

Resolução de pronomes em posição de sujeito em orações completivas na retoma de antecedentes em SNs complexos: estudo comparativo entre Português Brasileiro e Português Europeu

Katharine da Hora[♦], Paula Luegi^{✧1}, Marcus Maia[♦] & Armanda Costa[✧]

[♦]Universidade Federal do Rio de Janeiro, LAPEX, [✧]Universidade de Lisboa, CLUL

Abstract:

In this study we tested, in EP and in BP, complement clauses with a null or an overt pronoun in subject position that is forced by number agreement to retrieve an antecedent within a complex subject NP in the main clause. With an eye-tracking while reading paradigm, we analysed the impact of structural position on pronoun resolution, investigating if the bias described for null-subject and overt-object pronoun resolution for null subject languages replicates for antecedents with different structural positions: null-NP1, the highest structural entity, and overt-NP2, the lowest structural entity. Moreover, we especially investigate the impact of structural position in BP, where c-command relations are considered to be of great relevance for null subject resolution. Results indicate that structural position impacts on pronoun resolution in the predicted way: null-highest NP, overt-lowest NP. Also, BP results reveal that, not only is the null form more constrained by c-command relations, preferentially referring to the c-commanding antecedent (NP1), but also that the overt pronoun does not show a clear bias also when considering structural position (as does not for syntactic function, as shown in previous studies).

Keywords: *Language processing, correferential resolution, null and overt pronominal subjects, European Portuguese and Brazilian Portuguese*

Palavras-chave: Processamento da linguagem, resolução da correferência, sujeito pronominal nulo e pleno, Português Europeu e Português Brasileiro

1. Introdução

O estudo da resolução de expressões referenciais anafóricas é um tema central em diferentes áreas da linguística, desde a linguística computacional à psicolinguística e suas sub-áreas, nomeadamente, em aquisição de L1 (língua materna) e L2 (segunda língua), mas também no estudo de populações com comportamento típico e atípico. O foco tem sido o de identificar que fatores contribuem para a atribuição de proeminência às diferentes entidades referidas no discurso tornando-as, conseqüentemente, mais ou menos acessíveis em memória e, assim, passíveis de serem retomadas com expressões mais ou menos reduzidas. De acordo com a escala de acessibilidade de Ariel (1990), as entidades mais proeminentes no discurso, e que por isso se tornam mais acessíveis em termos de memória, são mais facilmente recuperáveis e geralmente referidas com formas linguisticamente mais reduzidas, como as formas nulas, nas línguas que as admitem.

1 Katharine da Hora e Paula Luegi contribuíram de igual modo para o presente trabalho, tendo-se optado por colocar os nomes por ordem alfabética.

Este estudo foi desenvolvido com fundos da Fundação para a Ciência e a Tecnologia, através de uma bolsa de pós-doutoramento atribuída a Paula Luegi (SFRH/BPD/84138/2012), e por fundos da Capes e do CNPq (bolsa de doutoramento) atribuídos a Katharine da Hora.



Diferentes trabalhos têm demonstrado que, especificamente, a resolução de formas pronominais nulas e plenas em posição de sujeito, em línguas de sujeito nulo, é fortemente apoiada na função sintática do antecedente, considerando que a função sintática é o fator determinante na atribuição de proeminência às entidades referidas no discurso. Consequentemente, formas pronominais nulas são preferencialmente interpretadas como retomando o antecedente em posição de sujeito e formas pronominais plenas como retomando o antecedente que não esteja em posição de sujeito, seja o objeto ou um referente extralinguístico (Costa, Faria e Matos, 1998; Carminati, 2002). No entanto, para além da função sintática, outros fatores têm sido identificados como relevantes na resolução de expressões referenciais anafóricas, nomeadamente, a informação semântica dos verbos e os papéis temáticos atribuídos aos antecedentes (Garvey e Caramazza, 1974; Costa 2003/2005; Costa, Faria e Kail, 2004; Morgado, 2012), ou, ainda, a ordem de referência das entidades no discurso (Järvikivi *et al.*, 2005; Luegi, 2012).

Neste estudo avaliamos se a posição estrutural ocupada pelas entidades referidas no discurso, neste caso, elementos constituintes do SN sujeito, com funções sintáticas e estruturais distintas das habitualmente testadas, tem também impacto na resolução de, especificamente, formas nulas e plenas em posição de sujeito em Português Europeu (PE) e em Português Brasileiro (PB).

Ao contrário de trabalhos anteriores, que testam cadeias correferenciais sobretudo em frases independentes (apostas ou coordenadas) ou no domínio intrafrásico de construções subordinadas (na maior parte dos casos adverbiais), neste estudo testamos condições em que os potenciais antecedentes para uma expressão anafórica pronominal integram um sintagma nominal (SN) complexo com função de sujeito. O SN complexo contém dois SNs: o SN que designaremos de SN1, que contém o núcleo do SN complexo, e o SN2, integrado no SP complemento do núcleo. Assim, as duas entidades referidas distinguem-se pela posição estrutural ocupada: SN1, mais alto e, por isso, supostamente mais proeminente; SN2, mais encaixado e, por isso, supostamente menos proeminente. Porque estudos anteriores demonstraram que a relação entre orações é também um fator determinante na resolução de expressões correferenciais (Costa, Faria e Matos, 1998; Corrêa, 1998; Carminati, 2002), optámos por testar orações completivas, em que o verbo da oração matriz seleciona uma oração como seu argumento interno, cujo sujeito pronominal deve buscar o seu antecedente no domínio sintático superior. Assim, temos uma relação sintática entre domínios oracionais de forte coesão que limita cognitiva e gramaticalmente o campo de busca do antecedente.

Esta manipulação, ou seja, o uso de orações com uma relação estrutural mais forte (matriz-completiva), é particularmente relevante para o contraste entre PE e PB. Estudos de Duarte (1995) e Duarte, Barbosa e Kato (2005), com dados de *corpora*, demonstraram que a ocorrência de formas nulas e plenas em PE e em PB é distinta, sendo a ocorrência de formas nulas mais reduzida em PB. Estudos recentes confirmaram que, de facto, também na resolução de formas anafóricas nulas e plenas em posição de sujeito existem diferenças entre as duas variedades (Luegi, Costa e Maia, 2014; Fernandes *et al.*, 2018): em PE confirma-se a preferência nulo-sujeito e pleno-objeto, mas em PB essa assimetria não se verifica, mantendo-se a preferência para o nulo mas não se registando uma preferência clara para a forma plena. Mais ainda, os estudos demonstram também que, para além de menos frequente, a forma nula é, também, mais restringida pelo contexto sintático em que ocorre: a forma nula em PB deve ser c-comandada pelo antecedente que refere (Barbosa, Duarte e Kato, 2005; Holmberg, Nayudu e Sheehan, 2009; I. Duarte e Figueiredo Silva, 2016).

Nas estruturas que agora selecionamos, o SN complexo no seu todo c-comanda o sujeito da completiva, mas o SP complemento que alberga o SN2 não. Esta manipulação permite-nos, assim, identificar possíveis efeitos das relações de c-comando na resolução de formas nulas em posição de sujeito. Semanticamente, o núcleo do SN complexo é um nome dependente (*estagiário, vizinho, amigo...*) que seleciona um complemento. Assim o SN complexo constitui-se como um nome dependente complementado (Raposo *et al.*, 2013:742 – vol I). Esta condição vai situar gramatical e semanticamente os dois SNs em diferentes patamares cognitivos de acessibilidade: SN1 contém o núcleo do SN complexo e a sua referência é semanticamente



dependente, o que poderá contribuir para que ganhe proeminência estrutural (é ele que determina a concordância verbal) e ganhe potencial como foco de atenção (o acesso lexical e semântico fica na expectativa do complemento que o especifica), em contraste com o SN₂, integrado no SP complemento e, consequentemente, menos proeminente.

Em termos metodológicos, desenvolvemos uma experiência com o paradigma de *eye-tracking*, na leitura de frases como as exemplificadas de (1) a (4). Os dados foram recolhidos em Portugal e no Brasil, com falantes nativos de cada variedade no seu país de origem, com as mesmas frases, de modo a comparar os efeitos das variáveis manipuladas em PE e em PB.

- (1) No final da noite, os estagiários do arquiteto informaram que discutiam irritadamente com frequência.
- (2) No final da noite, o estagiário dos arquitetos informou que discutiam irritadamente com frequência.
- (3) No final da noite, os estagiários do arquiteto informaram que eles discutiam irritadamente com frequência.
- (4) No final da noite, o estagiário dos arquitetos informou que eles discutiam irritadamente com frequência.

Em síntese, neste estudo, pretendemos verificar se, primeiro, a distribuição complementar entre nulo e pleno, identificada para as retomas de sujeito e de objeto (nulo-sujeito e pleno-objeto), se regista também para a retoma do núcleo do SN e seu complemento (nulo-SN₁ e pleno-SN₂), e, segundo, se este efeito é semelhante em PE e em PB ou se, pelo contrário, em PB também não se encontram preferências para a forma plena, como verificado nos estudos anteriormente referidos. Por fim, pretendemos também verificar se, num contexto mais restritivo, com estruturas subordinadas completivas, em que as relações de c-comando são mais fortes, se registam diferenças entre PE e PB, sendo que, nesta última variedade, se espera um efeito mais claro da relação de c-comando na resolução de formas nulas.

Nas próximas secções apresentamos uma revisão dos estudos sobre o tema e, posteriormente, descrevemos o trabalho experimental desenvolvido, apresentando a experiência realizada em PE e em PB. Por fim apresentamos os resultados obtidos, a sua discussão e as possíveis conclusões que podemos extrair dos resultados obtidos.

1.1. Pressupostos teóricos

A Teoria da Acessibilidade de Ariel (1990) propõe que cada expressão anafórica codifica o grau de acessibilidade do antecedente que retoma: quanto mais acessível o antecedente menor é a probabilidade de ser retomado por uma expressão anafórica com um índice de informatividade alto. Assim, considerando as formas pronominais nulas e plenas, prevê-se que expressões anafóricas mais informativas como o pronome pleno retomem antecedentes menos acessíveis e expressões anafóricas menos informativas, como o pronome nulo, retomem antecedentes muito acessíveis dentro do discurso.

Diferentes estudos (Costa, Faria e Matos, 1998; Corrêa, 1998; Carminati, 2002; Alonso-Ovalle *et al.*, 2002) demonstram que, nas línguas de sujeito nulo, existe uma distribuição complementar entre as formas referenciais nulas e plenas em posição de sujeito. Um destes estudos foi desenvolvido por Carminati (2002) para o Italiano. Com base nos resultados, Carminati (2002) propõe a Hipótese da Posição do Antecedente, uma proposta em que se destaca o papel da função sintática do antecedente. Segundo Carminati, em Italiano, os pronomes nulos são preferencialmente interpretados como retomando os antecedentes em Spec de IP enquanto que os pronomes plenos são interpretados como retomando os antecedentes que não estão em Spec de IP. A proposta de Carminati (2002) desenvolve-se a partir da proposta de Ariel (1990) no que concerne à relação entre a forma da expressão anafórica e a acessibilidade do antecedente, mas assume que a



proeminência do antecedente é estabelecida essencialmente com base em fatores sintáticos, especificamente, com base na função sintática do antecedente.

No entanto, outros fatores têm também sido identificados como relevantes na resolução de diferentes expressões referenciais, tais como a informação semântica dos verbos e os papéis temáticos atribuídos aos antecedentes (Garvey e Caramazza, 1974; Costa 2003/2005; Costa, Faria e Kail, 2004; Morgado, 2012), a ordem de referência por que são introduzidas no discurso as diferentes entidades (Järvikivi *et al.*, 2005; Luegi, 2012) ou mesmo a relação entre orações, ou seja, em domínios inter e intrafrásicos (Corrêa, 1998; Costa, Faria e Matos, 1998, Carminati, 2002).

Por exemplo, Morgado (2012) verificou que a relação existente entre as estruturas (justapostas vs. concessivas) em que os referentes e expressão referencial ocorrem influencia a escolha do antecedente, mas também que a relação existente entre as formas nulas e o sujeito e as formas plenas e o complemento do verbo não é tão estável como proposto pela Hipótese da Posição do Antecedente. Para Morgado (2012), os papéis temáticos atribuídos pelo verbo também interferem na resolução de formas nulas e plenas em posição de sujeito: o pronome pleno é preferencialmente interpretado como retomando um antecedente que não seja semanticamente proeminente, como o objeto direto Tema, em frases ativas (*Bruno*, em *O Samuel agrediu o Bruno no pavilhão. Horas mais tarde, ele discutiu o assunto com preocupação.*), e o sujeito Tema, em frases passivas (*Samuel*, em *O Samuel foi agredido pelo Bruno no pavilhão. Horas mais tarde, ele discutiu o assunto com preocupação.*), não dependendo assim exclusivamente da função sintática do antecedente. Morgado (2012) conclui que, na linha de outros estudos, as diferentes expressões anafóricas não são sensíveis aos mesmos fatores e que, sobretudo, a proeminência de uma entidade resulta da combinação de diferentes fatores. Esta linha segue a proposta *Multiple Constraints Approach* (Hipótese Multi-Fatorial) de Kaiser e Trueswell (2008), em que se assume que diferentes fatores contribuem para a proeminência de uma entidade. A Hipótese Multi-Fatorial propõe que não só diferentes níveis de informação linguística são considerados na resolução de expressões referenciais como, também, diferentes formas são sensíveis a diferentes fatores e de modo distinto.

Järvikivi *et al.* (2005), num estudo em que testam a resolução de formas plenas em Finlandês na retoma de antecedentes em estruturas com ordem sujeito-verbo-objeto (SVO) e com ordem objeto-verbo-sujeito (OVS)², em que o objeto é marcado com o caso partitivo com o sufixo “-ia” (*Tony Blair_{SUJ} katteli George Bushi_{OBJ} valkoisessa talossa.*³ vs. *George Bushi_{OBJ} katteli Tony Blair_{SUJ} valkoisessa talossa.*⁴), verificam que o pronome é preferencialmente interpretado como retomando o sujeito, mas, sobretudo, quando o sujeito é a primeira entidade referida no discurso. Este trabalho demonstra que não só a função sintática, mas também a ordem de referência (Gernsbacher e Hargreaves, 1988), ou a posição estrutural, contribuem para a atribuição de proeminência a uma entidade tornando-a o antecedente preferencial de uma forma reduzida em posição de sujeito.

Resumidamente, diferentes estudos demonstraram que a função sintática não é a única informação usada na resolução de expressões referenciais e que a escolha de um antecedente é o resultado da combinação de diferentes fatores.

² Em Finlandês, as funções sintáticas são marcadas morfossintaticamente, o sujeito tipicamente no caso nominativo e o objeto no caso partitivo, marcado com o sufixo “-ia”.

³ Tony Blair_{SUJ} apertou a mão a George Bush_{OBJ} na casa Branca.

⁴ A George Bush_{OBJ} apertou a mão Tony Blair_{SUJ} na casa Branca.



1.2. O sujeito nulo em PE e PB

Tem sido descrito em diferentes trabalhos que a distribuição das formas nulas e plenas em posição de sujeito é distinta em PE e em PB, apesar de serem duas variedades da mesma língua. Atualmente, enquanto se considera que o PE é uma língua de sujeito nulo consistente, o PB tem sido classificado com uma língua de sujeito nulo parcial (Barbosa, 2011; I. Duarte e Figueiredo Silva, 2016), uma vez que se tem registado, entre outros fenómenos, uma redução da ocorrência de formas nulas referenciais, registando-se, em contrapartida, um aumento de ocorrência de formas plenas em posição de sujeito (Duarte, 1995; Barbosa, Duarte e Kato, 2005).

Barbosa, Duarte e Kato (2005) fizeram um estudo comparativo entre PB e PE com dados de *corpora* escritos extraídos de jornais⁵ acerca da distribuição dos pronomes plenos e nulos na terceira pessoa. As autoras observaram que em PB, para além de uma menor ocorrência de formas nulas em posição de sujeito, com um consequente aumento do uso de formas plenas, a relação de c-comando existente entre o antecedente e a expressão referencial nula constitui o contexto que mais favorece o sujeito nulo. Assim, com base nos dados, Barbosa, Duarte e Kato (2005) verificam que, tanto em PB quanto em PE, os antecedentes que ocupam a posição de sujeito na oração principal são os que mais favorecem a retoma pelo pronome nulo, mas que em PB a relação de c-comando é crucial para a ocorrência de formas nulas: em PB a forma nula é essencialmente legitimada por antecedentes que a c-comandem.

Estudos recentes da comparação das duas variedades na resolução de formas nulas e plenas em posição de sujeito confirmam esta diferença entre PE e PB a nível de processamento. Luegi, Costa e Maia (2014; Fernandes *et al.*, 2018) verificam, numa experiência de questionário e numa tarefa de *self-paced reading*, que em PB não existem preferências claras na interpretação da forma plena: o pronome pleno é interpretado tanto como retomando o sujeito como o objeto da oração precedente. Mais ainda, Luegi, Costa e Maia (2014) verificam que, na tarefa de *self-paced reading*, em que se testaram apenas formas nulas retomando antecedentes em orações com ordem SVO (*A Sara escreveu para o Emanuel em Maio quando foi jubilada/o pela faculdade no ano passado*) ou OVS (*Para o Emanuel escreveu a Sara em Maio quando foi jubilada/o pela faculdade no ano passado.*), em PB, o nulo é mais facilmente interpretado como correferente com o sujeito ou com o antecedente mais próximo que o c-comanda, uma vez que se registaram tempos mais baixos na retoma do sujeito independentemente da sua posição estrutural. Este resultado está de acordo com as propostas de Barbosa, Duarte e Kato (2005), Holmberg, Nayudu e Sheehan (2009) ou I. Duarte e Figueiredo Silva (2016) que propõem que os contextos de ocorrência de sujeitos nulos são mais restritivos em PB: (i) o sujeito nulo tem de ter um antecedente e (ii) o antecedente do sujeito nulo tem de o c-comandar.

Assim, diferentes trabalhos demonstram que tanto ao nível da produção como ao nível do processamento existem diferenças entre PE e PB no que diz respeito às formas nulas e plenas em posição de sujeito⁶. Ao contrário do que ocorre em PE, em PB o uso de formas nulas é mais reduzido, sendo estas formas, quando ocorrem, mais restringidas pelo contexto sintático. As formas plenas, sendo mais frequentes em PB, tendem a retomar mais o sujeito do que em PE, não se registando por isso uma preferência clara pelo objeto, como se verifica em PE (e em outras línguas de sujeito nulo consistente, como o Italiano).

Na secção seguinte apresentamos o trabalho experimental desenvolvido explicando de que modo esse trabalho se desenvolve a partir do que foi exposto na presente secção.

⁵ Edições de domingo do jornal de Lisboa (*O Público*) e do Rio de Janeiro (*Revista Domingo*) de 1999 e 2000.

⁶ Estes resultados corroboram assim a distinção entre PE e PB enquanto línguas de sujeito nulo consistente e parcial, respetivamente.



2. Presente estudo

Com base em estudos anteriores (Costa, Faria e Matos, 1998; Corrêa, 1998; Carminati, 2002; entre outros), considera-se que a função sintática dos antecedentes é um fator preponderante na atribuição de proeminência a uma entidade referida (sobretudo em contextos intrafrásicos) e que essa informação é considerada na resolução de expressões referenciais, nomeadamente formas pronominais nulas e plenas, em posição de sujeito, sendo as primeiras preferencialmente interpretadas como retomando o sujeito da oração precedente e as segundas o objeto. No entanto, várias propostas têm sido apresentadas no sentido de se considerar outros fatores na resolução de formas nulas e plenas em posição de sujeito, por exemplo, a ordem de referência dos constituintes (Gernsbacher e Hargreave, 1988; Järvikivi *et al.*, 2005; Luegi, 2012).

Neste estudo avaliamos se a proeminência das diferentes entidades pode também ser influenciada pela posição estrutural que ocupam e se, conseqüentemente, a posição estrutural influencia também a resolução de formas nulas e plenas em posição de sujeito. Pretende-se assim avaliar se a distribuição complementar descrita para forma nula e plena para as línguas de sujeito nulo consistente se regista quando as duas entidades (potenciais antecedentes) fazem parte do mesmo constituinte mas ocupam posições estruturais distintas. Considerando as diferenças entre PE e PB, avalia-se também a resolução de formas nulas e plenas em posição de sujeito por falantes nativos de PE e de PB, fazendo-se uma análise comparativa entre as duas variedades.

Para avaliar o efeito da posição estrutural dos antecedentes, criámos condições em que o sujeito da oração matriz é um SN complexo, composto por um SN (SN1) cujo núcleo nominal é complementado por um SP que integra o SN2. Assim em termos de entidades disponíveis para identificar um pronome subsequente, existem dois SNs: SN1, a entidade estruturalmente mais alta, e SN2, estruturalmente mais encaixado. Esta estrutura permite distinguir a influência da proeminência estrutural em termos de hierarquia de constituintes, não entre funções sintáticas principais como tem sido estudado (sujeito vs. objeto), mas sim entre funções dentro do mesmo constituinte. A primeira hipótese de trabalho é a de que, se a posição estrutural influencia a proeminência das entidades e, conseqüentemente, tem um efeito na resolução das formas nulas e plenas em posição de sujeito, então a resolução de nulos e plenos será distinta, sendo os nulos preferencialmente interpretados como retomando o SN estruturalmente mais proeminente, ou seja, o SN1, e o pleno o SN estruturalmente menos proeminente, o SN2. Esta hipótese prevê uma distribuição (quase) complementar semelhante à verificada para a oposição sujeito/objeto mas, neste caso, para SNs em posições estruturais distintas, ainda que internos ao mesmo constituinte.

Para além da manipulação relativa à posição estrutural das entidades referidas, optámos por testar estruturas particularmente sensíveis à influência de fatores estruturais, ou seja, estruturas em que a oração matriz é seguida de uma oração subordinada completiva. Nestas estruturas existe uma relação forte entre oração matriz e subordinada. Apesar de não existirem restrições sintáticas (como não existem nas estruturas tipicamente estudadas), dadas as suas características, estas estruturas impõem maiores limitações⁷, particularmente relevantes para o contraste entre PE e PB.

Considerando as diferenças identificadas entre PE e PB, contrastamos estas duas variedades a fim de testar, em condições sintaticamente mais restritivas, as previsões de trabalhos anteriores. Como referido anteriormente, PE e PB distinguem-se não só quanto à ocorrência de formas nulas e plenas como também aos contextos de ocorrência destas formas. Em PB existe uma ocorrência mais elevada de formas plenas do que em PE (Duarte, 1995; Barbosa, Duarte e Kato, 2005) e essa diferença tem impacto na resolução dessas formas, não se identificando uma preferência clara na interpretação do pronome pleno em PB (Luegi, 2012; Fernandes *et al.*, 2018). Com base nos trabalhos anteriores, formulamos a hipótese de que, caso a primeira

⁷ [E]vita-se o pronome sujeito das completivas finitas quando a interpretação pretendida é a da co-referência do sujeito da frase superior e do da frase completiva (Raposo *et al.*, 2013:610).



hipótese deste trabalho se confirme, a complementaridade prevista para nulo e pleno na retoma de antecedentes em posições estruturais distintas seja mais forte em PE do que em PB.

Em termos de restrições, como referido anteriormente, vários autores propõem que os contextos de ocorrência de sujeitos nulos são mais restritivos em PB (Barbosa, Duarte e Kato, 2005; Holmberg, Nayudu e Sheehan, 2009; I. Duarte & Figueiredo Silva, 2016). Nas estruturas testadas, o SN complexo, no seu todo, c-comanda o sujeito da completiva, assim como o SN1, mas o SN2 não o faz. A terceira hipótese deste estudo é, então, a de que nas duas variedades a forma nula seja preferencialmente interpretada como retomando o SN1, ou seja, mais proeminente, e que em PB essa preferência seja mais acentuada uma vez que este SN é o SN que c-comanda o sujeito nulo da oração completiva. Considerando que em PB o sujeito nulo é opcional quando correferente com o sujeito da oração matriz e que nesses casos o antecedente deve c-comandar o sujeito nulo (Figueiredo Silva, 2000; Modesto, 2000; Rodrigues, 2004), será de esperar que a forma nula tenha maiores restrições em PB do que em PE, sendo preferencialmente interpretada como retomando o SN1 da oração precedente.

2.1. Metodologia

2.1.1. Participantes

Participaram neste estudo 24 falantes nativos de Português Europeu e 24 falantes nativos de Português Brasileiro, todos estudantes da Universidade de Lisboa e da Universidade Federal do Rio de Janeiro, respetivamente. Os participantes assinaram um termo de consentimento aceitando participar, de forma voluntária, no estudo e autorizando a divulgação dos resultados, sob anonimato, para fins de investigação científica.

2.1.2. Materiais

Foram criados 24 itens experimentais manipulando-se Pronome (nulo vs. pleno) e Antecedente retomado (SN1 vs. SN2). Todas as frases tinham a estrutura matriz-subordinada sendo a subordinada uma oração completiva selecionada pelo verbo da oração matriz. O sujeito da oração matriz era sempre um SN complexo composto por dois SNs, o SN1 que integra o núcleo do SN complexo e o SN2 interno ao SP complemento do núcleo do SN complexo. Os dois SNs tinham sempre informação de número diferente: singular vs. plural (SN1-singular/SN2-plural; SN1-plural/SN2-singular). O sujeito da oração completiva poderia ser um sujeito pronominal nulo ou pleno. A informação de número no pronome pleno e/ou no verbo da oração subordinada forçavam a retoma de SN1 ou de SN2, como se pode ver nos exemplos de (5) a (8).

- (5) No final da noite, os estagiários do arquiteto informaram que eles discutiam irritadamente com frequência.

Pergunta: Quem discutia irritadamente? (a) os estagiários (b) o arquiteto

- (6) No final da noite, o estagiário dos arquitetos informou que eles discutiam irritadamente com frequência.

Pergunta: Quem discutia irritadamente? (a) o estagiário (b) os arquitetos

- (7) No final da noite, os estagiários do arquiteto informaram que discutiam irritadamente com frequência.



Pergunta: Quem discutia irritadamente? (a) os estagiários (b) o arquiteto

(8) No final da noite, o estagiário dos arquitetos informou que discutiam irritadamente com frequência.

Pergunta: Quem discutia irritadamente? (a) o estagiário (b) os arquitetos

Todas as frases foram construídas de modo a que pudessem ser testadas tanto em PE como em PB tendo sido, no entanto, necessário fazer algumas adaptações a nível de léxico para cada variedade. Nenhuma das adaptações foi feita em regiões relevantes para a análise.

Para além das frases experimentais, foram apresentadas 48 frases distratoras com estruturas sintáticas distintas das das frases experimentais.

2.1.3. Procedimento

Neste estudo utilizou-se o paradigma de *eye-tracking*, registando-se os comportamentos oculares dos participantes durante a leitura silenciosa de frases. As frases foram apresentadas isoladamente no ecrã do computador e depois da leitura de cada frase o participante tinha de responder a uma pergunta de interpretação sobre a frase lida (ver “Pergunta” nos exemplos de (5) a (8)). Os itens experimentais foram apresentados de forma pseudoaleatória intercalados por frases distratoras. A ordem pseudoaleatória permitiu garantir que, apresentando os itens distratores por uma ordem fixa, as frases experimentais, apresentadas de modo aleatório, fossem sempre intercaladas por pelo menos uma frase distratora.

Antes do início da experiência os participantes passavam pelo processo de calibração do equipamento e, posteriormente, realizavam um pequeno treino, com a leitura de cinco frases, seguidas de perguntas de interpretação, para se adaptarem à tarefa a realizar. Após a fase de treino, o equipamento era novamente calibrado e a calibração validada e apenas se aceitavam os valores de calibração quando o erro/desvio não era superior a 0.5° de ângulo visual, considerando as coordenadas verticais e horizontais. Quando o valor era superior a 0.5° de ângulo visual, o processo de calibração e de validação era repetido.

Em PE os estímulos foram apresentados no programa *Experiment Center* desenvolvido pela *SensoMotoric Instruments (SMI)*. Para registar os movimentos oculares utilizámos o equipamento de registo *iVewX Hi-Speed* da SMI, a uma velocidade de registo de 1250Hz em modo monocular registando sempre apenas um dos olhos. Os dados foram extraídos com o programa *BeGaze*, ambos desenvolvidos pela SMI.

Em PB os estímulos foram apresentados no programa *EyeTracker 0.7.10m* desenvolvido pela UMASS. Para registar a movimentação ocular, utilizámos o equipamento de registo *EyeLink 1000*, a uma velocidade de registo de 1000Hz em modo monocular registando sempre apenas um dos olhos. Os dados foram extraídos com os programas *EyeDoctor* e *EyeDry*, ambos desenvolvidos pela UMASS.

Embora os equipamentos utilizados em Portugal e no Brasil sejam diferentes, são altamente compatíveis, sendo os dois equipamentos semelhantes quanto à velocidade de registo.

2.1.4. Análise

Como variáveis dependentes optámos por analisar (i) o tempo da primeira leitura, First Pass (FP), que reflete o processamento mais imediato, indiciador de processos lexicais e sintáticos locais, dando-nos indicações sobre os efeitos da manipulação introduzida, em fases iniciais do processamento; e (ii) o tempo total de leitura, Total Pass (TP), a soma do tempo da primeira leitura da região e de todas as releituras que tenham ocorrido, que reflete todo o tempo gasto no processamento da região em análise, e por isso é indicativo de processos sintáticos menos locais, de integração de informação descontínua, como a das cadeias referenciais. Analisámos ainda as respostas às perguntas de compreensão realizadas após a apresentação de cada frase.



Para as análises estatísticas, utilizámos os modelos lineares mistos no *software* R (R Core Team, 2014) com o pacote lmer (Bates *et al.*, 2015) e a função binomial para a análise das respostas (variáveis categoriais). Usámos o modelo máximo justificado pelo *design* (Barr *et al.*, 2013), com todas as variáveis independentes e suas interações como fatores principais e com participantes e itens como fatores aleatórios. Todas as variáveis independentes foram centradas.

Em cada frase experimental foram delimitadas áreas de análise relevantes: o pronome, nas condições com pronome pleno, o verbo e o advérbio. Como região crítica definimos a região do pronome + verbo, uma vez que nas condições com forma nula é a flexão do verbo que permite recuperar o possível antecedente. Sendo a área crítica diferente em extensão entre condições e de modo a minimizar os efeitos das diferenças entre os equipamentos usados em Portugal e no Brasil, optámos por residualizar o tempo de leitura de FP e de TP na região crítica e também na região do advérbio, como sugerido por Trueswell, Tanenhaus e Garnsey (1994). Foi assim aplicada uma regressão linear considerando o tempo em função do número de caracteres de cada região. Este valor foi posteriormente subtraído aos tempos reais para obter o tempo residual de leitura. Um valor negativo significa um tempo de leitura mais rápido do que o previsto (*predicted*) e um valor positivo um tempo de leitura mais lento do que o previsto (Townsend e Bever, 2001).

Antes da residualização, eliminámos todos os valores abaixo de 100ms e, posteriormente, por região, todos os valores acima da média mais 2.5 desvio padrão.

2.2. Resultados

2.2.1. Respostas

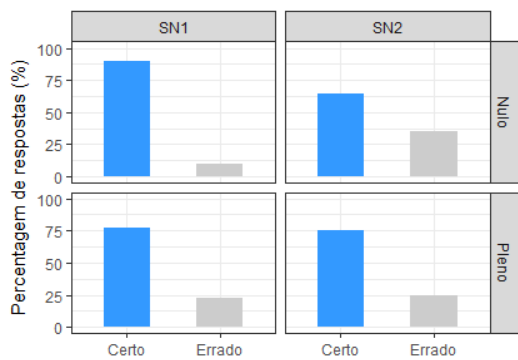
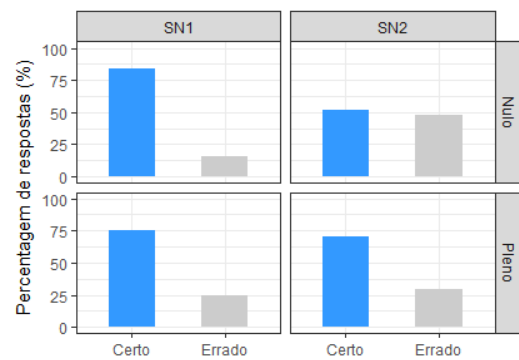
Começamos por apresentar os resultados de acerto na pergunta final. Como se pode ver na Tabela 1 e no Gráfico 1, em PE registou-se um efeito principal de antecedente, com mais erros⁸ na retoma de SN2, e um efeito de interação entre pronome e antecedente, com mais erro com nulo quando retoma SN2. Em PB, como se pode ver na Tabela 1 e no Gráfico 2, registou-se um efeito principal de antecedente, com mais erros na retoma de SN2, e um efeito de interação entre pronome e antecedente, com mais erro com nulo quando retoma SN2. Na comparação entre PE e PB, os resultados vão no mesmo sentido: efeito principal de antecedente (Est.=-1.096; SE=0.367; *t*-value=-2.983; *p*-value=0.003) e interação entre pronome e antecedente (Est.=1.909; SE=0.376; *t*-value=5.074; *p*-value=<0.001).

⁸ Escolha de SN1 quando a informação de número forçava a retoma de SN2 e vice-versa.



Tabela 1 – Resultados da análise LMM (respostas).

	PE				PB			
	Est.	SE	z-value	p-value	Est.	SE	z-value	p-value
(Intercept)	1.802	0.302	5.973	<0.001	1.419	0.290	4.897	<0.001
Pronome	-0.344	0.307	-1.121	0.262	0.172	0.408	0.422	0.673
Antecedente	-1.400	0.502	-2.791	0.005	-1.325	0.597	-2.218	0.027
Pronome:Antecedente	1.907	0.569	3.354	0.001	2.063	0.600	3.436	0.001

**Gráfico 1** – Taxas de acerto e de erro nas respostas à pergunta de interpretação em PE.**Gráfico 2** – Taxas de acerto e de erro nas respostas à pergunta de interpretação em PB.

2.2.2. Tempos de leitura

Apresentamos nesta secção os resultados referentes aos tempos de leitura, começando por apresentar os tempos da primeira leitura (*First Pass*) e posteriormente os tempos totais de leitura (*Total Pass*).

First Pass

Em PE, não se regista nenhum efeito significativo, nem, como se pode ver no Gráfico 3, na região do pronome e do verbo (Tabela 2) nem, como se pode ver no Gráfico 4, na região do advérbio (Tabela 3). Em PB, como se pode ver na Tabela 2, na região do pronome e do verbo, regista-se um efeito principal de pronome, com mais tempo na condição com nulo (Gráfico 4). Na região do advérbio (Gráfico 6), regista-se um efeito principal de pronome (Tabela 3), com mais tempo na condição com nulo, e um efeito de interação entre pronome e antecedente, com mais tempo no nulo na retoma de SN2.

No contraste entre as duas variedades, na região do pronome e do verbo, regista-se um efeito principal de pronome (Est.=-145.406; SE=38.691; t -value=-3.758; p -value=<0.001), com mais tempo no nulo do que no pleno, e um efeito de interação entre pronome e variedade (Est.=349.662; SE=60.296; t -value=5.799; p -value=<0.001), com mais tempo no nulo em PB do que em PE. Na região do advérbio, há uma interação entre pronome e antecedente (Est.=-49.177; SE=21.166; t -value=-2.323; p -value=0.0202), com mais tempo na condição em que nulo retoma SN2 e menos na condição em que pleno retoma SN2, e uma interação entre pronome e variedade (Est.=69.103; SE=28.445; t -value=2.429; p -value=0.015), com mais tempo com nulo em PB e mais tempo com pleno em PE.



Tabela 2 – Resultados LMM, região do Pronome e do Verbo em PE e em PB, durante a primeira leitura (FP).

	PE				PB			
	Est.	SE	t-value	p-value	Est.	SE	t-value	p-value
(Intercept)	0.529	8.409	0.063	0.950	-2.317	26.316	-0.088	0.930
Pronome	20.679	15.418	1.341	0.180	-325.541	76.439	-4.259	<0.001
Antecedente	-2.616	17.451	-0.150	0.881	57.341	59.307	0.967	0.334
Pronome:Antecedente	-50.479	28.132	-1.794	0.073	141.007	105.437	1.337	0.181

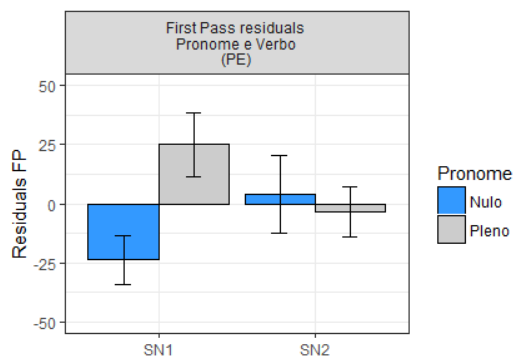


Gráfico 3 – Tempos residualizados em PE na região do pronome e do verbo.

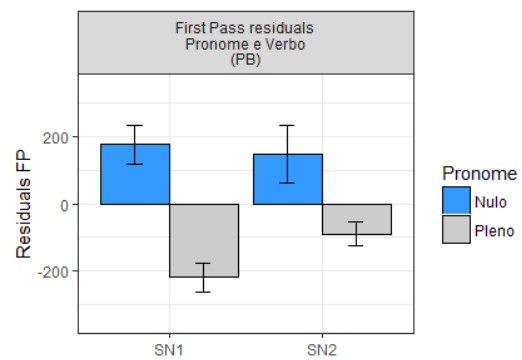


Gráfico 4 – Tempos residualizados em PB na região do pronome e do verbo.

Tabela 3 – Resultados LMM, Advérbio em PE e em PB (FP).

	PE				PB			
	Est.	SE	t-value	p-value	Est.	SE	t-value	p-value
(Intercept)	-0.263	7.730	-0.034	0.973	-0.374	9.447	-0.040	0.968
Pronome	10.436	21.664	0.482	0.630	-59.968	22.637	-2.649	0.008
Antecedente	0.099	21.724	0.005	0.996	15.380	17.694	0.869	0.385
Pronome:Antecedente	-31.099	30.917	-1.006	0.314	-67.666	26.782	-2.527	0.012



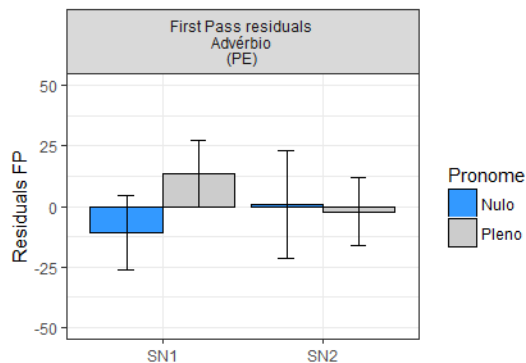


Gráfico 5 – Tempos residualizados em PE na região do advérbio.

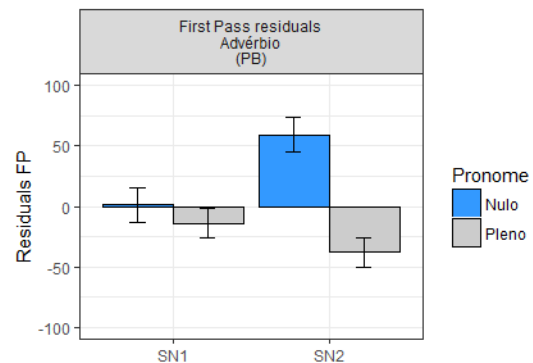


Gráfico 6 – Tempos residualizados em PB na região do advérbio.

Total Pass

Como se pode observar na Tabela 4, em PE regista-se um efeito principal de antecedente, com mais tempo no SN1, e um efeito de interação entre pronome e antecedente, com menos tempo na condição em que o pleno retoma o SN2 e mais quando retoma SN1 (Gráfico 7). Na região do advérbio não se registam efeitos significativos. Em PB, na região do pronome e do verbo (Tabela 4), pode-se verificar que existe um efeito principal de pronome, com mais tempo na condição com nulo (Gráfico 8), e um efeito principal de antecedente, com mais tempo na retoma de SN2. Na região do advérbio não se registam efeitos significativos.

No contraste entre as duas variedades, na região do pronome e do verbo, regista-se um efeito principal de pronome (Est.=-185.240; SE=46.151; t -value=-4.014; p -value=<0.001), com menos tempo na condição com pleno, uma interação entre pronome e antecedente (Est.=-173.661; SE=79.805; t -value=-2.176; p -value=0.030), com mais tempo na condição em que o nulo retoma o SN2, uma interação entre pronome e variedade (Est.=297.385; SE=79.857; t -value=3.724; p -value=<0.001), com mais tempo com nulo em PB, e também uma interação entre antecedente e variedade (Est.=-329.770; SE=79.813; t -value=-4.132; p -value<0.001), com mais tempo em SN2 em PB. Na região do advérbio regista-se uma interação entre antecedente e variedade (Est.=-85.458; SE=44.417; t -value=-1.924; p -value=0.054), com mais tempo em SN1 em PE e mais tempo em SN2 em PB.



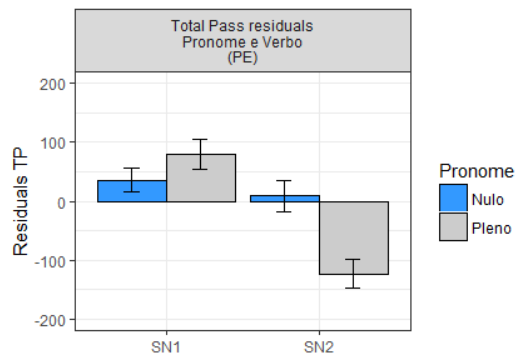


Gráfico 7 – Tempos residualizados em PE na região do Pronome e do Verbo (TP).

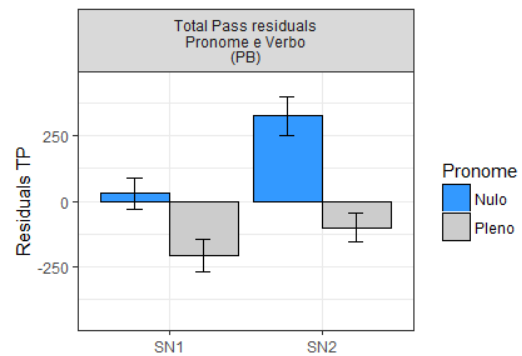


Gráfico 8 – Tempos residualizados em PB na região do Pronome e do Verbo (TP).

Tabela 4 - Resultados LMM, Pronome e Verbo em PE e em PB (TP).

	PE				PB			
	Est.	SE	t-value	p-value	Est.	SE	t-value	p-value
(Intercept)	0.612	15.389	0.040	0.968	0.132	37.534	0.004	0.997
Pronome	-42.494	37.181	-1.143	0.253	-338.760	89.199	-3.798	<0.001
Antecedente	-116.620	34.377	-3.392	0.001	208.738	75.833	2.753	0.006
Pronome:Antecedente	-176.242	61.569	-2.863	0.004	-173.086	150.290	-1.152	0.249

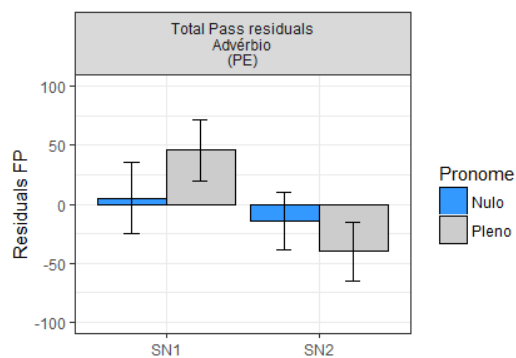


Gráfico 9 – Tempos residualizados em PE na região do advérbio (TP).

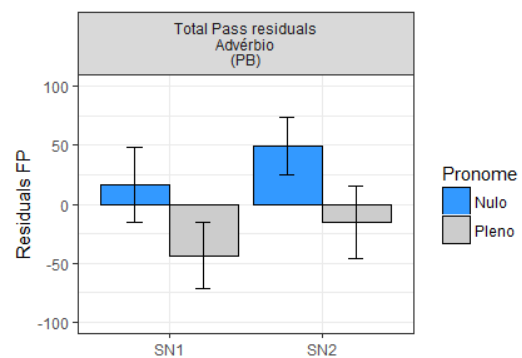


Gráfico 10 – Tempos residualizados em PB na região do advérbio (TP).



Tabela 5 - Resultados LMM, Advérbio em PE e em PB (TP).

	PE				PB			
	Est.	SE	t-value	p-value	Est.	SE	t-value	p-value
(Intercept)	1.509	31.867	0.047	0.962	0.342	14.033	0.024	0.981
Pronome	6.030	35.076	0.172	0.864	-69.544	46.371	-1.500	0.134
Antecedente	-52.811	30.938	-1.707	0.088	34.338	29.934	1.147	0.251
Pronome:Antecedente	-71.829	61.575	-1.167	0.243	-0.850	55.942	-0.015	0.988

2.3. Discussão dos resultados

Na medida *off-line* do presente estudo, ou seja, na taxa de acerto às respostas das perguntas finais realizadas, os resultados de cada uma das variedades, PE e PB, assim como da comparação entre ambas, foram muito semelhantes. De um modo geral, registaram-se mais erros na condição com nulo quando é retomado o SN2. O erro significa que os participantes escolheram o SN1 como resposta certa, ou seja, como antecedente do sujeito nulo, nas condições em que o nulo era forçadamente ligado ao SN2 pela informação de número. Parece-nos assim que estes resultados indicam uma penalização para as retomas de SN2 com a forma nula.

Este efeito é significativo e igualmente forte nas duas variedades, o que indica que em ambas as variedades parece haver um efeito penalizador para a retoma com a forma nula de um antecedente que não c-comanda o sujeito nulo. Esta penalização, flagrante nas duas variedades, pode ser explicada pela estrutura das frases testadas uma vez que criam uma relação mais forte entre as orações e, portanto, uma maior penalização nas condições em que a relação de c-comando não é cumprida.

Nas medidas de tempo, em PE, os efeitos não são muito claros, sobretudo na primeira leitura, ou seja, em FP, em que se regista apenas um efeito marginal ($p=0.073$) de interação entre pronome e antecedente. Em TP, ou seja, no tempo total de leitura, regista-se um efeito de antecedente e de interação entre pronome e antecedente que reflete sobretudo tempos mais rápidos na condição em que o pleno retoma o SN2. Assim, apesar de não ser imediatamente claro, a combinação entre estes dois resultados, com tempos mais baixos na retoma de SN1 com nulo na primeira leitura e tempos mais baixos na retoma de SN2 com pleno na leitura total da região, permite-nos concluir que, em PE, se regista uma distribuição tendencialmente complementar entre forma nula e plena, confirmando-se a primeira hipótese.

Considerando em conjunto os efeitos *on-line* e *off-line* de PE, parece ficar claro da análise global que existe uma distinção entre as formas e que a distribuição complementar entre nulos e plenos descrita para a retoma de sujeito e de objeto também se reflete em antecedentes com posições estruturais distintas internas ao mesmo constituinte⁹.

Em PB, na primeira leitura verifica-se, como nos resultados das respostas às perguntas de interpretação, uma penalização para a retoma de SN2 com nulo e, também, um aumento generalizado dos tempos nas

⁹ Como referido por um revisor anónimo do presente trabalho, a penalização é mais clara para a forma nula, o que se verifica também em outros trabalhos: de um modo geral, as preferências são geralmente mais claras e mais fortes, ou seja, menos variáveis, com as formas nulas; por exemplo, em condições interfrásicas (orações independentes) as preferências de retoma da forma plena são menos claras (as retomas de sujeito com pronome pleno crescem), mantendo-se as preferências da forma nula (Carminati, 2002; entre outros).



condições com nulo. No tempo total de leitura, em PB, o efeito de penalização da retoma com nulo mantém-se, registrando-se tempos mais elevados nas condições com nulo e em que é retomado o SN2, sendo os tempos elevados de retoma de SN2 sobretudo explicados pelos tempos das condições com nulos. Os resultados de PB, quer *off-line* quer *on-line*, permitem afirmar que, de um modo geral, em PB a forma plena é mais facilmente processada do que a forma nula, verificando-se ainda a existência de uma penalização quando a forma nula retoma o SN2. Estes resultados confirmam assim a nossa terceira hipótese, a de que em PB a forma nula é preferencialmente interpretada como retomando o antecedente que a c-comanda, sendo este resultado mais acentuado e claro (logo desde FP) do que em PE (efeito marginal em FP). Os resultados não nos permitem, no entanto, confirmar a nossa segunda hipótese, ou seja, de que seriam de esperar diferenças entre PE e PB quanto à forma plena.

Contudo, numa análise *post-hoc* às condições apenas com pleno, registamos um efeito estatístico marginal de antecedente (Est.= 127.09; SE= 67.356; *t*-value= 1.887; *p*-value= 0.059) em PB na FP na região do pronome e do verbo, com menos tempo na retoma de SN1. Na mesma análise aos dados de PE, não se regista nenhum efeito em FP (como nas análises anteriores), mas regista-se um efeito significativo de antecedente (Est.= -204.241; SE= 45.307; *t*-value= -4.508; *p*-value<0.001) em TP na região do pronome e do verbo, com tempos mais baixos na retoma de SN2. Esta análise parece assim indicar que, de facto, as preferências de interpretação da forma plena, também na retoma de antecedentes com diferentes posições estruturais, são diferentes em PE e em PB: em PE existe uma clara preferência pela retoma do SN2, a entidade menos proeminente, enquanto que em PB a preferência é pouco clara ou no sentido da retoma do SN1, o antecedente mais proeminente. Estes resultados permitem assim confirmar a nossa segunda hipótese e também que, de um modo geral, formas nulas e plenas têm comportamentos distintos em PE e em PB.

Na análise comparativa entre as duas variedades, os efeitos de PB são tão persistentes que são os efeitos que acabam por se refletir na diferença entre as duas variedades. Na região do pronome e do verbo, os resultados refletem essencialmente os efeitos de PB em que se regista um aumento de tempo nas condições com nulo e de retoma de SN2 com nulo, ou seja, regista-se, nas condições com nulo, um efeito penalizador generalizado. Os efeitos de interação entre pronome e variedade e antecedente e variedade refletem, respetivamente, a penalização em PB pela retoma pelo nulo e a preferência em PB por SN1, que referimos acima. Na região do advérbio regista-se uma diferença de tempos entre PE e PB, resultado de tempos mais baixos em PE na retoma de SN2 e em PB na retoma de SN1.

Os resultados do presente trabalho permitem concluir que a proeminência sintática em termos de hierarquia de constituintes, nomeadamente de c-comando, influencia e é um fator relevante na escolha do antecedente das expressões referenciais, no caso, nulas e plenas em posição de sujeito. Assim, consideramos que a posição que os antecedentes ocupam na estrutura da frase contribui para a definição da sua proeminência. Antecedentes em posições estruturais mais altas são mais proeminentes e, conseqüentemente, os antecedentes preferidos de formas mais reduzidas, como as formas nulas, como previsto pela Teoria da Acessibilidade de Ariel (1990). Em PE, a distribuição complementar entre as duas formas aparece como um efeito de interação, ainda que marginal, na primeira leitura da região, o que nos permite corroborar a nossa primeira hipótese, ainda que apenas para PE. Em PB esta distribuição não se verifica, ou não é tão clara, o que, na verdade, permite corroborar a nossa segunda hipótese, em que assumimos que a distinção entre nulo e pleno não seria tão forte em PB, uma vez que, de acordo com estudos anteriores (Luegi, 2012; Fernandes *et al.*, 2018), em PB a forma plena não apresenta, geralmente, uma preferência clara pelo antecedente a retomar. Assim, os resultados de PB parecem confirmar não só o que foi proposto por Duarte (1995) e Barbosa, Duarte e Kato (2005) sobre a improdutividade do Princípio Evitar Pronome em PB, mas, também, os resultados dos estudos (Luegi, Costa e Maia, 2014; Fernandes *et al.*, 2018) que demonstram que em PB não existe uma preferência clara por antecedentes com maior ou menor proeminência.



O que se destaca também nos nossos resultados são os tempos e também as taxas de erro nas condições com forma nula em PB, essencialmente nas condições em que nulo é forçadamente ligado ao SN2. Estes resultados, que confirmam a nossa terceira hipótese, estão de acordo com o descrito na literatura para o PB: as formas nulas são condicionadas essencialmente por restrições sintáticas, nomeadamente de relações de c-comando entre sujeito nulo e antecedente, como proposto por Barbosa, Duarte e Kato (2005), Holmberg, Nayudu e Sheehan (2009) ou I. Duarte e Figueiredo Silva (2016). Para além disso, os tempos elevados nas condições com forma nula e a ausência de preferência clara na interpretação da forma plena reforçam a proposta de que o PB tem vindo a perder a oposição entre formas nulas e plenas, estando a forma plena a aproximar-se do comportamento da forma nula.

3. Conclusões

No presente trabalho pretendeu-se avaliar a influência da hierarquia estrutural na resolução de formas nulas e plenas em posição de sujeito, verificando se a complementaridade esperada para formas nulas e plenas na retoma de antecedentes com diferentes funções sintáticas se estende a condições em que os antecedentes, apesar de não terem funções sintáticas claramente distintas, têm posições estruturais diferenciadas dentro do mesmo constituinte. Mais ainda, considerando as restrições que as estruturas selecionadas impõem, pretendemos comparar a resolução de formas nulas e plenas em posição de sujeito em orações completivas em PE e em PB prevendo que em PB a resolução de formas nulas seja mais restringida pelas relações de c-comando. Quanto às formas plenas, considerando também a literatura anterior sobre o tema, esperavam-se diferenças entre as duas variedades, prevendo-se uma preferência clara em PE ao contrário do que se espera para PB.

Os resultados do presente estudo permitem-nos confirmar as nossas três hipóteses. Primeiro, os resultados demonstram que a distribuição complementar entre formas nulas e plenas em posição de sujeito se estende à retoma de antecedentes em diferentes posições estruturais internas ao mesmo constituinte. Assim, para além da função sintática desempenhada pelos antecedentes, da sua ordem de referência, do seu papel temático, entre outros, também a posição estrutural que os antecedentes ocupam desempenha um papel relevante na resolução de expressões referenciais. Este resultado enquadra-se na proposta Multi-Fatorial de Kaiser e Trueswell (2008), uma proposta em que se assume que diferentes fatores contribuem para a atribuição de proeminência a uma entidade discursiva.

Em segundo lugar, os nossos resultados permitem confirmar o que vem sendo demonstrado por estudos anteriores: em PB a forma plena não apresenta, ao contrário do que acontece em PE e na maioria das línguas de sujeito nulo consistente, uma preferência pela entidade menos proeminente. Neste estudo, corroboramos essa perspetiva na análise da posição estrutural e não apenas da função sintática, como avaliado nos trabalhos anteriormente realizados. Esta extensão à análise da posição estrutural, parece-nos, reforça a ideia de que, de facto, em PB a forma plena se tem vindo a aproximar da forma nula quer em termos de produção, como verificado em Duarte (1995), quer em termos de compreensão, como confirmado em Luegi (2012) e Fernandes *et al.*, (2018).

Por fim, os resultados permitem ainda reforçar as propostas que consideram que em PB a forma nula é essencialmente condicionada por restrições de c-comando, sendo preferencialmente interpretada como retomando antecedentes que a c-comandem.

O nosso estudo contribui assim não só para a identificação da posição estrutural enquanto um dos diferentes fatores que contribuem para a atribuição de proeminência a uma entidade discursiva, mas, também, para a análise comparativa entre PE e PB. Assim, o efeito da posição estrutural, à semelhança do efeito da função sintática, tendo impacto na resolução de expressões referenciais, nomeadamente, de formas nulas e plenas em posição de sujeito, contribui para a sua distribuição (quase) complementar: formas nulas são



preferencialmente interpretadas como retomando o SN estruturalmente mais alto, tanto em PE como em PB, e formas plenas como retomando o SN estruturalmente mais baixo, sobretudo em PE.

Quanto ao contraste entre PE e PB, o nosso estudo dá indícios claros de que as duas variedades se distinguem não só quanto à resolução de formas plenas em posição de sujeito, mas, também, de formas nulas. Assim, na sequência de trabalhos anteriores, o presente estudo confirma que PE e PB se distinguem quanto à assimetria entre forma nula e plena também na retoma de potenciais antecedentes dentro do mesmo constituinte sintático, mas com posições estruturais distintas. Enquanto em PE a forma nula tende a ser interpretada como retomando a entidade estruturalmente mais proeminente, a forma plena é preferencialmente interpretada como retomando a entidade menos proeminente, replicando a assimetria nulo-sujeito e pleno-objeto. Em PB, apesar de se manter a preferência de interpretação para a forma nula, a forma plena não registou uma preferência clara, existindo, no entanto, uma preferência marginal pela retoma do antecedente estruturalmente mais proeminente, o SN1. Este resultado parece indicar que, em PB, se confirma a perda de princípio evitar pronome e, conseqüentemente, a complementaridade entre as duas expressões referenciais estando, aparentemente, a forma plena a ocupar a função designativa da forma nula.

Mais ainda, quanto à forma nula, os nossos resultados indicam que esta, como proposto em diferentes trabalhos, é essencialmente limitada em termos de ocorrência e de resolução pela relação de c-comando que estabelece com o seu antecedente: a forma nula é preferencialmente interpretada como retomando o antecedente que a c-comanda, sendo este resultado mais acentuado em PB nos dados *on-line*. Estes resultados contribuem assim para a discussão não só sobre a classificação de PB enquanto língua de sujeito nulo parcial vs. PE enquanto língua de sujeito nulo consistente (Barbosa, Duarte e Kato, 2005), mas, também, possivelmente, sobre a classificação da categoria de sujeito nulo. No entanto, não está no âmbito do presente trabalho essa discussão, apesar de considerarmos que os resultados podem contribuir para investigações futuras nessa área.

Por fim, consideramos que seria interessante, em trabalhos futuros, testar a retoma de entidades em posições estruturais distintas, mas, desta vez, em posição de objeto, uma função sintática menos proeminente.

Referências

- Alonso-Ovalle, L., Clifton, C., Frazier, L., & Solera, S. F. (2002) Null vs. Overt Pronouns and The Topic-Focus Articulation in Spanish. *Journal of Italian Linguistics*, 14:2, pp. 151-169.
- Ariel, M. (1990). *Assessing Noun-Phrase Antecedents*. Routledge.
- Barbosa, P. (2011) Partial pro-drop as null NP anaphora. *Glsa Publications*.
- Barbosa, P., Duarte, E., & Kato, M. (2005) Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4:2, pp. 11-52.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3), 10.1016/j.jml.2012.11.001. doi:10.1016/j.jml.2012.11.001
- Bates, D, Maechler, M, Bolker, B., & Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48, doi:10.18637/jss.v067.i01
- Carminati, M. (2002) *The processing of Italian subject pronouns*. Electronic Doctoral Dissertations for UMass Amherst.
- Corrêa, L. (1998) Acessibilidade, paralelismo na interpretação do pronome sujeito e o contraste pro/pronome em português. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 24, pp. 295-392.
- Costa, M. A. (2005) *Processamento de frases em Português Europeu. Aspectos cognitivos e linguísticos implicados na compreensão da língua escrita*. Lisboa: FCG [ed. tese de doutoramento de 2003]



- Costa, A., Faria, I. H., & Matos, G. (1998) Ambiguidade referencial na identificação do Sujeito em estruturas coordenadas. *Textos seleccionados do XIII Encontro Nacional da Associação Portuguesa de Linguística* (pp. 173-188). Lisboa: Colibri.
- Costa, A., Faria, I., & Kail, M. (2004) Semantic and Syntactic Cues' Interaction on Pronoun Resolution in European Portuguese. In Branco, McEnery, & Mitkov (Ed.), *DAARC 2004, 5th Discourse Anaphora Resolution Colloquium* (pp. 45-50). Lisboa: Colibri.
- Duarte, E. (1995). *A perda do princípio "evite pronome" no português brasileiro*. Tese de Doutorado, Universidade Estadual de Campinas.
- Duarte, I., & Figueiredo Silva, M. C. (2016) The Null Subject Parameter and the Structure of the Sentence in European and Brazilian Portuguese. In Wetzels, Menuzzi & Costa (eds.), *Handbook of Portuguese Linguistics*. Wiley-Blackwell.
- Fernandes, E.G., Luegi, P., Correa Soares, E., de la Fuente, I., & Hemforth, B. (2018) Adaptation in pronoun resolution: Evidence from Brazilian and European Portuguese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol 44(12), pp. 1986-2008.
- Figueiredo Silva, M. C. (2000) Main and Embedded Null Subjects in Brazilian Portuguese. In M. Kato, & E. Negrão, *Brazilian Portuguese and the Null Subject Parameter* (pp. 127-146). Vervuert – Iberoamericana.
- Gernsbacher, M.A., & Hargreaves, D.J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, 27, pp. 699–717.
- Holmberg, A., Nayudu, A., & Sheehan, M. (2009) Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, 63: 59-97. doi:10.1111/j.1467-9582.2008.01154.x
- Järvikivi, J., Van Gompel, R. P., Hyönä, J., & Bertram, R. (2005) Ambiguous pronoun resolution: Contrasting the rst-mention and subject-preference accounts. *Psychological Science*, 16(4), 260-264. doi: 10.1111/j.0956-7976.2005.01525.x
- Kaiser, E., & Trueswell, J. (2008) Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, 23(5), 709-748. doi: 10.1080/01690960701771220
- Luegi, P. (2012) *Processamento de sujeitos pronominais em Português: efeito da posição estrutural dos antecedentes*. Tese de Doutorado. Universidade de Lisboa.
- Luegi, P., Costa, A., & Maia, M. (2014) Processamento e interpretação de sujeitos nulos e plenos em Português Europeu e em Português do Brasil. *Cadernos de Letras da UFF* (49), pp. 67-88.
- Modesto, M. (2000) Null Subjects without 'Rich' Agreement. In M. Kato, & E. Negrão, *Brazilian Portuguese and the Null Subject Parameter* (pp. 147-176). Vervuert – Iberoamericana.
- Morgado, S. (2012) *Processamento da co-referência pronominal: informação sintáctica e semântica*. Tese de Mestrado apresentada à Faculdade de Letras da Universidade de Lisboa.
- R Core Team (2014) R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL <http://www.R-project.org/>.
- Raposo, E., Nascimento, M., Mota, M. A., Seguro, L., & Mendes, A. (2013) *Gramática do Português*. Fundação Calouste Gulbenkian.
- Townsend, D. J., & Bever, T. G. (2001) *Language speech and communication. Sentence comprehension: The integration of habits and rules*. Cambridge, MA, US: The MIT Press.
- Trueswell, J.C., Tanenhaus, M.K., Garnsey, S.M. (1994) Semantic Influences On Parsing: Use of Thematic Role Information in Syntactic Ambiguity Resolution, *Journal of Memory and Language*, Volume 33(3), 285-318, doi.org/10.1006/jmla.1994.1014.

