

Alternâncias linguísticas na comunidade portuguesa residente em Andorra: um estudo sociolinguístico

Language alternations in the Portuguese community living in Andorra: a sociolinguistic study

Joana Aguiar¹, Carolina Bastida Serra²

¹ Centro de Investigação Transdisciplinar em Educação e Desenvolvimento (CITeD) – Instituto Politécnico de Bragança

² Grup de Recerca en Llengües –Universitat d'Andorra

Abstract

According to the Department of Statistics of Andorra (2025), there are 8510 Portuguese people living in Andorra, which represents 9.7% of the population. The objective of this paper is to characterise the language alternations present in the oral speech of informants from the Portuguese community in Andorra. More specifically, based on Poplack's (2020) proposal, our aim is to verify how and in what contexts language alternations (Portuguese-Catalan/Portuguese-Spanish) are present in the oral discourse of speakers who have Portuguese as their native language and as a heritage language. In order to study the influence of Catalan and Spanish on the oral discourse of Portuguese residents in Andorra, we interviewed 18 native speakers of Portuguese living in Andorra for more than four years and 9 Portuguese heritage language speakers. The results show that the social variables *schooling* and *years of residence in Andorra* are determining factors in explaining the frequency of occurrence of linguistic alternations and their characterisation. The role of other variables, such as language use in work contexts, in the patterns of alternation displayed by speakers is also discussed.

Keywords: Portuguese community in Andorra, grammatical convergence, code-switching, lexical borrowing, language attitudes.

Resumo

De acordo com o Departament d'Estadística d'Andorra (2025), residem, em Andorra, 8510 portugueses, o que constitui 9,7% da população. O objetivo deste trabalho é caracterizar as alternâncias linguísticas presentes no discurso de informantes da comunidade portuguesa em Andorra. Mais especificamente, a partir da proposta de Poplack (2020), pretende-se verificar de que forma, e em que contextos, as alternâncias linguísticas (português-catalão/português-espanhol) estão presentes no discurso oral dos falantes que têm o português como língua materna (PLM) e como língua de herança (PLH). De forma a estudar a influência do catalão e do espanhol no discurso oral dos portugueses residentes em Andorra, foram entrevistados 18 falantes de português língua materna, residentes há mais de quatro anos em Andorra, e 9 falantes de português língua de herança. Os resultados mostram que, no caso dos falantes de PLM, a variável *escolaridade* e o tempo de residência em Andorra são fatores determinantes para explicar a frequência de ocorrência das alternâncias linguísticas e a sua caracterização. É ainda discutido o papel de outras variáveis, como o uso das línguas em contexto de trabalho nos padrões de alternância exibidos pelos falantes.

Palavras-chave: comunidade portuguesa em Andorra, convergência gramatical, code-switching, empréstimo lexical, atitudes linguísticas.

1. Introdução

Este trabalho apresenta o resultado de um projeto de investigação, realizado no âmbito da Cátedra Camões da Universitat d'Andorra (AD03 - UdA-2023/2024), sobre a língua portuguesa em Andorra e as atitudes linguísticas da comunidade portuguesa. A recolha de amostras de fala foi realizada através de entrevistas informais semiestruturadas a falantes de português língua materna (PLM) e de português língua de herança (PLH). Os objetivos gerais deste projeto são: i) contribuir para o estudo da diversidade linguística em Andorra, com enfoque na comunidade portuguesa; ii) identificar, caracterizar e analisar as instâncias de alternância linguística; iii) identificar eventuais processos de variação e mudança; iv) analisar as atitudes linguísticas em relação ao português e ao catalão. Neste artigo, debruçar-nos-emos sobre as instâncias de alternância linguística no discurso oral da comunidade portuguesa e afluiremos as atitudes linguísticas em relação ao português e ao catalão.

Este artigo está organizado da seguinte forma: na secção 2, é feita uma descrição da evolução sociodemográfica da comunidade portuguesa em Andorra; na secção 3, é feita uma breve descrição das línguas em contacto e dos seus usos sociais. É, também, feita referência aos sistemas educativos vigentes em Andorra e às implicações que estes têm na manutenção do multilinguismo. A secção 4 é dedicada à caracterização das alternâncias linguísticas segundo a proposta de Poplack (2020) e na secção 5 é explanada a metodologia adotada. As secções 6 e 7 são dedicadas à apresentação e discussão dos resultados, respetivamente.

2. A comunidade portuguesa em Andorra

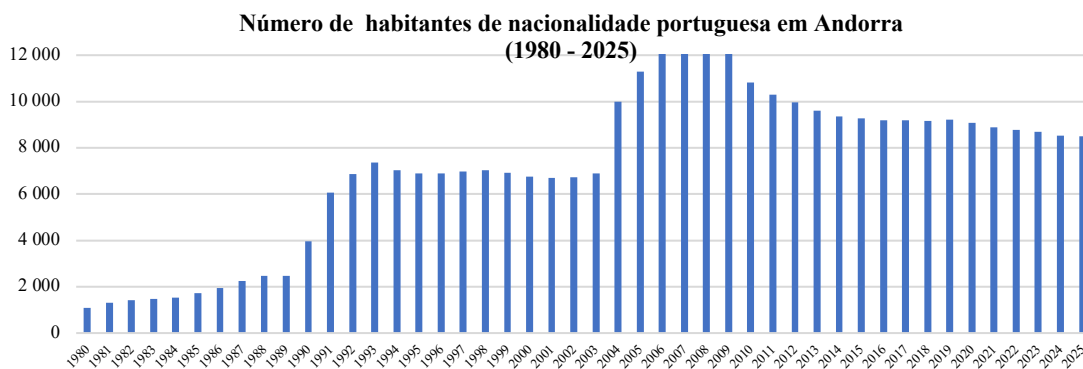
De acordo com informação do Departament d'Estadística do Govern d'Andorra, residem, em Andorra, 8510 portugueses, o que constitui 9,7% da população¹. A maioria da população portuguesa residente em território andorrano concentra-se nas paróquias de Encamp, Andorra La Vella e Sant Julià de Lòria (Departament d'Estadística d'Andorra, 2025).

Apesar de o estabelecimento de portugueses no Principado ter tido início em XX (Carvalho, 2007), só a partir da década de 80 do século XX é que a presença da comunidade portuguesa começou a ser expressiva. Nesse sentido, de acordo com a tabela relativa à evolução do número de portugueses em Andorra (Figura 1), podemos identificar três principais movimentos migratórios: (i) o início dos anos 80, (ii) o início dos anos 90 e (iii) o início do século XXI (entre 2003 e 2006).

¹ Dados de fevereiro 2025 (Departament d'Estadística d'Andorra, 2025).

Figura 1

Comunidade Portuguesa em Andorra em Número de Habitantes de 1980 a 2025



Nota. Fontes: Carvalho (2007); Departament d’Estadística d’Andorra (2025); Matias (2008); Observatório da Emigração (2010). Elaboração própria.

De uma forma geral, a comunidade portuguesa é relativamente homogénea no que diz respeito à região de origem. A maioria é proveniente do Alto Minho e Trás-os-Montes (Matias, 2008). A escolha de Andorra como país de acolhimento resulta de um processo de “migração em cadeia” (Matias, 2008), motivado por familiares ou amigos já estabelecidos em Andorra.

Na década de 80, os imigrantes portugueses em Andorra eram provenientes de zonas rurais (Carvalho, 2007) e trabalhavam principalmente em trabalhos pouco qualificados, nomeadamente na construção e em serviços domésticos ou hotelaria. A comunidade portuguesa a residir em Andorra nessa época é descrita por Comas e Pujadas (1999) como a menos qualificada, razão pela qual ocupava a posição mais baixa na escala social: “As the newest and the least qualified group they occupy the lowest position on the social scale: the men mainly work as day labourers in agriculture and construction; the women are mainly shop assistants or cleaners” (Comas & Pujadas, 1999, p. 260).

A partir do início do século XXI, a caracterização sociodemográfica dos portugueses que emigram para Andorra sofre alterações: são mais jovens e provenientes de centros urbanos (Matias, 2008). Em termos de escolaridade, terminaram o ensino secundário ou um curso superior em Portugal, facto que facilitou a sua integração na sociedade andorrana (Carvalho, 2007).

3. Línguas em contacto

Andorra é um território pequeno, com uma população reduzida, onde, devido aos sucessivos fluxos migratórios e sazonais, confluem comunidades e línguas distintas. A partir da segunda metade do século XX, assiste-se a um crescimento de migrantes, provenientes principalmente de Espanha e de Portugal. Em resultado das sucessivas vagas migratórias, o perfil linguístico da população adulta residente em Andorra é diverso (Govern d’Andorra, 2023): 44,1% da população adulta é falante nativo de catalão, 40,3% de espanhol e 13,5% de portugueses.

A língua oficial no Principado de Andorra é o catalão. Apesar de o espanhol ser comumente usado, o catalão é a língua das instituições e dos serviços públicos. Não obstante os esforços para a promoção da aprendizagem e do uso do catalão, nomeadamente a disponibilização de cursos de língua catalã gratuitos, a publicação da *Llei d’ordenació de l’ús de la llengua oficial* (Govern d’Andorra, 2000) e a recente aprovação da

*Llei 6/2024, de la llengua pròpia i oficial*² (Govern d'Andorra, 2024), há fatores que comprometem o uso da língua oficial. Por um lado, podemos identificar dois fatores demográficos relevantes, como explanaremos de seguida. Por outro lado, ao contrário do que observamos noutros países, como a Alemanha, por exemplo, o uso de outras línguas, como o espanhol ou o português, em detrimento da língua oficial, não é discriminatório (Ballarín, 2006). A proximidade geográfica de Andorra com a Espanha e a França permitiu a coexistência do catalão com o francês e, principalmente, do catalão com o espanhol.

Relativamente aos fatores demográficos, 54,9% dos residentes em Andorra não têm nacionalidade andorrana. Adicionalmente, verifica-se uma grande flutuação de residentes devido ao trabalho sazonal (Bastida & Nicolau, 2012) na denominada “temporada da neve”. Esta necessidade de mão de obra é maioritariamente preenchida por falantes nativos de espanhol, provenientes da América do Sul, principalmente da Argentina (Matias, 2008). Concorre, ainda, para o uso frequente do espanhol o facto de 42,6% dos habitantes de Andorra terem nacionalidade espanhola ou serem provenientes de países onde se fala espanhol (Departament d'Estadística d'Andorra, 2025). Para além das comunidades provenientes de Espanha e Portugal, residem em Andorra 3.920 franceses (4,5%) e 14.801 cidadãos de outras nacionalidades (16,9%), provenientes maioritariamente de países da América do Sul. O contacto entre línguas, entendido como a situação em que os membros de uma comunidade de fala acedem regularmente a duas ou mais línguas no decorrer das suas interações linguísticas (Poplack, 2020, p. 47), é evidente. O país caracteriza-se, assim, pelo seu multilinguismo individual e social (Hawkey & Horner, 2022). Por esta razão, no entender de Comas e Pujadas (1999), o uso do espanhol em Andorra não tem, para o falante, um valor identitário ou simbólico, mas um carácter instrumental: “[Spanish] has an instrumental rather than a symbolic or identifying value” (Comas & Pujadas, 1999, p. 260). Neste sentido, é possível afirmar que o espanhol se assemelha a uma língua franca (Hawkey & Horner, 2022), utilizada não só em trocas comerciais por agentes de turismo e turistas, mas também pela maioria dos residentes não andorranos. O espanhol é, também, comumente utilizado por portugueses, principalmente por aqueles que (ainda) não aprenderam catalão. A este propósito, convém salientar que a maioria dos portugueses residentes em Andorra provém de regiões raianas em Portugal (Matias, 2008). Como tal, o contacto com o espanhol ou o galego era frequente antes de residirem em Andorra. Muitos não aprenderam formalmente o espanhol e falam uma mistura de português com espanhol, um linguajar informalmente denominado de *portunhol*, como refere o informante 14M3SE³ (1):

(1) Mas, claro, vir aqui e começar a trabalhar e tudo isso foi uma aventura. E eu, talvez dos poucos que depois nas entrevistas vai ver, que primeiro adaptei-me à Andorra, cursos de catalão, participar em associações andorranas, ou seja, integração total, praticamente total. **O que eu acho é que muita gente acomodou-se ao castelhano, e eu vou dizer um exagero, é o portunhol, muitas pessoas com quem faleste encontrei isso, gente que mistura o castelhano com o português e alguma palavra em catalão**, se é que se lembra, mas muito pouco. [14M3SE]

O inquérito ao conhecimento e usos linguísticos de 2023 (Govern d'Andorra, 2023) revela que o catalão é atualmente a língua materna da maioria da população residente em Andorra. O mesmo inquérito revela que 63,7% dos inquiridos referem que o catalão é a língua que habitualmente usam para comunicar com outras pessoas. Nos setores da construção e do comércio, há uma maior presença de espanhóis e portugueses, o que permite que parte da população empregada nestes segmentos profissionais comunique principalmente na sua língua materna (português ou espanhol) (Margarit et al., 2010). Adicionalmente, o português é a língua de comunicação em contextos familiares em que pelo menos um dos progenitores é português. Contribuem, ainda, para o uso do português a existência de associações culturais e recreativas, com uma presença ativa na sociedade

² A *Llei 6/2024* prevê a exigência de um conhecimento mínimo de catalão para pedidos e renovações da autorização de residência, assim como a garantia de um serviço público em catalão.

³ Para mais informações relativamente ao código alfanumérico adotado para codificação dos informantes, veja-se a secção 5.2.

andorrana, como o Grupo de Folclore Casa de Portugal e a Associação de Portugueses do Alto Minho. São frequentes as exposições e os eventos culturais e etnográficos alusivos à cultura e tradições portuguesas (o dia de S. Martinho, o cantar das janeiras, o mercado tradicional, entre outros). Estas manifestações culturais promovem o contacto entre falantes e a perpetuação da cultura portuguesa e da identificação coletiva da comunidade portuguesa em Andorra.

Relativamente ao conhecimento explícito das línguas, o inquérito ao conhecimento e usos linguísticos de 2023 (Govern d'Andorra, 2023) indica que, apesar de o conhecimento explícito do catalão ter aumentado, a maioria dos entrevistados considera ter mais competências linguísticas a espanhol do que a catalão. Os jovens residentes em Andorra, entre os 16 e os 20 anos, são predominantemente bilingues (espanhol e catalão), havendo, no entanto, casos de monolinguismo e plurilinguismo (Margarit et al., 2010). Os autores referem, ainda, que, apesar de, em contexto familiar, predominar o monolinguismo, é comum o uso de duas línguas para comunicar na escola ou com os amigos. O inglês é a língua que os adolescentes dominam menos, o que se deve à sua introdução tardia no currículo escolar (Margarit et al., 2010).

O multilinguismo de Andorra está patente no sistema educativo público andorrano, que compreende três tipologias de ensino, estreitamente associadas à língua adotada para veicular o conhecimento: o andorrano, o francês e o espanhol⁴. Apesar de existirem acordos específicos com Espanha e com França, todos os sistemas são regidos pela *Llei qualificada d'educació* (de 3 setembro de 1993) e pela *Llei d'ordenament del sistema educatiu andorrà* (de 09 de junho de 1994). A escolha do sistema de ensino condiciona, por isso, não só a exposição formal ao catalão, mas também os contextos de uso da língua (Bastida & Nicolau, 2012). No sistema espanhol, o ensino é feito em espanhol. No sistema andorrano, a lecionação é feita em catalão. As escolas do sistema francês são financiadas pelo governo francês, seguem o *curriculum* elaborado pelo Ministério da Educação francês e o ensino é feito em francês. No sistema francês, a disciplina de português está disponível como optativa, desde o ensino primário até ao *lycée*, e é lecionada após as restantes atividades letivas ou durante o intervalo letivo para almoço.

A opção por um sistema em detrimento dos demais prende-se principalmente com fatores familiares e geográficos (Matias, 2008). No caso dos filhos de portugueses que frequentam ou frequentaram a escola andorrana, Matias (2008) considera que fatores como uma maior identificação cultural e o desejo de uma maior integração na sociedade andorrana poderão ser relevantes para explicar esta escolha. O trabalho de Hawkey (2024) aponta, ainda, como fator determinante para a escolha do sistema educativo a política linguística familiar. Segundo o autor, as famílias que privilegiam a aprendizagem e uso do português matriculam os filhos no sistema francês.

4. A alternância linguística

A alternância de códigos linguísticos diz respeito ao uso de duas ou mais línguas, pelo mesmo falante, no mesmo ato comunicativo. A alternância linguística corresponde a uma estratégia comunicativa, com motivações diversas e cuja extensão é variável, do ponto de vista sintático. Concorrem para as alternâncias entre duas línguas, entre outros, os seguintes aspetos (Poplack, 1980):

- a idade e o nível de escolaridade;
- a forma e o *timing* de aquisição das línguas envolvidas: idade de aquisição, período de exposição à(s) língua(s), contacto com a(s) comunidade(s) de fala e tipologia das redes sociais às quais os intervenientes pertencem e com as quais se relacionam;
- a identidade étnica, a profissão e o local de trabalho.

Do ponto de vista linguístico, as alternâncias linguísticas permitem que o locutor molde a sua interação de acordo com o contexto, o tópico e o interlocutor e possibilitam o preenchimento de possíveis lacunas

⁴ Em Andorra, existem, também, escolas privadas (por exemplo, o *British College of Andorra* e a *Agora International School*) em que o ensino é feito em inglês.

gramaticais ou de vocabulário (Flores, 2004). É também comum o uso de outra língua “quando esta dispõe de componentes mais económicas para expressar um dado conteúdo verbal” (Flores, 2004). O uso de outra língua permite, ainda, enfatizar ou clarificar o que é dito (Androutsopoulos, 2005) e substituir expressões vernáculas e calão.

Neste trabalho, é adotada a proposta de Poplack (2020) para análise das alternâncias linguísticas. Baseada nos pressupostos metodológicos da linguística variacionista (Labov, 1966, 1972), a autora propõe uma metodologia de análise das *manifestações linguísticas decorrentes do contacto linguístico* (Poplack, 2020, p. 47). Assim, dependendo da sua extensão e das suas propriedades morfossintáticas exibidas, é possível distinguir as seguintes estratégias: (i) o empréstimo lexical (*lexical borrowing*), (ii) o code-switching (*multi-word intrasentential code-switching*) e (iii) a convergência gramatical (*grammatical convergence*), que explanaremos de seguida.

O empréstimo lexical consiste na incorporação de palavras da língua doadora (L_D), o espanhol ou o catalão, no discurso da língua recipiente (L_R), o português. Prevalece a morfologia e a sintaxe da L_R .

Considerem-se os exemplos (2) e (3):

(2) Para mim é uma das línguas mais feias do mundo. Prefiro o **chino**. [24F3SU]

(3) [16F2SE] Nesta ponte, estes são os túneis novos que fizeram que ligam, digamos, Encamp a Massana.

Quando estavam a fazer, aí **sim que** houve um grande acidente, morreram vários portugueses.

[Entrevistadora] Mas caíram pedras, foi?

[16F2SE] Não, derramou-se. Estavam a pôr **hormigón** e tudo isso, algo não estava bem feito [16F2SE]

Nos exemplos (2) e (3), os termos do espanhol *chino*⁵ e *hormigón*⁶ são inseridos no interior de uma frase em português, antecedida de outra, também produzida em português. Não há alteração da morfologia nem da sintaxe do português. Em “aí **sim que** houve um grande acidente” (exemplo 3) podemos identificar o uso da locução adverbial do catalão e do espanhol *sí que*, usada coloquialmente para designar *certamente, de verdade*⁷.

O termo *code-switching* tem sido utilizado por diferentes autores para designar, *lato sensu*, a justaposição, dentro da mesma interação discursiva, de elementos pertencentes a dois (sub)sistemas gramaticais distintos (Gumperz, 1982). De acordo com Poplack (2020), o termo *code-switching* designa a alternância de línguas ou variedades em contacto dentro da mesma frase. Normalmente, a extensão é superior à palavra (no original, *multi-word intrasentential code-switching*). No *code-switching*, nenhuma das línguas é tida como a recipiente, uma vez que se verifica a manutenção da morfologia e da sintaxe das línguas envolvidas, como ilustram os exemplos (4) e (5).

(4) Claro que **estás bien, has descansado un rato**. [6F2SE]

(5) **Y porque hay un chico** que está de manhã e ele faz o turno de manhã, mas precisavam alguém para fazer o turno de tarde. [28F1SE]

Note-se que, ainda que alternâncias de itens lexicais isolados sejam tendencialmente etiquetadas como *empréstimo lexical*, a diferença substancial entre *empréstimo lexical* e *code-switching* não é a extensão da alternância, mas a identidade da gramática subjacente a estas duas estratégias (Poplack, 2020).

O termo *convergência gramatical* designa as alternâncias linguísticas em que se observa a alteração da gramática da L_R por influência da L_D . Nesta tipologia, incluem-se propriedades ou processos que diferem, quer

⁵ Em português, *chinês*.

⁶ Em português, *betão armado*.

⁷ Real Academia Española (s.d.).

pela presença quer pela ausência, dos observados na variedade não exposta ao contacto e que emergiram em resultado do contacto linguístico com a L_D (Poplack, 2020, p. 47). Observem-se os seguintes exemplos:

(6) Fui para Barcelona, comecei a estudar tradução e interpretação, com base no francês e no russo, e depois de um tempo descobri que não era aquilo que eu gostava e que queria fazer, com o qual depois mudei **a** Humanidades. [21M2SU]

(7) Os apartamentos estão caríssimos. (...) E depois, **te fica** muito pouco para comer, ou para te vestires, ou para.... [24F3SU]

No exemplo (6), o sintagma preposicional *a Humanidades* é introduzido pela preposição *a*. Em português, o verbo *mudar*, nos contextos em que designa alteração da área científica ou de curso superior, selecciona a preposição *para*, ao passo que, em espanhol e catalão, a preposição seleccionada pelos verbos *cambiar* e *canviar*, respetivamente, é *a*: *cambiar a Humanidades/ canviar a Humanitats*⁸.

No exemplo (7), verificamos a anteposição do clítico (**te** fica). Em português europeu, na ausência de atratores de próclise ou estruturas sintáticas que permitam ou exijam a próclise, o clítico ocorre em ênclise, ou seja, a estrutura seria: *E depois, fica-te muito pouco para comer*. Ao invés do que ocorre em português europeu, em espanhol e catalão, é privilegiada a próclise. Estes dois exemplos ilustram, assim, a influência do espanhol e do catalão (as línguas *doadoras*) no português (a língua *recipiente*).

5. Metodologia

Nesta secção, é descrita a metodologia adotada para a recolha e tratamento de dados linguísticos. É, ainda, apresentado o perfil sociolinguístico dos informantes entrevistados.

5.1 A entrevista

De forma a verificarmos a influência do catalão e do espanhol no discurso oral dos portugueses residentes em Andorra, foram recolhidas amostras de fala, em formato de entrevista informal semiestruturada. As entrevistas foram conduzidas por uma falante nativa de português europeu. Apesar de ter sido desenhado um protocolo de entrevista semiestruturada, de forma a uniformizar, sempre que possível, os temas em torno dos quais se desenvolvia a entrevista, pretendeu-se que o discurso fosse o mais espontâneo e informal possível, tendo o curso da interação sido adaptado ao entrevistado. Procurou-se, assim, desviar a atenção do informante para outros aspetos que não o processo comunicativo, de forma a facilitar a produção espontânea de formas vernáculas utilizadas no dia a dia (Labov, 2001; Tagliamonte, 2012). A análise da variação linguística a partir do modelo variacionista pode ajudar a esclarecer alguns aspetos relacionados não só com o funcionamento da linguagem, mas também com a forma como a linguagem está representada mentalmente e socialmente (Henry, 2002). Os temas principais foram a história de vida do entrevistado e as vivências em Andorra, assim como o uso do português em contextos familiares, sociais e laborais. As entrevistas foram realizadas em locais escolhidos pelos entrevistados, minimizando o impacto do momento da entrevista na vida quotidiana do informante. Frequentemente, as entrevistas decorreram em esplanadas, bancos de jardim, convívios associativos ou em deslocações de ou para o local de trabalho. Todos os informantes assinaram uma declaração de consentimento autorizando o uso da entrevista para fins científicos. No caso dos informantes menores, a declaração de consentimento foi assinada pelo encarregado de educação ou tutor legal.

Para a seleção de informantes e posterior entrevista foi determinante o estabelecimento de uma rede de contactos, quer institucionais (o Consulado Honorário de Portugal em Andorra, o Instituto Camões, a

⁸ De referir que também os verbos *pasar* (espanhol) e *passar* (catalão) seleccionam a preposição *a* quando designam mudança de área científica ou percurso académico: *pasarse a Humanidades* (espanhol) / *passar-se a Humanitats* (catalão).

Universitat d’Andorra) quer informais (a Associação de Portugueses do Alto Minho e o Grupo de Folclore Casa de Portugal). Os amigos de infância e familiares de antigos alunos foram, também, determinantes para o contacto com a comunidade portuguesa em Andorra. As entrevistas foram realizadas entre fevereiro e junho de 2024, durante duas deslocações ao Principado de Andorra.

5.2 Os informantes

As 27 entrevistas foram realizadas seguindo os pressupostos da Sociolinguística Variacionista. A amostra está estratificada de acordo com as seguintes variáveis sociais: escolaridade (até 12 anos de escolaridade/ ensino superior), idade (entre 16-25, 26-44 e >44) e género. O tempo de permanência em Andorra, expresso em anos, foi, também, considerado para efeitos de tratamento estatístico e posterior análise quantitativa. Para além desses dados, foram, ainda, registados os seguintes elementos: o concelho e o distrito de origem dos informantes, a atividade profissional exercida e os contextos de uso das línguas.

Tabela 1

Número de Informantes Entrevistados por Variável Social (Escolaridade, Intervalo Etário, Género)

Escolaridade	Até ao 12.º ano						Licenciatura			
Idade	16–25		26–44		>44		26–44		>44	
Género	F	M	F	M	F	M	F	M	F	M
Informantes PLM	--	--	3	3	3	3	3	--	3	--
Informantes PLH	3	3	--	--	--	--	--	3	--	--

Os informantes entrevistados podem ser agrupados em duas categorias. Por um lado, temos 18 falantes nascidos em Portugal, cuja língua materna é o português, e que emigraram para Andorra depois dos 16 anos. Quando se lhes perguntou se haviam aprendido catalão em contextos de ensino formal, 9 falantes de português língua materna responderam que sim e 9 informantes responderam que não. Os falantes que aprenderam catalão trabalham em serviços públicos e serviços administrativos ou técnicos em empresas privadas. Quatro são do sexo masculino e cinco do sexo feminino. Os falantes que não aprenderam catalão trabalham na construção civil, no atendimento ao público em estabelecimentos comerciais e em serviços de limpeza. Sete são do sexo feminino e dois do sexo masculino. Destes, só uma informante tem formação superior.

Para além dos falantes de PLM, foram entrevistados 9 informantes nascidos em Andorra, filhos de portugueses. Consideram-se falantes de herança os que crescem em contexto de migração e adquirem, a par da língua dominante, a língua de origem da sua família (a língua de herança) desde a nascença, normalmente em contexto familiar e informal (Rinke et al., 2019).

Os falantes PLH que entrevistámos são informantes trilingues (catalão, espanhol e português), têm entre os 16 e os 25 anos, do sexo feminino (3 informantes) e masculino (3 informantes). Foram, ainda, entrevistados informantes do sexo masculino (3 informantes), com um curso superior, entre os 26 e os 44 anos. Esta bipartição da amostra resultou da impossibilidade de entrevistar falantes de PLM com menos de 26 anos. Note-se que o objetivo inicial do projeto era entrevistar somente falantes de português língua materna. No entanto, um dos constrangimentos sentidos, ao longo da recolha, foi o contacto com informantes com as características sociolinguísticas desejadas, pelo que se optou por entrevistar falantes de português língua de herança e considerar esta variável na análise dos dados. Independentemente de terem o português como língua materna ou de herança, não foi possível identificar e entrevistar informantes do sexo masculino com um curso superior e com mais de 44 anos.

5.3 Transcrição, codificação e análise

As entrevistas foram transcritas preservando as instâncias de alternância linguística e eventuais regionalismos presentes no discurso espontâneo dos informantes. Para garantir o anonimato das entrevistas, a cada informante foi atribuído um código alfanumérico. Considere-se o informante 4F3SE. O primeiro dígito corresponde ao número da entrevista. A letra F refere-se ao género do informante (feminino). O número seguinte diz respeito ao intervalo etário (1 para os informantes entre 16 e 25 anos, 2 para os informantes entre 26 e 44 anos e 3 para os informantes com mais de 44 anos). O código termina com a indicação do grau de escolaridade: SE indica informantes com um grau de escolaridade até ao 12º ano e SU indica que o informante é detentor de um curso superior.

Após a transcrição, as entrevistas foram compiladas na ferramenta *Sketch Engine* (Kilgariff et al., 2014⁹). Esta ferramenta de análise de *corpus* permite extrair informação metalinguística e obter dados de frequência. Possibilita, ainda, buscas por termos, categoria gramatical e contexto sintático, entre outras funcionalidades. Em paralelo, foram identificadas e codificadas as alternâncias linguísticas de acordo com as variáveis sociais e linguísticas consideradas relevantes para o presente estudo. A análise quantitativa foi realizada no software IBM SPSS Statistics. Considerando que os dois grupos de informantes (PLM e PLH) não partilham as mesmas características sociodemográficas, os resultados não são diretamente comparáveis e, como tal, a análise estatística foi realizada em separado.

As variáveis sociais consideradas pertinentes foram: o género, a idade, a escolaridade, os anos em Andorra, a idade do informante quando emigrou para Andorra, a língua de trabalho e a aprendizagem do catalão em contextos formais de ensino. No caso dos informantes nascidos em Andorra, considerou-se, ainda, o sistema educativo que frequentou ou frequenta em Andorra.

As variáveis linguísticas consideradas foram: a tipologia de alternância (Poplack, 2020), a extensão da alternância (item lexical, sintagma, oração, expressão fixa), a categoria gramatical (aplicável apenas aos itens lexicais isolados) e a língua doadora envolvida na alternância.

Após codificação, os dados foram revistos por uma falante nativa de espanhol e catalão.

6. Resultados

Nesta secção, são apresentados os resultados da análise às alternâncias linguísticas. É, ainda, feita uma breve reflexão relativamente às perceções linguísticas.

Considerando que a amostra é constituída por informantes nascidos em Portugal e que têm o português como língua materna (PLM) e por informantes nascidos em Andorra, filhos de pais portugueses, que têm o português como língua de herança (PLH), os dados não podem ser analisados conjuntamente. Assim, os resultados serão apresentados fazendo referência aos dados recolhidos em entrevistas realizadas a falantes de português língua materna, que, por conveniência, etiquetamos de subcorpus I, e aos dados recolhidos em entrevistas realizadas a falantes PLH, aos quais nos referimos como subcorpus II.

O subcorpus I (falantes PLM) é constituído por 18 entrevistas. Foram identificadas 532 ocorrências de alternâncias linguísticas. Não há diferenças significativas a observar na distribuição das alternâncias por género. No que diz respeito ao subcorpus II (falantes PLH), foram recolhidas 9 entrevistas e identificadas 464 ocorrências de alternância linguística.

No que diz respeito às línguas das alternâncias (cf. Tabela 2), verificamos que o padrão de ocorrências das alternâncias de acordo com a língua é semelhante no discurso de falantes do PLM e do PLH, ainda que haja ligeiras diferenças percentuais. Assim, no caso dos falantes de PLM, 44,5% das alternâncias envolvem itens lexicais comuns ao catalão e ao espanhol, 38,9% das alternâncias são feitas com o espanhol, 10,9% com o catalão, 4,9% com o inglês e apenas 0,8% com o francês. No discurso dos falantes de PLH, 45,9% das alternâncias são feitas com o espanhol e 43,5% com termos lexicais ou expressões comuns ao espanhol e o

⁹ Disponível em <http://www.sketchengine.eu>

catalão. Apenas 8,4% das alternâncias ocorrem com termos exclusivos do catalão. Uma minoria das alternâncias envolve o inglês (2,2%). Não se registam alternâncias com o francês neste grupo de falantes.

Tabela 2

Número e Percentagem de Ocorrência das Alternâncias, de Acordo com a Língua e o Subcorpus em Análise

Línguas de alternância	Subcorpus I - PLM		Subcorpus II - PLH	
	N.	%	N.	%
Espanhol	237	44,5%	202	43,5%
Espanhol + Catalão	207	38,9%	213	45,9%
Catalão	58	10,9%	39	8,4%
Inglês	26	4,9%	10	2,2%
Francês	4	0,8%	0	0
Total	532	100	464	100

No que diz respeito às alternâncias em inglês, trata-se de itens lexicais cujo uso está convencionado e, em alguns casos, dicionarizado. Observem-se os seguintes exemplos:

- (8) Sim, também hay un tema que também é os *youtubers* [25M3SE]
 (9) Porque a segurança social dá-te o dinheiro do *ticket* do gasol. [6F2SE]
 (10) Lá está, se eu estivesse a estudar fora, uma coisa que eu não queria, para fazer o meu *máster*, ia ser muito pesado e ia desistir. [20M1SE]
 (11) Vivi num *camping*, num *bungalow*. [4F3SE]

A integração destes termos do inglês no discurso é ilustrativa do que De Decker e Vandekerckhove (2012) apelidam de alternâncias “não intencionais”, na medida em que este uso não resulta de uma escolha deliberada, pois não existe uma tradução do termo em português, espanhol ou catalão ou o seu uso não é produtivo, como se verifica com *youtubers* ou *bungalow*. Outros termos em inglês, como *máster* ou *camping*, estão dicionarizados em espanhol e em catalão¹⁰.

No caso das quatro alternâncias com o francês, verifica-se que se registam no discurso de uma informante que relata um episódio ocorrido em França (exemplo (12)) aquando do registo de nascimento do filho mais velho. Por esta razão, estes exemplos foram excluídos da análise de resultados.

- (12) Escreveram-nos uma direção num papel. Incorreta. Lá fomos nós para a França com aquela direção. Rua para aqui, rua para acolá. Amostravamos o papel... à *droite*, à *gauche* e à *droite* e à *gauche*. E para aqui..., E lá andávamos... E Deus... eu acho que é Deus... nos orientou lá. E registamo-lo lá. [4F3SE]

De igual modo, considerando as especificidades da inclusão dos termos em inglês no discurso, aliadas à percentagem residual de ocorrência, os exemplos de alternâncias linguísticas envolvendo o inglês foram excluídos da análise que doravante se apresenta.

¹⁰ O uso de itens lexicais provenientes do inglês no espanhol e no catalão não é raro. A este propósito, veja-se Lorenzo Criado (1996) e Freixa Aymerich (2019).

Atendendo apenas às alternâncias com o espanhol e o catalão (Figura 2 e

Figura 3), existe uma preponderância de empréstimos lexicais (85,5% nos falantes de PLM e 70,7% nos falantes de PLH). Apenas uma minoria das alternâncias envolve a inserção de constituintes frásicos ou oracionais (5,3% no discurso dos falantes de PLM e 7,3% no discurso dos falantes de PLH). Um dado relevante é a percentagem de ocorrência de convergência gramatical. Esta tipologia de alternância representa 22% do total de alternâncias no discurso dos falantes de PLH e apenas 9,2% no discurso dos falantes de PLM.

Figura 2

Tipologia de Alternância Linguística no Discurso dos Falantes PLM

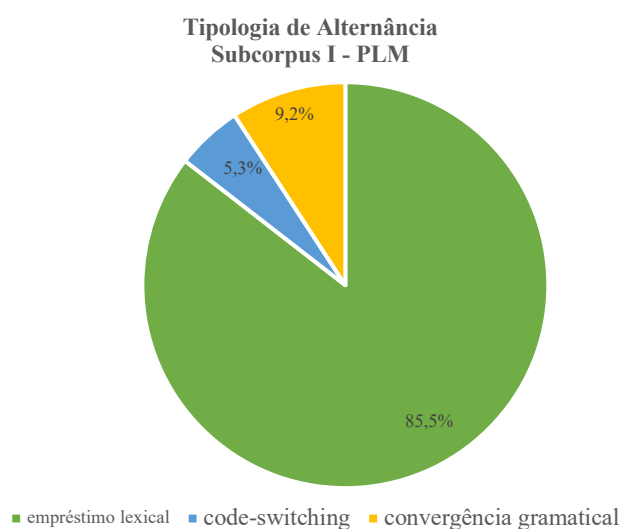
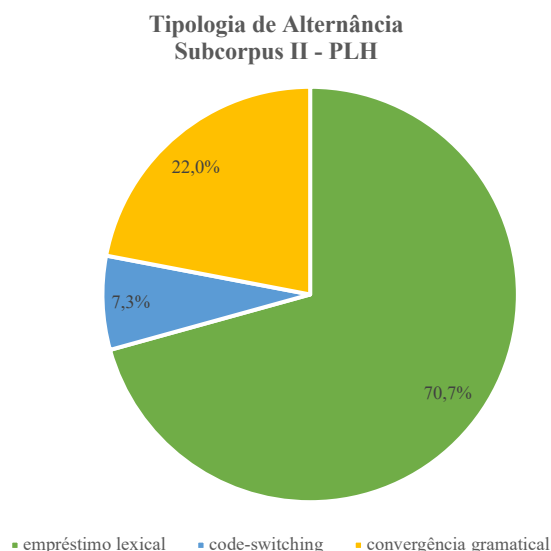


Figura 3

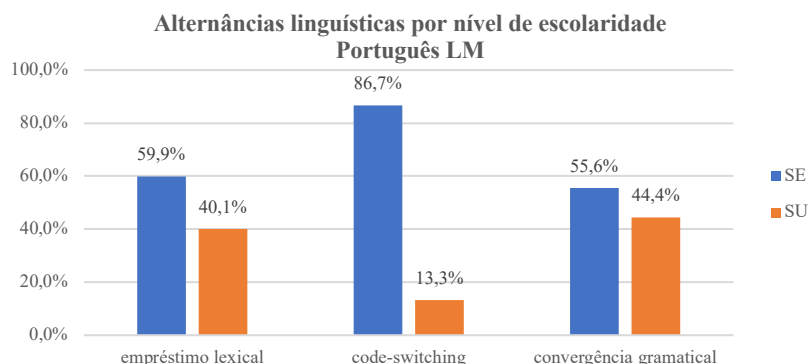
Tipologia de Alternância Linguística no Discurso dos Falantes PLH



Ainda no que diz respeito à tipologia de alternância linguística no discurso dos falantes de PLM, analisámos a sua distribuição por nível de escolaridade (falantes com escolaridade até ao 12º ano (SE) vs. falantes com um curso superior (SU)). Considerando a distribuição das entrevistas pela variável *nível de escolaridade*, as amostras foram equilibradas. Verificamos que, em termos percentuais, as alternâncias linguísticas são mais frequentes nos falantes com escolaridade até ao 12º ano, independentemente da tipologia de alternância (cf. Figura 4). Adicionalmente, é nestes falantes que a percentagem de *code-switching* é mais frequente (86,7% vs. 13,3%).

Figura 4

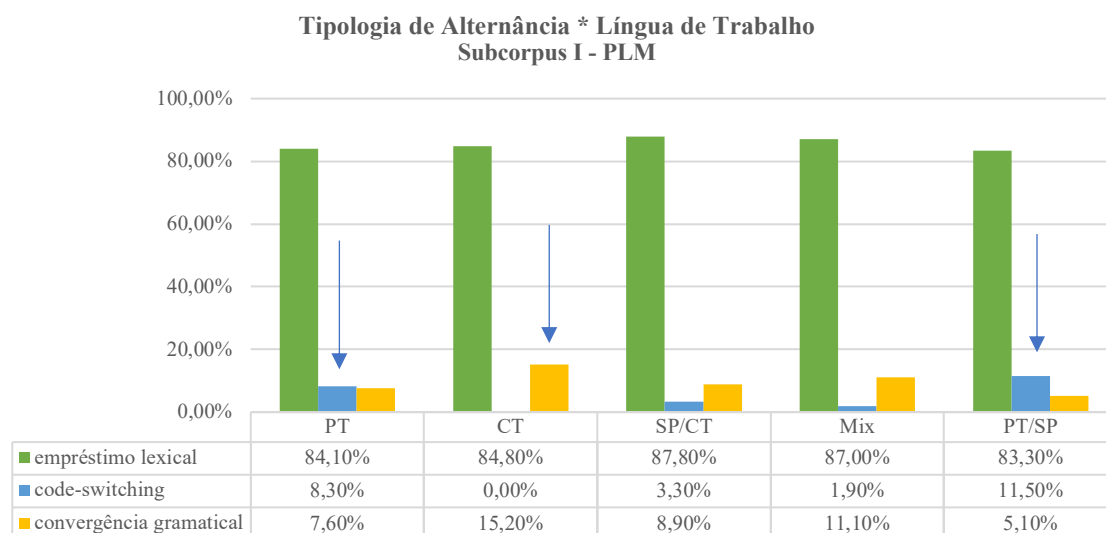
Tipologia de Alternâncias Linguísticas por Nível de Escolaridade (SE – até ao 12.º Ano, SU – com um Curso Superior) no Discurso de Falantes de Português LM



Outro aspeto analisado foi a influência da língua de trabalho. A Figura 5 ilustra a distribuição das três tipologias de alternância linguística de acordo com a(s) língua(s) que os informantes falam em contexto laboral (PT – português, CT – catalão, SP/CT – espanhol e catalão, Mix – português, espanhol, catalão e/ou francês, PT/SP- português e espanhol).

Figura 5

Tipologia de Alternâncias Linguísticas de Acordo com a Língua de Trabalho (Falantes de PLM).

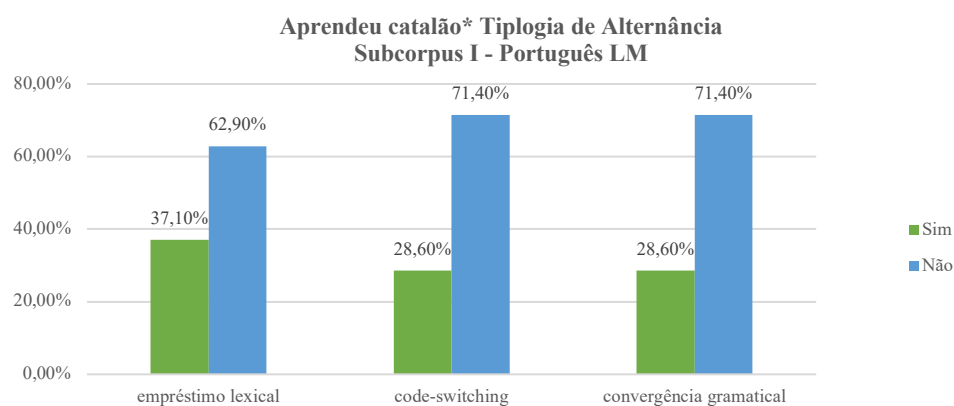


Independentemente da língua de trabalho, as instâncias de empréstimo lexical são a tipologia de alternância mais frequente. Este dado é decorrente da frequência de ocorrência desta tipologia de alternância e a língua de trabalho dos informantes não é relevante para explicar a sua preponderância. Há, no entanto, aspetos que merecem referência. A percentagem de ocorrência de *code-switching* é maior no discurso dos informantes que usam o português como língua de trabalho (8,3% no discurso dos informantes que usam português como língua de trabalho e 11,5% no discurso dos informantes que usam tanto o português como o espanhol como língua de trabalho). Outro dado relevante é a percentagem de ocorrência de instâncias de convergência gramatical no discurso dos informantes que falam exclusivamente catalão no seu local de trabalho (15,2%). De referir, ainda, que a variável *língua de trabalho* está correlacionada com a profissão exercida e com a aprendizagem do catalão. Uma vez que o catalão é a língua usada nos serviços públicos, incluindo os de saúde, os informantes que exercem funções públicas ou trabalham no hospital ou em outros serviços de saúde são falantes de catalão.

A Figura 6 diz respeito à distribuição das alternâncias linguísticas de acordo com a aprendizagem formal do catalão (considerando apenas os falantes de PLM).

Figura 6

Tipologia de Alternâncias Linguísticas de Acordo com a Aprendizagem Formal do Catalão (Falantes de PLM)



Os dados revelam que os falantes que não aprenderam o catalão em cursos de línguas exibem uma maior percentagem de alternâncias linguísticas independentemente da tipologia. Os dois grupos em análise (os que aprenderam catalão e os que não aprenderam) exibem, no entanto, o mesmo padrão no que diz respeito à distribuição das alternâncias, não sendo possível, do ponto de vista estatístico, afirmar que a aprendizagem formal do catalão influencia a tipologia de alternância presente no discurso ($\chi^2(2) = 2.102$, $p = 0.35$)¹¹.

Considerando as características da amostra (secção 5.2) e a correlação entre a profissão exercida e o domínio da língua catalã, não é possível afirmar categoricamente se os resultados obtidos são influenciados pela aprendizagem formal do catalão ou pela profissão exercida.

Analísámos, ainda, a influência da variável “anos em Andorra”. A

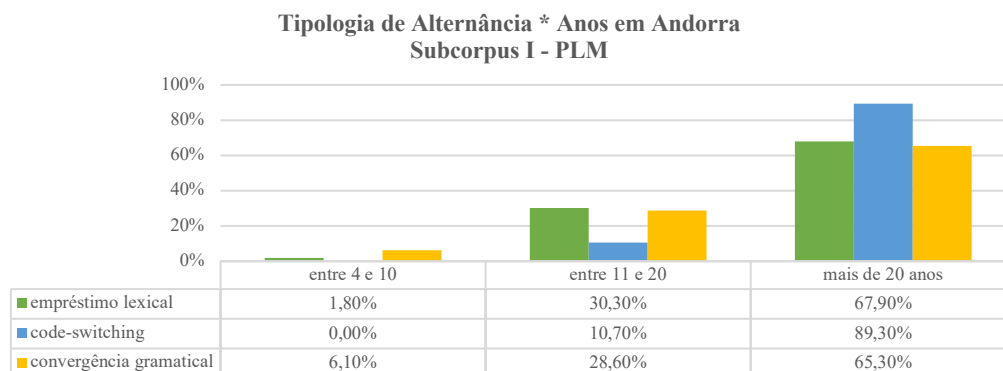
¹¹ De forma a aferir se a distribuição da tipologia de alternâncias pode ser associada às características sociodemográficas dos informantes considerados, usámos o teste do Qui-Quadrado. Para efeitos de clareza na apresentação dos resultados, é referido o valor de prova (p-value). Quando o valor de prova é superior a 5% (>0.05), as variáveis são independentes.

Figura 7 mostra a distribuição das tipologias de alternância de acordo com três intervalos de tempo: entre 4 e 10 anos de permanência em Andorra, entre 11 e 20 anos, e mais de 20 anos, independentemente do grau de escolaridade¹². Os dados mostram que o tempo de residência em Andorra é um fator relevante para explicar a presença de alternâncias linguísticas no discurso, principalmente no que diz respeito ao aumento de ocorrências do tipo *code-switching*.

¹² O cruzamento das variáveis *tempo em Andorra* e *escolaridade* não permite extrair conclusões robustas do ponto de vista estatístico, uma vez que a maioria dos informantes que residem há mais tempo em Andorra não possui curso superior. Esta característica da amostra decorre das especificidades dos diferentes fluxos migratórios que caracterizam a comunidade portuguesa em Andorra (cf. secção 2).

Figura 7

Tipologia de Alternâncias Linguísticas de Acordo com os Anos de Residência em Andorra (Falantes de PLM)



No que diz respeito às variáveis linguísticas, as categorias gramaticais mais frequentes nos empréstimos lexicais são: nomes (13) (14), (17), verbos (15), interjeições (16), adjetivos (17) e advérbios (18).

(13) É, azeite, **almendra**. É uma **salsa** especial mesmo para **calçots**¹³. [1M2SE]

(14) Ou cantamos alguma coisa em **català**, se por exemplo for o padre que começa a cantar e a gente segue. [2F3SE]

(15) Agora, se **parlas** o catalão perfeito, se falas o catalão perfeito, sim, é melhor. [23M2SU]

(16) **Bueno**, é um país pequeno.

(17) O mal dos espanhóis é que dobravam as **películas**, ou seja, não eram **subtituladas**. [14M3SE]

(18) Ao ser pequeno, nos conhecemos **casi** todos. [28F1SE]

Quando analisamos as instâncias de empréstimo lexical nos dados recolhidos, verificamos que, independentemente do subcorpus em análise, é possível traçar uma tendência no que diz respeito à categoria gramatical alvo de alternância (cf. **Figura 8**). Verifica-se que os nomes ocorrem mais frequentemente como empréstimo lexical (49,7% no corpus I – PLM e 39,3% no subcorpus II – PLH), seguindo-se a categoria verbo (15,5% no corpus I – PLM e 20,1% no subcorpus II – PLH). A categoria “Outros” compreende locuções adverbiais (19) (20) e expressões (21).

(19) Às vezes misturo um bocadinho. Acho que uma palavra de cada idioma, depende de como...**ao melhor**¹⁴ acabas por misturar e eu digo, ah, mas já nem sei o que é que eu estou a falar. [16F2SE]

¹³ *Calçots* são cebolas brancas, tradicionalmente comidas assadas. No entender de Blas Arroyo (2021), as alternâncias linguísticas envolvendo termos para designar comida ou bebida, expressões idiomáticas ou fórmulas de saudação, correspondem a *tag-switches*. Funcionam, no entender do autor, a uma estratégia de acomodação, solidariedade ou sinal de identidade coletiva, podendo inclusive ocorrer no discurso de monolíngues com uma competência passiva na língua ou variedade dominante.

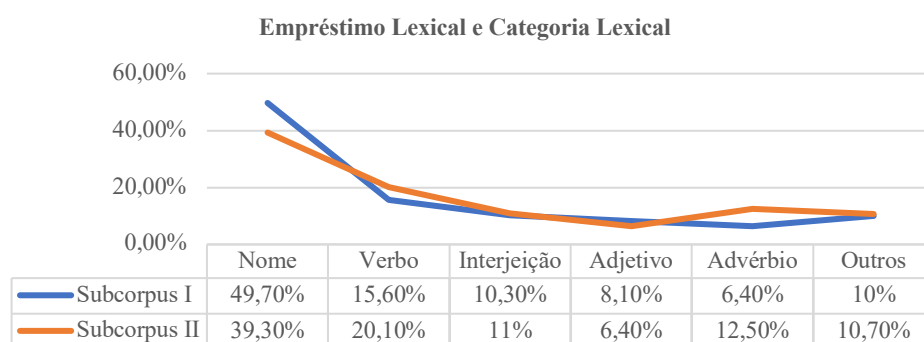
¹⁴ Do espanhol *a lo mejor*. Locução adverbial coloquial com o significado de *talvez* (Real Academia Española, s.d.)

(20) **Hoje por hoje**¹⁵ eu já vejo muitas mulheres que estão [a trabalhar] em fábricas, enfermeiras. Vejo que [Portugal] evoluiu muito, está bastante mais avançado do que quando eu estava lá. Quando eu estava lá era o campo, campo e campo. [2F3SE]

(21) Aí tive [estive] a ir a trabalhar 20 anos! **E nada**¹⁶... e é assim [4F3SE]

Figura 8

Empréstimo Lexical: Percentagem de Ocorrência das Categorias Lexicais nos Dois Subcorpora



No que diz respeito à convergência gramatical, os exemplos mais frequentes no *corpus* em análise compreendem:

i. a próclise em contextos típicos de ênclise (10 ocorrências no subcorpus I e 25 no subcorpus II):

(22) Eu **me identifico** mais como *andorrà*. [28F1SE]

(23) E depois, **te fica** muito pouco para comer [24F3SU]

ii. processos que envolvem o sintagma preposicional (9 ocorrências no subcorpus I e 34 no subcorpus II), nomeadamente a seleção da preposição (24) e a determinação do nome no interior do sintagma preposicional (25):

(24) O único susto, estava a vir de treinar futebol e caiu um javali do monte **ao** meio da estrada. [8M1SE]

(25) A minha primeira viagem a Portugal, fi-la eu a última vez que fui, **em Páscoa**, que aliás conduzi a metade da viagem, 5 horitas. [20M1SE]

iii. a concordância de género (5 ocorrências - subcorpus II)

¹⁵ Do espanhol *hoy por hoy*. Locução adverbial usada para indicar que algo é ou acontece de uma certa forma agora, mas que pode mudar mais tarde (Real Academia Española, .d.).

¹⁶ Concorreu para a classificação de empréstimo lexical a alta frequência de ocorrência da expressão *y nada* em espanhol, a par de *nada* e *pues nada* (Engels & Tanghe, 2019), principalmente no discurso dos falantes mais jovens, assim como a pronúncia do informante ao produzir este segmento do discurso.

(26) eu quando vou lá em Portugal, eu tenho a sensação de que a gente reparam para mim como aquele estrangeiro que chegou lá e, pronto, eu faço **aquelas pontapés** no dicionário e tudo. [26M2SU]

(27) [Café com gelo] É **uma costume muito espanhola**, de quando vemos o primeiro raio do sol, o que significa verão ou calor, toma-se assim. [20M1SE]

No exemplo (24), o informante 8M1SE faz referência a uma situação em que um javali caiu (das montanhas) para o meio da estrada. Neste exemplo, verifica-se a influência do catalão, língua que, neste contexto, seleciona *al* (contração da preposição *a* e do artigo *el*): "va caure **al** mig del carrer". No exemplo (25) está em causa a determinação do nome, através da ausência do determinante artigo. Neste caso, o uso da preposição simples *em* em vez da preposição contraída *na* é o resultado da influência do espanhol, língua em que o nome de festividades religiosas como o *Natal* ou a *Páscoa* é antecedido da preposição simples *em*, em vez da forma contraída como se verifica em português.

Por fim, os exemplos (26) e (27) ilustram a influência da LD no que diz respeito à atribuição de um valor de género indevido. Nestes casos, o falante, por influência do catalão e/ou espanhol, altera o género dos elementos que concordam com o nome. No exemplo (26), o nome masculino *pontapés* concorda com o determinante demonstrativo feminino *aquelas*, resultado da influência dos nomes femininos *patadas* (em espanhol) e *puntadas* (catalão). No exemplo (27), por influência de *costumbre* (nome feminino) do espanhol, o nome masculino *costume* concorda com o determinante artigo feminino *uma* e com o adjetivo feminino *espanhola*.

Para além da análise das alternâncias linguísticas, este trabalho aflora, ainda, as percepções da comunidade portuguesa em Andorra em relação ao uso do português e ao uso das línguas dominantes: o catalão e o espanhol. Os relatos recolhidos reiteram o facto de a comunidade portuguesa em Andorra ter orgulho em falar em português e nos costumes portugueses. Os imigrantes portugueses em Andorra vêm a Portugal várias vezes por ano, normalmente durante as férias escolares, ouvem estações de rádio portuguesas, veem televisão em português e sentem orgulho no facto de os filhos saberem falar português, como relatam os informantes 1M2SE (28) e 16F2SE (29):

(28) Mas connosco em casa [os filhos falam] sempre português, porque [o catalão] foi uma das coisas que eu nunca entendi muito bem. Eu também não falo muito bem espanhol. [1M2SE]

(29) Em casa sempre, em casa sempre falamos [português]. De feito, os meus filhos vão a Portugal e todo mundo me diz, aí os teus miúdos, que lindos estão, porque eles falam português e eles entendem-nos e nós entendemos bem eles. Porque há muitos, por exemplo, eu tenho primos da França, que eles não falam mais que francês para os miúdos. [16F2SE]

Os falantes de PLM consideram vital o uso da língua portuguesa no contexto familiar. A língua portuguesa é, assim, tida como um bem simbólico (Bourdieu, 1991), um legado imaterial que pais transmitem aos filhos. O facto de os filhos falarem português é motivo de orgulho, apesar de eles próprios terem consciência de que já misturam o português com o espanhol e o catalão (exemplos (30), (31) e (32)).

(30) Por exemplo, há palavras, que eu sou sincera, e há palavras em português que eu agora já não sei o significado delas. [2F3SE]

(31) Os meus filhos falam português, sempre falaram português. Também já... mas nem eu nem eles falamos um português já correto, correto. Também não há... em Portugal também...o minhoto também é diferente. [4F3SE]

(32) Às vezes misturo um bocadinho. Acho que uma palavra de cada idioma, depende de como... ao melhor, acabas por misturar. E eu digo, ah, mas já nem sei o que é que eu estou a falar. Não sei se estou a falar de uma coisa, se estou a falar de outra. [16F2SE]

No caso dos falantes PLH que já concluíram os seus estudos em Andorra ou em Espanha e que já constituíram família em Andorra, a língua portuguesa funciona como uma ligação simbólica às origens (33).

(33) vamos todos os anos para Portugal, porque eu acho que é importante que os meus filhos tenham aquela ligação, não é? (...) A língua portuguesa lá vai, está no ar, não é todos os dias, mas, às vezes, procuro para eles, para que eles tenham aquela ligação com Portugal, com o português, não é sempre, mas procuro. [26M2SU]

Os portugueses de 2ª e 3ª gerações adquiriram o português como língua de herança (e não como L2¹⁷). O *input* foi maioritariamente oral e informal, pelo que alguns informantes, como o informante 26M2SU, relatam não terem “domínio absoluto” da língua (34):

(34) vou falar consigo em português, faz-me ilusão, mas não é uma língua que eu tenha a sensação de... tenha a perceção de domínio absoluto. [26M2SU]

Apesar de todos os informantes PLH terem frequentado aulas de português, oferecidas pelo Instituto Camões como disciplina extracurricular, tipicamente até aos Níveis Comuns de Referência B1 ou B2, gostariam de ter podido aprofundar mais o conhecimento linguístico da língua e o domínio da produção escrita, principalmente o registo formal (35).

(35) [Informante - 30F1SE] – Eu, quando era mais pequena, sim, que fui a aulas de português e se não me engano, tenho o B2. É muito importante pôr um nível mais alto, porque parece que somos já... tipo, filhos de portugueses, o A1 não nos serve. Eu gostei muito [das aulas de português], mas também acho que não aprendi grande coisa, porque, tipo, sou uma pessoa que fala bastante português, que tem toda a sua família portuguesa, então falo muito português. Mas se fossem, tipo, níveis mais altos, sim, que iria aprender. Já não o vocabulário, porque o vocabulário praticamente temos muito. Sino... tipo, a forma de falar, se calhar, ou de referir-se a pessoas mais importantes.

[JA] O português mais formal.

[Informante - 30F1SE] Sim.

7. Discussão

A análise das três tipologias de alternâncias (empréstimo lexical, *code-switching* e convergência gramatical) permite concluir que falantes de português LM e LH exibem padrões de alternância linguística semelhantes, apesar de a frequência de ocorrência ser diferente. O cruzamento da frequência de ocorrência das alternâncias linguísticas e das variáveis sociais permite enquadrar os resultados obtidos. Mais especificamente, nos falantes de PLM, o tempo de residência em Andorra e a aprendizagem formal do catalão são fatores relevantes para explicar os padrões de alternância observados. O nível de escolaridade é, também, uma variável determinante, pois falantes com um curso superior exibem uma percentagem menor de alternâncias no discurso. De referir que, considerando as características da comunidade portuguesa em Andorra, as variáveis sociais *anos*

¹⁷ Veja-se Rinke et al. (2019) a propósito das implicações que as diferentes fontes de *input* têm no desenvolvimento da gramática mental do falante.

de residência em Andorra, nível de escolaridade e profissão estão muitas vezes interligadas, sendo necessário alargar a amostra para obter resultados significativos do ponto de vista estatístico. Assim, os informantes que residem há mais tempo em Andorra têm tipicamente menos anos de escolaridade e trabalham na construção civil, restauração ou em serviços de limpezas. Estes informantes comunicam normalmente em português ou “portunhol” e não aprenderam o catalão em contextos formais, como refere o informante [1M2SE]:

(36) A base é português. Nós aqui em Andorra, verdadeiramente, se não quisermos falar outra língua durante o ano, não precisamos. E então não há necessidade mesmo de aprender. (...) Que deveríamos, deveríamos. Eu é uma coisa que tenho assente, não faço, mas deveríamos. Porque estamos num país que não é o nosso, devíamos falar a língua do país. [1M2SE]

É, também, no discurso dos informantes a residir há mais anos em Andorra que identificamos uma maior percentagem de alternâncias lexicais. Este resultado reitera as afirmações impressionistas decorrentes de um trabalho sociológico realizado na comunidade portuguesa em Andorra (Matias, 2008). Segundo a autora, “Estas variações linguísticas, na introdução de palavras numa língua diversa daquela com que se produz o discurso, estão presentes sobretudo nos emigrantes com mais anos de residência ou com menor nível de estudos” (Matias 2008, p. 62). Por outro lado, os informantes com um curso superior trabalham em serviços públicos ou empresas privadas e aprenderam catalão (apenas uma informante diz não ter aprendido catalão em contextos formais). Estes informantes são críticos relativamente às gerações anteriores e ao facto de não dominarem a língua do país de acolhimento.

No que diz respeito aos *empréstimos lexicais* e às categorias lexicais envolvidas, o padrão observado tanto no discurso dos falantes de PLM como nos falantes de PLH foi já identificado noutros trabalhos sobre alternâncias linguísticas. Veja-se, por exemplo, Torres Cacoullous e Travis (2020, p. 259), segundo os quais “single-word insertions are disproportionately nouns”. Distribuição semelhante é observada em Poplack (2018), num trabalho sobre os empréstimos lexicais do inglês no discurso em francês canadiano: nomes (64%), verbos (14%), interjeições e expressões fixas (12%), adjetivos (8%). Ainda no que diz respeito aos empréstimos lexicais, verificou-se que, nos falantes de PLH, os empréstimos lexicais exclusivos do catalão ocorrem com mais frequência nas produções dos falantes que frequentaram a escola andorrana (49,1%) do que nas produções dos informantes que frequentaram a escola espanhola (10,3%) ou a escola francesa (36,8%).

Em relação à categoria convergência gramatical, como referimos na secção anterior, os processos mais frequentes nas entrevistas recolhidas são: (i) a ênclise em contextos de próclise, (ii) a seleção da preposição e a determinação do nome no interior do sintagma preposicional e (iii) a atribuição de um valor de género indevido e a consequente ausência de concordância de género. Estes resultados encontram eco no trabalho de Scetti (2016) sobre a comunidade portuguesa em Montreal. De acordo com o autor, é possível identificar sete “pontos de mudança” no português falado em Montreal, resultado do contacto com o francês e o inglês, entre os quais o uso da próclise em contextos tipicamente de ênclise e ausência de concordância de género em construções *N_{fem} + Adjetivo*. Neste contexto, o adjetivo surge frequentemente na forma masculina. De acordo com o autor, estes processos são mais frequentes em falantes mais jovens, que, no trabalho de Scetti (2016) correspondem a falantes de PLH).

No que diz respeito aos clíticos pronominais, como é sabido, em português europeu, a sua colocação depende principalmente de fatores sintáticos (Martins 2015, 2016), que sintetizamos:

- i. A polaridade da frase: em frases afirmativas, observa-se a ênclise; em frases de polaridade negativa, o clítico pronominal é atraído para posição pré-verbal;
- ii. A ocorrência, antes do verbo, de quantificadores, sintagmas QU- e alguns advérbios (*até, só, lá enfático, também, sempre, já, talvez*);

- iii. A oposição oração principal / oração subordinada finita: em estruturas subordinadas, o clítico ocupa a posição pré-verbal;

Assim, na ausência de atratores de próclise ou de estruturas sintáticas que permitam ou exijam a próclise, o clítico pronominal ocupa, em português europeu contemporâneo, a posição enclítica¹⁸. Em espanhol e catalão, a posição do clítico é condicionada pela morfologia verbal: a ênclise ocorre com o infinitivo, o gerúndio e o imperativo, ao passo que a próclise ocorre geralmente com as formas finitas (Martins, 2016).

Como refere Martins (2015, p. 84), “o conceito de proclisador/palavra proclisadora aplica-se ao português europeu (e ao galego), mas não às outras línguas românicas, porque nas orações principais afirmativas ocorre normalmente a ênclise e só em casos particulares a próclise (concretamente quando uma palavra proclisadora precede o verbo)”. Considerando os dados apresentados, assim como os trabalhos de Scetti (2016), podemos afirmar que, no que diz respeito à ordem, os clíticos pronominais parecem ser elementos instáveis e permeáveis a processos de interferência estrutural decorrentes do contacto de línguas. Ademais, é sabido que, do ponto de vista diacrónico, são elementos sujeitos a processos de variação¹⁹ (Martins, 1994 e seguintes).

Considerando os dados extraídos das entrevistas recolhidas, verifica-se que a próclise em contextos típicos da ênclise ocorre principalmente em falantes PLH (25 ocorrências). As dez ocorrências identificadas no discurso de falantes PLM são produzidas por quatro falantes (três do sexo feminino e um do sexo masculino) a residir em Andorra há mais de 13 anos. Nenhum aprendeu catalão formalmente. Segundo Poplack, os anos de contacto linguístico, assim como o grau de integração do falante na comunidade dominante, são fatores que poderão exponenciar a ocorrência de alternâncias linguísticas e a “interferência estrutural”. Como enfatiza a autora, “the longer and more intense the contact, and the fewer and more marginal the speakers, the greater the likelihood of language mixing and structural interference” (Poplack, 2020, p. 49).

Relativamente a processos envolvendo o sintagma preposicional, o trabalho de Rinke et al. (2024) mostra que a seleção da preposição (simples ou contraída) pode ser desafiante para falantes de PLH. O mesmo trabalho faz referência à dificuldade que os falantes bilingues apresentam na atribuição de género (Montrul & Ionin, 2010). Não obstante a complexidade linguística e a opacidade semântica de algumas das estruturas que envolvem preposições (Rinke et al., 2024), quer para falantes bilingues quer para falantes monolíngues, os exemplos aqui apresentados ilustram a influência da gramática das línguas dominantes (o catalão e o espanhol) no português.

Apesar de a escassez de dados não permitir extrair conclusões robustas relativamente ao efeito do ensino formal da língua catalã na ocorrência de instâncias de alternância linguística, nomeadamente as do tipo convergência gramatical, este é um aspeto que será aprofundado em publicações futuras.

Outro aspeto aflorado neste trabalho foram as perceções da comunidade portuguesa em Andorra em relação ao uso do português e do catalão. De uma forma geral, os falantes de PLM estão cientes da importância do uso do catalão e do prestígio social que este pode proporcionar. O domínio do catalão é, na realidade, um fator que determina o acesso a oportunidades económicas e sociais. A este propósito, Hawkey e Horner (2022, p. 212) realçam que as políticas monolíngüísticas oficiais colidem com a prática multilingüística da sociedade andorrana: “these official de jure state monolingual policies stand in contradiction with widely practiced de facto societal multilingualism.”. A *Llei d’ordenació de l’ús de la llengua oficial* (Govern d’Andorra, 2000) determina que o catalão seja a língua da comunicação oficial, assim como dos serviços públicos, incluindo os de saúde. O acesso a empregos públicos e aos quadros técnicos superiores depende, assim, do domínio da língua,

¹⁸ Ao longo da história da língua portuguesa, a colocação dos clíticos pronominais sofreu processos de variação e mudança. A este propósito, veja-se Martins (1994, 2011 e seguintes).

¹⁹ A este propósito, Martins (2011, 2015) propõe a existência, no português antigo, de duas variedades linguísticas: “o português antigo originou dois dialetos diferentes no que diz respeito à colocação dos pronomes clíticos: um em que a ênclise nunca deixou de ser dominante nos contextos de variação e que representa a linha de continuidade entre o português antigo e o português contemporâneo, e outro que representa uma linha evolutiva a que poderíamos chamar «pan-ibérica» e que veio a perder-se, embora pareça ter correspondido à variedade de prestígio no português quinhentista e seiscentista” (Martins, 2015, p. 84).

o que não significa que o catalão seja utilizado em exclusivo, como ilustram os resultados do relatório *Coneixements i usos lingüístics de la població d'Andorra* (Govern d'Andorra, 2023). De acordo com os testemunhos recolhidos, o espanhol substitui frequentemente o catalão nos contextos formais e administrativos. Mesmo em contextos sociais e situações laborais em que o catalão é exigido, o espanhol ou o português é frequente, como referem os informantes 15F3SU e 11M2SE:

(37) Quando estou de cara ao usuário, como sou enfermeira, obrigam-nos a falar em catalão, que é a língua oficial do país. Como trabalho em saúde laboral, há muita gente de muitos sítios e de muitas origens. E quando começo a ver alguém, por exemplo, português, que está ali a patinar um bocado e que não... pois... troco e falo em português. [15F3SU]

(38) Isto [falar espanhol] está tão mentalizado que mesmo aqueles mais ferrenhos, aqueles mesmo do catalão a sério, veem que se tu respondes em castelhano, eles começam a responder em castelhano. [11M2SE]

Relativamente às perceções dos informantes PLH entrevistados no que diz respeito ao uso e aprendizagem formal do português, o relato da Informante 30F1SE (exemplo 35) é paradigmático e espelha uma das conclusões do trabalho de Rinke et al. (2019). Segundo as autoras, o conhecimento fonológico e lexical de falantes de PLH é semelhante ao dos falantes de PLM. No entanto, falantes de PLH apresentam dificuldades em alguns processos morfológicos e sintáticos. O ensino formal a falantes de PLH deve, assim, incidir sobre processos linguísticos complexos, que ocorrem principalmente no discurso formal e que são, tipicamente, adquiridos mais tarde.

8. Conclusão

Este trabalho constituiu uma primeira abordagem ao mosaico linguístico constituído pela comunidade portuguesa em Andorra, centrada principalmente nas alternâncias linguísticas e na inevitável influência social que as línguas dominantes, o espanhol e o catalão, exercem. A adoção da proposta de Poplack (2020) permitiu observar padrões de alternância relevantes. Por um lado, falantes de PLM e falantes de PLH não exibem padrões de alternâncias distintos: as instâncias de empréstimos linguísticos constituem a maioria das alternâncias nos dois grupos de falantes, seguidas das instâncias de convergência gramatical. Por outro lado, verificou-se que, no discurso dos falantes de PLH, a percentagem de convergência gramatical é superior. No caso dos falantes PLM, de acordo com os dados recolhidos, o tempo de residência em Andorra e o menor grau de escolaridade são fatores que potenciam as instâncias de convergência gramatical. Por fim, a ocorrência dos processos identificados como convergência gramatical, nomeadamente a posição do clítico pronominal, a atribuição de um valor de género indevido e a consequente ausência de concordância de género e a seleção da preposição, não é aleatória, o que nos leva a colocar a hipótese de haver elementos mais permeáveis à influência exercida pelo contacto linguístico, entre outros fatores.

Espera-se que os resultados aqui apresentados tenham uma capacidade multiplicadora, na medida em que poderão constituir o ponto de partida para análises futuras no campo dos estudos sociolinguísticos e da política da língua, assim como para estudos sobre o português língua de herança e o português falado pela comunidade portuguesa no estrangeiro.

Agradecimentos

Agradecemos as observações e sugestões de melhoria propostas pelos revisores anónimos. Qualquer erro ou imprecisão remanescente é de inteira responsabilidade das autoras.

Um agradecimento especial aos informantes, sem os quais este trabalho não seria possível.

Financiamento

Investigação realizada no âmbito da Cátedra Camões (Universitat d'Andorra). A autora Joana Aguiar reconhece a bolsa de investigação recebida da Universidade de Andorra, com a referência AD03 - UdA-2023/2024. Este trabalho foi apoiado pela FCT - Fundação para a Ciência e Tecnologia no âmbito do Centro de Investigação Transdisciplinar em Educação e Desenvolvimento (<https://doi.org/10.54499/UIDB/05777/2020>).

Referências

- Androutsopoulos, J. (2005). Research on youth language. In U. Ammon, N. Dittmar, K. J. Mattheier, & P. Trudgill (Eds.), *Sociolinguistics: An international handbook of the science of language and society* (pp. 1496–1505). De Gruyter.
- Ballarín, J. (2006). Identitats culturals i usos lingüístics a Andorra. *Treballs de Sociolingüística Catalana*, 19, 183–220.
- Bastida, C., & Nicolau, M. (2012). Un projecte de prospectiva sociolingüística: El model sistèmic de l'evolució de l'ús del català a Andorra. *Llengua i Ús*, 51, 100–112.
- Blas Arroyo, J. (2021). Desenlaces del contacto de lenguas en la Comunidad Valenciana: Entre el cambio de código y el préstamo léxico. In M. Sanz Gil (Ed.), *Estrategias lingüísticas para la sociedad multilingüe* (pp. 13–24). Ediciones Octaedro, S.L.
- Bourdieu, P. (1991). *Language and symbolic power* (G. Raymond & M. Adamson, Trans.; J. B. Thompson, Eds.). Harvard University Press.
- Carvalho, J. (2007). *Portugueses em Andorra: Uma visão global*. Andbanc.
- Comas, D., & Pujadas, J. (1999). Living in/on frontier: Migration, identities and citizenship in Andorra. *Social Anthropology*, 7(3), 253–264.
- De Decker, B., & Vandekerckhove, R. (2012). English in Flemish adolescents' computer-mediated discourse: A corpus-based study. *English WorldWide*, 33(3), 321–351.
- Departament d'Estadística d'Andorra. (2025). *Nota de premsa A001. Estimacions de població. Març del 2025 i A003. Estadística dels censos parroquials*. https://sig.govern.ad/SIGDDE.Public/Files/Documents/Notes_premsa_noticies/A001_20250313_A.pdf
- Enghels, R., & Tanghe, S. (2019). On the interplay between historical pragmatics and sociolinguistics: The case of the Spanish pragmatic marker *nada* and its recent grammaticalization process. *ONOMAZEIN*, 44, 132–165. <https://doi.org/10.7764/onomazein.44.07>
- Flores, C. (2004). «Lá está ele a schmatzen!»: Particularidades do discurso de bilingues luso-alemães. *Diacrítica*, 19(1), 89–108.
- Freixa Aymerich, J. (2019). Augment constant de manlleus en la premsa catalana: Una conseqüència de la mundialització? *Treballs de Sociolingüística Catalana*, 29, 43–64. <https://doi.org/10.2436/20.2504.01.154>
- Govern d'Andorra. (2000). “Llei d'ordenació de l'ús de la llengua oficial”. *Butlletí Oficial del Principat d'Andorra*, 12(2), 66–72. <https://www.bopa.ad/Documents/Detail?doc=1B83E>
- Govern d'Andorra. (2023). *Coneixements i usos lingüística de la població d'Andorra. Situació actual evolució (1995–2022)*. Govern d'Andorra. <https://www.ari.ad/images/pdf/llengua.pdf>
- Govern d'Andorra. (2024). “Llei 6/2024, del 25 d'abril, de la llengua pròpia i oficial”. *Butlletí Oficial del Principat d'Andorra*, 59. https://www.bopa.ad/Documents/Detail?doc=CGL_2024_05_21_09_12_55
- Gumperz, J. (1982). *Discourse strategies*. Cambridge University Press.
- Hawkey, J. (2024). “Aquestes arrels que queden sempre et criden una mica”: Actituds i ideologies envers la llengua portuguesa a Andorra. *Treballs de Sociolingüística Catalana*, 34, 51–66.
- Hawkey, J., & Horner, K. (2022). Officiality and strategic ambiguity in language policy: Exploring migrant experiences in Andorra and Luxembourg. *Language Policy*, 21(2), 195–215. <https://doi.org/10.1007/s10993-021-09602-3>

- Henry, A. (2002). Variation and syntactic theory. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *Handbook of language variation and change* (pp. 267–282). Blackwell.
- Kilgariff, A., Baisa, V., Busta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1, 7–36.
- Labov, W. (1966). *The social stratification of English in New York City*. Center for Applied Linguistics.
- Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.
- Labov, W. (2001). *Principles of linguistic change: External factors*. Blackwell.
- Lorenzo Criado, E. (1996). *Anglicismos hispánicos*. Gredos.
- Margarit, E., Micó, J., & Monné, A. (2010). La interrelació dels joves amb les llengües i els grups lingüístics presents al Principat d'Andorra. In O. Guasch & M. Milian (Eds.), *L'educació lingüística i literària en entorns multilingües: Recerca per a nous contextos* (pp. 331–346). Universitat Autònoma de Barcelona - Servei de Publicacions.
- Martins, A. M. (1994). *Clíticos na história do português*. [Dissertação de Doutorado, Universidade de Lisboa].
- Martins, A. M. (2011). Clíticos na história do português à luz do teatro vicentino. *Estudos de Lingüística Galega*, 3, 83–109.
- Martins, A. M. (2015). Variação sintática no português quinhentista: a colocação dos pronomes clíticos. *Estudos De Lingüística Galega*, 7, 83–94. <https://doi.org/10.15304/elg.7.2373>
- Martins, A. M. (2016). A colocação dos pronomes clíticos em sincronia e diacronia. In A. M. Martins & E. Carrilho (Eds.), *Manual de linguística portuguesa* (pp. 401–430). De Gruyter.
- Matias, M. (2008). *A que sabe um Sumol a 2.200 metros de altitude?*. Pagès Editora.
- Montrul, S., & Ionin, T. (2010). Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition*, 13(4), 449–473. <https://doi.org/10.1017/S1366728910000040>
- Observatório da Emigração. (2010). *Portugueses em Andorra*. <https://observatorioemigracao.pt/np4/1667.html>
- Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics*, 18, 582–617.
- Poplack, S. (2018). *Borrowing: Loanwords in the speech community and in the grammar*. Oxford University Press.
- Poplack, S. (2020). A variationist perspective on language contact. In E. Adamou & Y. Matras (Eds.), *Routledge handbook of language contact* (pp. 46–62). Routledge.
- Real Academia Española. (s.d.). *a lo mejor*. In *Diccionario de la lengua española*. Acedido a 23 abril 2025, em <https://dle.rae.es/a%20lo%20mejor>
- Real Academia Española. (s.d.). *hoy por hoy*. In *Diccionario de la lengua española*. Acedido a 23 abril 2025, em <https://dle.rae.es/hoy%20por%20hoy>
- Real Academia Española. (s.d.). *sí que*. In *Diccionario de la lengua española*. Acedido a 23 abril 2025, em <https://dle.rae.es/sí%20que>
- Rinke, E., Flores, C., & Santos, A. L. (2019). Heritage languages at school: Implications of linguistic research on bilingualism for heritage language teaching. In C. Gabriel, J. Grünke, & S. Thiele (Eds.), *Romanische Sprachen in ihrer Vielfalt: Brückenschläge zwischen linguistischer Theoriebildung und Fremdsprachenunterricht* (Romanische Sprachen und ihre Didaktik) (pp. 211–232). Ibidem-Verlag.
- Rinke, E., Flores, C., & Torregrossa, J. (2024). How different types of complexity can account for difficult structures in bilingual and monolingual language acquisition. In M. Polinsky & M. T. Putnam (Eds.), *Formal approaches to complexity in heritage languages* (pp. 43–71). Language Science Press. <https://doi.org/10.5281/zenodo.12090434>

- Scetti, F. (2016). The Portuguese language in the particular context of the "Portuguese community" of Montreal. In R. Muhr (Ed.), *Pluricentric languages and nondominant varieties worldwide: The pluricentricity of Portuguese and Spanish* (pp. 261–274). Peter Lang.
- Tagliamonte, S. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Wiley-Blackwell.
- Torres Cacoullos, R., & Travis, C. (2020). Code-switching and bilinguals' grammars. In E. Adamou, & Y. Matras (Eds.), *The Routledge handbook of language contact* (pp. 252–274). Routledge. <https://doi.org/10.4324/9781351109154>