	2.2	0	2	0	2	2	1	0	1	1	1	1.0	
Total												2.4												2

Note. T1 = Time 1. T2 = Time 2. St = Student. L = Level. S = Sentence. M = Mean. Light gray shadow = Improvement after treatment in the individual scores. Dark gray shadow = Improvement after treatment in means.

Appendix D: Comparison of unfilled pauses – t-test results

Comparison of number of pauses across Time (T1 vs T2) and Group (Control vs Experimental) – t-tests results

Sentence	T-test		df	T	p-value	Cohen's d
1	Paired	C T2-T1	20	.370	.715	.0814
		E T2-T1	20	-		-.250
	Independent	C T1-E T1	40	.302	.765	.078
		C T2-E T2	40	1.227	.227	.382
2	Paired	C T2-T1	20	-.484	.634	-.106
		E T2-T1	20	-		-.278
	Independent	C T1-E T1	40	.209	.836	.061
		C T2-E T2	40	.766	.448	.240
3	Paired	C T2-T1	20	-		-.365
		E T2-T1	20	1.673	.110	
	Independent	C T1-E T1	40	.000	1	.000
		C T2-E T2	40	1.870	.069	.584
4	Paired	C T2-T1	20	-.852	.404	-.186
		E T2-T1	20	-		-.304
	Independent	C T1-E T1	40	1.395	.178	
		C T2-E T2	40	-.117	.908	.036
5	Paired	C T2-T1	20	-.241	.812	-.052
		E T2-T1	20	.678	.505	.148
	Independent	C T1-E T1	40	-.164	.871	.048
		C T2-E T2	40	-		.402
6	Paired	C T2-T1	20	1.302	.201	
		E T2-T1	20	-.024	.981	-.005
	Independent	C T1-E T1	40	-		-.244
		C T2-E T2	40	1.116	.278	
7	Paired	C T2-T1	20	-.240	.811	.074
		E T2-T1	20	.816	.419	.247
	Independent	C T1-E T1	40	-		-.254
		C T2-E T2	40	1.164	.258	
8	Paired	C T2-T1	20	-		-.400
		E T2-T1	20	1.826	.083	
	Independent	C T1-E T1	40	.724	.473	.219
		C T2-E T2	40	.830	.412	.253
8	Paired	C T2-T1	20	.000	1	.000
		E T2-T1	20	-		-.477
	Independent	C T1-E T1	40	2.186	.041*	
		C T2-E T2	40	.416	.680	.130

		C T2-T1	20	.181	.858	.040
	Paired	E T2-T1	20	2.855	.010**	-.623
				-		.768
9	Independent	C T1-E T1	40	2.480	.017*	
		C T2-E T2	40	-.255	.800	.074
		C T2-T1	20	1.563	.134	.341
	Paired	E T2-T1	20	1.188	.249	-.260
10	Independent	C T1-E T1	40	-.144	.886	.047
		C T2-E T2	40	1.869	.069	.568

Note. C = control group. E = experimental group. T = time. Paired = paired-samples t-test. Independent = Independent-samples t-test. Cohen's *d* = Effect size.

* $p < .05$.

Sobre as formas de tratamento no Português em Angola: o caso de você¹

Queneth Pires

pireskenneth21@hotmail.com

Universidade do Porto [Portugal]

Instituto Superior de Ciências de educação da Huíla [Angola]

ABSTRACT.

The starting point for this study is the observation that, by generalizing its use in the Angolan context, the treatment form *you* will acquire, as is registered in the diatopic and diastratic varieties of PE, distinct semantic values, pragmatic and discursively determined. It is because of the context (formal, informal) and the relationship (vertical, horizontal) that *you* registers considerable rates of occurrence in Angola, including in contexts in which, according to the standard norm, such use is discouraged. Data were obtained by applying a questionnaire with twenty questions, all simulating real communication situations. The questionnaire was distributed and filled out online by 115 participants of different educational levels, provinces of residence, age group and sex. From the data analysis, which follows an eminently qualitative approach, it was possible to identify different types of use of *you*: *you of intimacy*, *you of intimacy-affection*, *you of respect-affection*, *you of superiority*, *you of inferiority*, *you of detachment*; *you interpolate*, *you derogatory*. The discursive context, the type of expressive act, of praise or criticism, formal or informal, the relationship between the interlocutors, of power, solidarity, or proxemics, are determining variables for the association of one of those semantic values to *you*. The conclusion is the following: when a positive semantic value is associated with *you*, it occurs in an expressive speech act of praise and the pronominal treatment presents percentage data (very) close to the data related to the nominal treatment; when, on the contrary, it occurs in expressive acts of criticism, it reveals itself to be considerably threatening from the negative face, which is reflected in the reduction of its use, translated into (very) low percentage rates, compared to percentage rates related to nominal treatment. Therefore, the less used the *you* will be, the more threatening it is to the negative face of the addressee.

KEYWORDS.

Angolan Portuguese; *you*; expressive act; positive face; negative face.

RESUMO.

O ponto de partida para este estudo é a constatação de que, generalizando-se o seu uso no contexto angolano, o *você* vai adquirindo, a exemplo do que se regista nas variedades diatópica e diastrática do PE, valores semânticos distintos, pragmática e discursivamente determinados. É em

¹ Agradecimentos especiais às professoras doutoras Maria Alexandra de Araújo Guedes Pinto e Isabel M. Duarte, pelos importantes contributos.

virtude do contexto (formal, informal) e da relação (vertical, horizontal) que você regista índices consideráveis de ocorrência em Angola, incluindo nos contextos em que, de acordo com o padrão, tal uso é desaconselhado. Os dados foram obtidos pela aplicação de um questionário com vinte perguntas, todas simulando situações reais de comunicação, o qual foi distribuído e preenchido *online*, por 115 participantes de distintos níveis de escolaridade, províncias de residência, faixa etária e sexo; da análise dos dados, que segue uma abordagem eminentemente qualitativa, foi possível identificar diferentes tipos de uso de você: *você de intimidade*, *você de intimidade-afeto*, *você de respeito-afeto*, *você de superioridade*, *você de inferioridade*, *você de distanciamento*; *você interpolar*, *você depreciativo*. O contexto discursivo, o tipo de ato expressivo, de elogio ou de crítica, formal ou informal, a relação entre os interlocutores, de poder, de solidariedade ou proxêmica, são variáveis determinantes para a associação de um daqueles valores semânticos ao você. A conclusão é a seguinte: quando se associa ao você um valor semântico positivo, este ocorre num ato de fala expressivo de elogio e o tratamento pronominal apresenta dados percentuais (muito) próximos dos dados relativos ao tratamento nominal; quando, ao contrário, ocorre em atos expressivos de crítica, revela-se consideravelmente ameaçador da face negativa, o que se reflete na redução do seu uso, traduzido em índices percentuais (muito) baixos, comparativamente aos índices percentuais relativos ao tratamento nominal. Portanto, tanto menos usado será o você, quanto mais ameaçador se revelar para a face negativa do alocutário.

PALAVRAS-CHAVE.

Português em Angola; você; ato expressivo de elogio; ato expressivo de crítica; face positiva; face negativa.

0. Introdução

Os falantes do português têm, para distintas finalidades – alocução, elocução, interlocução –, um conjunto diverso de formas de tratamento, assim categorizadas: nominais, pronominais e verbais; a opção por uma destas muitas possibilidades, de acordo com Nascimento (2020: 2702), é um exercício para o que intervêm, além das opções pessoais, fatores de natureza sociocultural: i. o contexto – formal ou informal – em que se insere o discurso; ii. a hierarquia dos papéis sociais e profissionais, de que constituem exemplo as relações entre dirigentes e subordinados, professores e alunos, pais e filhos; iii. os traços particulares dos interlocutores – além da idade, das competências linguística e comunicativa e do género, inclui-se nesta categoria o grau de conhecimento mútuo; e iv. “as tradições próprias de certos grupos sociais ou familiares e de certas regiões” (*idem, ibidem*).

Para o falante comum, a escolha de uma ou de outra forma de tratamento

(FT) nem sempre se revela um exercício fácil; no contexto angolano, a dificuldade pode ilustrar-se por recurso ao comentário seguinte²:

- (1) “Professor, se eu usar você com um senhor que encontrar na rua, vai parecer desrespeitoso; o tu não posso mesmo usar; mas se eu usar senhor, também posso parecer mal educado!”.

Junte-se-lhe estes, retirados de Teixeira (2008: 11):

- (2) “Eu acho, eu entendo que (você) é um termo de inferiorizar alguém”.
(3) “A um amigo, chamo de uma maneira particular e digo-lhe mesmo: você.”
(4) “Eu gosto muito desse termo “você”. “Tu” na minha boca não.”

No comentário (1), revela-se a dificuldade em selecionar a FT adequada às características do contexto e do alocutário; no (2), o valor hierarquizante, ofensivo de *você*; no (3) e no (4), a consciência dos participantes relativamente aos usos preferenciais de *você*, em detrimento de *tu*; nos quatro comentários, talvez menos evidentemente no (4), o saber atinente aos valores hierarquizante, igualitário e/ou proxémico associados quer a *você*, quer a *tu*, do que resulta a sua interpretação como duas FT potencialmente ameaçadoras da face negativa, se utilizadas em contextos marcados por relações de poder e/ou distanciamento entre os interlocutores (cf. comentário (1)). Assim, se o objetivo for igualmente o de não “parecer desrespeitoso”, ou, conforme o segundo comentário, não “inferiorizar alguém”, portanto, preservar a face positiva, a estratégia deverá consistir em evitar o tratamento por *tu* ou por *você*.

Há, com efeito, usos merecedores de atenção:

- (5) “Professor, vamos ficar como, sem você?”³
(6) “Mãe, aqui mesmo você é a única que me entende!”.⁴

2 Feito por um estudante, durante uma aula de Português I, unidade curricular do primeiro ano do curso de Licenciatura em Ensino da História, ministrado pelo Instituto Superior de Ciências de Educação da Huíla.

3 Idem.

4 Idem.

Professor e mãe situam-se, à partida, em contextos distintos – formal para o primeiro e informal para o segundo; denotam, entretanto, um mesmo tipo de relação, assimétrica, e ocorrem num mesmo tipo de ato de fala, o expressivo (de elogio). O recurso a *você*, nestes dois exemplos, pode assim explicar-se: está a generalizar-se o uso do *você* no contexto angolano (cf. Teixeira 2008: 9); mas este uso, como ficará demonstrado no decurso do presente estudo, não ignora nem as especificidades do contexto discursivo, nem as do alocutário; é, pois, provável que, em virtude do seu uso generalizado, o *você* ocorra associado a distintos valores semânticos, discursiva e pragmaticamente determinados, i.e., como índice de deferência, de afetividade e, por isso, a interpretar-se, para só citar estes casos, como um “*você* de respeito”, descrito em Bazenga (2017: 3); para tanto, será relevante a categoria do ato de fala (expressivo de elogio) em que se insere a FT, como ficou ilustrado, pelos comentários (5) e (6).

Interessava, assim, pelo que ficou demonstrado até aqui, i. verificar os valores associados à utilização do *você*, no contexto angolano; ii. identificar as variáveis pragmático-discursivas (ato de fala, contexto de elocução, tipo de relação, faixa etária) determinantes para que o falante adote ou evite o tratamento por *você*.

1. Da instabilidade do *você* no português europeu

Considerando-se o nível supradialetal, o padrão europeu⁵, *você* tem uso muito restrito entre os mais velhos, mas também entre os falantes mais cultos; a condição básica é a cortesia verbal, a preservação da face. Ocorre, por isso, em contextos marcados por graus mais ou menos acentuados de formalidade, assimetria, proximidade, denotando entre os intervenientes “uma relação social de igualdade (por exemplo, entre colegas de trabalho sem grande intimidade) ou de superioridade (de patrão para empregado)” (Nascimento 2020: 2711), donde advém o uso desaconselhado quando em referência a um destinatário superiormente localizado, na relação de horizontalidade, ou a um destinatário que, situando-se ao mesmo nível

5 Norma de referência para o português que se fala em Angola.

na estrutura hierárquica, mantém com o locutor um certo distanciamento. Convém, entretanto, notar o seguinte: mesmo nos casos em que, como dito previamente, a entidade referida pela FT é um subordinado, o risco de o seu uso se revelar pouco cortês, portanto, ameaçador da face negativa, é imanente, pelo que se tornam preferíveis as “formas nominais alternativas, como *o senhor, a senhora*⁶, o nome próprio do ouvinte, etc.” (*idem, ibidem*); o sujeito nulo (com o verbo na terceira pessoa do singular) é, a par do tratamento nominal, uma estratégia eficaz.

Com efeito, e porque uma língua não se realiza uniformemente, o tratamento pronominal por recurso a *você* pressupõe variações – diastrática e diatópica, não faltando referências e exemplos de variação diafásica. Disto resulta a complexidade do seu uso em PE: Matoso (2003), Gouveia (2008) Duarte (2010), Cavalheiro (2016). É também em virtude destas variações que se vai registando uma diversidade de valores semânticos, discursiva e pragmaticamente determinados; ou seja, não estando totalmente definidos estes valores (Marques 2008: 286), é na interação discursiva que se atualiza, delimitando-se cada valor conforme a intencionalidade e a interpretação que se lhe atribui, a relação que se estabelece entre os interlocutores, as características do espaço e do tempo discursivos, etc. Fala-se, assim, de um certo *você* no Norte, de um *você* no Sul; assim como se fala de um *você* entre a classe alta, e um outro entre a classe média; Marques & Duarte (2019) dão conta de um *você* por cujas propriedades contextuais resultaria ofensivo; Marques (*idem*, 286) alude a um tipo que, pelas mesmas razões, poder-se-ia considerar um “*você depreciativo*”; Lešková (2012: 30-31), a dois outros: um “*você de intimidade*” e um outro, que se poderia considerar de neutralidade, como é o que se usa nas rádios e televisões, por jornalistas, apresentadores, modelos publicitários, etc., para se dirigir ao seu público. Junte-se-lhes estes outros, retirados de Bazenga, citando Hammermüller (1993):

o “*você*”⁷ de respeito; o “*você*” de igualdade; o “*você*” de inferioridade; o “*você*” que elimina a ambiguidade⁸ (nas situações ambíguas onde não queremos ser informais, tratando a pessoa por

6 Grifo da autora.

7 As aspas são da autora.

8 Para facilidade de leitura, será denominado “*você interpolar*”.

“tu”, ou ser muito formais, tratando alguém por “o/a senhor/a”); o “você” afetivo, com pessoas conhecidas dentro da família; o “você” de distanciamento, quando o locutor quer manter a distância com a pessoa a quem se dirige; e, por fim, o “você” metalinguístico, ou seja, quando usado para referir um “você”, como em “Você é um pronome” ou “Você é o sujeito da frase”, sendo desprovido de significado

(Bazenga 2018: 3)

Tantos valores para uma mesma palavra acarretam consequências; Gouveia (2008: 94) refere-se a uma “quase generalização do uso de você em vez de o *Senhor*, a *Senhora*, etc.”; para Nascimento *et al.* (2018: 251), que o situam principalmente entre os mais jovens e os menos cultos, o fenómeno consiste no uso indiscriminado de *você* direcionado a qualquer pessoa; Marques & Duarte (2019: 247) apontam para o mesmo, tendo demonstrado, entretanto, que “uma crescente ocorrência e aceitação por parte de muitos falantes coexiste com esta dimensão negativa”.

Não raras vezes, este uso indiscriminado cria ruído na comunicação, i.e., incide negativamente sobre a informação veiculada; noutras vezes, em que a consequência se revela mais severa e mais prejudicial também, gera-se uma espécie de tensão e/ou conflito de proporções mais ou menos graves, como o demonstraram Marques & Duarte (*idem, ibidem*), de quem provém o primeiro dos três comentários seguintes: “convém evitar que se repita. Isso desvaloriza e amarfanha o conteúdo de qualquer declaração, mesmo entre a classe operária” (Marques & Duarte: *idem, ibidem*); “Se me tratam por você, inadequadamente, isto é, alguém que por convenção social eu considere inferior a mim, eu fico ainda mais surdo do que sou.” (Lešková 2012: 30-31); “Tal como me irrita quando vejo/oiço um jornalista tratar por você alguém com o dobro da idade dele.”⁹

Fica assim evidente uma bipolarização; de um lado, os que o aceitam naturalmente – são 36,36%, segundo os dados de Lešková (*idem*, 59); de outro, os que se insurgem – são 63,4%, segundo os dados da mesma autora. Fica também evidente que as consequências de um uso inapropriado de

9 Excerto de um artigo de opinião publicado em <https://www.jn.pt/opiniaio/alice-vieira/senhoras-donas-por-favor-1018968.html>; data de consulta: 26 de janeiro de 2022.

uma forma de tratamento, mesmo que de uso tão generalizado, como é o caso de *você*, estão longe de se circunscrever ao domínio linguístico. Tem, portanto, implicações mais ou menos evidentes nas relações humanas, desde as socioculturais, profissionais, religiosas, comerciais, etc.

A literatura relaciona a generalização do *você* no PE a fatores e/ou agentes de ordem diversa: i. a incapacidade de a escola proporcionar uma sólida formação linguística (cf. Duarte 2010) aos seus estudantes, estabelecendo-se, assim, entre o ensino e a competência linguística e/ou comunicativa uma relação de causa e consequência; ii. a exposição de falantes do PE ao PB, mas não só, donde resulta uma “eventual influência do PB, do castelhano e do inglês, mais “igualitários”¹⁰ do que o PE” (*idem, ibidem*); iii. Matoso (2003: 278) refere-se-lhe como um índice de “má educação” do indivíduo. Uma outra causa para essa generalização é a necessidade de adequação discursiva, i.e., optar por *você* nos contextos em que o bom senso desaconselha quer o *tu*, quer o *você*; portanto, “é possível que o uso mais frequente de *você* corresponda à necessidade de suprir a ausência de um pronome de 2.^a pessoa a usar em conversas informais, mas sem intimidade, opondo-se, portanto, ao pronome *tu*” (Nascimento *et al.* 2018: 251). Note-se, entretanto, que estas razões não se podem compreender senão pela íntima relação que mantêm entre si, mas também com esta outra, provavelmente a primeira e mais eminente entre todas, e que tem que ver com uma reconfiguração em curso na sociedade portuguesa contemporânea, assim descrita por Gouveia:

Diria que estamos presentemente a sofrer as transformações da passagem de um sistema de face e solidariedade de base hierárquica e deferencial para um sistema de base igualitária e de envolvimento. No primeiro, o de base hierárquica e deferencial, que diria tradicional, os actores sociais, por um lado, reconhecem e respeitam as diferenças sociais que os colocam em posições relativas no quadro das relações interpessoais, e, por outro, usam estratégias de distanciamento e de deferência relativamente ao outro. No segundo, o de base igualitária e de envolvimento, que diria pós-tradicional, os actores sociais não sentem que existam diferenças de poder e de envolvimento entre as

10 As aspas são da autora.

peçoas com quem interagem e codificam as suas relações em termos de proximidade e de familiaridade

(Gouveia 2008: 97)

Em síntese: mais do que um facto de língua, as FT constituem “uma categoria pragmática com impacto direto nas relações estabelecidas por cada comunidade social e linguística” (Marques & Duarte 2019: 236), devendo o seu uso obedecer ao conjunto dos fatores pragmáticos, discursivos e sociais previamente indicados. *Você* é uma FT; não deve, por conseguinte, constituir exceção.

2. O caso do PA¹¹

Dos autores que sabemos dedicarem-se às FT no PA, merece destaque Teixeira, que, interessada nas similaridades entre o PB e o PA, desenvolveu um trabalho investigativo cujos resultados, descritos em “O Pronome *Você* no Português de Luanda”, foram apresentados no *I Simpósio Mundial de Estudos da Língua Portuguesa (SIMELP)* – em 2008; é a primeira que sabemos dedicar-se ao assunto no contexto angolano, tendo-se cingido ao uso das formas pronominais *tu* e *você*, não obstante a referência, muito breve, às formas nominais, sobre que assevera: “Entre os usuários dessa norma, somente na relação jovem X velho observa-se um tratamento respeitoso ou cortês. As formas corteses mais usadas são: a tia, paizinho, o avô Fulano, entre outras do mesmo tipo” (Teixeira 2008: 9). Esta questão será retomada anos mais tarde, por outros autores: Balsalobre (2015, 2017), Silva (2020), Naege (2021).

Cruzando os dados apresentados em cada um dos trabalhos supracitados, é possível definir algumas tendências:

- i. Em virtude dos graus mais elevados de delicadeza que se lhes associa, as FT nominais têm largo uso no contexto angolano, sobrepondo-se quer às pronominais, quer às verbais;

¹¹ Português em Angola.

- ii. Quando o tratamento se dá pronominalmente, usa-se *tu* ou *você* indiferentemente;
- iii. Não obstante a paridade, *você* sobrepõe-se a *tu*, um facto de importância assinalável para o uso generalizado no contexto nacional (cf. Teixeira 2008; Balsalobre 2015).

Atente-se ainda para este dado: é frequente a combinação de *você* com o verbo na segunda pessoa (“...mas como *você* crescestes na cidade...”¹²), intimamente relacionada ao grau de escolarização e à língua materna do participante, donde resulta a assunção de que, quanto mais instruído, menor é a probabilidade deste tipo de ocorrência no seu discurso – não se ignore, entretanto, os 38% relativos ao fenómeno entre os estudantes universitários que participaram da pesquisa; os 80% atinentes à prática entre os falantes que têm como L1 uma língua de origem africana demonstram situar-se neste grupo o maior índice de ocorrência do sujeito pronominal da segunda pessoa com o verbo na mesma pessoa.

Segundo os dados apresentados por Silva (2020: 173), só 10,1% de jovens (inquiridos pelo autor) até 25 anos usa uma forma pronominal (*tu*, *você*) para se dirigir ao seu progenitor; entre os 26 e os 30, o número sobe para 23,7%, baixando para 10,3%, se se considerar o intervalo entre 31 e 35 anos; nos demais grupos etários, os números são os seguintes: 36-40 (3,6%), 41-45 (0%), acima dos 46 (5,3%). Em sentido inverso, i.e., de pais para filhos, os valores correspondentes ao tratamento pronominal (32,5%) estão muito próximos dos valores referentes ao tratamento nominal (31,6%). Fica, assim, evidente uma clara tendência para o tratamento nominal de filhos para pais (*tu*, *você*, *pai*, *papá*, *mãe*, *senhor*, *kota*, *velha(o)*, *pai grande*, *mana(o)*, *chefe*, *dona*, *mamoite*, *papoite*, *boss*); quando em sentido inverso, o tratamento ocorre indiferentemente sob a forma nominal ou pronominal (*tu*, *você*, *filho(a)*, *querido(a)*, nome próprio). Na relação entre professores e alunos, a forma privilegiada é a nominal, quando do aluno para o professor (professor(a), senhor(a) professor(a)); se, ao invés, o tratamento ocorrer por uma FT pronominal (*você* (sucedido ou antecedido de *professor(a)*, *senhor(a) professor(a)*), os valores percentuais não chegam aos 10%; oscilam entre

12 O exemplo é de Teixeira (2008: 8); o grifo é nosso.

1,3% e 9,5%, considerando-se os diferentes grupos etários. Na relação entre crianças e adultos, dá-se o mesmo: predomina o tratamento nominal (76,3%), seguido de um tratamento simultâneo (nome e pronome) (17,5%) e, finalmente, pronominal (5,2%). Diferentemente do que é recomendável, segundo a norma padrão, a direção de um superior hierárquico para um seu subordinado se dá tendencialmente sob a forma pronominal (*chefe* (seguido ou não de nome próprio), *boss*, *doutor/engenheiro* (seguido ou não de nome próprio; *mano(a)*), demarcando muito nitidamente a posição subalterna do alocutário. A tendência é semelhante, quando considerada a relação entre colegas: os mais novos tendem a dirigir-se aos mais velhos por uma FT nominal (*colega* (seguido ou não de nome próprio), *tio(a)* (seguido ou não de nome próprio); os mais novos, aos mais velhos, pronominalmente (*tu*, *você*).

Os dados ora apresentados corroboram as conclusões de Teixeira (2008) e Balsalobre (2015), de acordo com as quais, no PA, privilegia-se o tratamento nominal, em virtude do seu elevado grau de cortesia. Não deixa, entretanto, de ser curioso o facto de se terem registado, nos estudos citados, tratamentos pronominais de filhos para pais, de subordinados para chefes, de mais novos para mais velhos (Boa tarde, senhora. Faz favor. Podes me indicar aqui alguma pastelaria, restaurante que possa apanhar alguma coisa, faz favor?)¹³; pode-se, simplesmente, associar o facto a um possível baixo nível de instrução; ou, simultaneamente, considerar estar em curso um processo de ressignificação, o qual consiste em associar ao *você* valores semânticos até agora não reconhecidos, donde resulta a manutenção da interdição do seu uso em determinados contextos.

Para concluir: a dificuldade em reconhecer, com o rigor que se impõe, os diferentes valores associados à utilização do *você* no PA é, por si mesma, reveladora de uma necessidade premente: desenvolver estudos, redirecionando o foco, que deverá então passar por identificar, descrever e sistematizar os valores semânticos que decorrem da variação no seu uso, bem como os fatores pragmático-discursivos que se lhe associam, incluindo aqueles usos em que *você* ocorre com o verbo na segunda pessoa.

13 Exemplo tal qual recolhido em Balsalobre (2015: 240).

3. Metodologia

A realização deste estudo obedeceu a um conjunto de etapas interdependentes, podendo dividir-se, fundamentalmente, em dois momentos: a conceção e a execução (parte empírica); neste último, procedeu-se à maior parte das tarefas, nomeadamente, a pesquisa bibliográfica, ao fim da qual se deu início à redação da parte teórica, coincidente com a elaboração e aplicação do questionário, a respeito do que importam as seguintes observações: para a sua elaboração, teve-se em conta as variáveis que se pretendiam testar, a saber, i. o tipo de ato expressivo, de elogio ou de crítica; ii. o contexto discursivo, formal ou informal; iii. a diferença de idade; iv. a relação hierárquica e/ou proxémica entre os interlocutores. Pretendia-se, assim, determinar a correlação entre a interpretação da FT pronominal em estudo e cada uma dessas variáveis. Procedeu-se à elaboração de perguntas para cujas respostas os participantes teriam de considerar, cumulativamente ou não, cada uma das variáveis referidas. Para cada pergunta, havia duas respostas possíveis, podendo o participante seleccioná-las em simultâneo; cada uma destas possibilidades de resposta constituía um ato de fala expressivo (de elogio ou de crítica) em que se inserem o *você* e uma FT nominal. Uma vez redigido o questionário, seguiu-se a distribuição, enviando-se, através do WhatsApp, o link gerado automaticamente pelo Google Docs, a pessoas e/ou a grupos; 115 pessoas responderam.

Encerrado o processo de respostas ao questionário, procedeu-se, simultaneamente, à análise e à interpretação dos dados, apresentados na secção correspondente; a abordagem é mista, i.e., qualitativa-quantitativa.

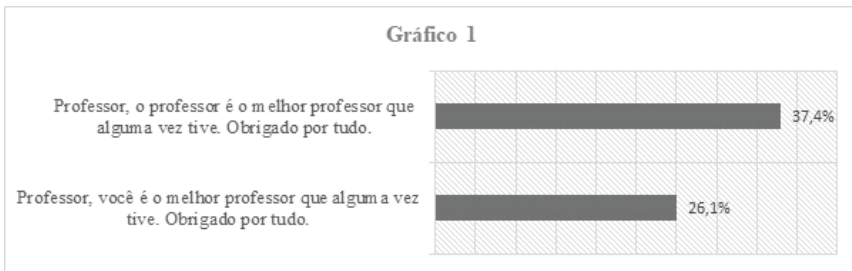
Para o acompanhamento da análise dos dados apresentados, considere-se o seguinte: não tendo sido possível distinguir, apenas pelas perguntas do questionário, com a precisão e o rigor necessários, entre o *você* de afeto, o *você* de intimidade e o *você* de respeito, caracterizados na parte teórica deste estudo, optou-se por denominar *você de intimidade-afeto*, para a relação entre pais e filhos, e *você de respeito-afeto*, para a relação entre professores e alunos; as demais denominações mantiveram-se como na parte teórica deste estudo: *você de igualdade*; *você de inferioridade*; *você interpolar*; *você de distanciamento*; *você metalinguístico*; *você depreciativo*; *você de neutralidade*.

Apesar de ter havido no questionário perguntas que visavam testar a interpretação de *você* como um facto associado quer à posição (inicial ou final) do vocativo, quer à ocorrência ou não de atenuadores da força ilocutória, estes dados foram descartados, em virtude de a análise preliminar ter revelado insignificantes. Por último: optou-se por não cruzar as variáveis aqui exploradas e outras, como o sexo, o grau académico, a província de residência, considerando a extensão do estudo.

4. Apresentação e discussão dos resultados

Pergunta 1 - Imagine-se na última aula do ano, com o professor de quem mais gosta; que frase, entre as sugeridas abaixo, usaria para que melhor expressasse o seu sentimento de gratidão e/ou admiração por este professor?

GRÁFICO 1 – Resultados da pergunta 1

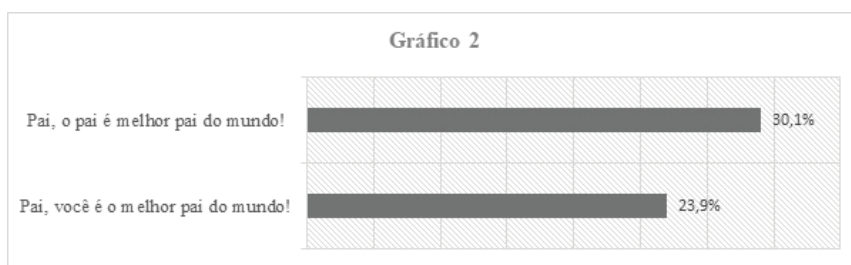


Os dados apresentados neste primeiro gráfico são referentes à primeira pergunta, com que se pretendia aferir, em relação ao PA, a ocorrência do *você* com valor positivo, inserindo-o num ato de fala expressivo de elogio, potencialmente favorecedor de tal ocorrência. O contexto discursivo para o qual remetem tanto a pergunta quanto as possibilidades de resposta pressupõe uma relação de hierarquia caracterizada pela deferência e pela relativa proximidade entre os intervenientes; era, por isso, expectável que, em respeito a um princípio elementar da cortesia, a proteção da face positiva do alocutário, a diferença entre as duas FT comparadas se revelasse expressiva. Tal não foi o caso, uma vez que, como demonstram os dados do

gráfico, a diferença situa-se na ordem dos dez pontos percentuais, o que se pode explicar pelo valor positivo associado à FT, donde, em decorrência, a legitimação do seu uso num contexto marcado por um grau de formalidade considerável. Justifica-se, assim, uma interpretação segundo a qual o *você*, naquele enunciado, é percebido como um *você de respeito-afeto*.

Pergunta 2 - Imagine que tenha recebido um belíssimo presente do seu pai; que frase, entre aquelas que se seguem, usaria para mais apropriadamente descrever a sua satisfação?

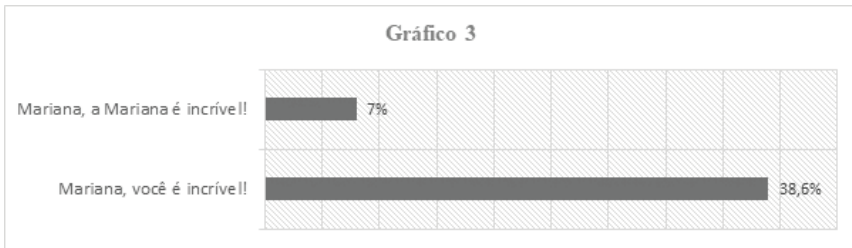
GRÁFICO 2 – Resultados da pergunta 2



Assim como no caso anterior, também aqui a diferença se constitui inexpressiva – são 30,1% para a FT nominal e 23,9% para a FT pronominal; conserva-se, entretanto, uma diferença que importa referir: partindo-se do princípio de que, pelas suas próprias especificidades, a relação entre pais e filhos, diferentemente da que se estabelece entre o professor e o seu aluno, é uma relação que torna muito mais propícia a criação de laços de intimidade, pode-se concluir que a ocorrência da FT *você* respeita a face positiva do seu interlocutor; e que, pelas mesmas razões, é interpretado, no caso dado, como o *você de intimidade-afeto*.

Pergunta 3 - Imagine que Mariana, sua subordinada, tenha executado muito bem uma tarefa. Selecione, por favor, entre as que se seguem, a frase que melhor descreveria a sua satisfação em relação à Mariana.

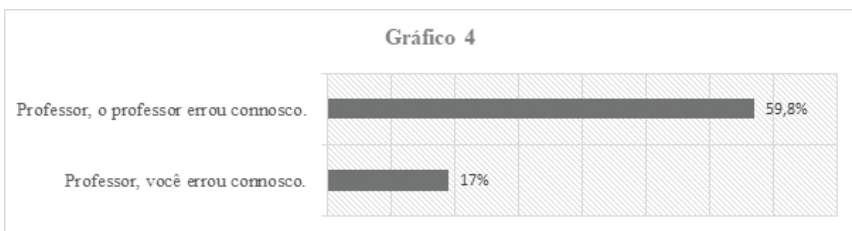
GRÁFICO 3 – Resultados da pergunta 3



Considerando-se que a pergunta que dá origem a este gráfico remete para uma relação assimétrica, caracterizada por um grau de relativa formalidade e proximidade, i.e., para uma relação nem muito formal, nem muito informal, nem muito íntima, nem muito afastada; que o ato de fala em que se insere o tratamento pronominal valoriza a face positiva do alocutário; que não se pode, por isso, associar à diferença expressiva entre o *você* (38,6%) e a FT nominal (7%) um tratamento descortês, mais precisamente, fortemente hierarquizante; será legítimo concluir que se trata de um *você* *interpolar*.

Pergunta 4 - Imagine que um professor seu tenha errado com a sua turma. Que frase, entre as sugeridas abaixo, usaria para melhor expressar o seu sentimento de descontentamento?

GRÁFICO 4 – Resultados da pergunta 4

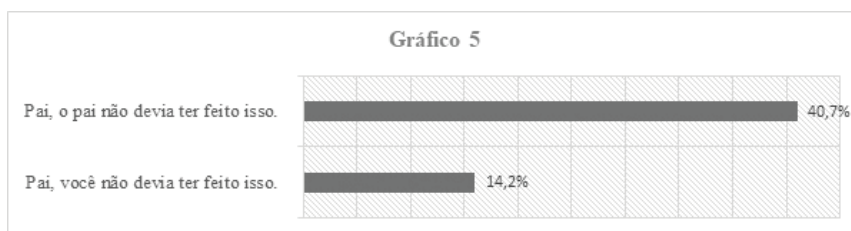


A quarta questão, assim como a primeira, remete para um mesmo contexto situacional e um mesmo tipo de relação entre os interlocutores; divergem, entretanto, na natureza dos enunciados apresentados como possibilidades de resposta, já que os da primeira constituem atos expressivos de elogio e os da segunda, atos expressivos de crítica; os dados obtidos a partir da primeira

revelam uma diferença não expressiva entre a FT pronominal e a nominal; ao contrário, aqueles que se obtêm a partir da quarta questão evidenciam uma disparidade expressiva entre as mesmas FT – são 17% relativos a *você* e 59,8% à FT nominal. Tal facto pode assim explicar-se: sejam ou não capazes de a explicitar, os participantes têm ciência de que, ocorrendo num ato como o expressivo de crítica, *você* constitui-se numa FT ameaçadora da face negativa; mais ainda quando usado para direção a um interlocutor que se situa no topo da hierarquia, como é o caso do professor; é, pois, considerando a diferença expressiva entre os dois tipos de FT, o tipo de ato de fala e a relação hierárquica, que se pode interpretar a forma pronominal como um *você depreciativo*, interpretação determinante para o baixo índice de ocorrência.

Pergunta 5 - Imagine-se na necessidade de repudiar o seu pai. Que frase, entre aquelas que se seguem, usaria para mais apropriadamente descrever a sua insatisfação?

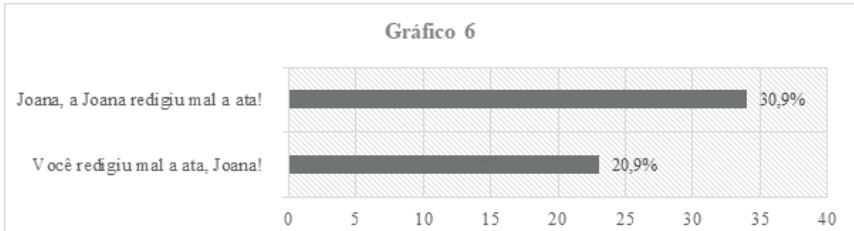
GRÁFICO 5 – Resultados da pergunta 5



Os 40,7% relativos ao tratamento nominal, contra os 14% relativos ao tratamento pronominal, evidenciam uma consciência presumivelmente sólida a respeito de um *você* por cuja interpretação se pode considerar ofensivo, ameaçador da face negativa, não obstante o grau de (relativa) intimidade que se estabelece entre o locutor e o alocutário, nas relações e nos contextos semelhantes aos que simulam os enunciados apresentados como possibilidades de resposta. Trata-se aqui também de um *você considerado depreciativo*, o que explica o baixo índice da sua ocorrência.

Pergunta 6 - Imagine-se triste com uma sua subordinada, uma senhora de idade avançada. Que frase, entre as sugeridas abaixo, usaria para se dirigir à senhora?

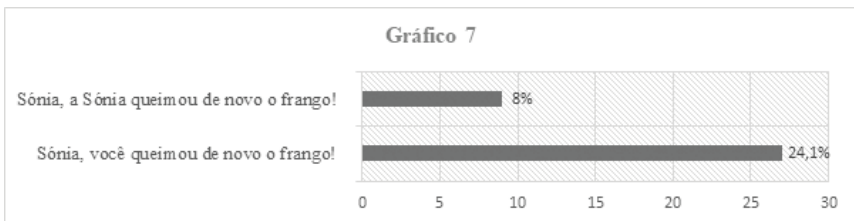
GRÁFICO 6 – Resultados da pergunta 6



Para a interpretação dos dados deste gráfico, é importante considerar que a pergunta remete para uma relação e um contexto em que, de acordo com o padrão, recomendável é o uso de uma FT nominal; neste sentido, o expeável era que um tratamento pronominal, a ocorrer, registasse níveis muito baixos. Tal não foi o caso, já que os 20,9% relativos à escolha de você para direção à subordinada constituem valores altos, se considerados os 30,9% relativos ao tratamento nominal. Tratar-se-á, por isso, de um *você* que, acentuando as diferenças entre ambos, pode interpretar-se como um *você de inferioridade*.

Pergunta 7 - Imagine-se triste com uma sua subordinada, uma jovem. Que frase, entre as sugeridas abaixo, usaria para se dirigir a esta jovem?

GRÁFICO 7 – Resultados da pergunta 7



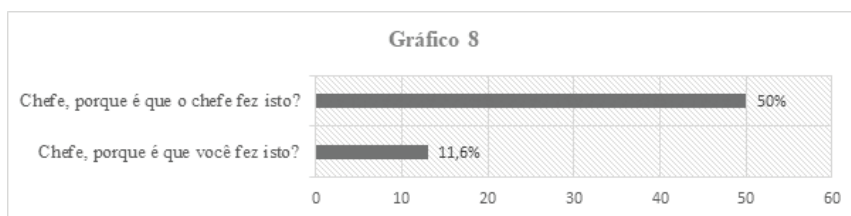
Apesar da identidade da relação e do contexto situacional simulados pela sexta e sétima perguntas, os dados desta remetem para uma diferença expressiva (24,1% para *você*, contra 8% para a FT nominal), contrastando com os da anterior; a justificação pode estar na faixa etária em que se situa a interlocutora do diálogo simulado pela sétima pergunta (e as possibilidades

de resposta). Trata-se, também aqui, de um *você de inferioridade*, o mesmo para o qual remetem as respostas à sexta questão, diferindo, no entanto, por ser, neste caso, fortemente hierarquizante; é em virtude disto que o tratamento nominal, recomendado pelo padrão, obteve índices percentuais tão baixos (8%), se considerados os 24,1% relativos ao tratamento pronominal.

Fica, assim, comprovada a pressuposição inicial, a partir da qual se compôs o par constituído pelas duas perguntas em questão: além da relação de poder, a diferença de idades determina a interpretação de *você*.

Pergunta 8 - Imagine que o seu superior hierárquico tenha feito algo que lhe tenha desagradado. Que frase, entre as sugeridas abaixo, usaria para lhe mostrar o seu descontentamento?

GRÁFICO 8 – Resultados da pergunta 8



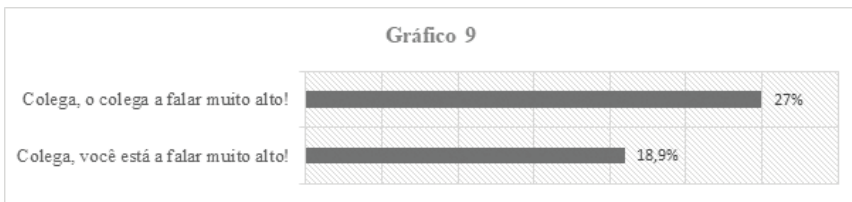
A interpretação dos dados relativos a esta questão pode fazer-se pela sua correlação com a quarta. Primeiro, porque as duas, assim como as suas possibilidades de resposta, simulam relações de assimetria; segundo, porque, apesar das diferenças decorrentes das suas especificidades, remetem para contextos formais, teoricamente marcados por um distanciamento em maior ou menor grau entre os interlocutores. Finalmente, porque os dados relativos às questões e às respetivas respostas apontam para o mesmo resultado, i.e., para uma interpretação que atribui ao *você* um valor negativo, em virtude do que se registam, para ambos os casos, baixos índices percentuais relativamente ao *você* – são 11,6% para esta questão e 17% para a quarta, contra os 50% referentes ao tratamento pronominal nesta e os 59,8% relativos àquela.

Corroboram, portanto, os dados até aqui apresentados, e que se interpretam como a demonstração de que, generalizando-se o seu uso no

PA, o falante usa ou não o *você* determinado pelo valor positivo ou negativo que lhe atribui, para o que intervém igualmente o tipo de ato fala expressivo.

Pergunta 9 - Imagine que começa um diálogo com um colega (de trabalho ou de escola) que ainda não conhece, mas que aparenta ter idade superior à sua. Que frase, entre as que seguem, usaria para se dirigir a este colega?

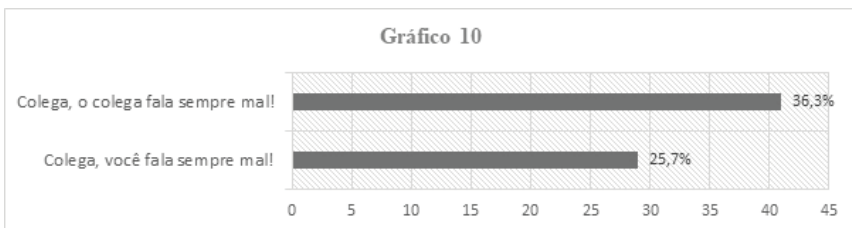
GRÁFICO 9 – Resultados da pergunta 9



Esta pergunta, bem como as suas possibilidades de resposta, simulam um tipo de simetria, caracterizado pelo distanciamento socio-afetivo entre os interlocutores, o que se pode constituir num fator determinante para a interpretação dos dados obtidos como apontando para um *você de distanciamento*. É isto o que explica os valores percentuais relativamente baixos (18,9%), tendo em conta os 36,9% referentes ao tratamento nominal.

Pergunta 10 - Imagine que começa um diálogo com um colega (de trabalho ou de escola) já conhecido, de idade superior à sua. Que frase, entre as que seguem, usaria para se dirigir ao colega?

GRÁFICO 10 – Resultados da pergunta 10



Os 25,7% referentes ao tratamento pronominal não são consideravelmente distintos dos 36,3% atribuídos ao tratamento pronominal; isto pode pressupor uma interpretação que resulta na atribuição de um valor proxémico, o qual permite, neste caso, a identificação de um *você de intimidade*.

5. Considerações finais

O ponto de partida para este estudo é a ideia de que, generalizando-se o seu uso no contexto angolano, o *você* vai adquirindo, a exemplo do que se regista nas variedades diatópica e diastrática do PE, valores semânticos distintos, pragmática e discursivamente determinados. Assim, tendo por base o valor associado a esta FT como um fator determinante para o seu uso (ou não) em contextos específicos de elocução, incluindo nos contextos em que, de acordo com o padrão, tal uso é desaconselhado, procurou-se identificar os valores semânticos que se lhe associam, e, concomitantemente, os fatores pragmático-discursivos inerentes a este uso e a esta associação.

É possível concluir, mediante a análise e discussão dos dados, que está a ocorrer, no contexto nacional, um processo de ressignificação dessa FT, em virtude do que resultam valores positivos e/ou negativos determinados contextualmente. Portanto, tal como no PE, em relação às suas variedades diatópicas e diastráticas, pode-se falar, no que se refere ao PA, de distintos tipos de *você*: *você de intimidade*, *você de intimidade-afeto*, *você de respeito-afeto*, *você de superioridade*, *você de inferioridade*, *você de distanciamento*; *você interpolar*, *você depreciativo*.

O tipo de ato expressivo – de elogio ou de crítica –, a diferença etária entre os interlocutores, o contexto discursivo, formal ou informal, a relação assimétrica – de poder ou de solidariedade – ou proxémica, constituem-se em variáveis determinantes para a associação de um daqueles valores ao *você*; pragmática e discursivamente, esta associação traduziu-se nos números percentuais aproximados entre esta FT, pronominal, e uma nominal, quando em atos de fala expressivos de elogio; quando em atos expressivos de crítica, a disparidade entre os dois tipos de FT tende a revelar-se expressiva; dito de outro modo: quando se associa ao *você* um valor semântico positivo – carinho, deferência –, este ocorre num ato de fala expressivo de elogio

e o tratamento pronominal apresenta dados percentuais (muito) próximos dos dados relativos ao tratamento nominal; quando, ao contrário, ocorre em atos expressivos de crítica, revela-se consideravelmente ameaçador da face negativa, o que se reflete na redução da sua ocorrência, traduzida em índices percentuais (muito) baixos, comparativamente aos índices percentuais relativos ao tratamento nominal. Portanto, tanto menos usado será o *você*, quanto mais ameaçador se revelar para a face negativa do alocutário.

Finalmente: o centro da reflexão deixa então de ser o questionamento sobre se ocorre ou não o tratamento pronominal por *você*; ao invés, uma reflexão em torno desta FT no PA deverá dar origem a estudos que visam, entre outros, identificar, descrever e/ou explicitar os distintos valores semânticos, bem como os fatores pragmático-discursivos subjacentes ao seu uso ou evitamento.

REFERÊNCIAS

- Balsalobre, S. R. G. (2015). *Brasil, Moçambique e Angola: Desvendando Relações Sociolinguísticas pelo Prisma das Formas de Tratamento* [Tese de Doutoramento, Universidade Estadual Paulista]. Repositório Institucional UNESP. <http://hdl.handle.net/11449/127872>
- Balsalobre, S. R. G. (2017). Minhas Mães, Meus Pais, Minhas Tias e Meus Tios: A Teoria do Parentesco como Contribuição ao Estudo das Formas de Tratamento Nominais de Angola e de Moçambique. *Estudos Linguísticos*, 46(1), 252-268. <https://doi.org/10.21165/el.v46i1.1542>
- Bazenga, A. (2017). Formas de Tratamento. https://www.researchgate.net/publication/323366226_Formas_de_tratamento
- Cavalheiro, V. M. (2016). As Diferentes Regras de Uso das Formas Tu e Você e Suas Influências na Compreensão de Narrativas Literárias: PB e PE [Tese de Doutoramento]. Universidade do Porto. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiQmujeodb4AhVUhP0HHUsOBelQFnoECAoQAQ&url=https%3A%2F%2Fsigarra.up.pt%2Ffep%2Fen%2Fpub_geral.show_file%3Fpi_doc_id%3D121849&usq=AOvVaw0KH7cQmg2ntOZ9Psyn3Zh
- Duarte, I. M. (2010). Formas de Tratamento: Item Gramatical no Ensino do Português Língua Materna. In A. M. Brito (Org.), *Gramática: História, Teorias, Aplicações* (pp.

- 133-146). Faculdade Letras da Universidade do Porto. https://www.researchgate.net/publication/327424379_Formas_de_tratamento_item_gramatical_no_ensino_do_Portugues_Lingua_Materna_Formas_de_tratamento_item_gramatical_no_ensino_do_Portugues_Lingua_Materna
- Gouveia, C. A. M. (2008). As Dimensões da Mudança no Uso das Formas de Tratamento em Português Europeu. In F. Oliveira, & I. M. Duarte (Eds.), *O Fascínio da Linguagem – Atas do Colóquio de Homenagem a Fernanda Irene Fonseca* (pp. 91-100). Centro de Linguística da Universidade do Porto. <https://ler.letras.up.pt/uploads/ficheiros/6695.pdf>
- Hammermüller, G. (1993). *Die Anrede im Portugiesischen: eine soziolinguistische Untersuchung zu Anredekonventionen und Anredeformen des gegenwärtigen europäischen Portugiesisch*. Nov neuer Verlag.
- Lešková, J. (2012). As Formas de Tratamento em Português Europeu [Dissertação de Mestrado]. Universidade Palacký de Olomouc. <https://fliphtml5.com/pqpej/gdog/basic/51-73>
- Marques, M. A. (2008). Quando A Cortesia é Agressiva. Expressão de Cortesia e Imagem do Outro. In F. Oliveira, & I. M. Duarte (Eds.), *O Fascínio da Linguagem – Atas do Colóquio de Homenagem a Fernanda Irene Fonseca* (pp. 277-297). Centro de Linguística da Universidade do Porto. <https://ler.letras.up.pt/uploads/ficheiros/6710.pdf>
- Marques, M. A., & Duarte, I. M. (2019). Formas de Tratamento e Preservação da Face em Interações Verbais Online. *Revista da Associação Portuguesa de Linguística*, 9(5), 236-249. <https://ojs.apl.pt/index.php/RAPL/article/view/15>
- Matoso, A. (2003). *Dicionário da Gramática da Língua Portuguesa*. Quarteto.
- Nascimento, M. F. B. do (2020). Formas de Tratamento. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, A. Mendes, & A. Andrade (Orgs.), *Gramática do Português* (Vol. 3, pp. 2701-2730). Fundação Calouste Gulbenkian.
- Nascimento, M. F. B., Duarte, E., & Mendes, A. (2018). Sobre Formas de Tratamento no Português Europeu e Brasileiro. *Diadorim*, 20 (Especial), 245-262. <https://revistas.ufrj.br/index.php/diadorim/article/view/23276>
- Naege, J. M. (2021). As Formas de Tratamento no Português de Angola: Contributo Semântico-Pragmático. In A. A. Timbane, D. P. Sassuco, & M. Undolo (Orgs.), *O Português de/em Angola: Peculiaridades Linguísticas e a Diversidade no Ensino* (pp. 124-141). Opção Editora.
- Silva, A. K. da (2020). Formas de Tratamento no Português de Angola. Estudo Sociolinguístico [Tese de Doutoramento, Universidade de Évora]. Repositório da Universidade de Évora. <http://hdl.handle.net/10174/28355>

Teixeira, E. P. (2008). O Pronome Você no Português de Luanda. In M. C. Lima-Hernandes, M. J. Marçalo, G. Micheletti, & V. L. Rossi (Orgs.), *A língua portuguesa no mundo* (pp. 1-16). FFLCH-USP. https://dlcv.fflch.usp.br/sites/dlcv.fflch.usp.br/files/03_9.pdf

Particularidades semânticas das expressões temporais com *até* e desafios de tradução entre português e inglês¹

Telmo Mória

tmoia@letras.ulisboa.pt

Universidade de Lisboa, Faculdade de Letras,
Centro de Linguística da Universidade de Lisboa (Portugal)

ABSTRACT.

This paper addresses the semantic analysis of Portuguese temporal adjuncts headed by the polyvalent connective *até* and its English – also polyvalent – counterparts (chiefly, *until* and *by*), and discusses problems in translation between English and Portuguese, documented in the website *linguee.com*. Curious data obtained in machine translation services (Google Translate [GT], Systran Translate [ST] and DeepL Translator [DT]) is also briefly discussed. The semantic values under discussion include two modes of temporal location (durative and inclusive, the latter often marked by Portuguese *até*-phrases, but not by their prototypical counterparts, *until*-phrases), and a combination of temporal location with (explicit) quantification over eventualities. A seldom discussed discourse connective, *até que*, not exclusively temporal in meaning, is also explored, to a certain extent. The long-debated issue of the English construction *not...until* is revisited: a comparison with Portuguese data shows that this form of *until* does not have *até* as its perfect counterpart. The semantic analysis is made with the logic of the Discourse Representation Theory, elaborating on my previous work (Mória 1995, 2000, 2015, 2016).

KEYWORDS.

Até; until; translation; temporal adjuncts; temporal location; quantification.

RESUMO.

Este artigo apresenta uma análise semântica dos adjuntos temporais com o conector polivalente *até* e as suas contrapartidas – também polivalentes – inglesas (principalmente, *until* e *by*) e discute problemas de tradução entre o português e o inglês documentados no *website linguee.com*. Alguns dados interessantes obtidos em plataformas de tradução automática (Google Translate [GT], Systran Translate [ST] e DeepL Translator [DT]) são também brevemente discutidos. Os valores semânticos que estarão em análise incluem dois modos de localização temporal, durativa e inclusiva (este último frequentemente sinalizado com *até*, mas não com a sua contrapartida inglesa prototípica, *until*) e uma combinação de valores de localização temporal e quantificação (explícita) sobre situações. Um conector discursivo pouco discutido na literatura – *até que* –, que não tem um valor exclusivamente temporal, também é analisado

¹ Este trabalho foi financiado com verbas do projeto estratégico do Centro de Linguística da Universidade de Lisboa UIDB/00214/2020.

com algum pormenor. O tópico, muito discutido na literatura, da construção inglesa *not...until* é revisitado: a comparação entre dados ingleses e portugueses mostra que esta forma de *until* não tem *até* como contrapartida exata. As análises semânticas do artigo são feitas com a lógica de Teoria da Representação do Discurso (Discourse Representation Theory), desenvolvendo aspetos de trabalho meu anterior (Móia 1995, 2000, 2015, 2016).

PALAVRAS-CHAVE.

Até; until; tradução; adjuntos temporais; localização temporal; quantificação.

0. Introdução

Neste trabalho, são analisadas – numa perspetiva predominantemente semântica e com foco em questões de gramática comparada português-inglês – as expressões temporais encabeçadas pelo conector *até*, que corresponde prototipicamente a *until* no inglês. A análise tem especial relevância para a tradução, na medida em que, a par das situações em que se observa a correspondência básica *até-until* – como (1) –, há situações (potencialmente propiciadoras de problemas de tradução) em que não existe essa correspondência – como (2) e (3).

(1) O bilhete é válido **até** março.

The ticket is valid **until** March.

(2) Eles têm de sair **até às** 10 horas.

They have to leave **by** 10 am.

(3) She didn't wake up **until** 9.

Ela {**só** acordou / **não** acordou **senão**} às 9 horas.

#Ela **não** acordou **até às** 9 horas. [TRADUÇÃO INFIEL – cf. secção 4 adiante]

Para a análise sintático-semântica dos dados do português, utilizo essencialmente trabalhos meus anteriores (Móia 1995, 2000, 2015, 2016), com recurso aos instrumentos formais da Discourse Representation Theory (cf. Kamp & Reyle 1993), em combinação com gramáticas de referência e com (vasta) literatura especializada sobre o inglês (cf. e.o. Karttunen 1974, Mittwoch 1977, Declerck 1995, de Swart 1996, Giannakidou 2002). São

usados dados de *corpora* eletrónicos de ambas as línguas – nomeadamente, o CETEMPúblico, para o português, e o British National Corpora (doravante, BNC), para o inglês – e ainda dados ilustrativos de potenciais problemas de tradução extraídos de plataformas eletrónicas de tradução (em particular o *website* *linguee.com*) ou de tradução automática (Google Translate, Systran Translate, DeepL Translator, doravante GT, ST e DT, respetivamente).

Os desafios da análise semântica de *até* temporal e as consequentes dificuldades da sua tradução relacionam-se centralmente com o facto de este conector ser muito polivalente e expressar diferentes valores (essencialmente) do domínio da localização temporal. A polivalência é também uma questão central para as contrapartidas inglesas deste conector, em especial *until* e *by*. Seguidamente, será feita uma apresentação e análise gramatical de cada um dos valores relevantes a considerar – num total de cinco –, discutindo as suas contrapartidas em inglês e destacando, nos casos relevantes (especialmente nas secções 3 e 4), os problemas de tradução detetados.

1. *Até* e *until* como conectores durativos

Começemos pelo caso porventura mais prototípico: o *até* e o *until* durativos. Estes conectores podem expressar um valor de **localização temporal durativa**, em que se asseve que a situação descrita na estrutura matriz (ev), tipicamente atélica, se estende ao longo de todo o intervalo de localização (t), identificado pelo adjunto encabeçado por *até/until*² – i.e. em que se asseve a condição $[t \subseteq ev]$. Geralmente, há a sugestão – por implicatura cancelável – de que a situação descrita na matriz termina no (final do) intervalo referido³, mas trata-se de uma questão predominantemente

2 Assumirei, seguindo e.g. Móia (2000: 89-92), que os adjuntos temporais encabeçados pelo conector *até* (e, *mutatis mutandis*, *until*) identificam um intervalo não pontual (t) cuja fronteira final é explicitamente marcada pelo complemento do conector (COMPL) – formalmente, $[COMPL(t) \wedge [end(t) \subseteq t]]$ – e cuja fronteira inicial (beg (t)) é normalmente definida no contexto discursivo ou inferida por conhecimento do mundo (embora possa ser deixada indefinida). É este intervalo alargado t que funciona com intervalo de localização das frases com *até/until*.

3 Uma forma de sinalizar explicitamente que a situação atélica não é necessariamente interrompida no fim do intervalo é aplicar o operador *pelo menos/at least* ao conector *até/until*. Esta possibilidade parece estar disponível apenas para este valor durativo, não para os valores discutidos nas secções 2 a 5. Exemplos: (i) «(...) Fernando Nogueira continuará titular da pasta da Defesa **pelo menos até** final de Maio, quando se realizar em Lisboa a cimeira da organização europeia de defesa.» (CETEMPúblico, ext5450-pol-95a-1); (ii) «The project was initially for three years, and will therefore run **at least until** 1993.» (BNC).

pragmática a que aqui não darei destaque. Tanto o conector português *até* como o conector inglês *until* sinalizam frequentemente este valor durativo, funcionando como contrapartidas equivalentes (sem gerar – tanto quanto pude observar – especiais problemas de tradução). Veja-se o seguinte exemplo do BNC, seguido de tradução minha:

(4) «Oxford City centre is expected to *be sealed off until tomorrow morning.*»

Prevê-se que o centro de Oxford *permaneça vedado até amanhã de manhã.*

Neste exemplo, o complemento do conector é nominal (*tomorrow morning*), mas também é muito frequente *até* e *until* receberem complementos oracionais (infinitivos ou finitos, no caso do português), sendo por vezes designados como conjunções, ou parte de locuções conjuncionais. Eis dois exemplos do BNC, mais uma vez seguidos de tradução minha; no segundo, é de destacar a sobreposição de um valor resultativo ao valor temporal, num uso destacado na literatura com a designação de “result *until*” (cf. e.g. Quirk *et al.* 1985: 1072, 1081, 1083; Huddleston & Pullum 2002: 734):

(5) «How Gemma Gibson came to be left alone in a flat in (...) London seems likely to *remain unclear until she comes back.*»

Como a Gemma Gibson foi deixada sozinha num apartamento em Londres parece que *irá permanecer um mistério {até ela regressar / até que ela regressar}*.

(6) «Add the kilo of sugar and *stir water, sugar and wort until they are (...) dissolved.*»

Adicione o quilo de açúcar e *mexa {até o açúcar e o mosto estarem dissolvidos / até que o açúcar e o mosto estejam dissolvidos}*.

Uma variante deste operador ocorre com descrições de eventos pontuais – como *emprestar um livro (a alguém)*, *interromper funções* ou *deixar o atletismo* –, mas em que a localização opera sobre o estado consequente desse evento pontual e está, portanto, associada a um mecanismo de comutação aspetual – cf. designação **localização temporal durativa**

derivada em Móia (2000: 276-281).

- (7) He *lent* me the book **until Monday**. (Mittwoch 1980: 220, *apud* Móia 2000: 277)
Ele *emprestou-me* o livro **até segunda-feira**.
- (8) «O governador civil de Aveiro (...) vai *interromper* funções **até ao próximo dia 12 de Junho** para se dedicar à preparação da campanha eleitoral para o Parlamento Europeu.» (CETEMPúblico, ext212300-pol-94a-1)
- (9) «Bronze medallist Steve Backley has *quit* athletics **until next year**.» (BNC)

Uma outra variante ainda ocorre em argumentos (não em adjuntos opcionais, como em (4) a (9)) dependentes de certos predicados (e.g. predicados durativos como *durar* ou *prolongar-se* ou predicados não durativos com argumentos temporais, como *adiar*) – cf. Móia (2015: 54-56; 2016: 12-13).

- (10) «A vicious colonial war then broke out which was to *last* **until 1954** (...).» (BNC)
Deflagrou então uma feroz guerra colonial que iria *durar/prolongar-se* **até 1954**.
- (11) «O Governo espanhol decidiu *adiar* **até ao ano 2020** a entrada em serviço do primeiro cemitério nuclear para resíduos de alta radioactividade (...).» (CETEMPúblico, ext107054-clt-soc-94a-1)
- (12) «(...), she decided to *adjourn* the case **until June 28**.» (BNC)
Ela decidiu *adiar* o caso **até 28 de junho**.

2. Até e *until* (e *by*) como conectores de delimitação temporal da quantificação

Existe ainda a possibilidade de usar *até* como uma expressão de delimitação temporal da quantificação (cf. e.g. Móia 2000, 2004, 2006), i.e. identificando um intervalo que funciona como um limite temporal para a quantificação explícita sobre situações. Nestas estruturas, a frase matriz representa o somatório (EV) de todas as situações de tipo descrito (ev) que ocorrem dentro do intervalo definido pelo adjunto com *até* (t), i.e. sinaliza a condição formal [EV = Σev : [... [ev \subseteq t]...]]. Existe, neste caso, um requisito (pragmático) de que a fronteira inicial do intervalo seja identificável (o que não acontece necessariamente noutros casos), de modo a que a quantificação tenha um intervalo bem delimitado⁴. Seguem-se dois exemplos do *corpus* CETEMPúblico:

- (13) a. «[Virgínia Moura] Foi (...) presa pela primeira vez em 1949, tendo voltado à cadeia por 16 vezes **até à revolução de 25 de Abril de 1974.**» (ext1243090-soc-98b-1)
- b. «Sete pessoas foram contaminadas **até agora** pelo vírus H5 N1, que transpôs a barreira das espécies.» (ext4164-soc-97b-2)
[fronteira inicial deduzida: início da epidemia, primeiro caso de contaminação]

Esta possibilidade parece estar também disponível para o conector inglês *until*, embora não seja muito discutida na literatura. É referida brevemente em Móia (2000: 300-301), citando Declerck (1995: 80), que menciona a possibilidade de combinar *until* com descrições de situações télicas, desde que explicitamente quantificadas:

the existential interpretation [leia-se “valor não durativo”] is even more strongly marked in the case of *until* than in the case of *since*.

4 Note-se, em linha com esta observação, que enquanto o conector *até* durativo aceita bem operadores associados a imprecisão de localização, do tipo de *pelo menos* ou *no mínimo* – e.g. *o Pedro esteve a trabalhar pelo menos até às 17h, a norma estará suspensa pelo menos até dia 12, a suspensão durará no mínimo até 2025* (cf. nota 2) – o *até* de delimitação temporal da quantificação rejeita normalmente esta combinação – e.g. *o Pedro saiu de manhã bem cedo e fez seis paragens no caminho (*pelo menos) até ao meio-dia.*

Still it remains possible to get it, as in [*until now John has done the washing-up **only three times***] (...), even when we replace *now* by another indication of time: [*until yesterday John had done the washing-up **three times***] (...). It is clear, then, that *until* can indicate a timespan in which a situation is said to actualize a specific number of times, or once, or not at all

(Declerck 1995: 80, *apud* Móia 2000: 300-301, sublinhados meus)

Vejam-se os seguintes dois exemplos do BNC (seguidos de traduções minhas), que documentam a existência deste *until* de delimitação temporal da quantificação, claramente não durativo:

- (14) «But seconds later she stopped breathing again – and John (...) had to give her mouth-to-mouth resuscitation *four times* **until paramedics arrived.**»

Mas, alguns segundos depois, ela deixou novamente de respirar e o John teve de lhe fazer respiração boca-a-boca *quatro vezes* **até chegar o INEM.**

- (15) «**Until last November** the council met *90 per cent* of the 4,000 required by the band to rent rehearsal rooms in local schools (...).»⁵

Até novembro (passado), a câmara conseguiu *90%* das 4000 libras solicitadas pela banda para arrendar salas de ensaio em escolas do concelho (...).

Estes contextos mostram claramente que o conector inglês *until* (mesmo excluídos os casos de interação com a negação, que consideraremos adiante na secção 4) não é necessariamente sempre durativo, ao contrário do que por vezes se assume, de forma algo simplista, na literatura⁶.

5 Dadas as propriedades de *until*, esta frase inglesa não pode ter uma leitura inclusiva simples, em que a situação – percecionada como atômica – de conseguir o valor em causa (possivelmente de uma só vez) se situa algures dentro do intervalo de localização; só admite a leitura em que o somatório de subsituações de conseguir x dinheiro, possivelmente distribuídas ao longo do intervalo de localização, resulta num situação não atômica descrita pela sequência na estrutura matriz (cf. Móia 2000: 325-326). Ambas as leituras – inclusiva simples e de delimitação temporal da quantificação – estão disponíveis na frase portuguesa com *até*, embora a segunda pareça preferencial.

6 Cf. e.g. Giannakidou (2002) – “*until* modifies states or processes (...), i.e. durative eventualities. Achievements and accomplishments, on the other hand, are incompatible with *until*”; Quirk et al. (1985: 1083): «*Till* and *until* can

Muito curiosamente, e de forma que, tanto quanto sei, ainda não foi sublinhada na literatura, este conector *until* parece competir (especialmente em certos contextos prospetivos) com *by* (que é igualmente contrapartida de *até* noutros contextos), como alternativa muito próxima ou mesmo equivalente⁷. Vejam-se os dois exemplos do BNC a seguir (com os predicados tólicos *lose* e *open*) e as respetivas traduções, minhas:

- (16) «Luxury car-makers Mercedes-Benz AG is to lose more than 10,500 workers **by the end of the year** (...).»
A empresa fabricante de carros de luxo Mercedes-Benz AG deve dispensar mais de 10 500 trabalhadores **até ao final do ano** (...).
- (17) «Of the *nine* terminals that will be opened **by the summer of next year**, the majority are already owned by British Rail (...).»
A maioria dos *nove* terminais que irão abrir **até ao verão do próximo ano** já são propriedade da British Rail (...).

A par da correspondência típica *até-until* que vimos até aqui, há pelo menos duas situações de discrepância nas correspondências entre operadores temporais afins, em português e inglês, ambas discutidas na literatura, que serão tratadas nas duas secções seguintes⁸.

3. Até e *by* (mas não *until*) como conectores inclusivos simples

Começemos com o *até* inclusivo, que tem como contrapartida o conector *by* do inglês. Nas frases com **localização temporal inclusiva (simples)**⁹, afirma-se que a situação (tipicamente télica) descrita na estrutura matriz (ev) se situa algures dentro do intervalo definido pelo adjunto (t), não importa

only cooccur with durative verbs (...).»

⁷ Creio que uma possível particularidade distintiva do uso de *by* é que o estado resultante do conjunto de situações se sobreponha ao intervalo denotado pelo complemento da preposição: no caso de (16), por exemplo, isso significa que a Mercedes não só dispensou 10 500 trabalhadores ao longo do período em causa, mas também que essa diferença (de 10 500) se mantém no intervalo representado pelo complemento de *by* (ou seja, no final do ano em causa).

⁸ Giannakidou (2002) fala em “UNTIL-by” para o valor que será tratado na secção 3 e em “NP-UNTIL” para o valor que será tratado na secção 4 (a par de “durative/stative UNTIL”, para o valor discutido na secção 1).

⁹ Utilizo aqui o atributo “simples” para distinguir este valor inclusivo do valor que é referido na secção 2, em que se pode considerar também a presença de uma relação inclusiva, mas associada a abstração sobre subeventos.

onde, i.e. sinaliza-se a condição [ev \subseteq t]. Considere-se a equivalência entre as seguintes frases inglesas e portuguesas:

- (18) «I have to *leave the country* **by Monday**, but that's no problem.»
(BNC)
Tenho de *sair do país* **até segunda-feira**, mas não há problema.
- (19) «(...) the head of Cedars was making his entry list for September already, and would have to *submit it* **by the end of the week** (...)» (BNC)
O diretor da Cedars já estava a preparar a lista de inscritos para setembro e teria de *a entregar* **até ao final dessa semana**.

A sinalização deste tipo de localização temporal é relativamente comum com *até* (embora apresente algumas restrições gramaticais, sendo e.g. predominantemente usada em contextos de localização prospetiva – Móia 1995: 346-349, 2000). Está, porém, completamente vedada ao operador inglês *until*, sendo antes *by* o marcador típico do valor em causa. Curiosamente, há várias outras línguas, além do português, que marcam esta leitura com homónimos do conector durativo (cf. Giannakidou 2002, sobre o uso polivalente – durativo e inclusivo – dos conectores alemão *bis* e grego *mexri*). O conector *até* tem nestes contextos um valor bastante próximo de *antes* (cf. e.g. equivalência entre *tenho de sair do país até segunda-feira* e *tenho de sair do país antes de segunda-feira*).

Seguem-se vários exemplos do *corpus* CETEMPúblico com *até* inclusivo simples, que atestam a vitalidade da construção. Observe-se que a situação descrita na matriz é sempre tética (e, tipicamente, pontual). Para todos eles, seria inadequada uma tradução com *until*; teria sempre de se usar *by* (ou formas afins, como *before* ou *no(t) later than*).

- (20) a. «O Tribunal Constitucional deverá *pronunciar-se* **até ao fim da semana** sobre a proposta de referendo (...)» (ext539722-pol-95a-1)
b. «Resta agora a Carlos Queirós *encontrar* **até domingo** uma solução para o lado esquerdo do ataque português (...)» (ext196149-des-93a-1)
c. «Sublinhe-se que o problema das descargas poluentes terá de ser resolvido **até 1995**.» (ext73017-soc-92b-1)

d. «(...) Joaquim Correia (...) chegou (...) a ameaçar desvincular-se do compromisso com a autarquia, caso o processo não fosse *desbloqueado até ao final do ano.*» (ext305-soc-95b-1)

Note-se que a leitura inceptiva de estados, na medida em que envolve o ponto inicial de uma situação, cabe dentro desta mesma interpretação (inclusiva simples), com comutação aspetual. Esta leitura está ilustrada nos dois exemplos a seguir:

- (21) a. «Foi antecipado o prazo para conclusão de reapreciação dos exames do 12º ano (...). A decisão sobre os recursos (...) terá que *estar tomada até à próxima quarta-feira* (...).» (ext12143-nd-96b-1)
b. «O imbróglcio das dívidas da Associação (...) deverá *estar resolvido até ao fim do mês.*» (ext226736-soc-96a-1)

3.1. Problemas de tradução envolvendo até inclusivo simples

Consideremos agora a tradução de até inclusivo em *sites* como o *linguee.com*, para ver em que medida ela é fonte de problemas. Nesta plataforma, encontramos exemplos de traduções adequadas com *by* (ou com as formas afins, *before* e *no(t) later than*), como as seguintes:

- (22) «Espero convictamente que, **até ao fim da semana**, o comandante Blair *tenha evitado o icebergue e conduzido a Presidência britânica para um porto seguro* (...).»
«I strongly hope that, **by the end of the week**, Captain Blair *will have avoided the iceberg and steered the British Presidency to a safe harbour* (...).»
- (23) «O Conselho decide que a Conferência Ministerial de Doadores (...) só se realizará se os prisioneiros *tiverem sido libertos [sic] até ao fim da semana em curso.*»
«The Council agrees that the Ministerial Conference of Donors (...) will take place only if the prisoners *are released before the end of*

the current week.»

- (24) «A Concessionária (...) deverá *notificar* (...) o departamento competente (...), **até ao final do mês de Fevereiro do ano seguinte ao ano em questão**, da sua intenção de proceder a essa amortização diferida (...).»
«The Concessionaire (...) shall *notify* the competent entity (...) **not later than the end of February after the end of the year in question** of its intent to take such deferred depreciation (...).»

Mas encontramos também exemplos de traduções problemáticas, com *until/till*, ou mesmo – de forma completamente desajustada – com *in*, como se observa a seguir:

- (25) «Este retiro vai permitir ainda mais acentuada descida das taxas de banco para *ser revisto* [sic] **até ao Verão** (...).»
«This retreat will allow even more pronounced lowering of bank rates to *be reviewed* **until the summer** (...).»
- (26) «A camisola será enviada por correio **até ao final desta semana** para a morada indicada pelo vencedor!»
«The shirt *will be sent by post* **till the end of the week** to the adress [sic] given by the winner!»
- (27) «O Ministério do Meio Ambiente vai lançar este mês (...) uma nova lei para reduzir as queimadas usadas na colheita, buscando *acabar com elas* **até 2020**.»
«Brazil's Environment Ministry this month will launch (...) a new law to reduce the burnoffs applied before plantings, with the goal of *ending them* **in 2020**.»

Os sistemas de tradução automática também apresentam, por vezes, problemas semelhantes na tradução de *até* inclusivo. Veja-se o exemplo a seguir, bem traduzido pelo Systran Translate e pelo DeepL Translator, mas mal traduzido pelo Google Translate.

- (28) O jogador espera conseguir recuperar **até** ao próximo domingo.
 [INCLUSIVO: *recover by*]
 #The player hopes to recover **until** next Sunday. [GT]
 The player hopes to recover **by** next Sunday. [ST]
 The player hopes to be able to recover **by** next Sunday. [DT]

Marginalmente, observa-se uma incapacidade geral destes sistemas para dar conta de ambiguidades, interessando aqui em particular a ambiguidade durativa/inclusiva de formas com *até*, que é relativamente frequente em português (cf. e.g. Móia 1995). Assim, a frase (29) é ambígua entre treinar continuamente até ao final do ano (*train until*) e começar a treinar antes do final do ano (*train by*), mas os três tradutores automáticos considerados só captam a segunda leitura. O mesmo, *mutatis mutandis*, para a frase (30).

- (29) O jogador espera conseguir treinar **até** ao final do ano.
 [AMBÍGUO: *train until / train by*]
 The player hopes to be able to train **by** the end of the year. [GT = ST = DT]
- (30) Os computadores estarão aqui **até** ao final da semana.
 [AMBÍGUO: (*will*) *be here until / be here by*]
 The computers will be here **by** the end of the week. [GT = ST = DT]

É de destacar ainda a incapacidade destes sistemas para distinguir a leitura relevante com pistas contextuais que excluem uma das leituras:

- (31) Os computadores já foram levados para a sala e **estarão lá até ao final da semana**.
 [SÓ DURATIVO (DADA A FRASE INICIAL): (*will*) *be there until / *be there by*]
 #The computers have already been taken/brought to the room and *will be there by the end of the week*. [GT = ST = DT]

Para terminar esta secção, importa ainda sublinhar que nem sempre os adjuntos temporais ingleses com *by* podem ser traduzidos com *até*, o que também pode gerar falhas de tradução. Tal sucede frequentemente em contextos de localização não prospetiva (e.g. de sobreposição ou

retrospectiva), em que é mais adequada a tradução com *em*, como ilustrado a seguir (com exemplos do BNC e traduções minhas).

- (32) «**By the end of the decade** [the United States] (...) *was importing* 8-9 million barrels per day (...).»
No final da década, [os Estados Unidos] (já) *estavam a importar* 8 a 9 milhões de barris por dia (...) [??até ao final da década]
- (33) «**By the end of the war** the factory *had produced* more shells than any other of its size.»
No final da guerra, a fábrica (já) *tinha produzido* mais bombas que qualquer outra fábrica da sua dimensão. [??até ao final da guerra]
- (34) «I don't mind shopping on Saturday when I've got the money, but I don't like it during the week – I'm usually *broke by Monday* (...).»
(...) eu geralmente já estou *sem dinheiro na segunda-feira* [*até segunda-feira]

Ao tentar fazer a tradução da sequência relevante do último excerto em plataformas de tradução automática, verificamos que estas construções também nem sempre são adequadamente convertidas. Observe-se, ignorando outros possíveis problemas, a tradução de *by* por *até* no Systran Translate:

- (35) I'm usually broke **by Monday**(.)
Normalmente estou quebrado **na segunda-feira**. [GT]
Normalmente estou falido {à **segunda-feira** / **na segunda-feira**}. [DT]
#Eu {normalmente estou duro **até segunda** / geralmente estou duro **até segunda-feira**}. [ST]¹⁰

¹⁰ Curiosamente (e sem uma lógica compreensível), as plataformas DT e ST dão uma tradução diferente consoante a frase original, *input* da tradução, seja seguida de ponto final (*I'm usually broke by Monday.*) ou não (*I'm usually broke by Monday*). As duas traduções são dadas aqui dentro de chavetas em primeiro e segundo lugar, respetivamente.

4. A expressão (ou construção) inglesa *not...until*

Analisemos agora um caso especialmente interessante e muito discutido na literatura: o da expressão (ou construção, dependendo da análise) inglesa *not... until*. Veja-se o contraste em inglês entre a frase negativa com a descrição de uma situação télica pontual em (36), que é plenamente gramatical, e a sua contrapartida sem negação em (37), que é agramatical. Note-se que a frase afirmativa só seria gramatical se, na matriz, em vez de uma situação télica se representasse uma situação atélica – e.g. *the princess **slept until midnight*** –, caso em que obteríamos uma leitura durativa comum, do tipo analisado na secção 1.

(36) The princess didn't arrive until midnight.

(37) *The princess arrived until midnight.

Estas construções – com operador de negação (NEG) e *until* interdependentes – estão entre as mais discutidas na literatura (há quase cinco décadas), sendo a sua análise por vezes referida como “the puzzle of English *until*” (cf. e.o. Karttunen 1974, Mittwoch 1977, Declerck 1995, de Swart 1996, Giannakidou 2002, Condoravdi 2008). Genericamente, a questão essencial é que o operador inglês *until* não se combina normalmente com descrições pontuais (ou télicas, em geral) para sinalizar leitura inclusiva (cf. (37)), mas tal combinação é plenamente legítima em contextos de negação, como (36). A discussão na literatura gira essencialmente em torno de duas (ou se quisermos três) hipóteses:

I. Há apenas um *until* (sempre durativo), sendo a negação (sob escopo de *until*) responsável por uma interpretação atélica (das descrições a que o adjunto se aplica), i.e. atuando com um *Aktionsart shifter*, um estativizador; hipótese defendida por e.g. Klima (1964) ou Mittwoch (1977), entre outros.

II. Há dois operadores *until*,

A. sendo o *until* associado à negação de situações télicas – como em (36) – um genuíno item de polaridade negativa (ocorrendo, portanto, sob escopo de negação), diferente do *until* durativo que ocorre em e.g. *the princess **slept until midnight***; hipótese defendida

por e.g. Karttunen (1974), Giannakidou (2002), Condoravdi (2008),
ou

B. sendo *not until* um operador complexo (descontínuo) resultado de um processo de lexicalização que funde um valor temporal com o valor de um operador de exclusão (ou seja, *not until* = *only at*); hipótese defendida por Declerck (1995) e que me parece a mais bem fundamentada (embora não pretenda aqui apreciar comparativamente as diferentes hipóteses de análise).

Uma característica que alegadamente distingue – e torna único – este uso de *until* (inclusivo) interdependente da negação é a presença de uma inferência lógica (*entailment*) – crucialmente, não uma implicatura cancelável – de realização (ou ocorrência) da situação descrita no intervalo referido no complemento de *until* (cf. Karttunen 1974, Declerck 1995, Giannakidou 2002).

- (38) The princess didn't fall sleep until midnight.
[Inferência lógica: the princess fell asleep at midnight.]
#At midnight she got up, got dressed and went out for a walk.

Assim, uma frase como (38) tem como inferência lógica, não cancelável, que a princesa adormeceu mesmo à meia-noite, e, portanto, é incoerente continuá-la com, por exemplo, *at midnight she got up, got dressed and went out for a walk*. Karttunen (1974) – como destaca e.g. Giannakidou (2002) – afirma explicitamente que em frases deste tipo: “*until X does not mark the endpoint of any interval. Instead it designates the time when the event in question took place*”. Declerck (1995: 53) refere esta propriedade como “sense of actualization”, e ela distingue (o que ele considera ser) o operador complexo – ou “unidade estereotipada” – *not until* de outras possíveis coocorrências de *until* e negação em enunciados em que não é formada uma unidade gramatical¹¹.

Há vários argumentos linguísticos invocados na literatura para

¹¹ Por exemplo, frases com *until* fora do escopo da negação, como *until nine, John did 't wake up ou he didn't run away, until the police came (and it was too late to escape)* (Declerck 1995: 62-63), ou frases em que ocorrem certas expressões adverbiais (e.g. *any more, any longer*) que bloqueiam a formação da unidade complexa *not until*, como *he didn't sleep any more until morning ou he did not love her any longer until his death* (*ibid.*: 70).

fundamentar a especificidade das construções inglesas com *not... until*. Sem desenvolver muito a questão, destaco aqui três (um dos quais já mencionado acima):

(i) a impossibilidade de cancelamento da inferência (que mostra que se trata de um *entailment*), propriedade já referida – cf. ainda o exemplo (39) a seguir, semelhante a (38) acima, e o exemplo (40):

(39) #John didn't wake up until nine. In fact, I heard later that he didn't wake up at all.

(Declerck 1995: 62)

(40) #Nancy didn't get married until she died.
(vs. ^{OK}Nancy remained a spinster until she died.)

(Declerck 1995: 69)

A forte estranheza da frase (40) resulta de o uso de *not... until* sugerir, de forma que gera incoerência, que a Nancy se casou quando morreu.

Note-se ainda, marginalmente, que a inferência em causa distingue *not until* de *not before*. Com *not before* (cf. e.g. *the princess didn't fall asleep before midnight*) há, quando muito, uma implicatura, cancelável, de ocorrência do evento. Assim, a frase com *before* – ao contrário da frase com *until* – é compatível com a princesa não ter adormecido à meia-noite ou nem sequer ter adormecido (ou seja, a frase pode ter a mesma continuação da da frase (41) abaixo, dada entre parênteses retos).

(ii) a impossibilidade de ocorrência do sintagma com *until* fora do escopo da negação (mantendo-se a leitura relevante), mostrando que a interpretação em causa só é obtida na sequência *not... until*; a frase (41) – com o adjunto com *until* em posição inicial (fora do escopo da negação) – tem uma leitura distinta de (38), mais próxima da que seria obtida com o *até* português, que não tem o *entailment* de ocorrência referido. Veja-se a continuação possível dada a seguir à frase (41).

(41) Until midnight, the princess didn't fall asleep.

[LEITURA DISTINTA, DURATIVA COMUM]

[I don't know about later, for I left at midnight. In fact, I heard later that she did not fall asleep after midnight either.] (cf. Declerck 1995: 62)

(iii) a possibilidade de clivagem de *not until X* (estrutura em que a inferência de ocorrência da situação é porventura ainda mais evidente) e que pode ser visto como um fortíssimo argumento a favor da hipótese de lexicalização de Declerck (1995):

(42) It wasn't **until midnight** that the princess fell asleep.

Estas diversas propriedades mostram que *not until* (do tipo relevante) não tem como contrapartida exata a combinação portuguesa *não até* (que não apresenta o mesmo tipo de restrições e – crucialmente – de *entailment*). Com efeito, em português podemos dizer (43) e (44), não notamos diferenças do tipo observado em inglês com a mudança de ordem – cf. (45) – e não podemos clivar (*não*) *até X*, tendo de recorrer a (*não*) *senão X* para obter uma forma equivalente (e morfossintaticamente paralela) – cf. (46) vs. (46').

(43) A princesa não adormeceu até à meia-noite. (Por isso,) À meia-noite levantou-se, vestiu-se e foi dar uma volta.

(44) A Nancy não/nunca se casou até morrer.

(45) Até à meia-noite, a princesa não adormeceu. Depois dessa hora, não sei...

A princesa não adormeceu até à meia-noite. Depois dessa hora, não sei...

(46) *Não foi até à meia-noite que a princesa adormeceu.

(46') Não foi **senão** à meia-noite que a princesa adormeceu.

Podemos, de forma genérica, afirmar que as contrapartidas portuguesas mais aproximadas de *not until t* são *só EM t* ou *não senão EM t*¹², sendo que vários autores propõem equivalências semelhantes (noutras línguas), quando parafraseiam o valor de *not until* (cf. e.g. Declerck 1995 ou Giannakidou 2002). As traduções com *não até* – e bem assim *não antes* – são gramaticais, mas não garantem a inferência pretendida, pelo que não

12 A forma (em versaletes) *EM* representa um conector de localização temporal direta (cf. Móia 2000: 91), que pode ser materializado em português através de diferentes preposições – e.g. em *{só / não senão} em 1995*, a *{só / não senão} à meia-noite*, *durante* *{só / não senão} durante os primeiros anos do século XX* –, ou pode mesmo estar apenas subentendido (*{só / não senão} ontem*, *{só / não senão} a semana passada*).

são traduções totalmente fiéis¹³.

4.1. Problemas de tradução envolvendo *not...until*

Frequentemente, estas particularidades de *not... until* não são tidas em conta nas traduções para português. No *website* *linguee.com*, a par de traduções adequadas com *só EM (não senão EM e afins)* – cf. (47) –, encontramos traduções discutíveis com *até* e *antes* – cf. (48) e (49):

(47) «The decisive step in hydroponics (...) didn't happen **until the early 20th century** (...). »

«O passo decisivo na hidroponia (...) ocorreu **apenas no início do século 20** (...).»

(48) «This video was the least known (...) because it didn't leak onto internet **until late in 2007**. »

#«Esse vídeo foi o menos conhecido (...) porque **não** vazou na internet **até o final de 2007**.»

[vs. ... *porque só* “vazou” para a *Internet no final de 2007*.]

(49) «I suppose my journey to Marist High School began back in 1973, even though I didn't actually arrive **until 1987**. »

#«Eu penso que minha viagem até a Marist High School começou por volta de 1973, embora na realidade eu **não** tenha chegado ali **antes de 1987**.» [vs. ... *embora na realidade eu só* tenha chegado lá **em 1987**.]

As particularidades de *not... until* também parecem não ser consideradas em sistemas de tradução automática como Google Translate, Systran Translate ou DeepL Translator, onde a tradução da frase (38), por exemplo, é feita sempre (inadequadamente) com *não até*:

¹³ Eventualmente, em determinados contextos que favorecem fortemente a implicatura relevante, as expressões em causa podem ser equivalentes, para todos os efeitos práticos (cf. e.g., com ligação causal, *não me assustei até perceber que tinha perdido a chave / só me assustei quando percebi que tinha perdido a chave*), mas não vou desenvolver aqui a questão.

- (50) The princess didn't fall asleep **until midnight**.
#A princesa **não** adormeceu/dormiu **até meia-noite**. [GT = ST = DT]

Seguem-se exemplos de uso do *not... until t* no BNC (seguidos de traduções minhas), em que é clara a inferência de realização da situação em *t* e se vê bem a inadequação da tradução com *não até*.

- (51) «The fact that your uncle was missing wasn't reported to the police **until Sunday morning** – isn't that rather odd?»
O facto de o seu tio ter desaparecido {só foi reportado / não foi reportado **senão**} **no domingo de manhã** (...)
#O facto de o seu tio ter desaparecido **não** foi reportado **até domingo de manhã** (...)
[tradução infiel, porque não se infere necessariamente que houve denúncia do desaparecimento na manhã de domingo]
- (52) «We didn't know he'd won this award **until last month**. It was a lovely surprise.» {Só soubemos que ele tinha ganho este prémio / Não soubemos que ele tinha ganho este prémio **senão**} **o mês passado**.
#Não soubemos que ele tinha ganho este prémio **até ao mês passado**.

Note-se que, em contextos em que se mencionam situações passadas, a inferência (que Declerck refere como “sense of actualization”) pode ser formulada simplesmente como “a situação ocorreu em *t*”, como é usual na literatura. Mas dada a possibilidade de *not... until* ocorrer em contextos que mencionam situações futuras e/ou com marcação explícita de incerteza, como (53)-(55) abaixo, a formulação da inferência tem de ser modalizada: “a situação ocorreu, ocorrerá ou assume-se que ocorrerá em *t*”. O dicionário *Collins Cobuild English Dictionary for Advanced Learners* descreve o sentido de *until* em *not... until* com uma formulação abrangente que dá conta das diversas situações: “You use **until** with a negative to emphasize the moment in time after which the rest of your statement becomes true, or the condition which would make it true”.

- (53) «I'm off to Yorkshire tonight and I won't be back **until Sunday afternoon.**»
Vou para Yorkshire esta noite e {só regresso / não regresso **senão**}
no domingo à tarde.
#Vou para Yorkshire esta noite e **não** regresso **até** domingo à tarde.
- (54) «Chrysalids (...) will probably **not hatch until next spring** (...)»
Provavelmente, as crisálidas {só se metamorfosearão / não se metamorfosearão **senão**} **na próxima primavera.**
#Provavelmente, as crisálidas **não** se metamorfosearão **até** à **próxima primavera.**
- (55) «The full resonance of this production may **not be felt until next May.**»
Os efeitos desta produção {só se devem sentir em pleno / não se devem sentir em pleno **senão**} **em maio.**
#Os efeitos desta produção **não** se devem sentir em pleno **até maio.**

Termino esta secção referindo novamente a construção com *not until* t clivado¹⁴. É uma construção particularmente interessante porque nestes casos, as traduções com *até* (ou com *antes*) são normalmente muito inadequadas e por vezes até agramaticais¹⁵. Vejam-se os seguintes dois exemplos do BNC e as minhas traduções subseqüentes:

- (56) «Originally the working men's clubs were strictly teetotal and **it wasn't until 1860 that** beer was allowed on the premises (...). »
(...) e {só em 1860 é que / não foi **senão em 1860 que**} o consumo de cerveja passou a ser permitido nesses espaços.
?*/*(...) e **não foi até 1860 que** o consumo de cerveja passou a ser permitido nesses espaços.

14 Uma curiosidade: a estrutura ocorre com muito maior frequência em contextos retrospectivos, referindo situações que ocorreram de facto (com *simple past*: *it wasn't until X that*), como (56)-(58), do que em contextos prospectivos, referindo situações previstas (com *future tense*: *it won't be until X that*), como a que é dada a seguir; encontrei apenas 9 registos desta construção no BNC. Exemplo:

«The (...) group reckons **it won't be until the second half of next year that** Unix will achieve the degree of functionality needed to match proprietary systems (...)»

O grupo imagina que **só na segunda metade do próximo ano é que** a Unix conseguirá o grau de funcionalidade (...).

?*/O grupo imagina que **não será até à segunda metade do próximo ano que** a Unix conseguirá o grau de funcionalidade (...).

15 A procura de adjuntos com *até* clivados no *corpus* CETEMPúblico (pesquisa: “[N,n]ão” [lema = “ser”] “até” [] {1,5} “que”) não gerou qualquer resultado.

- (57) «**It wasn't until he produced the ring that** I took him seriously. »
Só quando ele mostrou o anel é que eu percebi que era a sério.
??/***Não foi até ele mostrar o anel que** eu percebi que era a sério.
Segue-se um exemplo semelhante, mas com uma tradução real (anómala, com até) do *linguee.com*:
- (58) «**It was not until the year 1546** at the Council of Trent **that** the Catholic Church decreed that the 14 books of the apocrypha should be included in the canon.»
??«**Não foi até 1546** no Concílio de Trento **que** a Igreja Católica decretou que os 14 livros da apócrifa deveriam ser incluídos no cânon.»
[vs. ^{OK}*Não foi senão em 1546 que...*]

5. Particularidades de até com complementos oracionais

Para terminar, importa considerar alguns aspetos particulares das construções em que até tem complementos oracionais, infinitivos ou finitos. Em muitos casos (porventura na generalidade deles), o estatuto oracional ou não do complemento de até não tem um impacto muito significativo na interpretação. Vejam-se os três exemplos abaixo, com até durativo, delimitativo da quantificação e inclusivo simples, respetivamente, e complementos nominais ou oracionais sem diferenças interpretativas relevantes:

- (59) A lei esteve em vigor até {2010 / o Parlamento a revogar, em 2010}.
[LOCALIZAÇÃO DURATIVA]
- (60) A guerra civil, que deflagrou em 1967, foi muito violenta. Morreu cerca de um milhão de pessoas até {1970 / o conflito ser dado como terminado}.
[DELIMITAÇÃO TEMPORAL DA QUANTIFICAÇÃO]
- (61) Quero acabar o relatório até {domingo / ir de férias, no final desta semana}.
[LOCALIZAÇÃO INCLUSIVA]

Uma situação semelhante parece verificar-se com o conector inglês *until*. Giannakidou (2002), por exemplo, afirma: “I will not discuss clausal *until*, but assume that what I say for phrasal *until* carries over to the clausal one with only minor adjustments”.

Porém, há casos em que os adjuntos oracionais com *até* (e possivelmente também com *until*) apresentam particularidades semânticas que os distinguem dos adjuntos não oracionais e merecem uma análise autónoma. Destacarei duas situações distintas, ilustradas em (62) e (63) abaixo e que discutirei em cada uma das duas subsecções seguintes, respetivamente.

- (62) Não me vou reformar tão cedo. Trabalharei **até** poder.
 (63) Ouviu-se um rugido na montanha, depois outro, ... **Até que**, de repente, o vulcão explodiu.

5.1. Até durativo equivalente a *enquanto*

Quando o conector *até* durativo se combina com complementos oracionais **infinitivos**, pode assumir uma de duas formas, sendo a primeira a menos marcada (ou *default*):

(i) **até**₁ – contrapartida de *until* (tipo analisado na secção 1): a frase complemento de *até* indica uma situação télica cuja ocorrência – cf. (64) –, ou uma situação atélica cujo início – cf. (65) –, marca o fim da situação (atélica) descrita na frase matriz

- (64) «Linda Tripp (...) foi funcionária da Casa Branca **até ser transferida para o Pentágono** (...).» (CETEMPúblico, ext380628-pol-98b-1)
 [cf. contrapartida inglesa: ... **until** she moved to the Pentagon]
 (65) «Quem não viu este fim-de-semana o Fio de Azeite terá de esperar **até poder fazê-lo**.» (CETEMPúblico, ext239503-soc-92b-1)
 [cf. contrapartida inglesa: ... **until** they can do it]

(ii) **até**₂ – contrapartida de (*for*) *as long as* ou *while*; menos comum, de uso bastante mais restringido gramaticalmente e tanto quanto sei não dicionarizado (cf. Móia 1995: 356-357, que discute brevemente este valor

particular de até): a frase complemento de até indica uma situação cujo **fim** é concomitante com o fim da situação (atélica) descrita na matriz; até₂ é, portanto, sinónimo de *enquanto* e até₂V é geralmente parafraseável por até₁ não V.

Seguem-se alguns exemplos do *corpus* CETEMPúblico que ilustram este valor de até:

- (66) a. «Felizmente, sou muito saudável. E trabalharei **até poder**.»
(ext557989-soc-98a-2) [cf. português: até₁ já não poder (mais), enquanto puder; inglês: (for) as long as I can / *until I can]
- b. «Auschwitz foi tornado lugar sagrado, ou lugar maldito, logo após o fim da guerra, pelas autoridades polacas, mas estas menosprezaram, **até poderem**, o facto de 90 por cento do milhão, ou milhão e meio, dos mortos do campo serem judeus.»
(ext551367-soc-95a-1)
- (67) a. «Huckleberry (...) não tinha de ir à escola nem à igreja, nem de receber ordens de ninguém; podia ir pescar ou nadar quando e onde lhe apetecesse, demorando-se **até querer**; (...).» (ext284714-nd-94b-1)
- b. «É uma refeição comunitária, resquício de hábitos pagãos (...); é, de facto, o assumir da abundância e até do esbanjamento, onde se come **até querer** e se bebe o que o corpo aguenta.» (ext939600-clt-95a-1)
- (68) «O avião, cuja partida para Singapura está prevista para hoje, continuará a viajar "**até ter vontade**, mas não se sabe até onde" (...).» (ext149142-soc-95b-2)

Como é referido em Móia (1995: 357), esta interpretação de até "é possível na presença de alguns subtipos de verbos que incluem pelo menos os que exprimem capacidade (*poder*) ou volição (*querer*, *apetecer*, *ter vontade*)", mas não se usa com naturalidade com a generalidade das expressões predicativas – cf. *?o Paulo escreveu até ter tinta na caneta*, *?o Paulo escreveu até estar bem-disposto*. Adicionalmente, observa-se que o complemento do verbo tem normalmente um carácter anafórico, sendo nulo ou realizado com recurso a um verbo anafórico como *fazer*

– cf. e.g. *trabalharei {até poder / até poder fazê-lo}*; numa frase sem essa dependência, como *trabalharei {até poder comprar uma casa / até poder ir para a reforma}*, o valor que se obtém é o de até₁.

Seguem-se agora alguns exemplos de (*for*) *as long as* ou *while* no BNC, que correspondem a este mesmo valor (não representável por *until*):

- (69) a. «I'll try to keep her here **as long as I can** if it will help.
 b. «Normally, Mickeen would detain you **as long as he could**, jabbering away about this, that and the other.»
 c. «He climbed a different shelf each day and hid **for as long as he could** behind piles of books, (...) »
 d. «Demonstrators (...) were urged by leading liberal Yeung Sum to cherish freedom of speech **while they could**, and to continue drawing attention to the denial of such freedom in China. »
- (70) a. «When a child has got the squeeze, he is allowed to hold on to it **as long as he wants** before passing it on to the next person.»
 b. «Eventually the carabinieri got bored and let Eric stay in the lavatory **for as long as he wanted**.»

Os sistemas de tradução automática Google Translate, Systran Translate e DeepL Translator não reconhecem este valor particular, gerando traduções incorretas:

- (71) Trabalharei **até poder**.
 #{I will / I'll} work **until I can**. [GT = ST = DT]

5.2. Até (*que*) como conector discursivo de Narração

Nesta secção, explorarei um valor de *até* (que surge apenas na forma *até que*), tanto quanto sei não analisado em pormenor na literatura (com exceção de uma breve observação em Lobo 2003: 91, a que voltarei adiante), em que discutivelmente já não estamos perante um conector temporal do plano (inter)frásico, mas perante um marcador discursivo (da relação discursiva) de Narração (no sentido de Lascarides & Asher 1993). Está ilustrado na frase

(63) acima, repetida a seguir:

(72) Ouvia-se um rugido na montanha, depois outro, ... **Até que**, de repente, o vulcão explodiu.

Importa começar por notar que, na combinação de *até* com complementos oracionais, podem surgir várias questões gramaticais específicas (além das questões meramente temporais), nomeadamente de modo/modalidade, de emergência de valores interproposicionais específicos (cf. referências ao “*result until*”, ilustrado acima em (6), em gramáticas inglesas) e de pressuposições ou inferências associadas ao uso dos adjuntos. Refiro aqui apenas uma das diferenças mais evidentes, que envolve o uso de modo indicativo associado a retrospectividade e factualidade/veridicidade – cf. (73a) –, em contraste com o uso do modo conjuntivo associado a prospetividade e eventualidade/não veridicidade – cf. (74a). Estes dois valores também podem estar presentes em orações infinitivas dependentes de *até*, mas sem uma distinção formal (morfológica) comparável – cf. (73b) e (74b). Observem-se os seguintes exemplos do *corpus* CETEMPúblico:

(73) a. «(...) andou de obra em obra à procura de trabalho, **até que soube** por mero acaso que tinha um padrinho a trabalhar nas minas da Panasqueira.» (ext833137-soc-93a-1)

[RETROSPETIVO, VERÍDICO / MODO INDICATIVO]

b. «Não dei importância à ameaça, **até saber** o que aconteceu com Solanas.» (ext1059482-clt-91a-2)

[IDEM COM INFINITIVO]

(74) a. «[O BPI] continua a ser alvo preferencial dos especuladores. E sê-lo-á pelo menos **até que se saiba** claramente o que vai fazer a Sonae para entrar na área financeira.» (ext1387743-eco-95a-2)

[PROSPETIVO, NÃO VERÍDICO / MODO CONJUNTIVO];

b. «Os Maxwell terão que esperar pelo menos seis meses **até saber** o veredicto de um júri composto por 12 pessoas que ontem prestaram juramento.» (ext469439-soc-95a-1)

[IDEM COM INFINITIVO]

A combinação de *até* com orações finitas com modo indicativo é especialmente interessante, porque parece envolver – pelo menos nalguns casos, como (72) – uma mudança interpretativa (possivelmente envolvendo um processo de gramaticalização), em resultado do qual *até que* deixa de atuar como um puro conector (frásico) de localização temporal adjunta (seja durativa, seja inclusiva) e passa a comportar-se como um **conector discursivo associado à relação retórica de Narração** (comparável com e.g. *e depois*, *e seguidamente*, ou outros como veremos melhor adiante). Esta análise está em linha com as observações de Lobo (2003: 91), que não considera as orações com este conector (e outros afins) como membros da classe das subordinadas adverbiais¹⁶.

Nos casos com predicções atélicas na matriz – como nos dois exemplos do CETEMPúblico em (75) abaixo –, a diferença entre localização temporal durativa (“a situação prolonga-se até um dado momento do tempo [em que, em princípio, é interrompida]”) – e marcação discursiva de Narração (“a situação x sucede-se à situação y”) é subtil e difícil de destringir com precisão.

- (75) a. «A família foi vivendo de ajudas e à “custa de muita fome” **até que** José (...) teve uma nova proposta de trabalho como pastor de vacas (...).» (ext14306-soc-96b-1)
- b. «O sistema funcionou bem **até que** um dia Londres mandou cobrar impostos que não tinham sido votados em Boston; os apertos reais levaram à separação pela guerra.» (ext69581-opi-96b-2)

16 Cf. Lobo (2003: 91): “Nas orações introduzidas por conectores temporais, existe um subconjunto de orações que se caracterizam por ocorrerem obrigatoriamente em posição final (...). Trata-se de orações introduzidas por *quando* (ou *eis senão quando*) ou por *até que* com o verbo no pretérito perfeito ou no presente (...). Este tipo de orações temporais descreve uma situação pontual que interrompe uma situação de carácter durativo (...). Pode dizer-se que têm uma função narrativa. (...) (158) a. O João brincou com o castelo de areia durante toda a manhã, até que subiu a maré e destruiu tudo... (...) Dadas as características muito particulares destas orações, nomeadamente a impossibilidade de serem antepostas, não as vou considerar na classe das subordinadas adverbiais.» Como podemos observar através dos exemplos desta secção, a descrição da autora “uma situação pontual que interrompe uma situação de carácter durativo” não cobre todas as possibilidades combinatórias de *até que*. As restrições de *Aktionsart* parecem menos categóricas. Na frase matriz pode, por exemplo, descrever-se uma sucessão de eventos, inclusive pontuais (cf. e.g. (79) adiante: *uma bala* [ev₁ – pontual], *depois outra* [ev₂ – pontual]) –, e na oração dependente de *até* pode igualmente descrever-se uma sucessão de eventos, incluindo não pontuais, ou situações atélicas (cf. e.g. as seguintes modificações de (79): ... *até que o homem caiu* [ev₁], *se contorceu em agonia durante quase um minuto* [ev₂ ∅ tónico não pontual] e *morreu* (ev₃); ... *até que o homem caiu* (ev₁) e, *quando olhámos, já estava morto* [ev₂ – atélico (não pontual)]).

Os casos mais interessantes – e claramente ilustrativos deste novo valor – são dados por estruturas matriz (i) com descrições claramente não atélicas (mostrando que a interpretação do adjunto não é durativa), ou (ii) com descrições atélicas mas sem o valor típico (de prolongamento até *t* e interrupção em *t*). Nas orações introduzidas por *até que* ocorre tipicamente o pretérito perfeito simples do indicativo, ou presente do indicativo narrativo, equivalente. Vejam-se exemplos do *corpus* CETEMPúblico (e atente-se na pontuação – com ponto final antes de *até que* – do último exemplo de cada um dos dois conjuntos a seguir¹⁷).

(i) *até que* F_{IND} com descrições de situações não atélicas

(76) «(...) estranho a resposta, porque pedi uma reunião ao SRE para discutir esta matéria, a qual nunca me foi concedida, **até que** decidi enviar-lhe os estatutos, sem esperar pela marcação dessa reunião.» (ext61101-soc-93a-2)

[interpretação: a decisão de envio sucede-se ao pedido de reunião e falta de resposta]

(77) «A partir daqui, o projecto de loteamento é analisado e sofre algumas alterações, **até que** a autarquia lhe concede o alvará número 22, a 10 de Julho de 1984.» (ext13623-soc-91b-2)

[interpretação: a concessão do alvará sucede-se à análise do projeto e à realização de alterações]

(78) «Foi adolescente no fascínio de James Dean (...), Elvis e o «rock». As primeiras canções que criou foram «rocks» em português, aos microfones da rádio, **até que** Daniel Filipe lhe deu a ouvir Brassens e Ferré.» (ext16738-clt-93a-1)

[interpretação: o contacto com Brassens e Ferré sucede-se à criação das primeiras canções]

(79) «(...) os atiradores furtivos alvejaram-no. Uma bala, depois outra, **até que** o homem se imobilizou por completo no pavimento e morreu.» (ext354647-pol-92a-1)

[interpretação: a imobilização e morte do homem sucedem-se aos

17 O ponto final sinaliza tipicamente uma construção do plano do discurso por oposição ao plano da frase complexa. É sintomático que esta marca de pontuação seja com bastante frequência escolhida para separar as expressões com *até que* do tipo relevante.

dois disparos (referidos por meios elípticos, não explicitamente oracionais)]

- (80) «Como o dinheiro não chegasse para comer, foi-lhe pedir aumento na fábrica. Mas o patrão não deu. Pediu outras vezes, sem sucesso. **Até que**, já desesperado, acabou por lhe apontar uma arma (...).» (ext13670-soc-92b-1)

[interpretação: a ameaça com arma sucede-se aos vários pedidos de aumento malsucedidos]

(ii) *até que* F_{IND} com descrições de situações atélicas e sem leitura durativa comum

- (81) «Para ele qualquer coisa não estava bem. Assim se foi passando o tempo, **até que** Sábado começou a explicar o tal de voto democrático, universal e secreto.» (ext4875-nd-95a-2)

[note-se que, naturalmente, o início da explicação do voto não interrompe a passagem do tempo]

- (82) «Após atingirem o Ilhéu de Cima, os marroquinos permaneceram no farol da ilha durante três dias, **até que**, para atraírem as atenções dos residentes de Porto Santo, optaram por desligar a rotativa do farol.» (ext1160-soc-92a-1)

[note-se que a opção por desligar a rotativa não interrompe a permanência dos marroquinos no farol]

- (83) «Tudo ocorreu rapidamente, a meio da tarde de quinta-feira, em plena rua, perante três testemunhas que não tiveram tempo de reagir: o casal, andando rumo à praia, discutia em voz alta, **até que** trava o passo.» (ext3286-nd-98a-2)

[note-se que o ato de travar o passo não sinaliza necessariamente o fim da discussão]

- (84) «Em 1934, construiu-se um estúdio-auditório com capacidade para 500 pessoas, mas a demanda foi tanta que pôs em risco a realização do próprio programa. Durante o Verão desse ano, não houve audiência. **Até que** alguém se lembrou de alugar o Hillsboro Theatre, com 2400 lugares.» (ext8672-nd-94b-2)

[note-se que a situação de não haver audiência, limitada pelo

adjunto temporal *durante o verão desse ano*, é temporalmente anterior ao aluguer do Hillsboro Theatre]

A sequência portuguesa – de registo relativamente informal (mas com 11 registos no *corpus* CETEMPúblico) – *até que, de repente* parece sinalizar normalmente apenas este valor.

- (85) a. «Não existia o mínimo controlo à entrada (...). Os fãs locais esperavam na pequena sala, com “tshirts” nos braços para serem autografadas. **Até que, de repente**, divisámos B.B. King a passear vagarosamente junto à piscina, de camisa estampada fora das calças (...).» (ext170288-clt-95b-1)
- b. «Seguiu-se um período sem êxodo, **até que, de repente**, em 1978-1979, cerca de 300 mil vietnamitas (...) foram mais ou menos expulsos do país.» (ext653116-nd-91a-1)
- c. «Os dias decorrem tranquilos **até que, de repente**, o céu começa a escurecer e um poderoso tufão assola a baía, varrendo tudo à sua passagem.» (ext1050484-clt-soc-94a-1)

A sequência inglesa *until suddenly* também parece poder ter um valor semelhante, como se verifica nos seguintes dois exemplos do BNC:

- (86) a. «Bitter disappointment caused her to merely pick at her breakfast **until suddenly** she was startled by the sound of his deep voice coming from behind her.»
- b. «We cried **until suddenly** my mother said, ‘We will not achieve anything with tears. (...)’.»

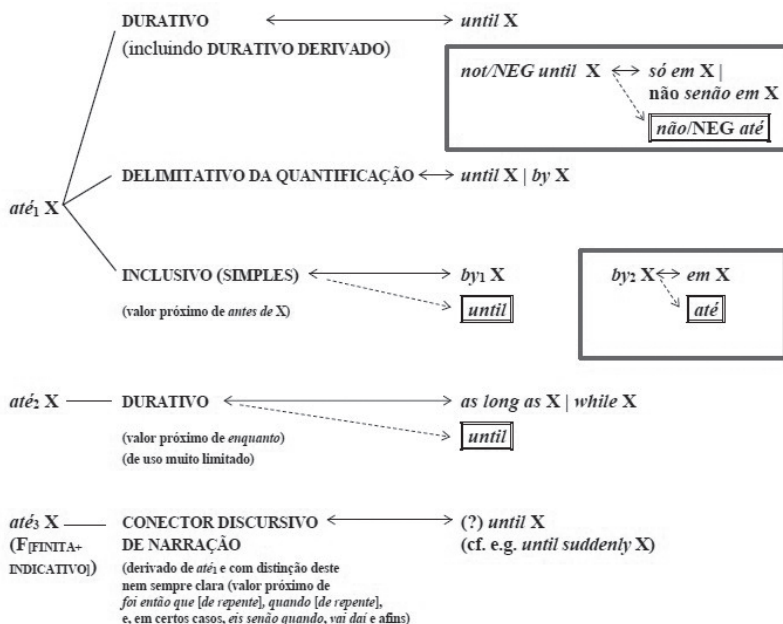
Várias sequências portuguesas sem *até*, como *foi então que, de repente* ou *quando, de repente* – ou mesmo (certos usos de) claros marcadores discursivos como *eis senão quando* (cf. Lobo 2003: 91) ou *vai daí* –, parecem ter um sentido equivalente ou muito próximo (e uma distribuição semelhante). Seguem-se exemplos (do *corpus* CETEMPúblico) de cada uma destas expressões em contextos em que são substituíveis sem alteração significativa de sentido por *até que*:

- (87) «Tentei outro. Não, duas lâmpadas fundidas no mesmo dia já não é normal. (...) **Foi então que, de repente,** caí em mim.» (ext664046-nd-93b)
[cf. *até que, de repente, caí em mim*]
- (88) «Kenny Chesney estava praticamente a terminar a sua actuação **quando de repente** e sem percebermos porquê, as pessoas na plateia começaram a debandar.» (ext855009-clt-95a-1)
[cf. *até que, de repente, (...) as pessoas (...) começaram a debandar*]
- (89) «A situação chegou a um ponto tal que, após eu ter alertado, sem sucesso, a Polícia, considerei a hipótese de mudar de casa. **Eis senão quando,** por obra do Espírito Santo, provavelmente acolitado por algum vizinho mais determinado, o bicho apareceu morto.» (ext560102-nd-91a-2)
[cf. *até que (...) o bicho apareceu morto*]
- (90) «Os homens do gang [sic] não o quiseram ouvir e insistiram na exigência. **Vai daí,** o visitante pegou na amante e pô-la fora da casa.» (ext157652-clt-98a-1)
[cf. *até que o visitante pegou na amante e pô-la fora de casa*]

6. Conclusões

Vimos neste trabalho que o conector temporal *até* pode ter – ou estar associado a – cinco valores distintos em português, correspondendo a pelo menos três expressões do inglês (*until*, *by* e as *long as*). Tanto o conector *until* como o conector *by*, por sua vez, possuem pelo menos um valor cada um em que não são contrapartidas de *até*, mas que é suficientemente próximo para dar origem a problemas de tradução. A Figura 1 sintetiza as correspondências aqui consideradas. Os vários problemas de tradução documentados estão sinalizados dentro de caixas com traço duplo, precedidas de setas a tracejado.

FIGURA 1 – Correspondências entre adjuntos temporais com até e suas contrapartidas inglesas (com sinalização dentro de caixas com traço duplo, precedidas de setas a tracejado, de problemas de tradução documentados no *site* linguee.com e/ou em sistemas de tradução automática)



REFERÊNCIAS

- Condoravdi, C. (2008). Punctual *Until* as a Scalar NPI. In K. Hanson, & S. Inkelas (Eds.), *The Nature of the Word: Studies in Honor of Paul Kiparsky* (pp. 631-654). MIT Press Scholarship Online.
- Declerck, R. (1995). The problem of *not... until*. *Linguistics*, 33, 51-98.
- Giannakidou, A. (2002). UNTIL, Aspect, and Negation: A Novel Argument for Two *Untils*. In B. Jackson (Ed.), *Proceedings of SALT XII* (pp. 84-103). Cornell University. <https://semanticsarchive.net/Archive/zA1Mzk1Y/until.salt12.pdf>

- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press.
- Kamp, H., & Reyle U. (1993). *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer.
- Karttunen, L. (1974). *Until*. In M.W. La Galy, R. A. Fox, & A. Bruck (Eds.), *Papers from the Tenth Regional Meeting. Chicago Linguistic Society* (pp. 284-297). Chicago Linguistic Society.
- Klima, E. (1964). Negation in English. In J. Fodor, & J. Katz (Eds.), *The Structure of Language: Readings in the Philosophy of Language*. Prentice Hall.
- Lascarides, A., & Asher, N. (1993). Temporal Interpretation, Discourse Relations, and Common Sense Entailment. *Linguistics and Philosophy* 16, 437-493.
- Lobo, M. (2003). *Aspectos da Sintaxe das Orações Subordinadas Adverbiais do Português* [Dissertação de doutoramento]. Universidade Nova de Lisboa.
- Mittwoch, A. (1977). Negative sentences with *until*. In W. A. Beach, S. E. Fox, & S. Philosoph (Eds.), *Papers from the Thirteenth Regional Meeting. Chicago Linguistic Society* (pp. 410-417). Chicago Linguistic Society.
- Mittwoch, A. (1980). The Grammar of Duration. *Studies in Language*, 4(2), 201-227.
- Móia, T. (1995). Aspectos da Semântica das Expressões Temporais com 'desde' e 'até' – Questões de 'Aktionsart'. In *Actas do X Encontro Nacional da Associação Portuguesa de Linguística (Évora 1994)* (pp. 341-358). Associação Portuguesa de Linguística.
- Móia, T. (2000). *Identifying and Computing Temporal Locating Adverbials with a Particular Focus on Portuguese and English* [Dissertação de doutoramento]. Universidade de Lisboa.
- Móia, T. (2004). Sobre a Delimitação Temporal da Quantificação. In T. Freitas, & A. Mendes (Orgs.), *Actas do XIX Encontro Nacional da Associação Portuguesa de Linguística (Lisboa, 1, 2 e 3 de Outubro de 2003)* (pp. 581-593). Associação Portuguesa de Linguística.
- Móia, T. (2006). On Temporally Bounded Quantification over Eventualities. *ZAS Working Paper in Linguistics*, 44(1), 225-238.
- Móia, T. (2015). The Durative Verbs of Portuguese. *Diacrítica*, 29(1), 27-59.
- Móia, T. (2016). Subclasses of Temporal and Spatial Phrases in Portuguese – Location vs. Mere Reference. *Journal of Portuguese Linguistics*, 15(1), 1-17.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. Longman.

de Swart, H. (1996). Meaning and use of *not... until*. *Journal of Semantics*, 13(3), 221-263.

Dicionários de inglês

Collins Cobuild. English Dictionary for Advance Learners, Glasgow, Harper Collins Publishers, 3.^a ed., 2001.

Dictionary by Merriam-Webster: www.merriam-webster.com

Lexico (feito em colaboração com o *Oxford Dictionary*): www.lexico.com

Longman Dictionary of Contemporary English Online: www.ldoceonline.com

Corpora

[BNC] British National Corpus, acessível em <https://corpus.byu.edu/>

[CETEMPúblico] *Corpus CETEMPúblico* 1.7 v. 11.2, acessível em <http://www.linguateca.pt/ACDC/>

Websites de tradução

[LINGUEE] www.linguee.com

[DT] www.deepl.com

[GT] www.translate.google.com

[ST] www.systran.net

