

REAL ACADEMIA DE CULTURA VALENCIANA

**SECCIÓN DE ESTUDIOS IBÉRICOS
"D. Fletcher Valls"**

**ESTUDIOS DE LENGUAS Y EPIGRAFÍA ANTIGUAS -
ELEA**

Núm. 19

**XXXV SEMINARIO DE LENGUAS Y EPIGRAFÍA
ANTIGUAS**

**PONENCIAS Y ESTUDIOS VA-
RIOS**



**RACV - 105 AÑOS AL SERVICIO DE VALENCIA Y LOS
VALENCIANOS**

**VALENCIA
2020**

ESTUDIOS DE LENGUAS Y EPIGRA-
FÍA

ANTIGUAS - ELEA

Número 19

X. BALLESTER

Universitat de València

LINGÜÍSTICA INDOEUROPEA Y SIMULACIÓN ROMÁNICA

Resumen: La legitimidad de la reconstrucción indoeuropea ha sido históricamente puesta en entredicho por motivos muy variados. De cara a una mejora del método reconstructivo, la comparación con el bien conocido grupo románico, a causa de su innegable relación con el latín, nos ofrece un banco de pruebas potencialmente muy útil y puede ayudarnos por analogía a ponderar la validez del método tradicional.

Palabras Clave: Indoeuropeo, lenguas románicas, latín, Lingüística.

Indo-European Linguistic and Romance Simulation

Abstract: The legitimacy of Indo–European reconstruction has historically been called into question for different reasons. A comparison with the well–known Romance linguistic group, because of its undeniable relationship with Latin, offers us a set of potentially very useful test cases for improving the reconstructive method. Additionally, and by analogy, such a comparison can help us to weigh the validity of the traditional method.

Keywords: Proto–Indo–European, Romance languages, Latin, Linguistics.

La legitimidad de la reconstrucción de la entidad lingüística indoeuropea ha sido históricamente puesta en entredicho por motivos muy variados. Uno de los problemas señalados es el hecho de que, aunque se parta de lenguas reales y bien documentadas como las diferentes lenguas históricas indoeuropeas, por definición siempre se reconstruirá una lengua no otramente documentada y, por tanto, hoy por hoy sin posibilidad real de verificación. De cara a una mejora y—por así decir—*refinamiento* del método reconstructivo, acaso la comparación externa, en este caso con el bien conocido grupo románico, pueda ayudarnos por analogía a ilustrar la validez del método tradicional y eventualmente a proponer otras fórmulas diferentes—más flexibles obviamente—de análisis que no conduzcan inevitablemente a un indoeuropeo irreal por sistemático y uniforme.

UN ROMÁNICO BANCO DE PRUEBAS: AFINIDADES Y DIFERENCIAS

El grupo de lenguas románicas, a causa de su innegable relación—esta sí avalada por hechos arqueológicos e históricos externos—con el latín, relación acreditada con una documentación y amplitud prácticamente únicas dentro de las lenguas indoeuropeas, nos ofrece un banco de pruebas potencialmente muy útil, ya que se acepta que una relación básicamente similar a la existente entre el latín y el romance debió darse para los grupos indoeuropeos y *el* indoeuropeo.

Aunque por desgracia no conocemos específicas simulaciones estadísticas y cuantitativas sobre la evolución desde un ficticio y supuesto prerromance hasta el romance, es posible al menos hacerse una lúcida idea de que, en primer lugar, encontraremos en el más seguro ámbito románico la misma tipología de inconsistencias reales y visibles irregularidades que parecen darse en el ámbito indoeuropeo y, en segundo lugar, de que, si bien ciertamente dichas inconsistencias se darían en una proporción muy probablemente menor en el caso románico, este hecho encuentra una justificación sólida en las diferencias temporales y locales que comporta un seguro menor margen cronológico y espacial para la evolución de las lenguas románicas—nunca más de dos milenios y para una parte de Europa—que para la evolución de los grupos indoeuropeos

—varios milenios y para un continente y medio—lo que supone que por parte indoeuropea es lógicamente previsible un *parte de incidencias* con un nivel mayor, quizá considerablemente mayor, de irregularidades.

A ello se suma consecuentemente el implícito factor de que, mientras en el caso románico comparamos y relacionamos de lengua a lengua, tipo latín \geq español, ya que no es posible establecer nítidos grupos intermedios entre ambos estadios lingüísticos, en el caso indoeuropeo empero, por mor de ese indudablemente mayor margen cronológico, comparamos y relacionamos en realidad tres entidades a causa de la situación intermedia que ocupan los grupos lingüísticos reconstruibles, tipo indoeuropeo \geq báltico| céltico| germánico... \geq lituano – letonio...| bretón – irlandés...| gótico – sueco... lo que evidentemente aumenta en considerable medida el esperable nivel de distorsión y rebaja las expectativas estadísticas de las correspondencias que se detecten. Y todo ello sin tener en cuenta el aspecto puramente cuantitativo de ambas comparaciones que serían de unas pocas decenas de lenguas por el lado románico y de un par de centenares por la parte indoeuropea. Esto significa que si constatamos fehacientes irregularidades en seguras correlaciones latín – romance, con mayor razón habremos de esperarlas en el más temporalmente duradero, espacialmente más vasto y lingüísticamente más numeroso ámbito indoeuropeo. Sinópticamente:



Tras estas precisiones veamos, pues, cuáles los tipos de series, regulares y no tanto, que en lo relativo a los aspectos fono-semánticos encontramos en romance en relación a su correspondencia vertical o externa con el latín y también en relación a su correspondencia horizontal o interna entre las propias lenguas románicas, sabiendo como sabemos que el latín es punto de partida—al menos un punto de partida seguro—de las lenguas románicas. No contemplaremos por el momento aspectos morfológicos y solo entraremos en pormenores del análisis semántico cuando se observen significativas diferencias de detalle o de conjunto.

RECIBIDOS, REGULARES, RECUPERABLES

Anticipemos que nos interesará principalmente distinguir en cualquier lengua las voces según estos tres básicos criterios:

- formas recibidas, cuñas patrimoniales, frente a formas irrecibidas, frente a creaciones autónomas o sobre todo frente a copias foráneas,
- formas reversibles, cuyo origen es recuperable, frente a formas irreversibles, cuyo origen no es recuperable, y
- formas regulares, de evolución previsible, frente a formas irregulares o de evolución en principio no previsible o no totalmente previsible.

Evidentemente las formas con la triple *r-* inicial: recibidas, regulares y reversibles son las más interesantes y potencialmente las más productivas para la reconstrucción lingüística. Las formas con *i-* inicial: irrecibidas, irregulares e irreversibles son, por el contrario, las menos apetecibles para el reconstructor de lenguas, aunque ello no significa, como veremos, que carezcan de valor testimonial.

En la práctica los citados criterios se manifestarían en romance según estos operativamente más útiles cuatro tipos principales:

- *estériles*: voces latinas sin resultado o casi sin resultado—regular o irregular—en romance,

- *prolíficas*: voces latinas con resultado regular y, por tanto, con normalmente abundancia de formas interrelacionadas dentro de una lengua románica y entre las diversas lenguas románicas,
- *fértiles*: voces latinas con resultado irregular y, por tanto, con normalmente escasez de correlaciones directas dentro de una lengua románica y entre las diversas lenguas románicas, y
- *bastardas*: voces no latinas pero con resultado—regular o irregular—en romance.

Sinópticamente:

	<i>patrimonialidad</i>	<i>reconstrucción</i>	<i>evolución</i>
<i>estériles</i>	recibida	irreversible	ir/regular
<i>prolíficas</i>	recibida	reversible	regular
<i>fértiles</i>	recibida	reversible	irregular
<i>bastardas</i>	irrecibida	reversible	ir/regular

IRRECUPERABLES O CASI: LOS SIN DESCENDENCIA

Pues bien, en primer lugar puede haber formas o simplemente elementos de la lengua latina que no se hayan conservado en ninguna lengua románica o que se hayan conservado apenas en alguna lengua románica o bien que tan solo se hayan conservado indirecta o colateralmente en alguna lengua no románica. Se trata de casos más bien excepcionales. El motivo más probable para la desaparición de aquellas palabras latinas es el de que estas hubieran desaparecido o estuvieran ya en trance de desaparecer ellas mismas del estadio de la lengua hablada real a partir de la cual se originaron las lenguas románicas. Tales formas pudieron ser substituidas por otras consideradas más expresivas o como producto de una moda o bien porque alguna característica fonética—ambigüedad sobrevinida, connotación negativa, escaso volumen silábico...—propiciara su desaparición. Si la evolución hubiese sido regular o no, es obviamente aquí algo secundario por intrascendente. Veamos unos ejemplos.

Previamente se notará que aquí marcaremos solo la cantidad o tensión de la vocal en los términos latinos cuando esta sea larga y además tónica, pues sábase que en latín tardío se produjo una neutralización de la cantidad en las vocales átonas, de modo que en posición átona las lenguas románicas no presentan tratamiento distinto alguno de las vocales que en latín clásico eran largas y breves. He aquí,

pues, un primer elemento estéril en la lengua latina—la presencia de vocales largas y breves en posición átona—ya que no dejó rastro alguno en romance. En el caso de los nombres partimos, en principio, del caso acusativo singular, base muy mayoritaria—nueva instancia de esterilidad y esta vez morfológica la de los demás casos—de los nombres románicos en este número gramatical, aunque prescindiendo de la *-m* final que por lo general—un tercer caso más de esterilidad—tampoco dejó vestigios en los nombres románicos. Veamos cinco ejemplos ilustrativos.

① estériles

latín	<i>ēsse</i>	<i>hic</i>	<i>pueru-</i>	<i>-que</i>	<i>ut</i>
español	<i>comer</i>	<i>este</i>	<i>niño</i>	<i>y</i>	<i>que</i>
francés	<i>manger</i>	<i>celui-ci</i>	<i>enfant</i>	<i>et</i>	<i>que</i>
italiano	<i>mangiare</i>	<i>questo</i>	<i>bambino</i>	<i>e[d]</i>	<i>che</i>
logudorés	<i>mandigare</i>	<i>custu</i>	<i>pitzinnu</i>	<i>e</i>	<i>che</i>
portugués	<i>comer</i>	<i>isto</i>	<i>menino</i>	<i>e</i>	<i>que</i>
rumano	<i>a mînca</i>	<i>acest</i>	<i>copil</i>	<i>și</i>	<i>că</i>
valenciano	<i>menjar</i>	<i>est[e]</i>	<i>chiquet</i>	<i>i</i>	<i>que</i>

DISPERSOS RESTOS

El verbo clásico en latín para ‘comer’ *ēsse*, de origen indoeuropeo, contaba al menos con tres motivos para no tener *descendencia* en romance: era irregular, se podía confundir en algún caso con *esse* ‘ser’, el verbo más frecuente en latín, y además pertenecía a ese grupo de significados que suelen competir con sinónimos o parasinónimos más coloquiales y expresivos, como nuestros *devorar*, *jalar*, *papear*, *zampar*... De hecho fue substituido de manera general por un verbo mucho más expresivo: *manducāre*, verbo de origen coloquial, pero que ya había sido recogido en las traducciones populares de la Biblia, por lo que contaba con cierta tradición escrita y literaria. En realidad las formas española y portuguesa *comer* contienen la forma *ēsse* o, mejor dicho, esa vetusta raíz, cuyo infinitivo pasó por analogía (*cfr. edō* ‘yo como’) a *ēdere*, forma que reforzada con *cum*—preverbio muy frecuente en latín tardío y consecuentemente muy presente en romance—había generado ya en latín el doblete *comēsse* y *comēdere*, procediendo de esta última forma las voces ibéricas.

Por supuesto, sin la documentación de la lengua de partida y por razones estrictamente fónicas sería seguramente imposible que el [proto-]romanista identificara en la *-e-* de *comer*, tras regular pérdida de /d/ intervocálica, la vetusta raíz *ed-* de *edō* ‘como’, *ēs* ‘comes’, *ēsse* ‘comer’, *ēdī* ‘comí’, *ēssum* ‘comido’, y también muy difícilmente hubiese

restituido la forma analógica *edō* ‘como’, *edis* ‘comes’, *edere* ‘comer’, *ēdī* ‘comí’, *ēssum* ‘comido’. La forma clásica y antigua hubiese constituido, en definitiva, una forma irrecuperable en la práctica si contáramos solo con el material románico, el cual apunta de hecho a una forma **mandigāre* o similar. Ciertamente las disidentes formas española y portuguesa constituirían apenas la única pista de que pudo haber alguna otra forma para decir ‘comer’ en latín, pero muy difícilmente nos hubiesen siquiera acercado a una forma, como *ēs-se*, que sí realmente hubo existido.

Caso distinto el del epidíctico *hic* ‘este’. Es la de los demostrativos en general una categoría muy peculiar en todo y también en sus aspectos evolutivos, ya que tienden a renovarse pero sobre todo precisamente a partir de otros temas demostrativos—como en los citados *este*, *isto*, *est[e]* que proceden del demostrativo de segunda epidixis latina *iste* ‘ese’—o también a partir de pronombres afines—como el español *ese* y el valenciano *eixe* ‘ese’, que remiten al pronombre latino no demostrativo *ipse* ‘él mismo’—y en uno y otro caso tienden con frecuencia a la aglutinación, es decir: a hipercaracterizarse. Es esto precisamente lo que vemos, por ejemplo, en el *celui-ci* francés y formas afines, donde los *ce-* y *-ci* remiten al elemento que también se aglutinaba ya en latín (*hi-c*), mientras que *lui* remite a su vez a otro demostrativo (*cfr.* latín *ille* ‘aquel’), lo que también sucede en *questo*, *custu* y *acest*, donde aquel mismo elemento ha quedado aglutinado al antiguo *iste* latino en su segmento inicial.

Adicionalmente el epidíctico latino *hic, hæc, hoc*, de flexión irregular, ha pervivido periféricamente en otras formaciones: español *hogaño*, logudorés *occannu* o valenciano *enguany* con la variantes *ong[u]any* y *eng[u]any* en el llamado *Llibre de la Cort d'Alcoy* del s. XIII (Diéguez 2012: 69 ^{regis-}tro¹⁵⁸ año¹²⁶⁴: SEMBRE \ LA MEYTAT/ ONGAN [...] REEBE ENGAN LA MEYTAT), formas todas que aglutinaron en su segmento inicial el viejo demostrativo al sustantivo *annu-* ‘año’. También el antiguo sintagma latino *hāc hōrā* ‘a esta ahora’ está en el origen del español antiguo *agora* y actual *ahora* o portugués *agora*. Similarmente *hanc hōram* dio el ‘todavía’ de franceses (*encore*), italianos (*ancora*) y valencianos (*enca-ra, encà*), en el último caso con posible interferencia de *ara* ‘ahora’, otro potencial factor este de la interferencia fonosemántica que puede enmarañar la esperable lineal evolución de una forma. Así mismo la antigua lengua de oc, como se sabrá, es así llamada por haber conservado el antiguo *hoc* ‘esto’ con el valor de ‘sí’, uso bien vivo todavía en el valenciano del s. XIII, verbigracia: DEMANAT FO SI SABIA PUS, DIX QUE HOC (Diéguez 2012: 69 ^r158 ^a1264 y *passim*). El mismo antiguo demostrativo neutro *hoc* sigue hoy asimismo perfectamente vivo en el pronombre valenciano *ho*, pronunciado de diversas maneras —básicamente [ew], [u] y [o]—según posición y contexto. En latín el elemento *-ce* aparecía también en otras formas, como en *tunc* ‘entonces’, variante reforzada de *tum* ‘entonces’, ello significa que también nuestro *entonces* contiene en su *-c-*, en última instancia, un residuo de aquel elemento epidíctico, como asimismo probablemente el rumano *atunci* ‘entonces’. En suma, aunque no son tan

escasos los rastros del demostrativo latino *hi-c*, son rastros difíciles de seguir desde la metodología tradicional de la Lingüística indoeuropea.

Así pues, la evidencia, corroborada por similares fenómenos en otras lenguas, sugiere que los demostrativos en forma y significado son poco estables, pero que su desaparición evolutiva se produce como—diríase—por explosión, por lo que suelen dejar numerosos vestigios diseminados por aquí, acá o acullá, otros tres adverbios, por cierto, con raíz demostrativa aglutinada en su segmento inicial. Puesto que, como sabemos, los demostrativos tienen una importancia céntrica como generadores de categorías morfológicas en las lenguas y además, como estamos viendo, suelen dejar un rastro grande aunque disperso, el colorario indoeuropeo de lo expuesto será que las formas demostrativas antiguas serán de difícil aunque no imposible recuperación. A esto, sin embargo, hay que sumar el esperanzador dato tipológico que nos advierte de que los demostrativos suelen venir formados por una secuencia de consonante y vocal básicas (*CV*).

SIN EL MENOR RASTRO

Las formas con fuertes connotaciones afectivas son también proclives a desaparecer en la evolución siendo reemplazadas por otras consideradas todavía más expresivas hasta que estas dejan de sentirse expresivas y se reinicia el ciclo substitutivo. El término para ‘niño’ en latín, *puer*, es, pues, uno de esos típicos nombres coloquiales que estaba poco menos que destinado a ser substituido en romance. Mientras algunas lenguas recurrieron a raíces que sí estaban en la lengua básica de partida, como el francés *enfant* (latín *infante*— etimológicamente ‘que no habla’) o portugués *criança* (derivado del latín *creare* ‘crear – criar’), otras lenguas propiciaron por sí mismas, seguramente a partir de secuencias onomatopéyicas o expresivas, nuevas formaciones: *niño* – *nene*, *bambino* – *bimbo*, *chiquet*... Así, en el sardo campidanés tenemos *pipiu* y en el logudorés *pitzinnu* o *pizzinu*, formas que podrían proceder de una base ideofónica /pi/ evocando sinestésicamente la pequeñez gracias a la agudeza y brevedad fono–acústicas de sus componentes.

Puesto que—y en circunstancias más bien excepcionales—las lenguas románicas básicamente proceden de una lengua, como el latín, con amplio cultivo escrito y que se mantuvo como tal en la religión, la literatura y la ciencia occidentales mucho más allá de su connatural y orgánico uso hablado, podemos encontrarnos bastantes veces con el

especial fenómeno de la presencia de cultismos o copias directas—es decir: sin evolución—de formas latinas, lo que justifica la presencia de voces como *puericia*, *puericultor* o *pueril* con aquella raíz *puer*—conservada, pero evidentemente es muy difícil suponer que ese mismo fenómeno cultista tuviera gran relevancia en los grupos lingüísticos indoeuropeos, donde eventualmente solo podríamos acaso contar con la posibilidad de la perduración de formas muy antiguas y que se hubiesen saltado el normal proceso evolutivo en las tradiciones—pero orales y casi nunca escritas—del arte verbal.

Se observará que la forma enclítica *-que*, seguramente la de mayor frecuencia de uso en latín clásico, desapareció, al parecer, sin dejar rastro alguno siendo substituida sobre todo por su sinónimo *et*. El rumano, como en tantas otras ocasiones, se aparta aquí del resto presentando una forma *și*, que se tiene por heredera del adverbio latino *sic* ‘así’ y donde, por cierto, encontramos de nuevo aquella marca demostrativa *-c*. La vetusta forma *-que*, de origen indoeuropeo, competía, pues, con *et* y *atque* en latín como conjunción copulativa y esta dura *competencia* pudo ser también una de las razones que llevaron a la solución más radicalmente económica: su desaparición total, acaecida, por tanto, probablemente antes de la emergencia de cualquier romance.

Similar era el caso de la también frecuentísima conjunción latina *ut* ‘[de modo] que’, de la cual tampoco ha quedado directamente rastro alguno en romance. Seguramente su escaso volumen silábico—pues reducida a [ʊ] de modo general en el habla popular por pérdida de la *-t* final—

fue un factor que coadyuvó a su esterilidad evolutiva. Su antigua privilegiada posición de superconjunción de subordinación fue reemplazada por los equivalentes románicos a nuestra también superconjunción relativa *que* y donde probablemente tengamos la fusión de al menos las antiguas formas del relativo *quæ*, *quem* y *quia*, esta última funcionando ya como conjunción causal en latín clásico. Como de costumbre, sardo y rumano manifiestan una mayor singularidad o independencia. Por ejemplo, el logudorés *che* presenta un significado más restringido: ‘como’ (Pittau 1991: 122), empleándose más bien *chi* como ‘que’. Para otros muy diversos valores que expresaba el *ut* latino las lenguas románicas presentan, por lo general ya a título individual, diversas formaciones substitutivas.

CARTOGRAFÍA LINGÜÍSTICA DE LOS RESTOS DEL NAUFRAGIO

Así pues, el hecho indiscutible es que desaparecieron patrimoniales palabras comunísimas en el curso de la evolución del latín al romance y, como era de esperar, a menudo también palabras muy raras y de uso parsimonioso. Sin embargo, simultáneamente algunas de estas rarezas se conservaron, como era de esperar, en alguna lengua, normalmente más marginal o periférica. A menudo el sardo y específicamente sobre todo el más conservador sardo logudorés ofrece ejemplos de testimonios únicos de supervivencia de formas antiguas. Así, a partir de su procedencia en la región de Caria, en Anatolia, el latín denominó *ficus Cārica* a una variedad, cuyo nombre como designación para el higo solamente ha pervivido en Cerdeña: logudorés *cāriga*. También el *modesto* logudorés está entre las pocas hablas que mantuvieron hasta el día de hoy el monosilábico *cras* ‘mañana’, preservado también en el compuesto *pusticras* ‘pasado mañana’, adverbio aquel substituido por otras formaciones en las *grandes* lenguas románicas, como las más transparentes formaciones del tipo español *mañana* y *pasado mañana* o valenciano *demà* con aglutinación en inicial de la preposición *de* (*cfr.* francés *demain*, italiano *domani*) y *despusdemà* (literalmente ‘después-[de-]mañana’). No solo la forma latina *pro* sino también sus diversos valores mantuviéronse mejor en el logudorés *pro* ‘a razón de – en lugar de – para – por – según’ que en español *por* y francés *pour* ‘para’.

Los mismos principios que vemos aplicados a los *grosso modo* aspectos fono-semánticos servirían también, *mutatis mutandis*, para los particulares aspectos fonológicos, morfológicos, sintácticos, léxicos y semánticos individualmente considerados.

En lo fonológico, por ejemplo, cabe observar que de modo general las lenguas románicas no presentan, como dijimos, restos de la *-m* final, salvo casos aislados como el español *alguien* (\leq latín *aliquem*) y *quien* (\leq latín *quem*), francés *rien* ‘nada’ o valenciano antiguo *ren* también ‘nada – algo’ (\leq latín *rem* ‘cosa’), así en en el *Llibre de Cort de Cocentaina* (Torró 2009: 35 ^r44 ^a1269: SI VOLEN REN DIR; 39 ^r58 ^a1269: AB PUYN NI ABREN...) o bien en formas donde la *-m* en realidad no aparece en final absoluto, como en las preposiciones latinas *cum* ‘con’ (español e italiano *con*, sardo campidanés *cun*, portugués *com*, pero rumano *cu*) o *in* ‘en’ (español, francés y valenciano *en*, italiano *in*, portugués *em*, rumano *în*...). Nótese aquí la singular neutralización de nasales a favor de *-m* en portugués.

Como ya se ha apuntado, las elementos más irregulares son proclives o a su eliminación o a su regularización. Morfológicamente el uso postposicional de *cum* ‘con’ con los pronombres personales—así *mecum* ‘conmigo’, *tecum* ‘contigo’, *nobiscum* ‘con nosotros’—constituía una excepcionalidad ya bastante insólita en latín. Buena parte del romance eliminó esa anomalía, pero en lenguas como el español (en los diacrónicamente pleonásticos *conmigo*, *contigo*) o el logudorés (en los también pleonásticos *cun megus* ‘conmigo’, *cun tegus* ‘contigo’) o en algún arcaísmo en otras lenguas (*u.g.* italiano *meco* ‘conmigo’, *teco* ‘contigo’) puede todavía detectarse su rastro. El español *cuyo* es uno de los pocos restos

románicos del genitivo del relativo latino *cuius*, pero más arcaizante es el uso interrogativo del logodorés *cujus pegasus est?* “¿de quién es la res?”, una frase que sería prácticamente latín (*cuius pecus est?*) evocando aquel criticado por su rusticismo *cuium pecus* del poeta VIRGILIO (*ecl.* 3,1).

En lo léxico notemos que la forma usual en latín para ‘saber’, *scīre*, fue substituida en romance de modo general por el verbo *sapere* ‘tener sabor’: español, portugués y valenciano *saber*, francés *savoir*, italiano *sapere*... pero se conservó en el rumano *ști* y aun mejor en el sardo *scire*. Restos marginales del lexema tenemos en formas *quasi* cultas como español *ciencia* (\leq latín *scientia*–) y homologables (francés *science*, italiano *scienza*, portugués *ciência*, rumano *știintă*...) o en el valenciano antiguo CIENT ‘a sabiendas – consciente[mente]’ (Diéguez 2012: 21^r18^a1263), procedente del participio latino *sciente*– etc.

VOCES PUTATIVAS ¿EXISTIDAS QUIZÁ?

Si, como acabamos de ver, hay voces latinas estériles, pues no tuvieron eco o apenas lo tuvieron en las lenguas románicas, aparentemente habría otras que, sin tener documentación alguna directa de su existencia, dejaron sonoras resonancias. La Romanística tradicional ha establecido que formas como el catalán *aquest* ‘este’ y *aquell* ‘aquel’ o catalán antiguo *aqueix* ‘ese’, español *aquel* o castellano antiguo *aqueste* ‘este’, italiano *questo* ‘este’ y *quello* ‘ese – aquel’, portugués *aquele* ‘aquel’ y afines resultarían de la—como vimos—habitual aglutinación de temas demostrativos o pronominales afines: los clásicos *iste* ‘este’, *ille* ‘aquel’, *ipse* ‘él mismo’ más aquel lexema epidíctico que todavía se hace visible en el latín escrito en formas como *HOCCE* (*C.I.L.* 11,301: *HOCCE PEREGIT OPVS*; *C.I.L.* 11,6246: *HOCCE SVB TVMVLO*) o en el adverbio *ecce* ‘he aquí’. Así para el logudorés *custu* ‘este’, *cussu* ‘ese’ y *cuddu* ‘aquel’ reconstruye BLASCO (2000: 21) unos **eccu istu-*, **eccu ipsu-* y **eccu illu-*. Lo cierto es que hasta ahora en la—en términos comparativos para su época—inmensa literatura latina nadie ha encontrado esa deducible forma **eccu-* ni, por tanto, las dichas secuencias.

Español, portugués, provenzal y valenciano con *afeerrar* o italiano y logudorés con *afferrare* postularían un derivado verbal del nombre latino *ferrum* ‘hierro’: **adferrare*, que, sin embargo, hasta ahora no está documentado. Igualmente el antiguo español *ahilar* y moderno *afilar*, francés *afilier*, italiano y logudorés *affilare*,

portugués *afiar*, provenzal y valenciano *afilar* apuntan a un latín no documentado **adfilare* derivado de *fīlum* ‘hilo’. Las correspondencias de español *alguno*, francés *aucun*, italiano *alcuno*, portugués *algum*, provenzal *alcun*, valenciano *algun* etc. apuntan a un **alicunu*– latino, sin embargo, no documentado (Meyer–Lübke 1992: 22 ⁿ339 s. u.).

Parecidamente podrían citarse muchos otros casos, como el derivado verbal de, en última instancia, el latín *altu*– ‘alto’, un **altiare* al que apuntan correlaciones cuales, entre otras, español *alzar*, italiano *alzare*, portugués y valenciano *alçar* o provenzal *ausar* (Meyer–Lübke 1992: 25 ⁿ385 s. u.). Esta vez se trata, pues, de formas que presentan una correspondencia regular y, por tanto, en teoría reversible al latín. Sin embargo, ante la falta de documentación hemos de considerar que se trata de formas directamente incorporadas o creadas—normalmente no de manera paralela—en fase ya románica.

Se trata, por tanto, de puras restituciones que nos advierten de que incluso en los procesos ficticios de reconstrucción de una [pre]lengua escrita es posible que hubiera formas cuya documentación no hayamos conservado y cuyos putativos ancestros podamos restituir desde sus restos. Como esta modalidad de palabras que cabría, pues, denominar *putativas* o *huérfanas* no pudo darse en la reconstrucción indoeuropea, ya que no poseemos documentación escrita de esta prelengua, no la tendremos aquí en especial consideración en nuestra revisión de la evolución latín – romance. Es más: en realidad estrictamente hablando todas las formas indoeuropeas

pertenecen a esta categoría de voces huérfanas, ya que no poseemos documentación escrita o directa de las hablas comunes indoeuropeas.

	<i>patrimonialidad</i>	<i>reconstrucción</i>	<i>evolución</i>
<i>putativas</i>	ir/recibida	ir/reversible	ir/regular

VOCES ESTÉRILES FRENTE A VOCES PUTATIVAS

Es, pues, posible que en el caso indoeuropeo este mismo tipo de correlaciones nos lleven a reconstruir como preformas voces que en realidad podrían constituir más bien, por ejemplo, desarrollos paralelos, pero que también en algún caso podrían remitir efectivamente a una pretérita lengua. Las voces putativas constituyen, pues, el reverso de las formas estériles. Estas con documentación solo en el estadio lingüístico de partida y aquellas con documentación únicamente en el estadio lingüístico de llegada.

<i>estériles</i>	<i>putativas</i>	
con documentación	sin documentación	<i>estadio de partida</i>
sin documentación	con documentación	<i>estadio de llegada</i>

Sin embargo, en la caso de las palabras putativas no puede taxativamente excluirse la posibilidad de que hayan efectivamente existido en la lengua matriz, solo que, al menos por el momento, no cuentan con acreditación ni registro alguno. Mientras el indoeuropeísta reconstruirá *ciegamente* una forma común en la lengua matriz en estos casos, el romanista tenderá a ver estas voces, en la expectativa de que eventualmente su futura documentación lo desmienta, como desarrollos paralelos o, más bien, como creación autónoma de una lengua románica y que fue copiada por algunas otras.

Así, junto a formas putativas nunca documentadas ni acreditadas en la lengua *madre* latina, pero que pudieron existir, como testimoniaría indirectamente el consenso románico, habría otras, las estériles, que, sin embargo, existieron en latín, pero de las cuales a veces no tenemos ni tan siquiera evidencia románica indirecta.

Muchas voces putativas permanecen y permanecerán en las tinieblas a no ser que un examen mejor del material románico o circunstancias nuevas, como la aparición de epígrafes con material lingüístico hasta la fecha desconocido, las iluminen. Probablemente el mayor contingente de estas formas esté compuesto por términos técnicos muy especializados o también por formas muy coloquiales que no se estimaron dignas de emplearse en la lengua escrita. El sentido coloquial de ‘además’ para español *encima* o su correspondiente valenciano *damunt* no han entrado todavía en el regular registro escrito de estas lenguas. El valencianísimo y tan noble y antiguo uso de *en* con el sentido de ‘con’, como en el latín bíblico—*uirgam, in qua percussisti flumen* “la rama, en la que golpeaste el río” dice la *Vetus latina* (Ex. 17,5)—está ya directa e indirectamente acreditado desde la primera documentación de la lengua en el s. XIII pero solo comenzará a aparecer intermitentemente a partir del s. XV en algunos escritores de registro más popular y compitiendo siempre con la forma culta y arcaizante *ab* (\leq latín *apud* ‘junto a’) hasta finalmente triunfar de manera arrolladora en el s. XIX. Sin embargo, desde mediados del s. XX y por imposición *normativa* de los politicastros lingüicidas de turno la forma está proscrita con tal sentido en el registro

escrito en favor de la reciente innovación catalana *amb*, que lógicamente jamás ha existido en el valenciano real y genuino. En suma, la documentación de una forma en el registro escrito no deja de ser en muchos casos un capricho del azar... o de los censores políticos.

REGULARES ¡FELIZ TRAVESÍA!

Además tenemos ciertamente un gran número de voces recibidas o patrimoniales con correspondencias regulares y prácticamente sistemáticas o biunívocas tanto respecto al latín cuanto entre las propias lenguas románicas, relaciones que podríamos llamar, por tanto, *pluribiunívocas* y que fácilmente nos reconducen al latín. En el análogo caso indoeuropeo dicho tipo de multilaterales interrelaciones coherentes nos remitirían también más directa y seguramente a la prelengua. Tal sería el caso de, por ejemplo, las siguientes voces románicas recibidas regularmente del latín *bōnu-* ‘bueno’, *cāsa-* ‘cabaña – choza’, *dēce-* ‘diez’, *fōcu-* ‘fuego’ o *lōcu-* ‘lugar’:

② prolíficas

latín	<i>bonu-</i>	<i>casa-</i>	<i>dece-</i>	<i>focu-</i>	<i>locu-</i>
español	<i>bueno</i>	<i>casa</i>	<i>diez</i>	<i>fuego</i>	<i>luego</i>
francés	<i>bon</i>	<i>chaise</i>	<i>dix</i>	<i>feu</i>	<i>lieu</i>
italiano	<i>buono</i>	<i>casa</i>	<i>dieci</i>	<i>fuoco</i>	<i>luogo</i>
logudorés	<i>bonu</i>	<i>caza</i>	<i>deghe</i>	<i>focu</i>	<i>logu</i>
portugués	<i>bom</i>	<i>casa</i>	<i>dez</i>	<i>fogo</i>	<i>logo</i>
rumano	<i>bun</i>	<i>casă</i>	<i>zece</i>	<i>foc</i>	<i>loc</i>
valenciano	<i>bo[n]</i>	<i>casa</i>	<i>dèu</i>	<i>foc</i>	<i>lloc</i>

Como vemos, muchos de los tratamientos regulares y más idiosincráticos de cada lengua románica son perceptibles a simple vista ya en la primera correlación, como el excepcional mantenimiento de /u/ para la antigua *u* breve latina incluso átona en el logudorés o como el general tratamiento de /ue/ e n español (*bueno*) para la [ɔ] breve y tónica

del latín y de /uo/ en italiano cuando además la sílaba es abierta (*buono*), tratamientos refrendados en otras series (*fuego – fuoco, luego – luoco*)... La antigua /n/ se mantiene en valenciano solo en el adjetivo ante su sustantivo (*un bon home* “un buen hombre”) pero no en los otros casos (*un home bo* “un hombre bueno”; *res de bo* “nada bueno”). Estas variaciones—al parecer, mucho más libremente—se daban también para los sustantivos en el s. XIII. El “Libro de Alcoy” presenta PER RAON y PER RAZO en la misma línea (Diéguez 2012: 17 ^{r2} ^a1263) y en todo el documento nos topamos con un uso aparentemente indistinto de *razon*, *razo* y *raon*. En consecuencia el arqueoglotólogo debe tener en cuenta la existencia de variantes sincrónicas, reguladas o no, cuando estudie los documentos escritos. Un útil resultado puede obtenerse si se procede con éxito a establecer la secuencia diacrónica: *razon* ≥ *raon* ≥ *raó*, esta última la forma moderna y justamente la única no documentada en el “primer libro” en valenciano. Corolario de ello es que *razo* podría representar un registro grafemáticamente arcaizante, una confusión del escriba, una variante diatópica o bien una—para el reconstructor de lenguas—siempre utilísima ultracorrección. Ha de recordarse que por motivos no arcaicos la escritura presenta regularmente muchas menos variantes que la lengua. Así pues, por razones de economía, salvo en unos pocos casos de más sencillo registro, no reflejamos en el cuadro sinóptico evolutivo de la correspondiente forma latina las con frecuencia diversas variantes fónicas.

Semánticamente se observa cómo una palabra expresiva cual ‘choza’ ha devenido ‘casa’ en la mayoría de las lenguas románicas. Congruentemente el sardo ha mantenido

la forma y significado del latín clásico, *domu*– ‘casa’ (sardo campidanés [ˈdɔmu]; Blasco 2016: 40), mientras que *casa*– ‘choza’ se ha mantenido en el sintagma logudorés ‘choza de abejas’ para ‘colmena’ (Meyer–Lübke 1992: 137 ⁿ1728 s. u.: «log. *kaza de abes*»).

DE DIEZ: REGULARIDAD Y CONGRUENCIA

La serie para los resultados de *dece*— evidencia también otro factor que puede ser de gran ayuda en la tarea de reconstruir una prelengua: los cambios fonológicos no solo pueden y hasta cierto punto suelen ser regulares, sino que además pueden y hasta cierto punto suelen ser congruentes, es decir: suelen mostrar patrones de congruencia. Así, si el español trataba como /ue/ la antigua *ō* breve latina y el italiano como /uo/ en sílaba abierta y muchas otras lenguas como /ɔ/, paralelamente la antigua *ě* breve latina es tratada como /ie/ en español (*diez*) e /iɛ/ en italiano y como /ɛ/ en otras lenguas (*dèu*). Fonológicamente, por tanto, el patrón evolutivo resulta armónico y congruente en estas lenguas para las vocales breves latinas, presentándose como diptongos crecientes con análoga estructura en unas lenguas. Por ello, aquellas evoluciones reconstruidas que manifiesten congruencias internas—que, dicho sea de paso, usualmente serán también las establecidas por la Fonotipología concerniente—ofrecen una mayor verosimilitud y son en consecuencia preferibles en los procesos de reconstrucción.

Dada la variedad dialectal y las fluctuaciones ortográficas de las diversas tradiciones, preferimos a veces, en casos de ambigüedad, ofrecer la notación fonética para el sardo logudorés. Mientras no avisemos de lo contrario, seguiremos la notación de PITTAU (1991). De nuevo, pues, restringimos las variantes reales existentes. Aquí unos convencionales *focu* o *foku* pueden de hecho representar fonéticamente en el logu-

dorés centro–oriental [ˈfoku] ≥ [ˈhoʔu] ≥ [ˈoʔu] (Blasco 2016: 25 y 57) y seguramente en el cronológico orden de desarrollo que hemos señalado. Obviamente cada forma es, en realidad, regular en su propio dialecto, evidenciándose así que muchas veces la irregularidad se presenta simplemente como la irrupción de la regularidad de un dialecto en otro.

Fenómenos apenas apuntados, como el tratamiento de la latina /l-/ inicial como /ʎ-/ en valenciano, mostrarían inmediatamente una congruente regularidad interna en aumento en cuanto consideráramos otros testimonios: latín *lāna*– ‘lana’, *libru*– ‘libro’, *longe* ‘lejos’, *lūna*– ‘luna’... ≥ valenciano *llana*, *llibre*, *llunt*, *lluna*...

PRODUCTIVA Y PREDICTIVA SEGURIDAD

La principal ventaja de las correspondencias regulares es la seguridad que ofrecen para la reconstrucción, de modo que repercuten directamente en los criterios de productividad y predictividad que resultan ser tan buenos índices de las teorías científicas acertadas. Regulares correlaciones del tipo español *puerto*, francés y valenciano *port*, italiano y portugués *porto*... nos llevarían, de no disponer de la evidencia de la lengua latina, a postular correctamente un antiguo **pɔrtu-* (= latín *pōrtu-*).

Un principio metodológico razonablemente operativo, máxime cuando una raíz se haya conservado en un amplio número de voces y lenguas, es suponer que siempre quedará al menos una lengua, un habla, un dialecto que sí haya conservado, prácticamente tal cual, la forma original, como el logudorés *chentū* y ‘ciento’ (≤ latín *centu-*) o *chena* ‘cena’ (≤ latín *cēna-*). Esta posición maximalista nos parece más realista y cercana a la verdad—y, por tanto, más legítima—que la maximalista posición contraria de suponer que ninguna palabra en ninguna lengua habría conservado jamás fielmente una voz de un estadio anterior, que viene a ser en la práctica la situación de la lengua matriz indoeuropea desde la perspectiva laringalista, ya que al no haber conservado ninguna lengua indoeuropea dicho supuesto segmento, nos veríamos obligados a limitar nuestra posibilidad de encontrar coincidencias perfectas únicamente a aquellas muy pocas raíces reconstruibles sin segmento laringal.

En formas francesas como *chaise*, con palatalización de la velar /k/ ante /a/ o *feu* y *lieu*, sin resto de la antigua consonante intervocálica, a diferencia en ambos casos de lo que sucede en las demás lenguas románicas arriba registradas, se perciben ya los primeros indicios del talante fuertemente evolutivo del francés. La regularidad tiene más valor en estos casos, ya que se ha mantenido superando mayor número de condicionantes que en el caso de lenguas más conservadoras, como el logudorés, que simplemente, perdida ya de modo general en polisílabos la *-m* final en fase prerrománica, prácticamente no ha experimentado ningún cambio en sus *chentú* y *chena*. Así pues, hay regularidades de diverso nivel y, por tanto de diferente fuerza probatoria.

IRREGULARES: SINGLADURA ACCIDENTADA PERO EXITOSA

Adicionalmente tenemos formas con evoluciones irregulares, pero que, aun así, por diferentes indicios apuntan a una preforma común que, en efecto y gracias a la documentación disponible, podemos luego corroborar como efectivamente existida. Tal sería, por ejemplo, el caso de las voces románicas provenientes del latín *alteru-* ‘otro’, *et* ‘y’, *facere* ‘hacer’, *sum* ‘soy’ y *suu-* ‘suyo’, para las que en la mayoría de los casos es observable una falta, mayor o menor, de regularidad en la evolución del latín al romance y consecuentemente una falta de regularidad interna en la lengua románica respectiva:

③ fértiles

latín	<i>alteru-</i>	<i>et</i>	<i>facere</i>	<i>ianuariu-</i>	<i>sum</i>	<i>suu-</i>
español	<i>otro</i>	<i>y/ e</i>	<i>hacer</i>	<i>enero</i>	<i>soy</i>	<i>su[yo]</i>
francés	<i>autre</i>	<i>et</i>	<i>faire</i>	<i>janvier</i>	<i>suis</i>	<i>son</i>
italiano	<i>altro</i>	<i>e[d]</i>	<i>fare</i>	<i>gennaio</i>	<i>sono</i>	<i>suo</i>
logudorés	<i>átteru</i>	<i>e</i>	<i>fághere</i>	<i>bennarzu</i>	<i>so[e]</i>	<i>suo</i>
portugués	<i>outro</i>	<i>e</i>	<i>fazer</i>	<i>janeiro</i>	<i>sou</i>	<i>seu</i>
rumano	<i>alt</i>	<i>și</i>	<i>a face</i>	<i>ianuarie</i>	<i>sînt</i>	<i>său</i>
valenciano	<i>atre</i>	<i>i</i>	<i>fer</i>	<i>giner</i>	<i>soc</i>	<i>son / seu</i>

La *-e-* del latín *áltĕru-* solo ha permanecido en el sardo, así logudorés (*átteru*) como campidanés (*àteru*);

nuevamente un llamativo arcaísmo. La forma italiana *altro* es regular, pero la *o-* del español *otro* y portugués *outro* son inesperadas. El antiguo provenzal presentaba ya el doblete *altre – autre*. Aunque se podría esperar que catalán y valenciano, con una lateral tan velarizada ([ɫ]) y propensa, por tanto, al paso a [w], siguieran la pauta de la variante provenzal *autre*, sin embargo, siguieron la de *altre*, mientras que sorprendentemente la solución del español y portugués comporta la solución *autr-*. Posteriormente el valenciano además perdió la /l/, hecho acaecido al menos en el s. XVIII. En toda la literatura popular del siglo siguiente—el siglo de plata: el ochocientos valenciano—ya no hay resto alguno de /l/ en esta forma. Lo interesante sobre todo a título ilustrativo es que esta irregularidad comenzó muy probablemente en posición átona y de allí se extendió a la tónica. Así, por ejemplo, el compuesto *atresi* ‘otrosí – además’ aparece ya sin /l/ en el s. XIII como testimonian los tres libros valencianos de la corte (Diéguez 2013: 30^r40^a1263: ATRESI; 93^r231^a1263: ATRESSI; Torró 2009: 173^r156^a1275: ATRESSI; 216^r289^a1275 y 279^r37^a1277: ATRESI; Guinot & al. 2008: 89^r85^a1280 *bis*: ATRESI; *cfr.* ítem provenzal *atresi* frente a italiano *altresi*).

Algo similar debió de suceder con el infinitivo valenciano *vore* ‘ver’, que coloquialmente puede presentarse con una ultracorrecta *-r* paragógica: *vórer*, ya que esta debió de generarse a partir de una secuencia /au/ en sílaba átona, tipo **vaurà* ≥ *vorà* ‘verá’ y luego se extendió analógicamente a posiciones tónicas, como el antiguo infinitivo *veure*, forma procedente de un **vedre*, con el típico tratamiento común al catalán y valenciano de /u/ por /d/ en posición implosiva (latín *sede-* ‘sede – ubicación’ ≥ valenciano *sèu* ‘catedral –

capital de diócesis'). El desencadenante de aquel cambio está en una pertinaz particularidad del valenciano: la neutralización de /a/ y /e/ protónicas en favor de la primera vocal, tendencia operativa desde la primera documentación escrita de la lengua, allá en el siglo XIII, y hasta la actualidad y factor, en definitiva, responsable del paso de tipo *veurà* a *vaurà*. Así pues, es constatable que una *irregularidad* puede originarse a veces *regularmente* en un contexto determinado, como aquí en posición átona, para luego extenderse, normalmente vía analógica, a otros contextos, como aquí a posición tónica. Importante también el hecho de que ciertas irregularidades pueden tener su origen en fenómenos de ultracorrección, como en el citado *vórer* y no pocos casos análogos en valenciano: *deféndrer* 'defender', *escriurer* 'escribir', *víndrer* 'venir'...

RECONSTRUYENDO A TROMPICONES

La serie de correspondencias para la antigua conjunción copulativa latina dejaría probablemente reconstruir una correcta forma latina **et*, sobre todo si uno cuenta con el concurso del testimonio gráfico y eventualmente fonético del aquí conservador francés. El tratamiento español /i/ es irregular y convive con la antigua y esperada solución /e/ que se mantiene por eufonía, aunque más en el registro escrito que en el uso hablado, ante [i] inicial en la palabra siguiente. En el español se optó gráficamente por la notación de la variante asilábica <Y>, que se da regularmente ante vocal otra que [i]. Los normativistas catalanes eligieron, como de costumbre, la opción contraria: la notación de la variante silábica <I>, si bien aquí /i/ podría constituir un castellanismo, ya que en los textos más antiguos es constante y más perdurable la grafía <E> para la conjunción copulativa. En lo ortográfico el valenciano, tanto el genuino como lógicamente el denominado *normativo*, ha[n] seguido, *nolis uelis*, el modelo catalán, si bien históricamente el valenciano ha empleado de modo indistinto <I>, <Y> e incluso <HI> para la notación de la copulativa. El hecho de que /et/ pueda reducirse a esa variante semivocálica mínima de [j] ilustra el fuerte proceso de reducción fonética que suelen experimentar las voces más frecuentes.

Verbo también frecuentísimo, el /'fakere/ latino aparece sin resto de la oclusiva velar en algunas lenguas románicas. Mientras en valenciano, para el cual en época

antigua se puede reconstruir cómodamente un infinitivo **faer* (cfr. *facienda*– ‘quehacer’ ≥ *faena*; ítem *faedor* ‘hacedor’, *faent* ‘haciendo’, *faeren* ‘hicieron’, *faes* ‘haces’ etc. en S. Pedro PASCUAL y otros textos del s. XIII), la desaparición de la antigua /-k-/ interior ante /i/ y ante /e/ es regular (latín *vicīnu*– ‘vecino’ ≥ *veí*; latín *decem* ≥ *dèu* ‘diez’...), ni en francés ni en italiano esperaríamos su desaparición, ya que de normal la antigua /k/ de una u otra forma deja su huella: francés *voisin* – *dix*, italiano *vicino* – *dieci*. El sardo logudorés en su variante de Nuoro presenta la misma forma fonética que el latín: *fáchere* [‘fakɛrɛ] (Pittau 1991: 116; Blasco 2016: 95).

Igualmente desde las correlaciones que encontramos para el ‘soy’ en las lenguas románicas sería difícil restituir el *sum* latino en su integridad. En italiano la –o final en *sono* es paragógica, dada la renuencia de esta lengua a presentar consonantes finales y se da en otras formas verbales en análogo contexto: latín *sunt* ‘son’ ≥ *sono*, *amant* ‘aman’ ≥ *amano* etc. También la –c de *soc* en valenciano se extendió por vía analógica desde posiciones donde era naturalmente esperable, como en los antiguos incoativos (tipo *agraixc* ‘agradezco’) a otras formaciones, como *lligc* ‘leo’, *perc* ‘pierdo’ o *vullc* ‘quiero’. En Pere PASQUAL, siglo XIII, la forma regular es todavía *som* ‘soy’, forma incomódamamente homofónica de *som* ‘somos’. Es obvio que la nivelación por analogía—y no solo por razones meramente cuantitativas de un mucho mayor número de formas en el paradigma—es un fenómeno mucho más frecuente en el verbo, el cual por su complejidad requiere una mayor sistematización, que en el nombre, el cual

implica habitualmente muchas menos categorías morfológicas.

Similarmente algunos posesivos románicos comportan a veces elementos que no tienen reflejo en el latín escrito que nos ha pervenido. Desde época antigua el valenciano conoce dos formas de adjetivo posesivo, tal como el antiguo provenzal, que distingue formas tónicas con, entre otras variantes, las masculinas *meu*, *teu*, *seu* y las correspondientes formas átonas *mon*, *ton*, *son* (Schultz–Gora 1973: 77), pauta lógicamente seguida por el valenciano. Así en el *Llibre d’Alcoy* tenemos PROCURADOR SEU [...] FRARE SEU “su representante [...] su hermano” (Diéguez 2012: 19^r10^a1263) junto a SON MARIT (*bis*) “su marido” y SON CORZ “su cuerpo” (Diéguez 2012: 24^r26^a1263). En el valenciano moderno estas últimas formas que regularmente precedían al sustantivo, han quedado de modo habitual restringidas a la expresión de referentes más cercanos o íntimos: *mon pare* ‘mi padre’, *ta mare* ‘tu madre’, *son germà* ‘su hermano’, *sa casa* ‘su casa’... mientras que el antiguo posesivo tónico reforzado con el artículo se emplea en los demás casos: *el seu procurador* “su procurador”, *el seu cos* ‘su cuerpo’ etc.

GASTAR HASTA DESGASTAR

Aunque las formas románicas no presentan una regularidad interna que permita una proyección externa y reconstruir con relativa certeza una preforma común, sí apuntan decididamente a la existencia de esta. Evidentemente si no dispusiéramos aquí de la lengua de salida, aquellas preformas comunes, a causa de la irregularidad de sus resultados, serían mucho más difícilmente recuperables para cualquier Lingüística reconstructiva en comparación con las más prolíficas formas del tipo ②. Una—seguramente la principal—característica de todas estas formas es que se encontrarían entre los 25 lexemas más frecuentes en latín. Ello da la razón a la repetida demostración por MAŃCZAK (*u. g.* 1969, 1977, 1987, 1992, 1995, 1999, 2001, 2005, 2009, 2011 etc.) de que en las lenguas lo más frecuente—*ergo*, de alguna manera, lo más importante—suele ser irregular, ya que acaba evolucionando de manera más rápida que el grueso de aquellas formas que presentan una menor frecuencia. «A menudo se ha hecho notar que las formas más frecuentes son las más irregulares»¹ reconocía ya también GREENBERG (2005: 68).

Así, lo más frecuente suele ser irregular y seguir su propio tratamiento—y no sólo fónico—frente al tratamiento de las formas con una frecuencia de empleo, digamos, más *regular*—en el sentido esencial de ‘menos frecuente’—y

¹ «It has been often noted that the most frequent forms are the most irregular».

frente a las formas de escaso empleo, como suelen ser las denominadas voces *cultas* en tantas lenguas. Un *método* de reconstrucción lingüística que no contemple que, por ejemplo, la raíz del latín *multum* ‘mucho’ pueda generar más de un resultado *regular*, más de un resultado *patrimonial*, cuales *muy* y *mucho* o los documentados *mu* y *muncho* en ámbitos dialectales o coloquiales además de la raíz conservada *mult-* en el cultismo *multitud* o similares, está evidentemente condenado al fracaso. Lo cierto es que los hablantes—y consecuentemente las lenguas—no emplean una única *marcha* de velocidad sino varias, dependiendo del *terreno* lingüístico en general. La pérdida *irregular* de /d/ intervocálica en las denominaciones valencianas de los internacionales platos de la *fideuà* y la *paella* (del latín *patella* ‘plat[ill]o’ y *sic* ya documentada en 1265, Diéguez 2012: 108 ^r278: UNA PAELLA) y tantas otras, frente a su conservación en formas menos populares como *cadella* ‘cachorra’, *escudella* ‘escudilla’ o *vedella* ‘ternera’, manifiesta con todas sus connotaciones el carácter más popular, frecuente e informal de las voces sin /d/.

Lo que más se gasta, más se desgasta. No puede sorprender que muchas formas monosilábicas estén ya en el punto de partida de la evolución... y también a menudo en el punto de llegada. Congruentemente con lo dicho nombres monosilábicos frecuentes en latín como *ui-* ‘fuerza – violencia’ prácticamente no han dejado ningún rastro en romance. Tampoco sería fácil restituir el *ego* ‘yo’ latino con los divergentes resultados de catalán *jo*, español *yo*, francés *je*, italiano *io*, logudorés *e[g]o*—la forma mejor colocada—portugués y rumano *eu*, provenzal *ieu* y *eu* o valenciano

yo. Para el antiguo logudorés encontramos, aparte de *eo* y *ego*, también *deo* y *dego* (Pittau 1991: 88). Una variante dialectal ha, por tanto, conservado pasmosamente la forma prácticamente original. Se aplicaría aquí, pues, el anticipado principio reestructivo siempre orientativo de que al menos algún dialecto—normalmente remot[ísim]o—ha preservado la forma ancestral.

POLISILABISMO, DOBLETES Y OTROS OBSTÁCULOS A LA REGULARIDAD

Congruentemente con lo dicho las formas polisilábicas suelen verse afectadas por este tipo de *irregular* reduccionismo. Los largos numerales rumanos *unsprezece* ‘once’, *doisprezece* ‘doce’, *treisprezece* ‘trece’ etc. se reducen en el habla coloquial a *unșpe*, *doișpe*, *treișpe*.

Un polisílabo típicamente reducido de modo irregular en las lenguas románicas es la voz latina para ‘ombligo’, pues del diminutivo latino *umbilicu-* proceden en última instancia los regulares español *ombligo*, italiano *ombelico* y también provenzal con *ombelic* pero no con su variante regular *embelic*. Por una o otra razón son, en cambio, irregulares los resultados de asturiano *embeligru*, catalán *llombrígol*, francés *nombril*, friuliano *umbriçon*, gallego *embigo*, gascón, parte del catalán y valenciano *melic*, italiano tradicional *bellico* y *bellicolo*, logudorés *imbíligu* o, por ejemplo, el también sardo [‘iqđiyu] en el dialecto de Talana en la Ogliastra (Blasco 2016: 43), occitano *embonilh* o *embonígol*, portugués *embigo* o *umbigo*, rumano *buric*, veneciano *bonigolo*... Se acepta que en el catalán normativo *llombrígol* puede deberse a una justificada interferencia del nombre para la lombriz que es precisamente *llombrígol* en valenciano, un diminutivo de *lumbricu-* ‘lombriz’, denominación aquella que estaría metafóricamente bien justificada por la semejanza del cordón umbilical a tal invertebrado vermiforme. Como parte

muy sensible del cuerpo y tan íntimamente ligada al parto, no puede tampoco descartarse una motivación tabuística en las irregularidades presentes en los nombres románicos para esta singular parte del cuerpo.

Las voces para ‘aguja’ proceden en sede románica ora de una raíz latina con vocal larga (*acūcula* ≥ español *aguja*, francés *aiguille*, portugués y provenzal *agulha*, valenciano *agulla*...) ora de una raíz con vocal breve (*acŭcula* ≥ bergamasco *goĝa*, parmesano *agoča*...; *vide* Meyer-Lübke 1992: 9ⁿ120 s. u.). Es bien posible que muchas veces los dobles tratamientos románicos se deban a dobletes ya existentes en el mismo latín. Así, en general el diptongo latino /ae/, que ya pronto alternó en latín con la realización monoptonga [ɛ:], suele presentar resultados oscilantes en romance. En la voz *æqualis* ‘igual – raso – liso’, por ejemplo, el antiguo diptongo dio los pocos esperables resultados de *igual* en los romances peninsulares y de *uguale* en italiano. El mismo viejo diptongo desde la voz *æramen* ‘cobre’ dio como resultado /a/ en catalán, provenzal o valenciano *aram*, español *alambre* y rumano *aramă*, pero no dejó rastro en el italiano *rame* o en el sardo campidanés *ràmini*, que a causa de la intolerancia a /r/ inicial bien documentada en la isla conoce también una variante *arràmini* pero con /a/ simplemente como resultado de una prótesis vocálica (Casciu 2001: 53 s. *arràmini*, *ràmini*).

Aceptémoslo con humildad: las lenguas presentan irregularidades por diversas causas estrictamente lingüísticas—desgaste por frecuencia, excesivo volumen silábico,

dobletes en la lengua de partida...—lo que no significa que la inexistencia de correspondencias regulares en las lenguas descendientes comporte la inexistencia de una misma fuente de origen.

RAZONES EXOFÓNICAS PARA LA IRREGULARIDAD

Pero es obvio también, tal como estamos viendo, que no todas las *desviaciones* del patrón evolutivo esperado se deben siempre a razones puramente lingüísticas o puramente fonéticas, es decir: relacionadas con la articulación o la audición de los sonidos; hay también factores exofónicos que tienen que ver sobre todo con la ideología o psicología de los hablantes. La analogía, como fenómeno fono–morfológico, podría considerarse una motivación mixta por su fuerte componente mental de tipo asociativo. Con todo, hay sin duda otras modalidades que ciertamente exceden los límites de la lengua propiamente dicha al menos en sus aspectos fono–acústicos.

Ya habido ocasión para ponderar el efecto de los diversos tipos de tabú. No hay ninguna *ley* fonética en valenciano que convierta una /o/ en /i/ y la presencia de esta vocal en *collins* o en un aun más eufemístico *collinses* no tiene otro objeto fuera del de evitar la rotunda sonoridad de una palabra malsonante.

Vimos también cómo el fenómeno de la interferencia léxica o semántica de otra forma intervenía en el no esperado resultado de valenciano *encara* ‘todavía’ o catalán *llombrígol* ‘ombligo’. Igualmente en el sardo logudorés *esser in flamore* para “estar enamorado” parece responder a un cruce de *amore* ‘amor’ y *flama* ‘llama’ (Meyer–Lübke 1992: 29ⁿ427 s. *amor*). En las comedietas valencianas del s. XIX se utiliza ampliamente el cruce lingüístico con fines humorísticos, especialmente el bilingüe valenciano–español, para retratar

cultural y socialmente a los personajes, por lo que podemos encontrarnos con creaciones como *colomisar* ‘economizar’ en vez de *economisar* por un cruce con *colom* ‘palomo’ (Balader 1872: 14: *volent colomisar*). El cruce del español *de balde* y su sinónimo valenciano *debadés* da lugar a un jocoso *de Valdívía* (Martí 2006: 67 s. *baldívía*) con interferencia asimismo del nombre propio *Valdivia*. En valenciano encontramos además *abrecoc* ‘albaricoque’ (Beltran & Segura-Llopes 2018: 185 y 271) junto a *albercoc* por influencia seguramente de la variante popular y usual *abre* ‘árbol’ (Beltran & Segura-Llopes 2018: 183_d). Asimismo un original sintagma *a boqueta nit* ‘al atardecer’ se transforma a menudo en un *a poqueta nit* (Beltran & Segura-Llopes 2018: 212_b)—literalmente “a poquita noche”—por obvio cruce con el diminutivo de *poc* ‘poco’. Son fenómenos, pues, cercanos a otra modalidad básica de interferencia lingüística: la caricatura, de la que nos hemos ocupado en otros lugares.

Cara, pues, a la reconstrucción hay que tener en cuenta que de diversa manera todo el léxico se puede ver afectado y, por tanto, *irregularmente* tratado por motivos como los afectos, la censura, la expresividad, el humor, la ironía, el miedo, la parodia, la pedantería, el pudor... Algunas clases léxicas y algunos campos semánticos—familia o parentela, obscenidad y sexo, referencias sonoras, superstición y religión...—pueden ser especialmente propensos a resultar afectados por aquellos factores. La casuística es muy amplia.

EXITOSAS PARODIA, CREATIVIDAD E IMPERICIA LÉXICAS

El lenguaje considerado culto—consciente o inconscientemente, voluntariamente o involuntaria—suele acabar siendo parodiado en ciertos registros. Una mezcla de pedantería, ironía, parodia y humor parece ser la responsable de que el culto greco–latinismo *metaphora* ‘metáfora’ haya generado la *matàfula* ‘engañifa – embrollo’ valenciana, forma popular como evidencian sus múltiples variantes: *matàfola*, *matàfula*, *mantàfula*, *matafaula*, *mantànfula*, *mentàfula* (Martines 2000: 205) o su productividad léxica, que ha dado derivados como *matafulleria* o incluso un verbo como *matafullar* con posible interferencia de *fuller* ‘fullero’. También el cultismo greco–latino *cacemphaton*, término de la gramática, para describir una secuencia de sonidos malsonantes o una expresión de doble sentido y con significado involuntariamente obsceno generó los españoles *gazafatón* y un aun más memorable *gazapatón* ‘expresión malsonante – disparate’ por su asociación con *zapato*. El valenciano *toletole* ‘rumor – guirigay’ y también ‘obstinación – cabezonería’ procede del *tolle, tolle!* bíblico (Jn. 19,15: *Clamauerunt illi: “Tolle, tolle, crucifige eum!”*), mientras que el *adefesio* español procede—por razones debatidas—del título *ad Ephesios* de una epístola de San Pablo a los habitantes de Éfeso, en Asia Menor.

Parecidamente creaciones colectivas y anónimas de compuestos en español cuales *patidifuso*, *puticlú* o

propuestas individuales como *abrazafarolas*, *chupóptero*, *locatigüisqui* o *maricomplejinis* responden más a la inventiva, ingenio, buen humor y receptividad de los hablantes que a universales *leyes* de la evolución lingüística.

A veces la irregularidad se debe a la imprecisión léxica y esta a la impericia en el dominio del idioma cuando practicada por hablantes de otras lenguas. A causa probablemente de la naturaleza viajera de sus hablantes y al habitual componente internacional de sus participantes el lenguaje náutico suele comportar también muchas irregularidades. Nuestra voz *bitácora* procede del francés *bitacle*, forma que a su vez tiene su origen en el ‘habitáculo – habitacioncilla’ del latín *habitaculum*. Una parte del léxico náutico es internacional, como nuestros *abor* o *estribor*, respectivamente procedentes del holandés y del francés, o *quilla*, que también llegó vía francesa. Del holandés vendrían también los verbos románicos equivalentes a nuestro *amarrar*.

Es notoria la laxitud que presentan las correlaciones entre los anemónimos o nombre de viento mediterráneos: el *ábrego* español remite con dificultad al *Africu-* o viento ‘africano’ de los romanos, el *llebeig* valenciano no ha llegado directamente del greco-latino *Libycu-* ‘líbico – africano’, como prueba, entre otras cosas, el hecho de que no haya mantenido la posición acentual originaria. Otros buenos ejemplos encontraremos en una lengua con tantos elementos mestizos como el maltés (*vide* Cassar 2011: 203). Obsérvese en el siguiente cuadro la buena correlación de sus anemónimos con

los de otros pueblos indoeuropeos del Mediterráneo (abreviaturas: c. = corso, e. = español, g. = griego, i. = italiano, m = maltés, sc. = serbocroata, s. = siciliano, v. = valenciano).

c.	<i>libecciu</i>	<i>livantu</i>	<i>maistrali</i>	<i>punenti</i>	<i>sciroccu</i>	<i>tramuntana</i>
e.	<i>lebeche</i>	<i>levante</i>	<i>mistral</i>	<i>poniente</i>	<i>siroco</i>	<i>tramontanna</i>
g.	—	λεβάντες	μαΐστρος	πουνέντες	σιρόκος	τραμουντάνα
i.	<i>libeccio</i>	<i>levante</i>	<i>maestrale</i>	<i>ponente</i>	<i>scirocco</i>	<i>tramontana</i>
m.	<i>lbić</i>	<i>lvant</i>	<i>majistral</i>	<i>punent</i>	<i>xlokk</i>	<i>tramuntana</i>
sc.	<i>lebić</i>	<i>levanat</i>	<i>maestral</i>	<i>pulenat</i>	—	<i>tramontana</i>
s.	<i>libìci</i>	<i>livanti</i>	<i>maistrali</i>	<i>punenti</i>	<i>sciroccu</i>	<i>tramuntana</i>
v.	<i>llebeig</i>	<i>llevant</i>	<i>mistral</i>	<i>ponent</i>	<i>eixaloc</i>	<i>tramontana</i>

Notemos también la frecuente existencia de variantes—sobre todo cuando su origen es más exótico—así, el español conoce también una forma *jaloque* y el siciliano un *sciloccu*. Este mismo viento puede ser en portugués *siroco* o bien *xaroco* y en valenciano *aixaloc* o *xaloc*. Incluso en la lengua de un territorio interior y sin costa, como Aragón, pueden resultar también familiares algunos de estos nombres. Así, en altoaragonés encontramos *saloc* y *tresmontana*, además de *garbín*, anemónimo conocido incluso en griego: γαρμπής y que es también el *garbí* de catalanes y valencianos. Buenas correspondencias encontraríamos así mismo en aragonés y español *gregal*, corso *grigali*, griego γράλιος o italiano *grecale*.

ELOGIO DE LA ULTRACORRECCIÓN

La falsa erudición, la petulancia, en suma, puede también constituir un canal por el cual la irregularidad penetra en las lenguas. El paso de /au/ a /o/ es uno de los fonótipos más frecuente y mejor documentados en muchísimas lenguas, de suerte que estamos en condiciones de afirmar que el contranatural paso inverso de /o/ a /au/ sobre todo en posición átona, cuando se presente, tendrá como principal fundamento la ultracorrección. Por varias razones se piensa que tal pudo ser el caso de *plaudere* ‘aplaudir’ en latín (*vide* Ernout & Meillet 1979: 513 s. u.), de *plaustrum* ‘carro’ (*vide* Ernout & Meillet 1979: 513 s. u.) o de *plautus* ‘de orejas gachas – orejotas’ (*vide* Ernout & Meillet 1979: 513 s. u.). Segura ultracorrección es la intencionada broma del emperador VESPASIANO referida por SÜETONIO. Resulta que un *Flo-ro* escritor reprendió al César romano por decir *plostra* en vez de *plaustra* y aquel se vio saludado al día siguiente como... *Flauro* (Suet. *Vesp.* 22: *ab eo plaustra potius quam plostra dicenda, postero die Flaurum salutauit*).

Las mejores ultracorrecciones son, no obstante, las no intencionadas, las espontáneas, naturalmente. El paso, muy probablemente por ultracorrección, de /o/ a /au/ es muy frecuente en valenciano, como testimonian numerosos dobles: *obrir – aubrir* ‘abrir’, *ofegar – aufegar* ‘ahogar’, *olor – aulor* ‘olor’, *orella – aurella* ‘oreja’, *osar – ausar* ‘osar’... El paso es antiquísimo: en vez de *ociure* ‘matar’ (≤ latín *occidere*) tenemos ya *auciure* en provenzal (Schultz–Gora 1973: 28 ^r47, 49 ^r77 y 101 ^r151) y esta misma forma en el *Llibre de*

Cocentaina (Torró 2009: 437 ^r73 ^a1290: VOLGRENLO AUCIURE) y repetidas veces en la obra de Pedro PASCUAL (Costa 1998: 94 ^{capítulo}26: *auciura*; 99 ^e28: *aucis*, 119 ^e38: *auciuran* [...] *auciure* [...] *auciuran*), en cuyos textos la forma aparece también a veces con /l/: *alciure* (Armengol 1907: II 23 ^r6) con doble ultracorrección, pues a la etimológicamente incorrecta corrección de *ociure* como *aucire* se ha sumado luego la de *alciure*, al entenderse que allí /u/ era el resultado de la velarización de la lateral tan propia del grupo románico *llemosí*. Así en provenzal tenemos paso de /l/ a /u/ ante /d t n/ o /s/ (latín *alteru-* ≥ provenzal *autre*, latín *dulce-* ≥ provenzal *dous*, latín *falsu-* ≥ provenzal *faus* etc.; Cremonesi 1967: 48; Schultz–Gora 1973: 63).

Las ultracorrecciones constituyen una bicoca para el lingüista y el filólogo, ya que de modo fehaciente documentan la segura existencia de un fenómeno contra el que se reacciona, mientras que ante un cambio sin más siempre puede quedar la duda de si se trata de un error escrito o *lappus calami* o bien hablado, el denominado *lappus linguæ*, por lo que necesitaremos mucha mayor documentación para convencernos de que el cambio realmente se ha verificado.

ESCOLIOS A UNA IRREGULARIDAD IMPLÍCITA

En todo caso, la proyección del fértil tipo ③ románico a la reconstrucción indoeuropea supone la probable mas no siempre detectable o detectada existencia de un contingente de formas que apoyarían la unidad lingüística en una fase anterior a las lenguas indoeuropeas, es decir: la legítima asunción de que existen correlaciones irregulares que empero remiten a una preforma real y tan válida o real como en el caso de correlaciones regulares. En la aplicación de dicha proyección al ámbito indoeuropeo cabría tener en cuenta al menos dos matizaciones: por el lado pesimista, el mayor lapso de tiempo en el caso indoeuropeo hace más probable que hayamos perdido mayor cantidad de material irregular potencialmente comparativo; por el lado optimista, contamos, en cambio, con más unidades—lenguas y grupos—por comparar, lo que aumenta las posibilidades de que en alguna de ellas esa relación sea biunívoca y, por tanto, recuperable. Tampoco, por otra parte, conviene olvidar que a mayor tiempo, más posibilidades habrá de que se introduzca por vía de la niveladora analogía una mayor aunque ficticia regularidad.

Lo cierto es que estadísticamente es muy difícil que todos los resultados sean regulares en todas las lenguas, por lo que necesitamos criterios adicionales a los de la regularidad fonética para poder determinar que voces de forma y significación parecidas conforman una serie efectiva y eficiente de correspondencias.

En todo caso, como se dijo, un principio metodológico razonablemente operativo, máxime cuando una raíz se haya conservado en un amplio número de voces y lenguas, es suponer que siempre quedará al menos una lengua, un habla, un dialecto que sí haya conservado, como el español *ombliigo*, el francés *et* o el italiano *e[d]* y el logudorés *ego* la evolución regular esperada, la evolución esperable en términos mayoritarios y sin especial condicionamiento contextual.

BASTARDOS Y MESTIZOS: LA RECONSTRUCCIÓN EN- GAÑOSA

Finalmente nos encontramos con posibles series románicas, a menudo regulares, que en realidad no tienen correspondencia alguna en el antiguo latín:

④ bastardas

latín	(*) <i>camminu</i>	(*) <i>cauallu</i>	† <i>cafe</i>	† <i>guerra</i>	† <i>tabaccu</i>
	—	—	—	—	—
campidanes	<i>camminu</i>	<i>cuaddu</i>	<i>cafei</i>	<i>gherra</i>	<i>tabacu</i>
español	<i>camino</i>	<i>caballo</i>	<i>café</i>	<i>guerra</i>	<i>tabaco</i>
francés	<i>chemin</i>	<i>cheval</i>	<i>café</i>	<i>guerre</i>	<i>tabac</i>
italiano	<i>cammino</i>	<i>cavallo</i>	<i>caffè</i>	<i>guerra</i>	<i>tabacco</i>
portugués	<i>caminho</i>	<i>cavalo</i>	<i>café</i>	<i>guerra</i>	<i>tabaco</i>
rumano	<i>drum</i>	<i>cal</i>	<i>cafea</i>	<i>război</i>	<i>tabac</i>
valenciano	<i>camí</i>	<i>cavall</i>	<i>café</i>	<i>guerra</i>	<i>tabac</i>

En esta cuarta opción nos encontramos con el caso real de que podríamos interpretar algunas formas románicas comunes como heredadas, así, por ejemplo, nuestro *camino*, que es en realidad una antigua y exitosa copia céltica. Sin embargo, la verdad es que estamos ante copias de diversa época y procedencia, las cuales en estos casos gracias a nuestra documentación por lo general podemos precisar. Básicamente, pues, tendríamos copias bastantes modernas, incluso recientes y que, por tanto, presentarían una gran coherencia intrarrománica o también muy antiguas, como el

citado *camino*, tanto que habrían entrado no ya directamente en las lenguas románicas sino desde el propio latín, por lo que en principio se comportarían prácticamente como otra forma patrimonial más en muchos casos. El indoeuropeísta seguidor del método histórico-comparativo, al no disponer de la misma buena documentación del romanista, corre evidentemente el riesgo de reconstruir formas que en realidad o se dieron solo para algunos lenguas por contacto con otra, como ^(*)*camminu-*, o reconstruir formas que simplemente nunca existieron propiamente en la prelengua, como [†]*tabaccu-*.

Poseemos algunos indicios — indicios solamente — para detectar copias *clandestinas* o encubiertas, principalmente:

- la existencia del término en otras lenguas ajenas a la entidad examinada,
- la mejor motivación semántica en dicha lengua ajena,
- la restricción semántica operada en la lengua de recepción,
- alguna particularidad fonotáctica,
- la pertenencia a un campo léxico propicio para la copia, como los inventos o avances tecnológicos,
- detalles históricos como la datación u origen geográfico del referente...

Las correspondencias para ‘caballo’ aparentemente representaría la típica reconstrucción regular, ya que la mayoría de las formas apuntan a un original latino **cavallu-*, forma tardía en su origen designando un caballo castrado y que sin duda debió de introducirse desde otra lengua, ya que

la forma normal en latín era *equus*. Otro aspecto interesante de la cuestión es la supervivencia parcial o periférica de la vieja raíz en el caso del femenino *equa* ‘yegua’, lógicamente no afectada por la castración, así el citado *yegua*, antiguo francés *ive*, logudorés *ebba*, portugués *egoa*, provenzal *ega*, rumano *iapă*, valenciano *egua*...

Nótese que orientativamente la extrema similitud de las correspondencias puede ser un buen indicio de su carácter reciente, así para ‘café’ apenas contamos con las mínimas variantes /ka'fe/ y /kaf'fe/ y el doblete en campidanés *caffè* – *caffei* (Casciu 2001: 96 s. *caffè[i]*), además del *cafea* rumano, lengua donde, a juzgar por su algo mayor densidad lexicográfica, el término pudo introducirse más tempranamente.

Con mínimas dudas puede afirmarse que el término para nuestra *guerra* fue en la mayoría del romance tomado de una voz germánica. El rumano para su divergente *război* siguió su tan común pauta de recurrir a una copia eslávica (*cfr.* ruso *разбóи* ‘pillaje’).

Si la guerra es, por desgracia, una práctica muy antigua, el vicio del tabaco es una cosa mucho más moderna, desconocida por los hablantes naturales del latín, tan afectos por lo demás a muchos otros vicios. Aquí unos elementales conocimientos históricos bastarían para descartar que tan regular correspondencia remontara al *proto-romance*.

Es evidente, por otra parte, que ya el latín incorporara numerosas formas, por ejemplo, del etrusco y del griego y que estas pasaron como patrimoniales al ámbito románico. Por matar dos pájaros de un tiro, citaremos el caso del latín

persōna inicialmente ‘máscara’ y que tradicionalmente es explicada como una copia del etrusco *persu* que sería a su vez una copia del griego πρόσωπον (*vide* Ernout & Meillet 1979: 500 s. u.). Por tanto, también la antigua lengua indoeuropea pudo haber incorporado, adaptado y naturalizado directa o indirectamente un buen número de formas de otras hablas ajenas. Con más razón, por haber más lenguas en juego y más tiempo disponible para la evolución, hay que contar con la presencia de voces bastardas, de origen ajeno, unas antiguas y casi o totalmente patrimonializadas y otras recientes en el antiguo común conglomerado indoeuropeo, máxime cuando, como a continuación veremos, son numerosas las vías por las que un elemento foráneo puede introducirse en cualquier lengua.

VIAJES VERBALES DE IDA

En efecto, dentro del general fenómeno de la copia puede haber diversos tipos. Habitualmente se toma en consideración el proceso más frecuente, el cotidiano uso de un término tomado directamente de otra lengua en razón del prestigio de esta, origen del referente denominado o cualquier otro motivo: una copia directa, de emisor a receptor y sin intermediarios, pero hay otras diversas modalidades de copia. La que acabamos de ver—griego πρόσωπον \geq etrusco *persu* \geq latín *persōna*—es la copia viajera o forma que llega a una lengua trámite una o más lenguas. En el caso de la reconstrucción indoeuropea este tipo de copias son más difíciles de seguir y verdaderamente presentarían muchos problemas para el lingüista forense. Una pena, ya que potencialmente tales copias son muy informativas al incluir más datos—culturales, económicos, [pro]históricos, lingüísticos...—en esa plurirrelación.

De ser correcta, sería, por ejemplo, muy interesante, la propuesta—bien articulada aunque muy hipotética—de ALINEI (2000: 949) de hacer derivar el término latino *seruus* ‘esclavo’ del etnónimo de los antiguos serbios—en serbocroata *Srbi* o Срби—pueblo eslávico situado allende el Adriático. Lingüísticamente, en efecto, el proceso fonético se dejaría explicar bien si tenemos en cuenta la posible intermediación de la lengua etrusca, carente de /b/ pero sí disponiendo de /v/, por lo que lógicamente habrían adaptado un supuesto **serbu* o similar como **serve*. A su vez, no disponiendo el latín de un fonema /v/, los hablantes del Lacio habrían

naturalmente usado su [w] para dar cuenta de la /v/ etrusca y añadiendo la desinencia esperable tendríamos un nominativo *seruus* ‘esclavo’. Señala también ALINEI (2000: 949) que en tal caso el proceso anticiparía «el paso semántico que en la alta Edad Media ha llevado del latín tardío (y griego) *s[c]lavus* ‘eslavo’ al latín tardío *sclavus* ‘esclavo’, como reflejo del enorme tráfico medieval de esclavos de origen eslávico»². Quizá no fuera esta la única palabra que el etrusco habría pasado al latín en el campo semántico de la esclavitud, ya que otras palabras importantes como *familia* ‘conjunto de esclavos y sirvientes – servicio’ o *famulus* ‘sirviente’ (Ernout & Meillet 1979: 215 s. *famulus* y 620 s. *seruus*) y *uerna* ‘esclavo nacido en casa’ (Ernout & Meillet 1979: 724 s. *u.*) se tienen también en latín por posibles etrusquismos.

Más viajero sería aún el término polaco *kanister* ‘bidón’, verosíblemente tomado del alemán *Kanister* ‘bidón’, pero esta forma a su vez debe remontar al latín tardío *canister* ‘cesto – canasta’, forma copiada a su vez del griego *kánastron* (κάναστρον) ‘cesto’, un derivado de *kánna* (κάννα) ‘caña’, término que los helenos, por su parte, tomaron de alguna lengua semítica—*cfr.* acadio *qanu*, hebreo *qanē*, púnico *qn*, ugarítico *qn*...—si bien a su vez los semitas habrían tomado dicha raíz del término sumerio *gin* (Chantraine 1999: 492–3 s. *κάννα*). Esquemáticamente:

² «il passaggio semantico che nell’alto Medio evo ha portato il tardo latino (e greco) *s(c)lavus* “slavo” al tardo lat. *sclavus* ‘schiavo’, come riflesso dell’enorme traffico medievale di schiavi di origine slava».

polaco *kanister* =≤ alemán *Kanister* =≤ latín *canister* =≤ griego *kánastron* =≤ acadio *qanu* =≤ sumerio *gin*.

EL VIAJE VERBAL DE IDA Y VUELTA

Hay todavía otra modalidad digna de reseñar y que comporta igualmente serias dificultades reconstructivas. Son palabras que podríamos denominar de *ping-pong* o también con el germánico término *Rückwanderer* ‘repatriado’ de otras tradiciones, porque van y vienen de unos grupos lingüísticos a otros para acabar, como el hijo pródigo, finalmente volviendo a casa por Navidad. Así, las románicas hablas peninsulares rerecibieron formas procedentes del latín o de otras lenguas románicas vía el semítico árabe, de modo que *alcázar* remite finalmente a lo mismo que el también hispánico *castro*: al latín *castrum* ‘alcazar – fortificación – ciudadela’. La forma fijada en nuestro topónimo *Alcañiz* procede de la voz árabe [al] *kanīsa[tum]* que remontaría al griego ἐκκλησία ‘iglesia’, como en última instancia remiten a esta misma forma los términos *el[e]iza* del vascuence o *knisja* del maltés. *Almonacid* de, por ejemplo, *Algimia de Almonacid* (Castellón) o *Almonacid de Zorita* (Guadalajara), llegó desde el latín *monasterium* ‘monasterio’ trámite el árabe. Numerosos latinismos incluso semánticos, como *remover* con el sentido de ‘apartar’ (*cfr.* inglés *remove* ‘apartar’) están regresando a la casa común románica por la influencia del idioma de los económicamente superpoderosos angloparlantes. El *campus* de nuestras universidades es nuestro *campo* sólo que tomado directamente en su día por los ingleses de la fuente latina, como muchos otros términos del pedante mundillo académico.

Un caso algo más complejo sería el ilustrado por el topónimo *Alcántara*—por ejemplo en *Valencia de Alcántara* (Cáceres)—pues el segmento arábigo *-cántara* podría venir directamente del griego κέντρον o indirectamente vía el latín *centrum* ‘centro’ “referido a la luz del arco de bóveda” (Castaño 2011: 315). En este último caso nuestros *Alcántara* o con diminutivo ya romance *Alcantarilla* (Murcia) serían al menos griego, [latín,] árabe y español. Nótese la tan frecuente conjunción lingüística de complejidad cuantitativa (3/4 estratos) por una parte y banalidad cualitativa (mezcla) por otra. Mientras que el inverso fenómeno tan propio de la Lingüística indoeuropea tradicional de sencillez cuantitativa y excepcionalidad cualitativa en realidad es eso: sencillamente excepcional...

Lo cierto es que frente al usual modelo de evolución lingüística única y rectilínea, tenemos—y no solo en toponimia, claro—asimismo evoluciones plurales ya rectilíneas o ya curvilíneas y tortuosas. En todo caso, respecto a las copias podríamos establecer tres básicas modalidades, susceptibles evidentemente de subdividirse, ya que pueden complicarse más y más en la segunda y tercera modalidades:

- directas, pues tomadas directamente de una lengua,
- indirectas, pues tomadas de una lengua no directamente sino trámite una o más lenguas, lo que implica de alguna manera copias múltiples, y
- redirectas, pues copias que en realidad tienen su origen en la propia continuidad

lingüística, desde la que pasaron de nuevo a la continuidad matriz.

ROMANISTAS E INDOEUROPEÍSTAS ANTE LA RE- CONSTRUCCIÓN

Por tanto y como es lógico, el romanista dispone de una situación de privilegio frente al indoeuropeísta, ya que aquel en realidad no reconstruye propiamente una lengua sino eventualmente la[s] otra[s] lengua—porque a veces no documentadas—que conformaron junto al latín la idiosincrasia de la respectiva habla románica. El romanista salta al vacío de la reconstrucción... con la red del latín. El oneroso componente arábigo del valenciano: *albelló* ‘albañal – alcantarilla’, *alifac* ‘achaqué’, *almagasent* ‘almacén’, *assucac* ‘callejón [sin salida]’, *fardacho* ‘lagarto’, *mà* ‘agua’ en el lenguaje infantil... o hasta posibles calcos, como el ya visto de *en*, del latín *in* ‘en’, con el sentido de ‘con’ son fácilmente detectables, ya que el árabe hispánico, el llamado *andalusí*, es una lengua suficientemente conocida.

Puesto que en latín propiamente dicho no encontramos ninguna raíz para nuestro *páramo* y similares y además tampoco vemos esa raíz en otras lenguas románicas, podemos deducir que fue un lexema incorporado al español desde una lengua distinta del latín y gracias a la información disponible podemos además precisar que dicha fuente era una lengua prerromana—y lógicamente también *corromana*—de naturaleza céltica. Aserto, entre otros datos, refrendado por la existencia de topónimos cuales las ptolemaicas Σηγοντία Παραμίκια (*geogr.* 2,6,49) y Σηγοντία Παράμικια (*geogr.* 2,6,65), la propia emergencia de una forma *PARAMI* en un ara votiva dedicada a Diana (*C.I.L.* II 2660) y hallada en la provincia

de León, la densidad léxica de la raíz documentada ya en época antigua gracias a derivaciones como el aparente patronímico *PARAMONIS* (genitivo) de una epígrafe también de León (Montaner 2001: 22 n24) y el *AMPARAMVM* de un pacto de hospitalidad en lámina de bronce encontrada en Herrera de Pisuerga (Palencia, Castilla y León). Con pocas dudas la raíz debe de manifestarse también adjetivamente en otros dialectos afines con variaciones fónicas cuales un dativo *PARAM[A]ECO* en dos inscripciones, una de Lugo (Galicia: *PARA/MAECO*) y otra de Asturias (*PARAMECO*) etc.

Mucho menos conocida nos es, en cambio, la lengua ibérica, de la que sobre todo tenemos alguna información más segura de carácter fonológico, pero lo ignoramos casi todo sobre la semántica, por lo que no podemos determinar con certeza si, por ejemplo, el fitónimo o nombre de planta *tàrrec* ‘salvia’ tiene algo que ver con el nombre, seguramente de etimología ibérica, de la ciudad de Tarragona (latín *Tarracō* e ibérico *TARACON-* en la leyenda de las monedas ibéricas de la ciudad). Y eso que por suerte aquí al menos, como vemos (*TARACON-*), contamos con información escrita.

Naturalmente, el romanista no tiene siempre tanta suerte en todos los casos, pero, de cualquier forma, su situación privilegiada no puede para nada compararse con la posición muchísimo más precaria del indoeuropeísta a efectos de posibilidad de refrendos reconstructivos.

UNIFORMIDAD OBLIGATORIA EN LA RECONSTRUCCIÓN INDOEUROPEISTA

Así pues, una de las principales conclusiones de los datos y hechos que venimos exponiendo es que los simulacros de reconstrucción *virtual* en los casos de lenguas seguras y reales, porque ya documentadas, ofrecen una casuística mucho más variada que la que propone la Lingüística indoeuropea tradicional, cuyo método—se diría—impone una irreal y plana uniformidad en la prelengua reconstruida. La reconstrucción conforme al método histórico-comparativo comporta, pues, una idea algo rígida y unitaria, ya que el indoeuropeísta puede

- tener dificultades para reconstruir una forma realmente existente pero cuyos resultados fueran poco regulares en las lenguas derivadas, como el *umbilicu-* ‘ombligo’ latino,
- no reconstruir formas antiguas y comunes que no se hayan conservado, como el ubicuo *-que* ‘y’ latino,
- erróneamente no admitir por su mínima representatividad en una única lengua o dialecto el testimonio, como el *cras* ‘mañana’ logudorés, que empero sí haya verdaderamente existido, o
- equivocarse reconstruyendo formas, como el **camminu-* ‘camino’ de un *proto-romanista*, que en realidad no pertenecieran propiamente a la común prelengua patrimonial.

Por tanto, un primer corolario de lo expuesto sería el de que en las reconstrucciones lingüísticas—y en el caso indoeuropeo casi cabría a veces hablar más bien de *resurrecciones* que de reconstrucciones—hay que aceptar humildemente el principio de que ni será todo lo que está, ni estará todo lo que fue. Normalmente no lo habremos reconstruido todo y normalmente habremos reconstruido cosas incorrectas. Por ello es tan importante el aporte cuantitativo y cualitativo en la reconstrucción, ya que cuanto mayor y mejor sea el material de datos disponibles, menor margen habrá para las inevitables lagunas o fallas reconstructivas.

Otro relevante corolario de lo expuesto es que aparentemente todas las palabras *habent sua fata*, por usar la expresión del gramático romano Terenciano MAURO (Keil VI 363 v1286: *pro captu lectoris habent sua fata libelli*), es decir: cada palabra tiene su propio destino, su particular historia y el itinerario de cada una de ellas no suele ser tan plano y lineal como al investigador le gustaría. Al protocolo tradicional, general y sistemático, de reconstrucción lingüística convendría, pues, añadir un escatocolo pormenorizado y, por así decir, artesanal. Así, lejos de ser una operación automática, algebraica o mecánica, la *resurrección* es más bien un fino, sutil, filológico trabajo de artesanía. En todo caso, no se trata de una mera operación intuitiva. El hecho de que su metodología no comporte una aplicación automática, mecánica y sistemática, tampoco significa que el proceso reconstructivo sea puramente arbitrario o caótico, como tampoco lo es el del cirujano. La Lingüística no es una ciencia exacta, tampoco lo es la medicina, pero disponemos de muchos

criterios orientativos que, empleados con inteligencia y buen tino, pueden conducirnos a resultados óptimos porque congruentes y explicativos.

Lo que no parece admisible es una reconstrucción tan sistemática y completa como... irreal ni tampoco proyectar una reconstrucción protohistórica o arqueológica del pueblo indoeuropeo exclusivamente a partir de esas sistemáticas correspondencias regulares cuando estas resultan forzadas y no aportan ninguna ulterior proyección explicativa. Será, en fin, la depuración de las técnicas de análisis, sumada a un mayor y mejor acopio de datos, lo que realmente pueda en el futuro hacer más exitosas las reconstrucciones.

REFERENCIAS

ALINEI Mario, *Origini delle lingue d'Europa. II. Continuità dal Mesolitico all'età del Ferro nelle principali aree etnolinguistiche*, Il Mulino, Bologna 2000.

ARMENGOL VALENZUELA Pedro, *Obras de S. Pedro Pascual. Mártir, Obispo de Jaén y Religioso de la Merced, en su lengua original, con la traducción latina y algunas anotaciones*, Imprenta Salustiana, Roma 1907 voll. II y III.

BALADER Joaquín, *Miseria y Compañía*, Librería de Juan Mariana y Sanz, Valencia 1872.

BELTRAN CALVO Vicent & SEGURA-LLOPES Carles, *Els Parlars Valencians*, Universitat de València, Valencia 2018₂.

BLASCO FERRER Eduardo, «La tipología lingüística del sardo», *Revista de Filología Románica* 17 (2000) 15–29. *Corso di linguistica sarda e romanza*, Franco Cesati Editore, Florencia 2016.

CASCIU Giovanni, *Vocabulariu sardu campidanesu –italianu*, Edizione Grafica del Parteolla, Dolianova 2001₂.

CASSAR Mario, «Maltese meteorological and calendar proverbs», J.E. Gargallo Gil coord., *I proverbi meteorologici. Ai confini dell'Europa romanza*, Edizioni dell'Orso, Alessandria 2011, 177–206.

CASTAÑO FERNÁNDEZ Antonio M. “Extremadura”, E. Casanova ed. & X.Ll. García Arias coord., *Toponimia hispánica. Origen y evolución de nuestros topónimos más importantes*, Denes Editorial, Valencia 2011, 301–318.

CHANTRAINE Pierre, *Dictionnaire étymologique de la langue grecque. Histoire des mots*, Klincksieck, París 1999.

COSTA CATALA Joan S.J. intr. i trans., *La Biblia Parva de Sant Pere Pasqual*, Lo Rat Penat, Valencia 1998.

CREMONESI Carla, *Nozioni di grammatica storica provenzale*, Istituto Editoriale Cisalpino, Varese–Milán 1967₃.

DIÉGUEZ M^a Àngels [& FERRAGUT Concha], *Llibre de la Cort del Justícia d'Alcoi*, Universitat de València – Acadèmia Valenciana de la Llengua, València 2012.

ERNOUT Alfred & MEILLET Antoine [& ANDRÉ Jacques], *Dictionnaire étymologique de la langue latine. Histoire des mots*, Éditions Klincksieck, París 1979₄.

GREENBERG Joseph H., *Language Universals with Special Reference to Feature Hierarchies*, Mouton de Gruyter, Berlín – N. York 2005.

GUINOT RODRÍGUEZ Enric, DIÉGUEZ M. Àngels & FERRAGUD Carmel, *Llibre de la Cort del Justícia de València (1280–1282)*, Universitat de València – Acadèmia Valenciana de la Llengua, Valencia 2008.

MAŃCZAK Witold, *Le développement phonétique des langues romanes et la fréquence*, Uniwersitet Jagielloński, Cracovia 1969. *Słowiańska fonetyka historyczna a frekwencja*, Uniwersitet Jagielloński, Cracovia 1977. *Frequenzbedingter unregelmäßiger Lautwandel in den germanischen Sprachen*, Ossolineum, Wrocław 1987. «Frequenzbedingter unregelmäßiger Lautwandel im Altpreußischen», W. Smoczyński & A. Holvoet edd., *Colloquium Pruthenicum*

Primum, Wydawnictwa Uniwersytetu Warszawskiego, Varsovia 1992, 23–29. «Dévolement phonétique irrégulier dû à la fréquence en arménien classique», *Revue des Études Arméniennes* 25 (1995) 9–16. «Le dévolement phonétique irrégulier dû à la fréquence en hittite», *Archiv für bulgarische Philologie* 3 (1999) 121–131. «Coup d’oeil sur le dictionnaire de fréquence du lituanien», *Poznańskie Studia Bałtystyczne* 1 (2001) 47–53. «Dévolement phonétique irrégulier dû à la fréquence dans les langues turques», *Studia Turcologica Cracoviensia* 10 (2005) 259–267. «50–lecie nieregularnego rozwoju fonetycznego spowodowanego frekwencją», *Bulletin de la Société polonaise de linguistique* 65 (2009) 237–246. «Etymologia przyimka dla a nieregularny rozwój fonetyczny spowodowany frekwencją», *Prace filologiczne* 60 (2011) 189–195.

MARTÍ MESTRE Joaquim, *Diccionari històric del valencià col·loquial (segles XVII, XVIII i XIX)*, Universitat de València, Valencia 2006.

MARTINES Josep, *El Valencià del segle XIX. Estudi lingüístic del Diccionari Valencià de Josep Pla i Costa*, Institut Interuniversitari de Filologia Valenciana – Publicacions de l’Abadia de Montserrat, Valencia – Barcelona 2000.

MEYER–LÜBKE Wilhelm, *Romänisches etymologisches Wörterbuch*, Carl Winter Universitätsverlag, Heidelberg 1992 [= 1935].

MONTANER FRUTOS Alberto, «El ara leonesa de Diana (*CLE*, 1526): constitución literaria y dimensión ritual», *Emblemata* 7 (2001) 9–77.

PITTAU Massimo, *Grammatica della lingua sarda. Varietà logudorese*, Carlo Delfino editore, Sásari 1991.

SCHULTZ-GORA Oskar, *Altprovenzalisches Elementarbuch*, Carl Winter – Universitätsverlag, Heideberg 1973.

TORRÓ Josep, *Llibre de la Cort del Justícia de Cocentaina (1269, 1275–1278, 1288–1290)*, Universitat de València, Valencia 1989.