

Sobre a origem das preposições ibero-românicas *hasta*, *ata* e *até*

Mário Eduardo Viaro

Universidade de São Paulo (Brasil)

gmhp@usp.br

Recibido o 06/07/2012. Aceptado o 29/10/2012

Resumo

A etimologia das preposições ibero-românicas para designar o limite de um movimento é bastante polêmica. Todas as soluções propostas privilegiam um étimo especial que explica uma das línguas, mas não outras. O elemento árabe, porém, é sempre aceito ao lado do românico. Diante do polimorfismo do antigo *ata*, também é comum que soluções parciais deixem um grande número inexplicado de variantes. Com vistas a uma explicação que integre essas preposições no elemento latino românico, dado que o aspecto formal das preposições é extremamente conservador e avesso a empréstimos nas línguas do mundo, será investigado o caso das concorrências semânticas do conceito limitativo em preposições de outros dialetos e línguas da Península Ibérica, bem como das semelhanças formais entre os dialetos e línguas em toda a extensão das línguas românicas. Também a flutuação entre as classes das preposições, advérbios de lugar e conjunções temporais será explorada, em busca da conservação de estruturas que à primeira vista não apareceriam numa análise etimológica tradicional, haja vista a idiosincrática preferência de algumas formas em detrimento de outras. A investigação, por meio dessa perspectiva indutiva, aponta para uma inusitada transformação do latim *intra*, que vem a ser o étimo mais provável das preposições estudadas.

Palabras chave

Filologia Românica, línguas ibero-românicas, preposições, etimologia

Sumario

1. Características da tradição linguística. 1.1. Língua e herança linguística. 1.2. Preposições e outras classes de palavra. 2. Características das preposições limitativas no ibero-românico. 2.1. O conceito de "até" em latim. 2.2. A proposta da origem árabe para *hasta*, *ata* e *até*. 3. Características gerais da classe das preposições. 3.1. Advérbios e preposições. 3.2. Hiperbibasmo em preposições. 4. Investigação do étimo de *hasta*, *ata* e *até*. 4.1. Étimos alternativos para *até*. 4.2. Equivalentes a *até* nas Cantigas de Santa Maria (CSM). 4.3. Um novo étimo para *até*. 4.4. Origem de *tra* ~ *ta* e *tro*. 4.5. Formas concorrentes de *intra* ~ *intro* nas línguas ibero-românicas. 4.6. A forma *fasta*. 5. Investigação para além da Península Ibérica: a presença de *intus*. 6. Conclusões.

Concerning the origin of the Iberian Romance prepositions *hasta*, *ata* and *até*

Abstract

Considerable controversy surrounds Ibero-Romance prepositions expressing the end-point of a movement. Each of the solutions proposed focuses on an etymon that explains the state of affairs in one language while failing to explain that in another. It is widely accepted that there is an Arabic element involved besides the Romance component. On account of the polymorphism of early *ata*, it is also common for partial solutions to leave unexplained a large number of variants. In an attempt to find an explanation which integrates this range of prepositions into the common Romance element, the present article delves into rival forms expressing the limitative notion in the prepositions of other languages and dialects of the Iberian Peninsula, as well as formal similarities between the dialects and languages throughout the Romance linguistic territory, since prepositions generally display conservatism in their form, while borrowing is rare among the languages of the world. Crossovers between the classes of preposition, place adverb and temporal conjunction will be explored, looking for conservative structures that at first glance would not appear in a traditional etymological analysis, given idiosyncratic preference for some forms. Adopting an inductive perspective, this research points to an unusual transformation of Latin *intra* as the most likely etymon of the prepositions in question.

Keywords

Romance Philology, Ibero-Romance languages, prepositions, etymology

Contents

1. Characteristics of language tradition. 1.1. Language and linguistic heritage. 1.2. Prepositions and other word classes. 2. Characteristics of limiting prepositions in Ibero-Romance. 2.1. The concept of "until" in Latin. 2.2. The proposed Arabic origin of *hasta*, *ata* and *até*. 3. General characteristics of the class of prepositions. 3.1. Adverbs and prepositions. 3.2. Stress shift in prepositions. 4. Investigation of the etymon of *hasta*, *ata* and *até*. 4.1. Alternative origins of *até*. 4.2. Equivalent forms to *até* in the Cantigas de Santa Maria (CSM). 4.3. A new etymon for *até*. 4.4. Origin of *tra* ~ *ta*. 4.5. Concurrent forms of *intra* ~ *intro* in Ibero-Romance languages. 4.6. The form *fasta*. 5. Research beyond the Iberian Peninsula: the presence of *intus*. 6. Conclusions.

1. CARACTERÍSTICAS DA TRADIÇÃO LINGUÍSTICA

1.1. Língua e herança linguística

Em questões etimológicas, pressupõe-se uma *língua* costumeiramente de duas formas: ou seguindo o modelo schleicheriano da árvore genealógica ou o schmidtiano (ou schuchardtiano), como ondas não-concêntricas que se sobrepõem. Esses dois modelos não são, como se imaginava inicialmente, totalmente antagônicos: a contradição aparente reside na ideia de que, no primeiro, as línguas são entidades discretas e no segundo, fluidas e formam um *continuum* (Viaro 2011). No que se refere aos estudos morfológicos, o modelo da árvore genealógica consegue ser bastante eficiente nas chamadas *classes fechadas*, ao passo que o modelo das ondas é muito mais aplicável às *classes abertas*. Classes fechadas como *preposições*, alguns tipos de *conjunções* e *advérbios*, *pronomes pessoais* e *demonstrativos*, *numerais* e *terminações flexionais* costumam determinar, ainda hoje, a família linguística à qual uma língua se afilia. No entanto, o mesmo não é válido para as classes abertas (*substantivos*, *adjetivos*, alguns tipos de *conjunções* e *advérbios*, *verbos* e a maioria dos *pronomes*), as quais costumam ser bastante inovadoras, a ponto de possibilitarem que neologias sejam presenciadas durante a vida de um indivíduo falante.

Em diacronia, o fenômeno da *herança* não depende do sistema. A herança linguística é, além de um movimento diacrônico, também algo diatópico e diastrático. Não existe, a não ser em modelos ideais, a transmissão integral de uma língua, sem erro de cópia, como um *backup* de um arquivo de computador. O que se transmitem são unidades linguísticas: sons, sistemas fonológicos, regras, palavras e os significados atribuídos, alterados e confrontados com os de outros falantes. Em suma, a transmissão linguística é um fenômeno preponderantemente *memético* (Dawkins 1999). No final, o falante acaba por acreditar que a língua, como um todo, foi passada nos seus primeiros cinco anos de vida. Se a compreensão e o funcionamento da linguagem se submetem facilmente a modelos platônicos, a transmissão dos elementos linguísticos, para além da sua funcionalidade, se explicaria melhor em um modelo, ainda não formalizado, com pressupostos heraclitianos. Nessa transmissão, os dados nos mostram que a *monomorfia* é tão inexistente quanto a *monossemeia*. A situação de um único significante para um único significado, como na Lógica, não tem realidade nos elementos linguísticos. Indivíduos consabidamente trazem em suas heranças linguísticas uma rede complexa e desigual de variáveis (sexo, idade, origem e trajetória de vida, grau de instrução e contato com outras línguas e variantes, ideologia política, religião e muitas outras a ser investigadas). Portanto, quem investiga o étimo de uma palavra deve sempre pressupor a inexistência da monomorfia numa dada sincronia pretérita, para não incorrer em erros grosseiros.

1.2. Preposições e outras classes de palavras

Comparando com a transmissão dos elementos lexicais e dos significados, a adoção maciça de elementos pertencentes a classes fechadas sempre é mais impactante. Como provam as línguas crioulas, a herança de preposições está intimamente ligada a mudanças tipológicas radicais do funcionamento da sentença e à transmissão de relações e de valores, e não a uma renomeação sem grandes consequências, como ocorre com as classes abertas. A forma de elementos de classes fechadas, sobretudo as preposições, costuma manter-se, reforçar-se ou desaparecer (quando cede lugar à sintaxe), mas não é substituída facilmente. Observa-se empiricamente que têm sobrevivido e uma maior fidelidade formal à origem do que os elementos de classes abertas. Isso não quer dizer que preposições nunca se alterem. Quando se tornam demasiadamente ambíguas, costumam reforçar-se por meio da aposição de outras e, se a forma original se desgasta, seu material fônico pode aumentar por meio da gramaticalização de substantivos, advérbios

e participios. O termo *gramaticalização* é entendido, neste trabalho, como em Heine & Kuteva (2002:2-5).

Esse caráter muito conservador da forma das preposições (maior, por exemplo, que dos numerais) é provado pelo fato de raramente serem empréstimos bem-sucedidos: o romeno apresenta no linguajar administrativo até o século XIX uma série de preposições de origem eslava (*bez, na, ot, po, za*), as quais sumiram em seguida e nunca foram adotadas pela língua popular (Avram 1973). O mesmo ocorreu com a preposição grega *katá*, usada em latim eclesiástico *cata* na linguagem comercial. Uma vez popularizada, essa palavra mudou de classe morfológica (Väänänen 1988³:206-207). De fato, nas línguas ibero-românicas, *cada* é um pronome indefinido e não uma preposição. Em português, assim como em todas as línguas românicas, o conjunto de preposições vem integralmente do latim, como prevê o modelo schleicheriano (Viaro 1994). O fato de as preposições serem especialmente imunes ao empréstimo permite a classificação do português como língua neolatina: *a* < lat AD; *ante* < lat ANTE; *após* < lat *AD POST; *com* < lat CUM; *contra* < lat CONTRA; *de* < lat DE; *des(de)* < lat *DE EX (DE); *em* < lat IN; *entre* < lat INTER; *para* < lat *PER AD; *perante* < lat *PER ANTE; *por* < lat PRO ~ PER > port ant *per*; *sem* < lat SINE; *sob* < lat SUB; *sobre* < lat SUPER; *trás* < lat TRANS.

2. CARACTERÍSTICAS DAS PREPOSIÇÕES LIMITATIVAS NO IBERO-ROMÂNICO

2.1. O conceito de “até” em latim

O latim, para a noção alativa, além do acusativo indo-europeu (como em *ibo rus* ‘irei ao campo’), valia-se dos reforços preposicionais *in ~ ad*, que podiam especificar o modo do movimento (respectivamente: para dentro de algo ou, mais vagamente, para perto de algo). Essa língua não dispunha nem de uma preposição específica para marcar somente a direção (como galego *cara a*, espanhol *hacia*, francês *vers*, romeno *spre ~ către*), nem para limites temporais e espaciais (como português *até*, galego *ata*, espanhol *hasta*, italiano *fino*, romeno *până*).

Para reforçar tais aspectos específicos do movimento, o latim clássico usava alguns advérbios associados à preposição. Indicava-se a ideia de limite frequentemente por meio do advérbio *usque*, que podia designar tanto o ponto de partida (“desde”) quanto o de chegada (“até”), dependendo da orientação específica da preposição seguinte: *usque a mare supero Romam proficisci* (Cíc *Clu*, 192); *usque ex ultima Syria* (Cíc *Verr* 5, 157); *hoc malum usque ad bestias pervenit* (Cíc *Rep* 1, 67); *usque in Pamphyliam* (Cíc *Pomp* 35). O mesmo advérbio também se encontra gramaticalizado e transformado em preposição regendo o caso geral acusativo: *usque Siculum mare* (Plin 3, 75); *usque tempora Alexandri* (Just 2, 4, 32). O advérbio *usque* se conservou apenas na antiga Gália e, em parte, no sardo, quase sempre reforçado (Wagner 1960-1964):

lat (INDE-)USQUE(-AD) > fr (en)jusque; prov enjusca; log usce ~ usca; camp usce ~ usca ~ isca ~ isce ~ isci.

Semelhante função tinha a posposição *tenus* seguida de genitivo ou ablativo, embora mais rara: *labrorum tenus* (Lucr 1, 940); *Tauro tenus* (Cic *Dej* 36).

No latim falado de outras regiões, o substantivo *finis* se gramaticalizou e substituiu em grande extensão o advérbio *usque*. Vemo-lo, sob a forma *fine* ou *fini* seguido de genitivo no latim literário, já em algumas construções pré-clássicas: *radicibus fine* (Cat *Agr* 28, 2); *osse fini* (PI *Men* 859); *fine genus* (Ov *M* 10, 536). Sobrevivências de *fine ~ finis* são testemunhadas em toda a Península Itálica, incluindo as ilhas, sul da França (na zona provençal) e atingiu a Península Ibérica:

lat (IN ~ DE ~ PER-)FINE/IS(-AD) > catalão *fins*; provençal *fin ~ fins ~ in fin*; friulano *fin*; romanche *fin* (engadino *infin ~ infin in*, sobremirano *fignen ~ anfignen*, subselvano *enfin*), fassano *fin*, italiano *fin ~ fin'a ~ fino ~ infín'a* (piemontês *fin-a*, lombardo *fin ~ fina ~ de fin a*, vêneto *fin ~ in fine a ~ enfina ~ defina ~ define a ~ infine a ~ perfine a*, juliano-dálmata *fin ~ infín ~ infina*, romanholo *fèna ~ infèna ~ infèma*; genovês *fin ~ finn-a*, lígure *infina*, toscano *infino a ~ per fino a*, umbro *fine a ~ enfino a ~ fima a*, senês *enfine a*, bolonhês *fén*, calabrês *finu ~ figna*, apuliano *fini a ~ figna*, napolitano *nfi a ~ fi a*, abruzzês *nfi a ~ fi a*, lucano *figna*, siciliano *fini a*); sardo ant *fini ~ infini ~ fina* (logudorês *fina*, campidanês *finza*).

A região da atual língua romena e dialetos valia-se provavelmente de outra forma de expressar a noção de “até”. Propõe-se costumeiramente que o valor limitativo de *ad* tenha sido reforçado com outro advérbio (*pæne* “quase”), mas as razões semânticas não estão claras nem há paralelos desse uso nos testemunhos escritos. Não se costuma, apesar da semelhança, associar o *până* romeno ao étimo *fine ad* (apesar de haver na Itália Meridional testemunhos de *f* > *p*- em calabrês, campano e apuliano):

lat [?]*PÆNE-AD* > rom *până ~ pără ~ pon ~ pire ~ pir*.

Na Itália, também há uma transformação irregular antiga *f* > *s* que é amplamente testemunhada: fassano *scin ~ enscin*; juliano-dálmata *sin ~ insin*, romanholo *sèna ~ sèna a*; toscano *sino*, antigo fiorentino *sino*, genovês *scin ~ scinn-a*, calabrês *sinua*, napolitano *nzi ~ nsino a*, siciliano *sin'a* (Rosamani 1990; Ercolani 1994; Fierro 1995; Amato & Pardo 1997; Tajina 1998).

2.2. A proposta da origem árabe para *hasta*, *ata* e *até*

As preposições *até* do português, *ata* do galego e *hasta* do castelhano, todas com o mesmo valor limitativo (Ferreiro 1995, 1997; Álvarez & Xove 2002), aparentemente seriam uma espécie de exceção. Diferentemente das demais, essas preposições não têm um étimo claramente observável a partir do latim. A forma mais antiga nos manuscritos, nas três línguas, é *ata*, a qual já vemos nas Glosas Emilianenses (c900-950), nas quais o *donec* latino corresponde à indicação *ata quando*.

Desde muito cedo, optou-se pela explicação da origem árabe dessas preposições (Seibold 1888:40): a forma *ata*, comum às três línguas nos primeiros textos, teria como étimo o advérbio árabe. À primeira vista, de fato, existem semelhanças fônicas, semânticas e sintáticas. Reforçada pela presença árabe na Península Ibérica e pela inexistência de alguma preposição semelhante fora do mesmo território, aceita-se esse étimo a despeito da já mencionada ausência de empréstimos na classe das preposições, sem que haja uma investigação que explique o seu surpreendente sucesso até mesmo em regiões onde o domínio árabe foi muito tênue ou mesmo inexistente.

As alegadas semelhanças, porém, são, na verdade, superficiais e não resistem a uma investigação fonética mais rigorosa. O som de *h* é uma faringal surda em árabe [ħ], distinta da larinal *h* [h] e da velar *h* [x]. O som [ħ] sistematicamente se transformou em *f* nas línguas ibero-românicas: árabe *al-hawz* > port *alfoz*, árabe *al-ḥājah* > port *alfaia*. Uma grafia como *hata*, considerada rara por Menéndez-Pidal (1926) em testemunhos antigos (Guadalajara 1098 “*hata* val de Vacas”) dificilmente representaria uma aspiração inicial, mas apenas uma variante gráfica de *ata*. De fato, o *h*- inicial nas grafias muito antigas muito provavelmente não representam nenhuma pronúncia, da mesma forma que a grafia *-tt-* dupla (como no valenciano antigo *hatti*) não é uma reprodução fiel de *-[tt]-*, mas de *-*[t]-*, pois é muito improvável que o hispano-árabe pronunciasse tais consoantes como geminadas (Corriente 1983, 2008).

Testemunhos de *fasta* e *fata*, contudo, existem, já no século XIII, o que comprovariam a transformação $h > f$ inexistente em *ata*. Pensou-se que a dissimilação se efetuou inicialmente por meio de um som interdental ($-tt- > -*[θt]- > -st-$). A situação intermediária seria grafada como *adta*, testemunhada já no século X e XI (como em 945 Cardeña Oríg. p. 374 “via qui discurrir tras Buetrone *adta* lacuna de Vitagan”), mas é mais provável que $-dt-$ seja uma grafia pseudo-etimológica para o som $*[t]$, inspirada no *ad* latino. Além disso, embora haja as grafias *asta* e *hasta*, inexistente a forma **fadta*.

Para aceitar o étimo árabe das formas *ata*, *asta* e *adta*, seria necessário imaginar, além da transformação $h > f$, também a aspiração e a aférese dessa consoante ($f > h > ø$) não só no castelhano, mas em todo o norte da Península Ibérica, contudo, essa aférese é regional, típica do castelhano, e, além disso, tardia (começa por volta do século XIV e só se completa no Renascimento, cf. Maia 1995:855-859; Penny 2004:599), pois é devida a substratos e adstratos específicos da região castelhana. A aférese do *f-* não teria, de modo algum, ocorrido tão cedo e atingido áreas tão remotas quanto o galego e, menos ainda, o português. Além disso, resquícios do *f-* tampouco são conhecidos nos dialetos meridionais e provavelmente não existam. Também seria anômala uma aférese diretamente do árabe sem a transformação em *f* (isto é, $h > ø$). Esse é um argumento de fonética histórica bastante forte contra o étimo árabe de *ata* a despeito de sua “semelhança” com *hattà*.

Ainda mais problemática, do ponto de vista fonético, é a solução do $-tt-$ árabe no castelhano. Se o som $*[tt]$, de fato, tivesse chegado ao norte da Península Ibérica, o mais corriqueiro seria uma degeminação ($-tt- > -t-$) e não uma dissimilação ($-tt- > -st-$). Ao afirmar-se que *ata* e *fasta* ~ *hasta* têm o mesmo étimo árabe, imagina-se que algumas áreas preservaram o $h-$ (sob a forma de *f-*), outras conservaram a geminação de $-tt-$ (sob a forma de $-st-$), enquanto outras não conservaram nenhuma das duas formas (e esse seria exatamente o caso de *ata*). Tudo isso é bastante controverso.

Usando a terminologia da Cladística, podemos dizer que, diante da análise cuidadosa desses caracteres, há mais argumentos para acreditarmos que *ata* seja a forma mais antiga (grafada *ata*, *hata*, *adta*) e que *fasta* seja uma inovação (ou uma apomorfia, cf. Amorim, 2002). A grafia $-d-$ de *adta* deve vincular-se à imitação gráfica de cultismos latinos iniciados pelo préverbo *ad-*. Se o $-d-$ de *adta*, de fato, fosse interpretado como $*[θ]$, estaríamos diante de casos como o do lat ACCEPTOREM $> azztore$ (941) $> adtor > aztor \sim azor$ (port. *açor*), mas, se isso tivesse ocorrido, o resultado final teria sido outro: cf. *hattà* $> *fazta > *faza$, ou seja, convergeria com a preposição *faza*, de significado distinto, sobre a qual falaremos oportunamente. A tese de que $-st-$ se desenvolveu reacionariamente para impedir essa convergência com o *faza* homônimo supõe que *hattà* tinha a pronúncia geminada, o que é improvável. Além disso, a necessidade de impedimento de convergência de **faza* com *faza* é bastante questionável, uma vez que convergências formais de preposições de étimos distintos são fenômenos comuns na România (cf. lat PER $>$ port *per* = port *por* < lat PRO, cf. Malkiel 1979; Riho 1979). Conclui-se que uma solução *hattà* $>$ *adta* $>$ *hasta* é bastante problemática. Para provar a possibilidade de $-tt- > -*[θt]- > -[st]-$, Menéndez-Pidal (1926) cita exemplos da linguagem coloquial, em que $-tm-$ $>$ $-[ðm]-$ e $-tl-$ $>$ $-[ðl]-$ (respectivamente em *atmósfera* e *atlas*), mas nesse caso, além de se tratarem de cultismos, estão diante de consoantes sonoras. Menéndez-Pidal cita ainda outras soluções, provavelmente regionais, a saber, $-[θm]-$ e $-[sm]-$. Seja como for, a passagem $-*[θt]- > -[st]-$ não parece mais problemática que a passagem $-tt- > -*[θt]-$.

Mais difícil de entender é *asta*, sem o *h-* em textos do século XI, embora bastante rara e em textos em latim romanceado, como em 1036 Liébana Cart. p. 100: “et facimus ad tibi cartam de illa *asta* dies que uisquieres”. Mesmo nesse caso, se a origem árabe estivesse certa, teríamos de levar em conta novamente os mesmos problemas acima: a dificuldade de explicar a queda precoce do *f-* < ár *h-* e a controversa solução $-s-$ para a dissimilação do $-tt-$. Podemos, contudo, provisoriamente desconsiderar essa forma *asta* e imaginar que seja um mero problema de crítica textual a ser investigado com mais detalhamento pela Filologia, para o entendimento da argumentação que se segue.

Nas línguas modernas, formas ibero-românicas com *f-* são raras hoje em dia, mas testemunha-se no asturiano *fasta*, ao lado de *basta* ~ *hasta* ~ *hata* (Prieto García 2004). Nesse caso, o *h-* é pseudo-etimológico e também textos galegos apresentam a antiga grafia *hata* em vez de *ata* (Cuveiro Piñol 1876 s.v.). Conclui-se que, se o *f-* não vem do árabe, será necessário explicar esse som labial em formas tanto do passado quanto do presente, o que faremos oportunamente.

Por fim, a distinção das vogais de *ata* e *até* é explicada um tanto livremente por um fenômeno muito comum nos dialetos modernos da língua árabe, conhecido como *imālah*, como faz Steiger (1932). No entanto, *até* é muito raro nos textos antigos (nem há formas como **haste* no castelhano), embora encontremos um enigmático *adte* em textos do século XI:

(d1050) et jactarunt eum jn terra *ad te* suos pedes de illo abbatte (Bezd. Orig. p. 26);

(1092) leuetis totam uestram causam ad caput de viii^o dies, que leuetis uestras portas et uestros uscios et illas kasas et illas bardas non desfacatis que *adte* capo de anno (Sahagún, 650^o Orig. p. 371)

Abandonado o étimo semítico, torna-se igualmente difícil entender a rara forma *fata*, que parece vir diretamente do árabe *ḥattà*, sem grandes complicações. Tal forma aparece, por exemplo, em “*fata* acabamêto da sua vida” (Cr. Gal. 239bR). Não se descarta, porém a hipótese de um erro de grafia de *fasta*. Por fim, a diferença acentual entre a forma portuguesa e as demais necessitaria melhor esclarecimento, o que se fará no capítulo seguinte.

O argumento sintático a favor da tese arabista é ainda mais fraco. Algumas construções sintáticas com *até* são de fato similares a construções arábicas. Uma oração típica do português moderno com “*até eu sei isso*”, na qual a palavra *até* é usada adverbialmente antes do sujeito, corresponde literalmente ao árabe clássico *ḥattà ‘anā ‘a‘rifu ḥāḡā*. Contrários a essa tese, porém, mostram-se os estudos dos universais. De fato, há construções coincidentes em línguas sem contato, o que pode servir de argumento para provarmos que *até*+SUJEITO em português seja uma inovação independente. Não há pesquisas que retratem a antiguidade dessa construção portuguesa (de fato, esses dados ou não existem por não remontarem ao período árabe ou, menos provavelmente, teriam sido sistematicamente preteridos pela norma escrita). Outras línguas usam suas equivalentes preposições indicadoras de limite, de distintas origens, para enfatizar tópicos oracionais (que se confundem sintaticamente com o sujeito e o objeto direto). Em italiano: “*fino gli amici non vengono più*” (Rohlf s 964, que cita também Alfieri: “*perfino gli adorati cavalli furono da me trascurati*”; igualmente os sinônimos *sino*, *persino* e *sinanche*); como objeto direto: “*ho perduto fin’all’ultimo centesimo*”. Em francês: “*ils ont tué jusqu’aux enfants*”. Em romeno: “*până și privirea crîcenă a învătătorului se îmblînzise*” (Sadoveanu). A ideia básica seria a seguinte elipse mental: *todos, desde x até y, sabem disso* → *até y sabe disso*.

Tudo indica que, entre a palavra árabe e as preposições românicas, apenas haja alguma coincidência de significado, uma vez que, dos pontos de vista fonético e sintático, revelam, perante uma análise mais detalhada, uma semelhança bastante superficial.

3. CARACTERÍSTICAS GERAIS DA CLASSE DAS PREPOSIÇÕES

3.1. Advérbios e preposições

O caso ablativo, em substantivos ou participios, permitia a transição, em latim, entre o nominal e o adverbial. Inicia-se, a partir daí, um percurso de crescente gramaticalização entre advérbios e as preposições altamente polissêmicas. De um lado do *continuum*, portanto, teríamos advérbios (sobretudo de lugar) e, do outro, preposições, algumas delas altamente gramaticalizadas

(como *em*, *de*, *a*), outras menos (como *sem*). Algumas preposições românicas ainda revelam sua origem nominal (cf. cat *damunt* “em cima de”, *davall* “embaixo de”). No meio desse percurso entre o advérbio e a preposição, encontra-se a palavra portuguesa *até*, a qual é simplesmente seguida de um sintagma nominal, como requerem as preposições, mas, não raro necessita de outra preposição, o que revela seu *status* ainda conservado de advérbio (cf. *até a*, sobretudo na variante lusitana). Ao longo da história do latim e das línguas românicas podemos verificar várias repetições desse mesmo processo (Viaro 1994).

Podemos dizer que advérbios, participios e substantivos diacronicamente se gramaticalizam e geram novas preposições. Muitas delas, sobretudo as mais recentes, ainda têm um valor adverbial bastante marcado e necessitam agregar preposições mais polissêmicas à direita (como os seus étimos já revelam ter ocorrido no passado). No caso das preposições românicas marcadoras de limite acima citadas, encontramos:

francês: *jusqu'à, jusqu'en, jusque chez*;
italiano: *fino a, fino in, fino da*;
catalão *fins a, fins en*;
romeno *până la, până în*.

Isso também ocorre fora do âmbito românico. Em alemão, por exemplo, a preposição equivalente a “até”, a saber, *bis*, é comumente seguida de outras preposições: *bis an, bis zu, bis auf* etc. O advérbio seguido por esse elemento à direita, pode ser sentido como uma nova preposição e o ciclo se reinicia: dado um elemento *x*, podemos ter as seguintes situações: (*x*)^{ADV}, (*x*)^{PREP} e (*x*+^{PREP})^{PREP}. Para enfatizar-se o valor adverbial, inversamente, é comum o acréscimo de preposições à esquerda da palavra em questão, ou seja, (^{PREP}+ *x*)^{ADV}, as quais acabam valendo como verdadeiros prefixos (ou prevérbios), cf. *de post* > *depois*.

As preposições latinas mais usadas, tanto à direita, quanto à esquerda como instrumento para a desambiguação são *in*, *de*, *ad* (e, em formas mais antigas, também *ab* e *ex* à esquerda). Muito da argumentação abaixo desenvolvida levará em conta esses pressupostos. Praticamente todas as preposições parecem submeter-se a esses reforços, por exemplo: “com, junto com”, nas línguas românicas, pode dizer-se não só *CUM*, mas, como em bergamasco, *cund* < lat **CUM-DE* (re-tomando assim seu valor preposicional), ou, como em umbro, *ncō* < lat **IN-CUM* (reforçando um valor adverbial).

3.2. Hiperbibasmo em preposições

Uma característica importante da preposição portuguesa *até* é a sua diferença acentual, face ao galego e ao castelhano: enquanto *até* é oxítona, *ata* e *hasta* são paroxítonas.

Aparentemente esse elemento não é decisivo. Argumenta-se que as preposições, mesmo as dissílabas, não apresentam acento fixo, como ocorre com outras classes de palavras, uma vez que se subordinam prosodicamente ao núcleo nominal imediato. Dessa forma, por exemplo, a preposição *para* tem ambas as vogais semi-abertas no português europeu [*peɾɐ*], a ponto de ser considerada, para muitos gramáticos, “a única palavra átona dissílaba”. Embora na pronúncia brasileira, tenda a ser paroxítona [*para*], como no galego e no castelhano, a reconstrução histórica mostra que deve ter havido flutuação de acento, isto é, *[*pa'ra*] existiu, em algum momento, senão não teria ocorrido a síncope do primeiro *-a-* em português falado e no galego: lat **PER-AD* > [*pe'ra*] > *[*pa'ra*] > *pra*. Da mesma forma, no asturiano se diz *hicia* ~ *hiciá* (cf. castelhano *hacia*).

No conjunto das preposições italianas, há algumas cujo étimo também aponta para um hiperbibasmo seguido de reinterpretação sintática (ou metanálise) em outras sincronias: lat *INTRA* > [*in'tra*] → **in tra* → ital *tra*; lat *INFRA* > *[*in'fra*] → **in fra* → ital *fra*.

Esse hiperbibasmo, que se estende para além da Península Itálica, como provam muitos documentos em outras línguas românicas, será de grande importância para entendermos o étimo das preposições investigadas. Abaixo veremos que há motivos para pensarmos que *hasta* e *até* mantiveram seus acentos tônicos originais e, ao contrário do que se pensa, a posição paroxítona do *ata* galego é uma inovação, por tratar-se de um hiperbibasmo de *atá*, forma atestada na língua falada. Também se encontra a forma *até* no galego moderno (Carballeira Anllo 2009; Carballo Calero 1979⁷: 252 nota 193).

4. INVESTIGAÇÃO DO ÉTIMO DE *HASTA*, *ATA* E *ATÉ*

4.1. Étimos alternativos para *até*

Silva Neto (1958), seguindo o raciocínio de José Leite de Vasconcelos, questiona o étimo árabe *ḥatta* no galego-português. O autor enfatiza que essa solução não é suficiente para esclarecer algumas variantes antigas da palavra, como *atêes*. Embora omita sua opinião com respeito ao castelhano, defende que *até* e *hasta* não têm a mesma origem. Com relação ao *hasta*, Corriente (1983, 2008) acredita ser árabe, contudo, com interferência analógica de uma hipotética forma **AD ISTA* (que explicaria, de certa forma, o *asta* já visto). Ambos os autores, apesar de se valerem de étimos latinos, tendem a aceitar que as formas *ata* e *atá* se explicam indiretamente pelo árabe *ḥattà*.

O verdadeiro étimo de *atêes* se encontraria, na argumentação de Silva Neto, no lat *tenus*, advérbio que, como vimos, tem o mesmo significado limitativo de *usque*. Prova disso, segundo seu modo de ver, seria a nasalidade dessa variante, inexplicável pelo étimo árabe, a qual proviria da síncope de um *-n-* intervocálico, pouco depois da nasalização da vogal antecedente. O *a-* inicial seria, como em tantos outros casos, provenientes do *ad* latino. Teríamos, portanto, inicialmente um étimo **AD-TENUS*. No entanto, para que o étimo seja de fato eficiente, *tenus* teria de ser substituído por **TENES* (senão teríamos como resultado das transformações fonéticas **AD-TENUS* > **atêos* e não **atêes*), forma para a qual Nunes (1945³:369) encontra respaldo no latim arcaico, mas que Silva Neto acredita ser analógica, por causa de *fine*. Resumindo, o raro advérbio *tenus* teria tido grande uso no latim falado da zona galego-portuguesa, a ponto de ter-se gramaticalizado e alterado foneticamente sob a forma de um étimo **AD-TENES*: lat *TENUS* > **TENES* → **AD-TENES* > *atêes* (na nossa indicação etimológica, cf. Viaro 2011, distinguimos as transformações fonéticas, marcadas pelo tradicional símbolo >, das transformações morfossintáticas, marcadas pelo símbolo →).

Silva Neto (1958:753 nota 2) considera tal etimologia plausível, apesar de admitir que a preposição *tenus* seja “rara na prosa clássica e só representada no português (que é, como se sabe, uma língua muito conservadora)”. O pressuposto básico para a aceitação desse étimo é o da conservação semântica. O problema desse e de vários outros étimos alternativos, porém, é a sua pouca verossimilhança, devido a um caráter excessivamente dedutivo. A partir de uma amostragem parcial dos dados de uma única língua, chega-se a uma reconstrução (sem qualquer outra prova em outras línguas). Acreditamos que um étimo sem visão integradora dos dialetos e línguas circundantes reforça a pressuposição equivocada (e tardiamente romântica) de que há sistemas estanques em sincronias muito antigas da língua, como se aquela população particular estivesse destinada a formar uma futura língua de um povo com cultura, índole e características próprias. Também é comum, nesse tipo de solução, uma grande parcialidade etimológica, que se alicerça na ideia preconcebida de que *até* e *hasta* tenham étimos completamente distintos sem que se faça uma investigação dialetológica abrangente para prová-lo. Além disso, se o advérbio *tenus* era raro e não testemunhado em outro lugar além do galego-português, aposta-se

numa espécie de preciosismo, o que estimula algum viés ideológico pouco claro. Propostas etimológicas que sofrem de excessiva arealização não pretendem vincular o étimo a uma extensão geográfica maior.

O -s final de *atêes* e sua nasalidade, porém, parecem favorecer o étimo de Silva Neto, no entanto, podem ter outra interpretação. Existem, entre as línguas românicas, muitas preposições e advérbios com -s, algumas com o significado de “até”: catalão *finses*, logudorês *finas* ~ *vinas*, campidanês *finzas* ~ *finzes* ~ *vinzes*.

Esse -s, que ocorre em tempos e locais diferentes, aparece sobretudo quando a palavra é um advérbio pertencente a uma locução conjuntiva: trata-se do famoso -s paragógico românico, o qual é poligenético, sabidamente presente em vários advérbios (especialmente nos temporais), quer por razões etimológicas (lat *cras* “amanhã”, lat *post* > port ant *pós* → *após* ~ *empós*), quer analógicas (lat *dum interim* > port ant *dementre* → *dementres*, cf. castelhano *mientras*; lat *ante* → port *antes*; lat *post* → sardo *pustis*).

Dessa forma, o -s de *atêes* e outras formas permite ser explicado melhor do que a conservação do -s de um *tenus* modificado. Inversamente, a queda do -s é igualmente comum: lat *foras* > port *foras* ~ *fora*, sardo *foras* ~ *fora*.

Trata-se, portanto, de um morfema que reforça a adverbialidade, a qual, de outro modo, seria ambíguo, haja vista que as funções de advérbios e preposições se intercambiam. Tem a mesma função expediente do reforço da preposição à esquerda, já mencionado acima. Obviamente, uma nova forma adverbial se gramaticaliza com o tempo e dá lugar a preposições com -s, num círculo que pode manter pancronicamente a sua vitalidade ou extinguir-se. Prova dessa vitalidade se vê no asturiano **asta* > *(h)*astas* > *hastasa* (> **hastasa* > *sa*). Esse -s, usado em tempos antigos, foi substituído amplamente por *-mente*, que também se gramaticalizou em várias regiões gerando advérbios e preposições bastante longas (juliano-dálmata *finamente* ~ *insinamente*; vicentino *finamentre*, corso *finamente* ~ *sinament’a da*, campinês *finzamènti* ~ *finamente* ~ *finamenti*).

Não só o -s de *atêes* permite explicar-se por meio do -s adverbializador em vez de uma conservação do som final de *tenus*, mas também a nasalidade pode encontrar outra explicação, como veremos no capítulo seguinte.

As passagens de *atêes* a *até*, propostas por Silva Neto, parecem corretas, as quais envolvem transformações como desnasalização, crase e apócope do -s, abundantemente testemunhadas pelo autor em sua rigorosa investigação em manuscritos. No entanto, temos uma interpretação distinta das sincronias reconstruídas anteriores à formação de *atêes*. Outros dados podem mostrar-nos que, de fato, o étimo **ad-tenes* é passível de questionamento e refutação.

4.2. Equivalentes ao *até* nas Cantigas de Santa Maria (CSM)

Para além de sistemas rígidos e ideais há, por vezes, a convivência, num determinado momento e local, de uma diversidade insuspeita de formas, que observamos na polimorfia das preposições equivalentes ao *até* português. A forma mais antiga e mais frequente nos textos antigos é *ata*, sobrevivente no galego moderno, a qual aparece não só no galego-português, mas também no castelhano. Na época das Cantigas de Santa Maria (CSM), a variante *até*, abonada pelo Gran Dicionario Xerais da Língua, como variante galega atual, em textos antigos (apesar de raro) parece ser uma variante meramente posicional (Mettmann 1959-1972):

(1) E o miragre sabudo foi ben daqui *ata* Suz (CSM 329:77)

(2) (...) e tonbar/ no mar foi e mergullar *ate* nos fundamentos (CSM 33:35-37)

Ata tem o valor de preposição no exemplo acima, ao passo que *ate* mantém sua primitiva adverbialidade e, para tal, necessita de uma preposição, como vimos. Em muitos textos observa-se claramente que *até* se forma por meio de uma assimilação regressiva (*ata en* > *ateen* > *atēen*) e que, não raro, há a aférese do *e*- da preposição articulada (*eno, ena*) à direita. Trata-se, portanto, de um fenômeno fonético inicialmente optativo:

- (3) *Ata en Monpesler (...)* mellor dela non havia (CSM 271:14-16)
- (4) E foi e achou-a chêa de farynna ben *ata na cima (...)* (CSM 203: 27-28)
- (5) Como hũa moller (...) non pude abrir as portas *atēen que sse* mãefestou (CSM 98:1-2)

Se o *-s* é, como dissemos, paragógico e predomina em locuções conjuntivas, tampouco é obrigatório, como mostra o exemplo 5 acima (cujas variantes, nos manuscritos T e To, apresentam *ata que*). Silva Neto (1958:760) testemunha assimilações regressivas similares em outros textos: *atoo* < *ata o*, assim como *ateesta* < *ata esta*. Tais soluções tornam desnecessária a explicação da alternância vocálica *ata* ~ *até* como oriunda na *imālah* do árabe. Também a nasalidade de *atēen* aparentemente não requer explicação no *TENES de Silva Neto:

- (6) *atães que* o angeo foi dali partido (CSM 411:101-102)
- (7) *atêes que* un crerizon sandeu furtou un cabrit' en e o comeu (CSM 52:30-31)

Demonstram-se assim a predominância e a antiguidade de *ata*, bem como a derivação de *até* ~ *atêe*. O caso *atães* do exemplo 6, apesar da nasalização, não se vale da assimilação regressiva, que era opcional nesse momento. Tal solução parece mais razoável do que imaginar que formas como *atães* sejam cruzamentos entre o *ata* árabe e as formas setentrionais *atêes* ~ *atê* ~ *atees*, como quer Silva Neto, o qual aposta na Beira como o território decisivo para a formação da presença portuguesa *até*.

Mesmo determinando que *ata* seja a forma mais antiga nas três línguas, ainda não apresentamos solução para o problema de sua origem. Em suma, descarta-se definitivamente o étimo de Silva Neto mas, ao mesmo tempo, não se consegue retornar ao étimo árabe, dada a precariedade na explicação dos fenômenos fonéticos do árabe às línguas românicas. Uma inovação pode sobrepujar numericamente uma forma mais antiga a ponto de obnubilá-la ou mesmo extingui-la. Teria sido esse o caso de *ata*?

4.3. Um novo étimo para *ata*

As Cantigas de Santa Maria revelam outras variantes que podem iluminar o problema. Uma delas, bastante comum, é *ta*, entendida frequentemente como forma aferética de *ata*, melhor dizendo, de **atá*. Por outro lado, nada além da sua frequência menor nos textos impede que vejamos *ata* como inovação, formada por prótese, e *ta* como a forma primitiva (sobre o uso do termo “prótese” em vez de “prótese”, cf. Viaro 2011:132). Nos textos antigos, a forma *ta* apresenta as mesmas variações posicionais de *ata* e encontramos, além dela, *te*, *tê* e *tês*.

- (8) Rey e ben des Compostela *ta* o reyno d'Aragon (CSM A:3-4)
- (9) viinnam pera pobrarem des Jenua *te en* Chartes (CSM 379:21)
- (10) aguarda-lo-ei *tê ena* noite escura (CSM 65:116 T) (mas: To: *te ena*, E: *têe la*)
- (11) *têes que* a foi fazer que sayu do moêsteiro (CSM 94:33 To)

Estabelecida a etimologia *ta* → *ata* → *ata en* > *ate en* → *atêes* perfeitamente abonada no mesmo *corpus* falta, contudo, explicar de onde viria a forma original *ta*. Observemos que, no

aparato crítico, entre as variantes do exemplo (11), encontram-se *tra en que* (manuscrito E) e *tro en que* (manuscrito T). Formas com *-r-* estão longe de serem incomuns nos manuscritos antigos e não se explicam facilmente por epêntese.

Normalmente essa forma *tro*, típica do navarro-aragonês e catalão antigo, é entendida como independente de *ata*, com quem concorre semanticamente (Gifford & Hodcroft 1966). Mas o mesmo não se costuma dizer de *tra*. Estipular que *ta ~ ata* tenha uma origem distinta de *tra* é algo bastante difícil de se sustentar, uma vez que o *-r-* na situação fonológica de segundo elemento de ataque silábico várias vezes tende a sincopar-se. Por exemplo, preposição *pra* > *pa* ocorre em diversos registros do português falado (e línguas crioulas derivadas, além de ocorrer no galego, no próprio espanhol dialetal, asturiano e aragonês). O *-r-* da preposição *entre* também cai no asturiano e no espanhol dialetal (inclusive no galego-português, cf. CSM 37:14, onde temos *ent'*). No napolitano, a perda do *-r-* como segundo elemento do ataque silábico é igualmente comum.

Se *tro* vem de *intro*, por que *tra* não viria de *intra*? Isso pode conduzir-nos a pensar num jogo *tra ~ ta ~ tro* herdado e mais primitivo do que *ta ~ ata* (cf. idêntica interpretação em Lliteras 1992:425-427). Há ainda, como nos exemplos acima, outras variantes nas Cantigas de Santa Maria: com prótese do *a-* (*atra*, *atro*), com preposição à direita (*troa*, *tro en* > *trõe*), com ambos (*atra en*, *atro en*, *atrõe*) e com paragoge de *-s* (*trões*, *atrões*):

- (12) ca falcon *tra* u se farte da caça que á fillada (CSM 366.61-62)
 (13) *tra en que* a foi fazer que sayu do moêsteiro (CSM 94:33; To *têes que*, T *tro en que*)
 (14) nunca comeu nem dormyu, *tro* o mōesteyro viu (CSM 94:76-77)
 (15) tod' encududo foi ben da cabeça *tro ena* verilla (CSM 19:27-28)
 (16) de sa casa fezera un gran portal ben *atro ena* ygreja (CSM 84: 18-19)
 (17) Deus, que viver-me leixou *trões* aqui (CSM 417:22-23)
 (18) (...) viveu mui gran tempo, *trões que* un bisp' y véo (CSM 67:62)

Todas elas seriam cruzamentos? Silva Neto, contudo, julga que sim: para ele, *atro ~ atra* seriam formas misturadas entre *ata* e *tro*, etimologia *ad hoc*, que não explica a sua presença na zona do castelhano, do leonês e do galego: 1175 León “e passa el rio al coto de Uenero *tro a* la presa d'Eragonti” (Menéndez-Pidal 1926).

Diante desses dados poderíamos, contudo, estabelecer as seguintes sequências etimológicas, a maior parte delas encontradas nas Cantigas de Santa Maria:

tra > *ta* → *ta en* > *te en* → *têes*
tra > *ta* → *ata* → *ata en* → *atães*
tra > *ta* → *ata* → *ata en* > *ate en* → *atêes*
tra > *ta* → *ata* (→ *atas*)
tra (→ *atra*)
tra → *tra en*
tro → *atro* (→ *atro en*) (→ *atrões*)
tro → *tro en* → *trões*

Diversos textos do século XIII e XIV (cf. Seco 2003), mostram que essas formações ocorriam com frequência. As que estão entre parênteses no quadro acima, aparecem, por exemplo, em:

- 1265 *atroen* Pááos (Salazar 49:15);
 1282 maldita *atroes* la setima giarazú (Salazar 85: 23);
 1257 *atra* dia de nadal (Salazar 26: 15);

A falta de evidências para **atra en*, **trâes* ou **atrâes* nos faz imaginar uma certa sequência reconstruível de sincronias pretéritas para a Península Ibérica. Na sequência abaixo, representamos por maiúsculas as formas supostamente mais frequentes nos atos de fala. Deixaremos *hasta* de lado, por enquanto:

- [SINCRONIA 1] Originalmente tínhamos apenas *TRO* ~ *TRA*;
- [SINCRONIA 2] Há o surgimento da simplificação do ataque silábico. Nesta época teríamos *TRA* ~ *TA* ~ *TRO*. Não há testemunhos de **to*;
- [SINCRONIA 3] Há o surgimento das formas prostéticas com *a-*. O sistema desse momento seria *TRA* ~ *TA* ~ *atra* ~ *ATA* ~ *TRO* ~ *ATRO*;
- [SINCRONIA 4] Há o surgimento das formas reforçadas com *en*. Podemos imaginar agora variantes como: *TRA* ~ *TRA EN* ~ *TA* ~ *TA EN* ~ *atra* ~ *ATA* ~ *ATA EN* ~ *TRO* ~ *TRO EN* ~ *ATRO* ~ *ATRO EN*. Nesse período também se poderiam computar as formas anasaladas (*tro en* > *trõe* etc.);
- [SINCRONIA 5] Surgimento das formas com *-s* paragógico nas formas mais produtivas e seguidas de *en*: *tra* ~ *tra en* ~ *TA* ~ *TA EN* ~ *TĀES* ~ *atra* ~ *ATA* ~ *ATA EN* ~ *ATĀES* ~ *TRO* ~ *TRO EN* ~ *TRĒES* ~ *ATRO* ~ *ATRO EN* ~ *ATRĒES*. Aparentemente as formas com *-o* cada vez mais começam a ser associadas com o *-r-* que as antecede, fenômeno que se antevê na sincronia 2 (em que há *tro* mas não **to*) e, conseqüentemente, uma crescente falta de associação do *-r-* com as forma com *-a* (que provocará um colapso das formas *tra* e derivadas);
- [SINCRONIA 6] Surgimento da assimilação regressiva nas formas com alta frequência de uso: *tra* ~ *tra en* ~ *TA* ~ *TA EN* ~ *TE EN* ~ *TĀES* ~ *TĒES* ~ *atra* ~ *ATA* ~ *ATA EN* ~ *ATE EN* ~ *ATĀES* ~ *ATĒES* ~ *TRO* ~ *TRO EN* ~ *TRĒES* ~ *ATRO* ~ *ATRO EN* ~ *ATRĒES*. Nesse caso, observa-se que *tra en* não gera **trēe*. Este quadro é bem caracterizado pelas Cantigas de Santa Maria.
- [SINCRONIA 7] Desaparecimento (após diminuição na frequência de uso) das formas com ataque silábico complexo e menor uso das formas nasaladas: *TA* ~ *TA EN* ~ *TE EN* ~ *tāes* ~ *tēes* ~ *ATA* ~ *atas* ~ *ATA EN* ~ *ATE EN* ~ *atāes* ~ *atēes*;
- [SINCRONIA 8] Desaparecimento das formas nasaladas e menor uso das formas não-prostéticas e paragógicas com *-s*, independência das formas com *-é*: *ta* ~ *ta en* ~ *te en* ~ *te* ~ *ATA* ~ *ATA EN* ~ *ATE EN* ~ *ATÉ*;
- [SINCRONIA 9] Desaparecimento das formas seguidas com *en*, aparecimento de formas seguidas com *a*: *ATA* ~ *ATÉ* ~ *ATÉ A*. O antigo *atá* tem testemunhos recentes de existência no português do Alentejo (Pires 1903-1905) e no Ceilão (Dalgado 1900¹). No Oriente, o chamado português noroeste testemunha um *atqui* < **ata que* (Dalgado 1906). As formas *atés* ~ *intés* ~ *intês* são atestadas no concelho de Melgaço (Vasconcelos 1928: 329).

Um dos argumentos para provar que na sincronia 3 o elemento *a-* é prostético (e não herdado) está no fato já comentado de que várias preposições que ainda estão no antigo grau adverbial dependem de preposições mais esvaziadas que os amparem e que se fixam necessariamente à esquerda, assim: lat *TRANS* > *trás* → *atrás*, lat *POST* > *pós* → *após*. Esse *a-* é poligenético, inspirado originalmente em formas do lat *ad* herdado. Outras preposições à direita são *es-* < *ex* e *en* < *in-*: lat *CONTRA* > *contra* → *escontra*, lat *POST* > *pós* → *empós*. O mesmo ocorre, supostamente, com *té* na forma ainda conservada do português popular *inté* (na verdade, **enté*). Isso mostra que o processo de prótese iniciado na sincronia 3 se estende até a sincronia 6 e tinha concorrência com formas em *en-* (fato mais evidenciado em outras línguas românicas, inclusive na Península Ibérica).

Deve-se observar que se *a-* é prostético, trata-se originalmente de sílaba átona, da mesma forma que ainda se conserva em *até*. Desse modo, em estágios mais antigos, *ata* devia pronun-

ciar-se *atá*, com a qual devia oscilar naturalmente com a forma paroxítona *ata*, devido ao hiperbibasmo natural das preposições dissílabas já apresentado. A acentuação na penúltima sílaba, porém, fixou-se apenas em estágios posteriores, provavelmente por analogia com o *hasta* e o *hacia* do castelhano. Da mesma forma, *atra* ~ *atro* pronunciavam-se *[a'tra], *[a'tro], o que afastam a ideia de serem formas híbridas de *ata* e *tra* ~ *tro*. Silva Neto cita: “des ho rrey atoo pyam” (Cancioneiro Geral, vol III, ed. Kausler, p. 473-474). Esse *atoo* deve representar uma pronúncia *[a'to] e não *[a'to], da mesma forma que a combinação entre preposição e artigo *pro* < *pra*+o na pronúncia europeia é [prɔ] e não, como na brasileira, [pru] ~ [pro].

Pesquisas mais detalhadas, em edições confiáveis, contudo, precisam ser feitas para confirmar-se a hipótese da antiguidade de *ta* ~ *tra*. O que pode parecer realmente contraditório às conclusões acima é a presença de supostos *até* sem o contexto fônico esperado em alguns textos muito antigos do galego-português (portanto, anteriores à sincronia 6), como na grafia *adte* já discutida.

4.4. Origem de *tra* ~ *ta* e *tro*

Raramente se investiga dialetalmente toda a Península Ibérica à procura de uma resposta de impasses etimológicos, todavia, o detalhamento de formas em línguas minoritárias atuais podem tornar a discussão bem mais complexa. Mesmo a forma portuguesa popular *inté* carece de uma investigação mais detalhada. Teria *inté* elementos conservadores insuspeitos? Para responder tal indagação, seria necessário entender o que teria acontecido *antes* da sincronia 1. De onde viria o *tra* ~ *tro* originais? Essa pergunta parece mais fácil de ser respondida à luz da Filologia Românica. Como no italiano, em que se formaram, por hiperbibasmo, *fra* e *tra*, respectivamente a partir do latim *infra* e *intra*, as formas ibero-românicas com certeza também haviam passado por semelhante processo fonético a partir de *intra* e *intro*. Em vez de postularmos étimos no árabe ou em etimologias regionais, a única peculiaridade, a nosso ver, nas formas galego-portuguesas, seria uma mudança semântica.

Essa única inovação, por outro lado, é previsível dentro das línguas e dialetos românicos e não precisa também ser imaginada como uma interferência semântica de *hattà* sobre *intra*. Kontzi (1970) defende parcialmente a tese arabista, pois, para ele, o aragonês *enta* tem um sentido estático (“em, perto de”) proveniente do *inda* árabe, enquanto o sentido dinâmico (“para, em direção a”) seria legitimamente românico. Igualmente poderíamos imaginar que *tra* ~ *ta* sejam românicos, mas que o significado de “até” seja árabe. Mas mesmo essa interpretação é desnecessária pois é improvável que *hattà* tivesse também afetado *tro*. A forma latinas *intra* e *intro* (de onde vieram *tra* ~ *ta* ~ *tro*) são advérbios que substituíram a preposição latina *in*, em ambos os sentidos estático e dinâmico (ou, dizendo de outro modo, inessivo e ilativo: “para dentro de” >> “dentro de” >> função locativa genérica) e do sentido dinâmico nascem tanto a especialização direcional (“em direção a”, como castelhano *hacia*) quanto a limitativa (“até o ponto que” ou “de algum ponto até”, como castelhano *hasta*).

O uso do clássico *usque* parece ter permanecido apenas no latim vulgar da Gália, mas foi substituído aos poucos por *finis*. O latim falado introduzido na Hispânia, porém, não traz traços evidentes de nenhuma dessas formas. Ali, o *intra* ~ *intro*, que conserva o sentido inessivo e ilativo em quase toda a România parece ter assumido o valor limitativo, que, aliás, não é exclusivo dessa região. A não aceitação de formas novas como *finis* mostra que há laivos arcaizantes (também presentes em menor grau no galo-românico e no reto-românico) na escolha de *intra* ~ *intro*. Se a fala ibérica era conservadora nesse sentido, por outro, privilegiou a inovação do ponto de vista semântico (esse quadro mostra uma intensa regionalização do latim vulgar, distintamente da uniformidade pretendida por Maurer Jr, 1951). Semelhanças fonéticas e lexicais mostram que um latim pré-clássico com características comuns ao usado na Itália meridional

serviu de substrato para um latim vulgar posteriormente introduzido na Península Ibérica (a confusão entre *l* e *d* em diversos étimos ibéricos, por exemplo, como *laxare* ~ **daxare*, *scala* ~ **scada* também é um argumento a favor dessa hipótese). Provavelmente o extinto romance das colônias no norte da África também se integravam nesse movimento.

O português ainda conserva uma parcimônia arcaica nas forma herdadas de preposições alativas, pois não tem outras além de *a*: *para*: *até* (às vezes *em*). Desse modo, o português não sente necessidade de distinção mais amiadada da especialização direcional acima mencionada, presente nas demais línguas românicas:

galego moderno: *a* ~ *para* ~ *ata* ~ *cara*,
 castelhano: *a* ~ *para* ~ *hasta* ~ *hacia*,
 francês: *à* ~ *en* ~ *chez* ~ *vers*,
 romeno: *la* ~ *în* ~ *pentru* ~ *către* ~ *spre* etc.

No tocante à semântica latina, o romeno conservou melhor a oposição *in*: *ad*, o que nenhuma outra língua neolatina faz com tanta precisão. Se os adstratos e superstratos do romeno (eslavo, húngaro, germânico, turco) atuaram nessa conservação semântica, o mesmo poderíamos pensar do *ḥattà* árabe com relação à predominância do sentido de “até”, da escolha das formas com *a*- prostético na Península Ibérica, etc. Mas se isso ocorreu, o pressuposto necessário e nem sempre provado seria o intenso bilinguismo numa fase talvez tardia da convivência com a língua árabe, no entanto, como já vimos, *ata* é frequente inclusive em regiões onde o contato árabe foi bastante tênue. A atuação do adstrato ou do superstrato, por fim, não parece ser decisiva na escolha da matéria fônica herdada de elementos de classes fechadas, como vimos, mas apenas na conservação ou alteração de valores semânticos e, quando muito, em algumas fixações sintáticas (como mostram os estudos de linguística areal em várias línguas dos Balcãs, cf. Sandfeld 1930).

4.5. Formas concorrentes de *intra* ~ *intro* nas línguas ibero-românicas

Do ponto de vista onomasiológico, a ideia básica presente em “até”, em etapas mais antigas que a sincronia (1), apresentava, nas línguas ibero-românicas, além dos étimos *intra* ~ *intro*, ainda outros.

Como visto, no leste da península e ao longo do Mediterrâneo espalhou-se também a inovação *finis*. Outras formas menos estudadas também existiam. O galego *deica*, o asturiano *dica* ~ *diquia*, o aragonês *diquiá* ~ *dica*, o aranês *dinca* ~ *denquia* ~ *enquia* costumam ser explicadas pelo latim **DE-HINC-AD*. Chama à atenção a semelhança dessa forma com preposições no território italiano com o mesmo sentido: gardenês *nchin* ~ *nchin a*, antigo luquês *di qui a*, lombardo *de qui a* ~ *da qui a*, vêneto *de chi a* ~ *enchi a*, juliano-dálmata *china*, padovano *inchin'a*, buranês *inchina*, presentes também em locuções conjuntivas: antigo lombardo *de chi a che* ~ *de qui a che*, vêneto *de qui che* ~ *inchina che* (Rohlf 1969: 222; Carrera 2007; Bernardi 1994, 1999). Certamente o étimo de ambos os grupos é assemelhado, no entanto, **DE-HINC-AD* não prevê *di qui a*, que leva consigo um *qui* semelhante ao explicado normalmente pelo ** (A)CCU-HIC*, se não se trata, obviamente, de uma analogia.

A presença de *hinc* nesse étimo é decisiva. Também se postula *inde* no étimo do francês antigo *enjusque* e provençal *enjusca* < lat *INDE-USQUE(-AD)*, cf. REW 4368. O mesmo *inde* > *en-* pode ter-se fundido com *en-* < *in*, por exemplo na forma popular portuguesa *inté*. No italiano, também pode explicar algumas ocorrências com *nd-* em áreas que se esperaria encontrar *nt-*, embora haja regiões (lucano, apuliano, salernitano, napolitano) em que as duas formas sejam indistinguíveis, pois o *-n-* promove uma assimilação progressiva idêntica à do grego moderno (e que prova-

velmente também devia ter alguma extensão no sul da Itália em épocas antigas). De qualquer modo, as formas *hinc* e *inde* combinam semanticamente com *usque* (ou equivalentes), tanto como hoje dizemos “daqui até...” ou “de lá até...” (ou, anaforicamente, “desse ponto até...”). Eis o que justifica pensar que as formas italianas e hispânicas equivalentes a *deica* também possam ser independentes.

Resumidamente, podemos dizer que o valor de “até” se representa nas línguas ibero-românicas apenas por *INTRA* ~ *INTRO* e por **DE-HINC-AD*, com poucos traços de *FINIS* e nenhum de *USQUE*. Mas *intra* ~ *intro* também tinham os sentidos inessivo e ilativo conservados.

O valor locativo, contudo, é explorado por outras preposições, cujo molde fônico *-nd-* parecem ser determinantes. Esse molde provém de duas fontes latinas: o interrogativo *unde* e a anáfora *inde*, ambos com valor semântico ablativo. O jogo *-nt-* ~ *-nd-*, comum no sul da Itália, também se encontra na Península Ibérica e em outras regiões italianas. No caso do galego, o molde *-nd-* foi reforçado pela existência de uma preposição *onde* ~ *onda* ~ *ondia* “junto a” < lat *UNDE*-(*AD*). Não se descarta a sobrevivência de *unde* preposicional, amplamente representada com valor locativo na Itália (calabrês *ndi*, siciliano *unni*, corso *nunde*), como também em alguns casos mal explicados de formas com *-nd-* na região atual do catalão e valenciano (o português dialetal *ontre* parece ter sido afetado também por esse molde e talvez o *fons* alicantino, a despeito da opinião de Spitzer 1920).

O latim *INDE*, reforçado por esse *UNDE*, explicam o corso *nde* ~ *inde* ~ *inded* (o lat *UBI* também atuaria no corso em *ndue* ~ *indue*), o piemontês *ind'* (também no corso, milanês, bergamasco, bolonhês, trentino, toscano, pisano, luquês, aretino, umbro, napolitano e calabrês, como também na Suíça, em Poschiavo e Ticino), o lígure *nd'a*, o ístrio *nd'* (como em Roma e na Sicília) e o siciliano *nni* ~ *nn'*. No ticinês, no ístrio, no emiliano e no calabrês meridional convivem *(i)nt'* ~ *(i)nd'* (Angiolini 1897; Bajini *et alii* 1997; Brero & Bertodatti 1993; Brero 2001). Do mesmo modo, na Península Ibérica, o jogo entre *-nt-* ~ *-nd-* se encontra em valenciano antigo nas formas *(h)anda* ~ *danda* (o *h* é apenas gráfico, por semelhança com *hasta*).

Em outras línguas ibero-românicas a presença das preposições latinas *intro* ~ *intra* (seguidas ou não de outras preposições à direita) é evidente no aragonês *tro* ~ *troa* < **INTRO AD*. A forma *troa* também aparece numa miniatura das Cantigas de Santa Maria: “dormeceu logo *troa eno sol rayado*” (CSM 135 T F), que corresponde ao verso que diz: “adormeceu ben *tro eno sol levado*” (CSM 135:85-86). Além de *troa*, o aragonês conta com as formas *enta* ~ *ta* ~ *entro* ~ *ento* ~ *anda* ~ *arda* ~ *anza* ~ *enzio* (Andolz 2004). A forma *ta*, seminal na nossa proposta etimológica, sobrevive modernamente no aragonês com o sentido antigo de “a, para, hacia” (Benítez-Marco 2007). Kontzi (1970) imagina, baseado em Corominas (1954), que o *enta* aragonês venha do lat *INDE* > *ent* → *ent+a* > *enta*.

A forma adverbial latina *intro*, quando conservada na Península Ibérica aparece quase sempre na forma adverbial com valor inessivo e reforço preposicional à esquerda: **DE-INTRO*, donde português e castelhano *dentro* ~ *drento* (como no italiano). Como previsto, o reforço à direita gera locuções prepositivas: português e castelhano *dentro de*, mas em textos antigos *dentro en*. Esse reforço adverbial com *de-* está também na base da preposição romena *din* ~ *dintr-*. A aglutinação do *de-* ocorre em outros advérbios e migram com facilidade de um valor ablativo para um locativo (cf. espanhol e português *onde* < *UNDE* → **DE-UNDE* > *donde*). Não há, porém, provas da existência de **DE-INTRA*, o que aponta para uma maior gramaticalização de *intra* do que de *intro* (daí, também o seu maior desgaste fônico em *tra* ~ *ta*, sem equivalências em *tro* no galego-português).

A preposição *tro* já tem o sentido de “até” no século XI, como confirma o seguinte trecho (cf. Seco 2003) de Orgañá 1080-1095: “que uos ad iuuetis ad mulier mea et ad filio meio per pled et per gera *tro* lor dret los en sia exid”. A forma *intro* aparentemente substitui *finis* em amplas re-

giões entre a Hispânia e a Gália. Entre as variantes para “até” em provençal antigo encontramos *entro fin* (Levy 1973^s s.v. *fin*). A forma *tro* é levada em consideração por Spitzer (1920) como a que teria contaminado o latim *finis* na formação do alicantino *fons* (e o valenciano antigo *finsa* ~ *funsas*). Desse modo, se *tro* adquiriu tão cedo o valor de “até”, por que o mesmo não ocorreria com *tra* ~ *ta*?

Em aranês, que é uma língua galo-românica falada na Península Ibérica, o hiperbibismo acima discutido é bastante evidente: *entà* ~ *entò* ~ *end'* ~ *enda* ~ *endê* ~ *ende* ~ *ena* ~ *na* (o mesmo ocorre ao se passar a fronteira política entre Espanha e França com o gascão *endà* ~ *andà* ~ *andà* ~ *enà* ~ *entò* ~ *to* ~ *entou* ao lado de *enta* ~ *nta* ~ *ta* ~ *enti* ~ *ent* ~ *ende* ~ *enda*, cf. Rohlfs 1935). Observe-se nessas formas a conservação da primeira sílaba de *intra* ~ *intro*. Também **INTRA-DE* se testemunha no aranês *entad* ~ *entada* ~ *tà* ~ *tad* ~ *tada* (também *tad* se documenta no gascão). Mais distantemente, na Península Itálica, está o calabrês *nta* ~ *ntad*, que apresenta a mesma configuração.

4.6. A forma *fasta*

Na Península Ibérica, as formas com *-d-* e *-s-* são anteriores às com *f-* (temos *asta* e *adta*, mas não temos **fadta*, apenas *fasta*). Como dissemos, uma síncope do *-t-* como em *aztor* > *azor* teria gerado **faza*, que é abonado em leonês com outro valor semântico, uma vez que equivale ao castelhano *hacia*, o qual jamais é associado a um arabismo (Malkiel 1979; Díez Calleja 2008).

O único étimo aceito para *hacia* é o lat *FACIEM AD* > **faze-a* > *facia* > *hacia*, o qual tem paralelos também na preposição *cara* em galego (e em castelhano antigo *cara* ~ *carra* ~ *carria*) e no *cap* catalão e aragonês, língua em que se atestam também *capa* ~ *cata* ~ *cant-* < lat *CAPUT* (com possíveis interferências acima mencionadas dos *-nt-* ~ *-t-*), apenas para citar as línguas da Península Ibérica (Viaro 2004). O romeno usa a expressão *față* como preposição que exige genitivo; em italiano *in faccia a* também se testemunha no calabrês *pacce a*, campano *mfacci a* ~ *mbacci a* e no apuliano *mbacć*. Para outras transformações – em várias línguas do mundo – do significado da palavra “cabeça” no valor espacial de “em frente de, para cima”, baseando-se em razões cognitivas, cf. Heine & Kuteva (2002:167-170) e Svorou (1994).

A escolha do étimo *FACIEM* com valor limitativo é uma apomorfia sem paralelos do Ibero-Românico setentrional. No entanto, o jogo *a: ta* ~ *tra* atingiu a mesma área dessa inovação, de modo que podemos imaginar a seguinte situação, para o castelhano, equivalente às sincronias já apresentadas pelo português:

[SINCRONIA 1] Originalmente tínhamos apenas *A* ~ *TRO* ~ *TRA*;

[SINCRONIA 2] Há o surgimento da simplificação do ataque silábico. Nesta época teríamos *A* ~ *tra* ~ *TA* ~ *TRO*. Não há testemunhos de **to*;

[SINCRONIA 3] Há o surgimento das formas prostéticas com *a-* e de formas com *faz-*. O sistema desse momento seria *A* ~ **FAZEA* ~ *tra* ~ *TA* ~ *atra* ~ **FAZT(R)A* ~ *ATA* ~ *tro* ~ *atro*;

Portanto, paralelamente a **faze-a*, imaginemos também uma forma **faz-tra* ~ **faz-ta*, inspirada na alternância *a: tra* ~ *ta* em tempos antigos. De **faz-ta* deriva-se facilmente o *fasta* do castelhano antigo (donde o atual *hasta*), por meio da mesma mudança fonética -*[θt]- > -[st]- prevista para *hattà*. Os testemunhos de *fasta* remontam ao século XI no castelhano (Seco 2003) e os do galego ao XIV, embora somente depois disso o encontremos com mais frequência, sendo ainda mais tardia a aspiração e posterior aférese do *f-* inicial (cf. *TMLG*). As formas com *-s-* em galego falado atual, como em outras áreas da Espanha: *fasta*, que aparece inclusive em textos modernos (cf. *TILG*), (*h*)*asta*, *esta* são encontrados nos dialetos galegos atuais (García & Santamarina 1995, mapa 353), ao lado de (*h*)*astra*, *estra*. A alternância *t* ~ *tr* nessas formas com *-s-* poderia ser entendida quer como uma conservação do já mencionado *tra* (nesse caso, *atra* > *astra* e *ata*

> *asta*, por cruzamento com o *hasta* castelhano), ou como uma inovação (e nesse caso, um empréstimo direto do castelhano *hasta* > *hastra*, pois é bastante comum nas línguas ibero-românicas a alternância *-st-* ~ *-str-*). Além disso, as formas *esta* ~ *estra*, no entanto, apontam para uma oscilação na preposição associada a *ta* ~ *tra* (assim como *inté* no território português), uma vez que, em vez de *a-*, temos *es-* < lat *ex-* (observemos que o étimo, nesse caso, **ex-tra*, não pertence à mesma sincronia do extinto *extra* do latim clássico). Dessa forma, o molde fônico [st], autóctone ou castelhano, tem impacto na formação de várias outras formas supostamente híbridas no território espanhol. No valenciano, gerou *hasda* < **anda*. Esse *hasta* também associa-se com o *de-* adverbializador em *dasta*. Também no aragonês, em vez de *dinca* já citado, por vezes se encontra *dista*, o que mostra a extensão desse novo molde. Isso não contradiz o que dissemos sobre a improvável empréstimo árabe, uma vez que todas essas línguas ibero-românicas são tipologicamente assemelhadas (quer sob o ângulo estrutural, quer sob o filogenético).

Apesar de *fasta* aparecer já em documentos do século XI, a substituição do *ata* ~ *adta* castelhano por *fasta* só ocorre no século XIII. A forma *hasta* é tardia (apenas no século XV, com pronúncia aspirada que sofre aférese aos poucos). Se imaginamos que *h-* não representa som aspirado nenhum antes do século XV, o grafema *h-* nas formas preposicionais estudadas é pseudo-etimológico. O mesmo ocorreria com a ausência de sinais de nasalização em alguns casos. Por exemplo, no galego-português, *atees* poderia representar a forma mais comum *atêes* (“et entra *atees* os ossos” Tratado Albeitaria BCMO XI, 547.31; Va pee *atês* a crastra” Graal I, 276) da mesma forma que *atroes* e *troes* nas cantigas (cf. CSM 64:74 E no lugar de *trões* nos manuscritos *T* and *To*).

Desse modo, *hacia* não estaria isolado no seu étimo por causa da rápida gramaticalização da inovação *faciem* na zona castelhana: a preposição *hasta* também traria o mesmo reforço, cujo uso destronaria *atá* da área castelhana. Além disso, a atração acentual para o substantivo gramaticalizado **fáze-a* e **fáz-tra*, gerou as paroxítonas *hacia* e *hasta* que atuaram analogicamente sobre formas sobreviventes de *atá*, moldando, por hiperbibasmo (*atá* > *áta*), a pronúncia atual do galego moderno padrão.

5. INVESTIGAÇÃO PARA ALÉM DA PENÍNSULA IBÉRICA: A PRESENÇA DE *INTUS*

Passando para a fronteira do galo-românico, onde *intra* ~ *intro* conservam também o seu valor de “(para) dentro de”, há diversos testemunhos da existência de outra preposição sinônima, *intus*. Fora desse território, porém, a queda do *-s* ou a sua paragoge, bem como a oscilação *-tr-* ~ *-t-* dificultam a determinação de um étimo *INTRO* OU *INTUS*. A única exceção que pende para *INTRO* é a existência de um *-r-* que não poderia ter surgido por epêntese em vários lugares ao mesmo tempo, no entanto, o contrário, isto é, a síncope *-tr-* > *-t-*, é bastante comum.

Testemunhos evidentes de *INTUS* estão no provençal *ins* e no francês antigo *enz*. À semelhança de **DE-INTRO*, deduz-se também, nesta área, um **DE-INTUS*, que também penetra a Península Ibérica (cat *dins de*): occitano moderno *dins*, antigo fr *denz* > *dans* → *dedans* (com novo reforço com *de-* à esquerda). O *-s* herdado de *intus* reforça o padrão adverbial até agora apresentado pelo *-s* paragógico e análogo poligenético. Sem o reforço em *de-*, em antigo provençal encontram-se *entro* ~ *entro a* ~ *entro en* ~ *tro* ~ *tro en* ~ *troi* ~ *truei* ~ *entroi* e mesmo *tro enta* (com evidente gramaticalização avançada de *enta*).

A forma **INTUS-IN* explica não só o fr ant *enz en*, mas também sobresselvano *enten* e valáder/puter *aint en*, embora, nesses dois últimos casos, seja possível também deduzir-se um **INTRO-IN*. Igualmente o sobresselvano *enta*, sobremirano *ainta*, valáder/puter *aint a* apresentam similar convergência formal, uma vez que é difícil determinar se vieram diretamente de *INTRA*, de **INTRO-AD* ou de **INTUS-AD*. A última forma, como tantos outros fenômenos, já se encontra no latim pré-clássico: *evoca aliquem intus ad te* (PI *Most* 675). Aliás, a região reto-românica suíça nos traz

algumas construções que corroboram muitos dos pontos aqui defendidos. Encontram-se no romanche suíço, para o sentido de “em, para, até” formas bastante complexas (Bezzola & Tönjachen 1944; Peer 1962; Sonders & Grisch 1970; Vieli & Decurtins 1975, 1995). A mudança semântica de “para dentro de” >> “até” em romanche confirma definitivamente a falta de necessidade do elemento árabe para justificar a mudança semântica de *intra* ~ *intro* na Península Ibérica. Observa-se que algumas formas do romanche indubitavelmente vêm de *intro* e não de *intus* e, ademais, possuem um [-k]- (grafado -*ch*-), que também aparece em advérbios de outras línguas, os quais se explicam pelo étimo *INTROQUE:

Lat *INTROQUE > ital *introcque*;

Lat *INTROQUE-IN > sobs *entochen*; subs *trocen* ~ *tocen*; sobm *antrochen* ~ *antochen*;

Lat *INTROQUE-AD > prov *entroca* ~ *entoga*; sobs *entocca*; subs *antoca* ~ *antroca* ~ *troca* ~ *toca*.

Chama a atenção o fato de o provençal *entroca* e outros, assemelhar-se estruturalmente ao aranês *enquia* e ao veneziano *enchia* (entre outros, já citados), supostamente com um étimo levemente alterado (ou seja, com *intro* no lugar de *in*, postulando-se um *QUE poligenético): *IN-QUE-AD > aranês *enquia*, veneziano *enchia*.

Talvez todos esses étimos necessitem ser refeitos por padecerem do mesmo problema do excessivo regionalismo das soluções para os correspondentes de *até* nas línguas ibero-românicas. Em abruzês *schin* parece valer-se de *ex* em vez de *de* (*EX-QUE-IN) e o mesmo parece aparecer reforçado na conjunção vêneta *deschia che* <*DE-EX-QUE-AD. No ístrio *incheinta* não temos *in* à direita, mas *IN-QUE-INTRA ou *IN-QUE-INTUS-AD.

O elemento *QUE está presente sobretudo na formação de conjunções por meio de advérbios. Podemos resumir essas mudanças de função da seguinte forma: Adv → Prep; Prep+s → Adv → Prep; Adv+*que → Conj → Adv → Prep; Conj+*s → Adv → Prep. Ou seja, como dito, uma mesma palavra como *até* pode ser preposição ou advérbio. Para caracterizar melhor a sua função adverbial, acresceu-se-lhe, às vezes, na história das línguas românicas, um -s. Para o valor conjuncional, acrescentou-se um *QUE. Conjunções, por serem mais recentes que os advérbios e as preposições, tendem também a ser usadas com suas funções. Regras pancrônicas e pantópicas como essas não são incomuns também em fonética histórica (por exemplo: -t- > -d- > [-ð]- > ∅ é representada, em etapas nem sempre completas, em tempos distintos, inicialmente no francês, depois no espanhol e atualmente no português europeu). As formas -s e *QUE tiveram produtividades distintas, diacronicamente (mas também diatópica e diastraticamente) e seus ecos analógicos mais remotos seriam os antigos *intus* e *usque* do latim, preservados sobretudo no galo-românico setentrional. Como vimos, tais formas surgem no galego-português *atães que*, *atães que*, *tães que*, *trões que*, etc. mas também se entreveem no antigo provençal (algumas com evidente metátese e duplicação): *tros que* ~ *troscas* ~ *truesque* ~ *entrogas* ~ *entrusca* ~ *entruscas* ~ *entruscas a* ~ *entrusques que*. Aparentemente, esses advérbios em -ca, -ga, -cas, -gas, -sca, -scas, -que, -ques, -sque, -sques atingem áreas extensas que vão do reto-românico ao ibero-românico, com exemplos esparsos no italo-românico.

O elemento *QUE aparece no apuliano, lucano e calabrês como preposição *fenca* ~ *finca* ~ *fenga* ~ *finca* ~ *finche*, às vezes com a mudança de f- > s- já mencionada (calabrês *sinca* ~ *nzinca* ~ *nsinca*). Em outras regiões mantém o valor de conjunção (calabrês *finca chi*, napolitano *per fin che* ~ *fi che*, abruzês *per fi che*, siciliano *per fina chi*). Na sua forma conjuntiva de “até que”, *finché* se alterna com *sinché* ~ *sí* no antigo lombardo e aparece em obras antigas, como no Decamerão (*non ristette sí fu a casa de lei* 8,2; *mai ristette sí fu in Firenze* 3,9; *non si ritenne di correre sí fu a Castel Guglielmo* 2,2) e em Dante (*che non guardasti in là sí fu partito* Inf 29, 30). No romeno também

temos *pânc* ~ *pâncă*. Não raro, nascem cruzamentos de formas, como campidanês *fisca* ~ *fusca* “até” (em que *finis* se mescla a *usque ad*).

Também nascem daí cruzamentos semânticos. Os descendentes de *intra* deixam o significado inicial “(para) dentro de” e passa a indicar um movimento como “para, em direção de” no aragonês e, por fim, especializa-se no sentido de “até” no galego-português, gerando, desse modo, um percurso semântico previsível e compatível com várias línguas do mundo (como ocorreu, por exemplo, no romanche), no entanto, na Península Itálica, *intra* às vezes se confundiu formalmente não só com *intro*, mas também com *inter* e com *infra*, adquirindo, assim, novos significados: ital *entro* ~ *tra* ~ *tra di* ~ *fra* ~ *fra di* “entre” (cf. *introque* “enquanto isso”); milanês *intrâ* ~ *tra*, triestino *tra* ~ *fra*, apuliano *ntra* ~ *ndra* (também no calabrês, napolitano e siciliano); sardo *francu* “entre” < *INFRA-IN-CUM, cf. Pittau 1991.

Em território italiano encontramos, INTUS ~ INTRO ~ INTER ~ INTRA ~ INDE ~ UNDE ~ *IN DE com o sentido de “dentro de, em”, seguidos ou não de novas preposições à direita ou à esquerda. Uma história de cada uma das formas citadas mereceria ser amplamente debatida, no entanto, apenas apresentaremos aqui um quadro que mostre o quão extensa é essa distribuição de preposições nas línguas românicas e a falta de sentido de propostas etimológicas que não levam em conta essa ampla distribuição, já presente no latim falado.

No área reto-românica italiana, vemos essas preposições no ladino *t-* ~ *te-* (ampezzano *inze*, o fodom *nt-*, marebano *ta-*). Na anfizona ladina, em Cibiana do Cadore, registra-se *inthe* (Croatto 1986; Col 1991; Menardi 1997; Rossi 1999). Também o friulano tem formas como *int-* ~ *inta* ~ *ta* ~ *t-* no sentivo locativo (Nazzi & Nazzi 1995). Na área italo-românica, entre outros dialetos (Pinguentini 1969; Heilmann 1978; Romito 1985; Toso 2006; Vitali & Lepri 2009), citem-se:

piemontês *inter* ~ *té* ~ *enta* ~ *ənt* ~ *ant* ~ *nt'a* ~ *nt-* ~ *nd-*;
 lombardo *entro* ~ *int(r)e* ~ *int* (milanês *int* ~ *in d*, cremonês *in de*, bergamasco *int* ~ *ind*);
 vêneto (*i*)*nt'* ~ *t'* (juliano-dálmata (*i*)*nt-* ~ *nt-* ~ *t-* ~ *ind-* ~ *nd-*; padovano *int*, triestino *int'*);
 marquesano *t-* ~ *nti* ~ *nte* ~ *ntro*;
 romanholo *ent(r)o* ~ *int* ~ *t-* (emiliano *ind*, bolonhês *int* ~ *ind*, samarinês *at* ~ *t-*);
 genovês *inte* ~ *int-* ~ *int'in*;
 toscano *entro* ~ *intra* ~ *inn'*;
 corso *ind'è* ~ *nde* ~ *nded* ~ *ndu*;
 romanesco *nd'*;
 umbro *entra* ~ *nto-* ~ *int-* ~ *nt-* ~ *ta-* ~ *t-*;
 napolitano *int(r)o* ~ *d(r)into a* ~ *pe dint* *a* ~ *indo* ~ (*d*)*int'a* ~ *ntra* ~ *ind* ~ *nt'*;
 salernitano *inda*;
 lucano *nd'* ~ *nn'*;
 apuliano *ind'a* ~ *jinda* (barês *ijnde* ~ *ijnde a* ~ *ijnd'a* ~ *nda*, salentino *ntra*);
 calabrês *dintru a* ~ (*i*)*ntra* ~ *nta(d)* ~ *nt-* ~ *nti* ~ (*u*)*ndi* ~ *ind'*;
 siciliano *intra* ~ *ntra* ~ *nta* ~ *ta* ~ *nt-* ~ *nd-* ~ *nni* ~ *n'a*.

Na Itália, inclusive na Sardenha, *-nt-* se encontra também dialetalmente nas formas com *-mente* já citadas, que talvez tenha sido reforçadas pela moldes fônicos com *-nt-* de INTRA, INTRO após FINIS ~ *SINE ou mesmo com *DE-HINC-AD:

juliano-dálmata *cheinte* ~ *feinta* ~ *finta* ~ *incheinta* ~ *inchinta* ~ *infinta*, toscano *sinenti*, Lácio *sinent'a* ~ *nsinent'a*, romano *insinent*, assim como marquesano e umbro *finent'a* ~ *finant'a*, marquesano *finente*, abruzês *finant*.

Se *intra* e *intro* aparentemente só desenvolveram valores alativos e locativos, a associação com *finis* é bastante eloquente na transição de um significado para o outro, tal como postulado neste trabalho para a Península Ibérica (cf. o curioso *assinates* Hilty 2005). O friulano dispõe de um interessante *fintremai* com o sentido de “até” (também no juliano-dálmata *fintamai* ~ *fintanamai* ~ *fintanemai*), no qual se entrevê claramente a palavra latina *intra*. No sardo, o logudorês testemunha as formas *intra* ~ *intro*. No dálmata, documenta-se a forma *ent* (Bartoli 1906). No romeno, além das formas preposicionais *intri-*, *intru-* (presente também em *pentru*), encontram-se advérbios reforçados por preposições apostas à esquerda (como ocorre com *dentro*): *înăuntru* ~ *năuntru* ~ *înlăuntru* ~ *înuntru* ~ *înluntru* < lat **IN-AD-INTRO* (algumas formas com analogia com a preposição *la*). No arromeno há *ntru* ~ *ntu* ~ *tu* ~ *tra*; no megleno-romeno testemunham-se formas como *tri* ~ *tra* ~ *tru*.

Seria oportuno observar que, quando o *in-* não cai por meio de falsa segmentação, toda a sílaba se mantém ou apenas o *n-*, mas nunca há síncope do *-n-* sem haver queda do *i-*. Desse modo o *-n-* depende do *i-* na formação da sílaba *in-* como um único elemento significativo da etimologia (ou um ‘caráter’ na linguagem da Cladística), ou seja, há formas com *int-* (*ent-*, *ant-*, *end-*), com *nt-*, com *t-*, mas não apenas com **it-*. Daí imaginarmos que a prótese *t-* > *at-* esteja correta tanto nas línguas ibero-românicas estudadas quanto em samarinês. O caso de *inté* permite dupla interpretação, como visto acima, uma vez que a nasalidade pode ser herdada ou simplesmente é uma nova prótese, o que é mais plausível, uma vez que se supõe que a vogal *-é* em vez de *-a* seja uma inovação, embora o piemontês *té* talvez derive de um hiperbibusmo de *inter* < **intr* < *INTRO*, nos outros casos, a vogal *-e* equivale, na maioria das vezes, a mudanças fonéticas independentes de *-a* (seria o mesmo caso do antigo e inexplicado *adte*?) ou a um étimo *INTER*, cf. ladino *te*, ampezzano *inze*, talian *te* ~ *n-te*, genovês *inte* e possivelmente megleno-romeno *tri*.

A posição do *t*, seja em início de sílaba, seja após coda consonantal, é favorável à conservação. Para além do território italiano meridional, gascão e catalão, há, como visto acima, várias outras regiões na Itália em que a sonorização do *-t* > *-d-* se efetuou, provavelmente independentemente, por causa de uma assimilação parcial progressiva do *-n-* que o antecede. Dialetoalmente o *-n-* também acompanha essa mudança com uma desnasalização (*-n-* > *-r-*), como em aragonês *anda* ~ *arda*. No antigo valenciano, essa desnasalização é afetada pelo *-s-* do castelhano *hasta*: *anda* ~ *danda* ~ *hasda*. Também, de maneira independente (a saber, no aragonês *anza* ~ *enzio* e no ampezzano *inze*), o *-t-* sofreu lenização para *-[θ]-*, grafado como *-z-*.

6. CONCLUSÕES

Kontzi (1970) defende a tese de que palavras que não refletem elementos culturais (no caso por ele estudado, as preposições) não têm sua forma afetada pelo bilinguismo, embora seu significado seja maleável. Acredita que o significado estático de “para, em direção a” do antigo aragonês *enta* seja românico, mas que o estático de “em, perto de” tenha tido alguma influência do árabe *‘inda*. No entanto, a não ser em casos evidentes, é muito difícil determinar se significados e formas provêm ou são intensificados por outra língua tipologicamente tão distinta numa situação de bilinguismo, mesmo quando essa situação se prolongue temporalmente e inclua grandes extensões. Por exemplo, vemos que o étimo do *hacia* espanhol equivale, como se fosse um decalque, ao do *cara ta* aragonês (Benítez-Marco, 2007: 16). O castelhano, porém, valendo-se de *ta* concorrente de *a* poderia ter criado um **faz ta* a partir do mesmo *faz-* que gerou *hacia*, uma vez que *ta* ~ *tra* eram onipresentes na România. O aragonês dispõe em seu leque de preposições, também de um *dica* para o sentido de “até”. No galego-português, leonês, castelhano, *tro*, *tra* e *ta* conviviam com *a*, *para* e, provavelmente, com alguma forma extinta, da mesma tradição

formal do *deica* galego. Essas formas *ta* ~ *tra* ~ *tro* entrariam em concorrência e, por fim, desapareceriam se não tivessem recebido reforços sintáticos e especializações semânticas. As formas reforçadas *a-tá* ~ *a-trá* ~ **fáz-ta* passaram a especificamente marcar limite, enquanto *fázi-a* ~ *cara a* se especializaram na marcação de direção. Ao mesmo tempo, *tro* e derivados desapareceriam e a presença de *deica* se restringiria apenas à variante falada do galego.

É fato que formas se conservam em terrenos distintos. Por exemplo, a expressão de finalidade de *pramode* no português coloquial (inclusive brasileiro) se encontra em galego (*por mor de*), no aranês (*pr'amor de*, *per'mor de*) e em dialetos italianos (lombardo *per mor de*, bolonhês *pr'amour de*, ístrio *per amur de*, antigo vêneto *permorde*). Não seria de fato estranho que *intra* desaparecesse em vastas áreas na Península Ibérica justamente onde teria atuado um arabismo como *hattà*? A transição semântica “para” >> “até” não é incomum, pois ocorreu também em romeno *până*, caso venha, de fato, de um *ad* reforçado por *pæne* (outras soluções remetem a idêntica solução, por exemplo: **PER-AD* > **pâră* > *până*, com nasalização semelhante a *SERENUM* > *senin*).

Diante desse amplo leque de formas que apresentamos, a indagação real seria: que vem a ser de fato o elemento semântico de uma preposição? Se formas latinas se conservam tão bem no aspecto formal, o mesmo não se pode dizer do significado. Por exemplo, a preposição *lat de* praticamente concentrou todo o valor original ablativo e incluiu valores relacionais antes atribuídos ao caso genitivo. Não se imagina que um substrato tenha sido responsável por isso, dada a sua extensão territorial (de fato, apenas o romeno prefere ainda uma construção com dativo-genitivo no lugar do uso vulgarizado do *lat de* seguido de ablativo ou do acusativo lexicogênico).

Por que teria sido o árabe o responsável pela modificação do significado de *intra*? A frequência de uso de *ta* ~ *tra*, que inicialmente significava apenas “para”, devia ser grande e não era conflituosa com *hattà* nas zonas bilingües. Se a palavra árabe foi a responsável pela criação da especialização não é possível provar, mas isso é bastante duvidoso em áreas em que o contato árabe não era intenso. Contrariamente a esse pressuposto, a mesma derivação semântica se vê no romanche. Quando *ta* ~ *tra* (na forma *ata*) começa a competir com *deica*, houve a arcaização paulatina dessa última forma. Se deixarmos de lado *INTRO*, veremos que a Península Ibérica tinha inicialmente herdado do latim as preposições alativas *AD*, **PER-AD*, **DE-HINC-AD* e *(*IN*)*TRÁ*. A princípio, o galego-português tinha as formas *a*, *pera*, *deica* e *tra* daí nascidas. A forma *ta* e *atá*, de mesma origem que *tra*, também são antigas, por serem comuns ao castelhano. Só mais tarde surgiram formas derivadas de *cara* ou *faciem*, do qual o português não participou e ainda não participa, mas apenas nesse momento é que surgiria o *fasta* no castelhano, por um processo sintático equivalente ao utilizado hoje pelo aragonês *cara ta*. O decréscimo do uso de *atá* no castelhano é que fez surgir a preferência de *hasta* < *fasta* e a sua cisão com o grupo ibero-românico ocidental. Por fim, a mudança acentual de *ata* no galego é ainda mais recente, posterior à mudança de *a* > *e*, que gerou o português *até*.

AGRADECIMENTOS

Agradeço à profa. Dra. Safa Alferd Abou Chahla Jubran, da Universidade de São Paulo, pela tradução da frase e confirmação dos usos de *hattà* em dialetos modernos falados da língua árabe (sírio, libanês, palestino e egípcio). Ofereço meus agradecimentos também aos pareceristas anônimos da revista ELG que contribuíram sobremaneira para a aprimoração deste artigo com suas observações muito relevantes.

REFERÊNCIAS BIBLIOGRÁFICAS

- Amato, Bruno/ Anna Pardo (1997): *Dizionario italiano-napoletano, napoletano-italiano*. Milano: Antonio Vallardi.
- Álvarez, Rosario / Xosé Xove (2002): *Gramática da lingua galega*. Vigo: Editorial Galaxia.
- Amorim, Dalton S. (2002): *Fundamentos de sistemática filogenética*. Ribeirão Preto: Holos.
- Andolz, Rafael (2004): *Diccionario aragonés*. Zaragoza: Mira.
- Angiolini, Francesco (1897¹/1967): *Vocabolario milanese-italiano coi segni per la pronuncia preceduto de una breve grammatica del dialetto e seguito dal repertorio italiano-milanese*. Bologna: Forni.
- Avram, Mioara (1973): "Prepoziții neologice în limba română contemporană", *Studii și cercetări lingvistice* 24 (3), 239-253.
- Bajini, Alessandro et alii (1997): *Dizionario italiano-milanese, milanese-italiano*. Milano: Antonio Vallardi.
- Bartoli, Mateo G. (1906): *Das Dalmatische: altromanische Sprachreste von Veglia bis Ragusa und ihre Stellung in der Apennino-Balkanischen Romania*, 2 vols. Wien: Alfred Hölder.
- Bernardi, Rut (1999): *Curs de gherdëina: trêdesc lezioni per mparé la rujeneda de Gherdëina - drei-zehn Lektionen zur Erlernung der grödnerischen Sprache*. San Martin de Tor: Micurà de Rù.
- Bernardi, Rut et alii. (1994): *Handwörterbuch des Rätoromanischen*. 3 vols. Zürich: Offizin.
- Benítez Marco, María Pilar (2007): *Pos ixo: materials ta aprender Aragonés*. Zaragoza: INO.
- Bezzola, Reto R. / Rud. O. Tönjachen (1944¹/1997⁵): *Dicziunari tudais-ch - rumantsch ladin*. Cuoira: Lia Rumantscha.
- Brero, Camilo (2001): *Dizionario italiano-piemontese, piemontese-italiano*. Milano: Antonio Vallardi.
- Brero, Camilo / Remo Bertodatti (1993): *Grammatica della lingua piemontese: parola, vita, letteratura*. [Torino]: Savigliano.
- Carballeira Anllo, Xosé María (coord) (2009²): *Gran dicionario Xerais da Lingua*. 2 vols. Vigo: Ediciones Xerais de Galicia.
- Carballo Calero, Ricardo (1966¹/1979⁷): *Gramática elemental del gallego común*. Vigo: Galaxia.
- Carrera, Aitor (2007): *Gramatica aranesa*. Lleida: Pagès.
- Col, Gemo da (1991): *L'idioma ladin a Cibiana di Cadore, il paese dei murales: grammatica e vocabolario*. Pieve d'Alpago: Nuove Edizioni Dolomiti.
- Corriente, Federico (1983): "La serie mozárabe-hispanoárabe *aḡālah, aḡāqal, aḡāšš...*, y la preposición castellana *hasta*", *Zeitschrift für romanische Philologie* 99, 29-32.
- Corriente, Federico (2008): *Dictionary of Arabic and allied loanwords: Spanish, Portuguese, Catalan, Gallician and kindred dialects*. Leiden / Boston: Brill.
- Corominas, Juan (1954): *Diccionario critico etimológico de la lengua castellana*. Madrid: Gredos.
- Croatto, Enzo (1986): *Vocabolario ampezzano*. Belluno: Cassa Rurale ed Artigiana di Cortina d'Ampezzo.
- Cuveiro Piñol, Juan (1876): *Diccionario gallego*. Barcelona: N. Ramires & Co.
- Dalgado, Sebastião Rodolfo (1900¹/1998): *Dialecto indo-português de Ceilão*. Lisboa: Comissão Nacional para as Comemorações dos Descobrimentos Portugueses.
- Dalgado, Sebastião Rodolfo (1906): "Dialecto indo-português do Norte", *Revista Lusitana* 9, 142-166, 193-228.
- Dawkins, Richard (2004): "Chinese junk and Chinese whispers", em Richard Dawkins, *The devil's chaplain*. Boston / New York: First Mariner [= prefácio de Blackmore, Susan (1999): *The meme machine*. Oxford: OUP].
- Díez Calleja, Beatriz (ed.) (2008): *El primitivo romance hispánico*. [Burgos]: Instituto Castellano y Leonés de la Lengua.
- Ercolani, Libero (1994): *Nuovo vocabolario romagnolo-italiano italiano-romagnolo*. Ravenna: Girasole.
- Ferreiro, Manuel (1995-1997): *Gramática histórica galega*. 2 vols. Santiago de Compostela: Laiovento.
- Freixeiro, Xosé Ramón (2000): *Gramática da lingua galega. II. Morfosintaxe*. Vigo: A Nosa Terra.
- Fierro, Aurelio (1995): *Grammatica della lingua napoletana*. Milano: Rusconi.
- García, Constantino / Antón Santamarina (dirs.) (1995): *Atlas Lingüístico Galego. Volume II. Morfoloxía non verbal*. A Coruña: Fundación "Pedro Barrié de la Maza, Conde de Fenosa".
- Gifford, Douglas J./ Frederick W. Hodcroft (1966): *Textos lingüísticos del medievo español*. Oxford: Dolphin.
- Heilmann, Luigi et alii (1976): *Dizionario del dialetto cremonese*. Cremona: Libreria del Convegno.
- Heine, Bernd / Tania Kuteva (2002): *World lexicon of grammaticalization*. Cambridge: CUP.

- Hilty, Gerold (2005): "Los arabismos del libro compendio", *Sacrum Arabo-Semiticum. Homenaje al profesor Federico Corriente en su 65 aniversario*. Saragossa: Instituto de Estudios Islámicos y del Oriente Próximo, 181-194.
- Kontzi, Reinhold (1970): "Ist die aragonische Präposition enta ein Arabismus?", *Zeitschrift für Romanische Philologie* 86, 372-381.
- Levy, Emil (1909¹/1973³): *Petit dictionnaire provençal-français*. Heidelberg: Carl Winter.
- Lliteras, Margarita (1992): "Procesos evolutivos del subsistema latino de preposiciones de movimiento", em Gerold Hilty et al (eds.), *Actes du XXème Congrès International de Linguistique et Philologie Romanes*, vol. 2. Bern: Francke, 417-430.
- Maia, Clarinda de Azevedo (1995): *História do Galego-Português. Estado lingüístico da Galiza e do Noroeste de Portugal desde o século XIII ao século XVI (Com referência à situação do galego moderno)*. Coimbra: Instituto Nacional de Investigação Científica.
- Malkiel, Yakov (1979): "Problems in the diachronic differentiation of near-homophones", *Language* 55(1), 1-36.
- Maurer Jr, Theodoro Henrique (1951): *A unidade da România Ocidental*. São Paulo: [FFLCH].
- Menardi, Silvio (org) (1997): *Vocabolario italiano - ampezzano*. Cortina d'Ampezzo: Cassa Rurale ed Artigiana di Cortina d'Ampezzo e delle Dolomiti.
- Menéndez-Pidal, Ramón (1926¹/1950³): *Manual de gramática histórica española*. Madrid: Espasa-Calpe.
- Mettmann, Walter (ed) (1959-1972): *Cantigas de Santa Maria* (Afonso X, o sábio), 4 vols. Coimbra: Universidade de Coimbra.
- Meyer-Lübke, Wilhelm (1911-1920¹/1992⁴). *Romanisches Etymologisches Wörterbuch* (REW). Heidelberg: Carl Winter.
- Nazzi, Gianni / Luca Nazzi (1997): *Dizionario italiano-friulano, friulano-italiano*. Milano: Antonio Vallardi.
- Nunes, José Joaquim (1919¹/1945³): *Compêndio de gramática histórica portuguesa*. Lisboa: Clássica.
- Peer, Oscar (1962¹/1995⁴): *Dicziunari rumantsch ladin - tudais-ch*. Cuoir: Lia Rumantscha.
- Penny, Ralph (2004): "Evolución lingüística en la baja edad media: evoluciones en el plano fonético", em Rafael Cano, *Historia de la lengua española*. Barcelona: Ariel, 593-612.
- Pinguentini, Gianni (1969): *Nuovo dizionario del dialetto triestino: storico, etimologico, fraseologico*. Bologna: Cappelli.
- Pires, Antônio Thomás (1903-1907): "Vocabulário alentejano", *Revista lusitana* 8: 92-98, 298-300, 1903-1905; 9:167-176, 1906; 10: 87-101, 238-254, 1907.
- Pittau, Massimo (1991): *Grammatica della lingua sarda*. Sassari: Carlo Delfino.
- Prieto García, Luís Alberto (2004): *Diccionario de sinónimos y equivalencias castellano-asturiano*. Oviedo: Trabe.
- Riiho, Timo (1979): *Por y para: estudio sobre los orígenes y la evolución de una oposición prepositiva*. Helsinki: Societas Scientiarum Fennica.
- RILG = Seminario de Lingüística Informática - Grupo TALG / Instituto da Lingua Galega (2006-): *Recursos Integrados da Lingua Galega*. Vigo: Universidade de Vigo, Santiago de Compostela: Universidade de Santiago de Compostela <<http://sli.uvigo.es/RILG/>>
- Rohlf, Gerhard (1966-1969): *Grammatica storica della lingua italiana i dei suoi dialetti*, 3 vols. Torino: Einaudi.
- Rohlf, Gerhard (1935): "Le gascon: études de philologie pyrénéenne", *Beihefte zur Zeitschrift für Romanische Philologie* 85, 1-190.
- Romito, Giuseppe (1985): *Dizionario barese*. Bari: Levante.
- Rosamani, Enrico (1990): *Vocabolario giuliano dei dialetti parlati nelle regione giuliano-dalmata quale essa era stata costituita di comune accordo tra i due stati interessati nel convegno di Rapallo del 12-XII-1920*. Trieste: Lint.
- Rossi, Hugo von (1999): *Ladinisches Wörterbuch - vocabolario ladino (brach) - tedesco con traduzione italiana a cura di Ulrike Kindl e Fabio Chiochetti*. Innsbruck: Istitut Cultural Ladin.
- Sandfeld, Kristian (1930): *Linguistique balkanique: problèmes et résultats*. Paris: Honoré Champion.
- Seco, Manuel (ed) (2003): *Léxico hispánico primitivo (siglos VIII al XII)*. Madrid: Fundación Menéndez-Pidal y Herederos de Rafael Lapesa Melgar/ Real Academia de la Lengua/ Espasa Calpe.
- Seybold, Christian F. (1888): "Die arabische Sprache in der romanischen Ländern", em Gustav Gröber, *Grundriss der romanischen Philologie*. Straßburg: Karl J. Trübner, 398-405.
- Silva Neto, Serafim da (1958): "História da preposição portuguesa até", em Hans Erich Keller (ed.), *Etyimologica - Walther von Wartburg zum siebzigsten Geburtstag*. Tübingen: Niemeyer, 751-761.
- Sonder, Ambros / Mena Grisch (1970): *Vocabulari da Surmeir rumantsch-tudestg tudestg-rumantsch*. Coira: Leia Rumantscha.

- Spitzer, Leo (1920): Resenha a Pere Barnils Giol (1913): *Die Mundart von Alacant: Beitrag zur Kenntnis des Valenzianischen*. Barcelona: Palau de la Diputació. *Zeitschrift für romanische Philologie* 40, 239-242.
- Svorou, Soteria (1994): *The grammar of space*. Amsterdam / Philadelphia: John Benjamins.
- Tajina, Alessandra Dellantonio (1998): *Dizionèr talian-fascian-talian*. Pilat: Trent.
- TILG = Antón Santamarina (dir.) (1986-): *Tesouro Informatizado da Lingua Galega*. Santiago de Compostela: Instituto da Lingua Galega. <<http://ilg.usc.es/tilg>>
- TMILG = Xavier Varela Barreiro, (dir.) (2004-): *Tesouro Medieval Informatizado da Lingua Galega*. Santiago de Compostela: Instituto da Lingua Galega <<http://ilg.usc.es/tmilg>>
- Toso, Fiorenzo (2006): *Dizionario italiano-genovese, genovese-italiano*. Milano: Antonio Vallardi.
- Väänänen, Veikko (1981¹/1988³): *Introducción al latín vulgar*. Madrid: Gredos.
- Vasconcelos, José Leite de (1928): *Opusculos*. v. 2: *dialectologia*, parte 1. Coimbra: Imprensa da Universidade.
- Viaro, Mário Eduardo (1994): *Das preposições latinas às do português e do romeno: estudo das derivações semânticas*. Dissertação de Mestrado. São Paulo: FFLCH.
- Viaro, Mário Eduardo (1995): "Sobre a presença de *tenuis* no ibero-românico", *Confluência*. Assis 1, 269-277.
- Viaro, Mário Eduardo (2004): "A third hypothesis for the etymology of Portuguese *até*", *Revista do Gel*. Campinas: GEL. 1(1), 91-100.
- Viaro, Mário Eduardo (2011): *Etimologia*. São Paulo: Contexto.
- Vieli, Ramun / Alexi Decurtins (1975¹/1994⁴): *Vocabulari romontsch deutsch-surselvisch*. Cuera: Ligia Romontscha.
- Vieli, Ramun / Alexi Decurtins (1962¹/1995⁴): *Vocabulari romontsch sursilvan-tudestg*. Cuera: Ligia Romontscha.
- Vitali, Daniele / Luigi Lepri (2009): *Dizionario italiano-bolognese, bolognese-italiano*. Milano: Antonio Vallardi.
- Wagner, Max Leopold (1960-1964): *Dizionario etimologico sardo*, 3 vols. Heidelberg: Carl Winter.