

## FORMA CANÓNICA Y CAMBIOS EN LA FORMA CANÓNICA DE LA LENGUA VASCA: HACIA LOS ORÍGENES DEL BISILABISMO

Joseba A. Lakarra

### 1. INTRODUCCIÓN<sup>1</sup>

Como he defendido en comunicaciones a ediciones anteriores del Congreso, la evidencia —fonológica, morfológica, léxica y, en menor medida, sintáctica— favorable a la forma canónica (FC) monosilábica en protovasco (PV) antiguo y más tarde resulta abrumadora. La reconstrucción interna de 2º grado (basada en la FC), apoyada por la tipología holística diacrónica, permite alcanzar un PV más antiguo y una explicación coherente de la evolución del vasco, así como el establecimiento de las bases de una etimología “formal”, —imprescindible para complementar y guiar los testimonios filológicos existentes—, tal reconstrucción nos lleva a postular para la prehistoria de la lengua una amplia y larga deriva desde una estructura aislante a otra aglutinante en el conjunto de la misma. Es dentro de esta deriva donde situamos la conversión de la FC de monosilábica en bisilábica y las múltiples consecuencias asociadas a este hecho.

Tras resumir la situación de la reconstrucción previa a la teoría de la raíz monosilábica en PV antiguo (= Lakarra 1995 y ss.), recogemos en §2 los fundamentos y primeros resultados de la teoría en la reconstrucción de la prehistoria de la lengua. En §§ 3 y 4 se discuten dos vías por las que se desarrolla el estudio de la forma canónica y de sus implicaciones: por un lado la relación entre FC y tipología holística diacrónica —presentando pruebas del cambio experimentado por la lengua a partir de los estados más antiguos reconstruibles—; por otro, la necesidad de establecer, dado el

---

<sup>1</sup> Este trabajo se inscribe en los proyectos de investigación *Monumenta Linguae Vasconum* II (HUM2005-0847) y III (FFI2008-04516) del Ministerio y del Grupo de Investigación Consolidado GIC07/89-IT-473-07 del Gobierno Vasco. Agradezco a Julen Manterola su ayuda con el resumen.

Leyenda: T = oclusiva, R = sonante, S = sibilante, V = vocal/verbo; otras abreviaturas como FC o PV son explicadas en el texto.

aislamiento genético de la lengua y de las relativamente escasas disponibilidades filológicas, una “etimología formal”, basada en el estudio de la estructura y evolución de las FFCC de las raíces.

En §5, a guisa de posibles lazarillos para nuestro ensayo, aportamos algunos ejemplos de cambios en la FC de los morfemas experimentados por diversas lenguas como las australianas, la china, las austronesias y algunas tibeto-birmanas, a añadir al caso más conocido de deriva descrito modernamente, cual es el de las lenguas munda. En §6 se da cuenta de una serie de cambios fónicos y morfológicos (metátesis consonánticas de varios tipos, asimilaciones vocálicas, asimilaciones y disimilaciones consonánticas, inserción de /g/ en hiatos, etc.), no estudiados previamente de manera conjunta, los cuales “conspiran” para fracturar antiguas lindes entre raíz y afijos y para consolidar las nuevas estructuras bisilábicas. En §7 investigamos los orígenes del conocido pero inexplicado polimorfismo histórico de la raíz verbal vasca, mostrando la existencia de antiguos prefijos y combinaciones de prefijos que, tras fosilizarse y fusionarse con la raíz CVC anterior, extendieron ésta llevándola a su estructura histórica conocida. En §8 mostramos que la investigación etimológica de los bisílabos nos ha permitido obtener más de medio millar de nuevas etimologías, tanto de voces patrimoniales como de nuevos préstamos. Respecto a estos nos preguntamos porqué la labor multisecular de la lexicografía vasco-románica no ha llegado a detectar tal cantidad de voces (cerca de un centenar); en nuestra opinión, es muy probable que haya sido precisamente la acomodación de tales préstamos a las FFCC más difundidas en la historia posterior de la lengua, —si no al papel estelar de tales préstamos en la adquisición de nuevas FFCC por aquella—, lo que haya causado que fueran tan difíciles de detectar para hablantes y aun para lingüistas.

En §9 se exponen algunos argumentos (escasez, transparencia, reducida extensión geográfica de compuestos y derivados, mayoría de préstamos en modelos bisilábicos) a favor de una datación tardía del cambio de la FC de monosilábico a bisilábico y en §§10 y 11 se reúnen las principales conclusiones y la bibliografía utilizadas; en Apéndice añadido tablas con datos y porcentajes —según el estado de la investigación hacia 2004-2005— de (A) voces documentadas, raíces de etimología desconocida y “fósiles”, de (B) préstamos y onomatopeyas así como de (C) variantes y compuestos o derivados de casi medio centenar de modelos radicales.<sup>2</sup>

---

<sup>2</sup> En Lakarra 2009a hemos presentado un amplio estudio sobre diversos aspectos y consecuencias de la teoría de la raíz monosilábica y su aplicación a la reconstrucción del protovasco, resumiendo en siete apartados otros tantos capítulos del mismo: (1) “Hacia un PV más antiguo”; (2) “Vasconic y reconstrucción del PV”; (3) “Etimología y reconstrucción: historia de paradigmas”; (4) “¿Bisílabos en PV?”; (5) “Reconstrucción y etimología: hacia un nuevo paradigma”; (6) “Prolegómenos a la reconstrucción interna de 2º grado: forma canónica de los morfemas y tipología holística diacrónica”; (7) “Sobre iniciales, dentales perdidas y

## **2. FUNDAMENTOS Y RESULTADOS INICIALES DE LA TEORÍA DE LA RAÍZ MONOSILÁBICA**

Desde hace ya más de una década venimos defendiendo la necesidad de revisar, ampliar y profundizar la reconstrucción estándar de Michelena (cf. 1957a, 1964, 1977a, etc.) mediante trabajos puntuales o de mayor alcance (cf. Lakarra 1995 y ss.). Más precisamente, lo que se defiende es la posibilidad de llegar a estados de lengua más arcaicos que el PV clásico dibujado por Michelena para la lengua de los últimos siglos de la Era anterior, basándonos en la primera —“A new interpretation of the material on hand”— de las cuatro razones que Mary Haas (cf. Lakarra 1997b) presentaba hace cuarenta años para reconsiderar anteriores reconstrucciones. Corresponde al análisis de la forma canónica de los morfemas, en concreto a la teoría de la raíz monosilábica, el núcleo del paradigma reconstructivo adoptado. Los resultados obtenidos hasta ahora no son, desde luego, definitivos, ni alcanzan a la totalidad del léxico, de la fonología, de la morfología y de, particularmente, la sintaxis protovascas, pretensión no sólo excesiva sino probablemente inútil y contraproducente para el avance de la investigación; creo, sin embargo, que conocemos más y mejor sobre la evolución prehistórica de la lengua que hace unos pocos años: i.e., somos conscientes de nuevas generalizaciones como **\*\*TVTV**, luego convertida en **\*\*CVCV** (cf. Lakarra 2004b) o de aspectos de una gramática más antigua como (cf. Gorrochategui & Lakarra 2001) la reduplicación y el uso de prefijos tanto en el verbo como en el nombre, además de centenares de nuevas etimologías que se van ampliando y consolidando en función de los nuevos hallazgos (v. Lakarra en prep.-1).

He reiterado ya en numerosas ocasiones (cf. Lakarra 2006a y véase el último cap. de Trask 1997) que no parece demasiado osado constatar que la comparación tradicional, dirigida a probar parentescos entre lenguas, y estándar —es decir, la evaluable con los mismos criterios y fundamentos utilizados en otras familias de lenguas, particularmente en las mejor establecidas— no ha aportado nada relevante al conocimiento de épocas anteriores a las documentadas en vascuence, ni a la evolución de la lengua, ni tampoco, *a fortiori*, al conocimiento del PV. Esto equivale a decir que, a pesar de constituir una masa bibliográfica enorme, muy por encima del resto de la producción sobre diacronía y lingüística vasca *tout court*, no ha cumplido su cometido fundamental, el mismo o similar al que tiene entre lingüistas profesionales en la familia IE, en la semítica, en la urálica, en la austronesia o en otras. Es más, podría argüirse que en más de una ocasión (cf. Trask 1997 y Gorrochategui-Lakarra 1996) la comparación a ultranza ha llevado a muchos “tratadistas” a encarar con orejeras muy particulares, cuando no a falsificar, la realidad y, en todo caso, a postergar y dificultar el

---

FFCC verbales”. Véase también Lakarra 2007b para una primera síntesis de la teoría de la raíz y de la investigación asociada a la misma.

análisis de problemas reales y relevantes de la diacronía del vascuence que podían y debían ser encarados desde la propia lengua.

Muy otra es la situación de la reconstrucción interna, ejemplificada en la obra de Michelena: guiada por el conocimiento de los datos reales y de la filología pertinente, y por la mejor teoría lingüística a disposición del investigador, ha permitido conocer aspectos del pasado de la lengua y de su evolución que de otra manera no hubieran podido alcanzarse. No es probable que la proporción entre los magros e inciertos resultados de la reconstrucción comparada “genética”<sup>3</sup> y los espléndidos y esperanzadores de la interna cambie radicalmente en un futuro próximo o lejano; más bien hay fundadas razones para sostener que las diferencias entre ambas —abrumadoras a favor de la segunda— pueden acentuarse más y más, siempre en función de la labor de sus cultivadores, naturalmente. Como esta descripción de la situación puede parecer discutible a alguno, no oculto que he dedicado bastante más tiempo y esfuerzo a la reconstrucción interna que a la comparada, la cual, por cierto, las menos de las veces acaba (al menos por lo que a la parte vasca se refiere) en nada parecido a una reconstrucción y en nada que pueda pasar por algún tipo de explicación.<sup>4</sup>

Lo anterior no añade gran cosa a lo que Michelena y Trask dejaron meridianamente claro para casi todos. En realidad, considero que la comparación genética desarrollada a partir de 1950 (= “De etimología vasca”) y, sobre todo, de 1961 (= 1ª ed. de la *Fonética*) no es sólo una monumental pérdida de tiempo propio y ajeno, como dijera a otros efectos, más generales, Hamp 1998, sino una labor completamente descaminada, al menos si de lo que se trata es del conocimiento de la prehistoria y, en general, del pasado y de parte del presente de la lengua. El “paradigma Michelena”, constituido entre las dos fechas arriba citadas y rematado con *Lenguas y protolenguas* (1963) y *Sobre el pasado de la lengua vasca* (1964) ya había marcado un rumbo bien distinto al amateurismo de unos y al diletantismo de otros: filología y reconstrucción interna. También los resultados fueron significativamente diferentes: a mediados de los ‘60 contábamos con una reconstrucción del PV reciente, con un conocimiento de todos los testimonios relevantes y con unas hipótesis para relacionar unos y

---

<sup>3</sup> De la otra, i.e., de la tipología holística diacrónica, cabe decir que es probablemente la mejor ayuda que el reconstructor ha recibido de la lingüística teórica los últimos 30 años; véase Lakarra 2005a y 2006a y más adelante en el texto.

<sup>4</sup> Véanse Lakarra 1991 y 1996 sobre varios casos diferentes en los que la lengua vasca es emparejada —por decir algo— con el ibérico o el europeo antiguo entendido *à la* Vennemann; en Lakarra 1997b y 1999 me referí también a lo poco que hemos ganado en el conocimiento del pasado de la lengua con “las modernas técnicas” de la macrocomparación o de la “megalocomparación”, por utilizar el término, tan exacto como poco discutido entre profesionales, de Matisoff 1990. Véanse ahora las divertidas consideraciones finales de Vovin 2009 en torno a los “Encounters of the Third King in Linguistics” sobre un fondo altaico; con todo, es posible que —de haber conocido la historiografía lingüística vasca— no hubiera reivindicado en su trabajo de 1994 el *top* de la fantasía a favor de lo escrito en la nipona.

otros anteriormente inimaginables, todo lo cual había transformado el escenario de manera radical para quienes estuvieran interesados por el pasado de la lengua y, cabría suponer, por cualquier otro pasado relacionado con el de aquella.

En 1995 publiqué el primero de varios intentos de explorar nuevas vías para reconstruir una fase de la prehistoria de la lengua vasca anterior a la alcanzada por Michelena. Partiendo de las regularidades morfélicas de las voces patrimoniales, las cuales guardan información sobre fenómenos y estados de la lengua anteriores a la entrada de préstamos latinos, —base de la reconstrucción estándar (cf. Martinet 1950, Michelena 1951a-b, 1957a, 1957b, 1964, *FHV*)—, se trata de obtener morfemas antes no reconocidos (por fósiles) en PV moderno y esquemas morfélicos que guíen una reconstrucción más profunda, y que sirvan para identificar con alguna seguridad los lexemas y formas gramaticales pertenecientes a tal estadio lingüístico; la labor, por supuesto, no está sino esbozada y no carece de riesgos y dificultades. Con todo, parece establecido (cf. Lakarra 1995, 1998a) que la raíz PV antigua era CVC y entendida, además, esta fórmula de manera mucho más estricta que en IE o en kartvélico, con sus respectivas variantes y ampliaciones para C-, -V- o -C (cf. Lakarra 1998b), a gran distancia del segundo con sus tres y cuatro consonantes prenucleares, núcleos no vocálicos —como otro conocido nuestro y supuesto pariente, el bereber—, armonizaciones varias y otros “matices” que la hacen imposible de reconocer en semejante estructura, por mucho que esto pese a Gamkrelidze e Ivanov (cf. Harris 1990). Ya en 1995 quedó clara la necesidad de C- en la estructura radical y, posteriormente, ha podido reducirse la -C(C) final a una simple -C, con la segunda y última consonante explicada como sufijo. Si -T era imposible en la sílaba (cf. Artiagoitia 1990), entonces tampoco cabía explicar ninguna CVCV como CVT-V —aunque hallamos CVR-V y CVS-V en *\*barr-u* ‘dentro’ y *\*larr-u* ‘piel, pellejo’, además de los más evidentes o conocidos *hez-i* ‘dom-ado’ o *gaz-i* ‘sal-ado’—, a no ser que propongamos sufijos en *\*\*TV* y raíces en *\*\*CV*.<sup>5</sup>

<sup>5</sup> De hecho, no parece haber nada así en vascuense moderno, ni resulta necesario reconstruirlo para un pasado más o menos antiguo ya que,

(1) Varios de los escasos CV (*lo* ‘sueño, dormir’, *ke* ‘humo’ y alguna otra) —en la medida que no corresponden a variantes de un CVC anterior (*su* ‘fuego’ < *\*sur*, cf. *surtan* ‘en el fuego’), son onomatopeyas o fonosimbolismos, como sus correspondientes en otros idiomas;

(2) En los casos de CVCC (*hortz* ‘can-ino’, *bortz* ‘cinco’, p.ej.) cuando -CC no correspondía a un sufijo (cf. *\*hor-tz*, *\*bor-tz*, etc.), el grupo -CC podía ser —sólo o con otros, recuérdense *baradizu* / *paraiso* y *zekürü* / *sekula* (cf. Michelena 1957b, 1964 y, sobre todo, 1974)— indicio de préstamo como en *pertz*;

(3) Si establecemos que no existen lexemas menores que CVC, entonces *lagun* ‘compañero’, *labur* ‘corto’, etc., no pueden ser compuestos normales de lexema con lexema, sino combinaciones de prefijo y raíz;

(4) Las restricciones *\*\*CV* y *\*\*VC* en los modelos de morfemas radicales están, sin duda, relacionadas con la inexistencia de compuestos en *\*\*CVCV*, *\*\*CVCVC* o *\*\*CVCCV*.

En 2002 nos ocupamos de una restricción estructural de la raíz (\*\*TVTV), —señalada por vez primera en 1995—, o, mejor, de su fórmula más amplia y significativa (\*\*CVCV) y de sus orígenes y consecuencias cercanos y lejanos, estudiando uno por uno todos los esquemas radicales potenciales que englobaba tal modelo; no parecía aceptable que la imposibilidad de combinar dos sílabas abiertas (combinación tan corriente entre las lenguas del mundo) debiera quedar sin noticia ni intento de explicación como hasta entonces. Merecía, por tanto, invertir unos cientos de horas en examinar todas y cada una de las voces documentadas<sup>6</sup> y determinar su posible antigüedad como raíces simples en la lengua. En 1995 la inexistencia en la lengua antigua de ese modelo radical fue relacionada con la abundante presencia de préstamos (*bake*, *bike*...) y compuestos o derivados (*bada* ‘pues, si es’, *begi* ‘ojo’, etc...) entre los miembros modernos de tal modelo, lo cual —dado el carácter cuasi universal de CVCV— iba en contra de la extendida creencia, que viene al menos desde Uhlenbeck, de que las antiguas raíces vascas eran bisilábicas.

A pesar del tiempo transcurrido, no veo razones para cambiar mi conclusión de entonces respecto al PV más antiguo; al contrario, creo que el monosilabismo radical es extensible a etapas mucho más recientes de la lengua. Aunque las investigaciones etimológicas avancen más lentamente que lo que algunos quisiéramos, bastantes polisílabos o bisílabos más pueden ahora ser reducidos a monosílabos o explicados como préstamos (v. Lakarra en prep-1); no son todos, ni mucho menos, pero se va abriendo una vía de investigación productiva que antaño no podíamos ni siquiera intuir, y —lo que es más— va proporcionando preguntas, problemas y generalizaciones antes desconocidas, imposibles o sin sentido, signo del desarrollo del nuevo paradigma reconstructivo.

Como hicimos notar en nuestra comunicación al Congreso de Barcelona (= Lakarra 2005a), la forma canónica de raíces y morfemas no ha sido estudiada hasta fechas recientes en la lengua vasca de manera exhaustiva ni aún casi de forma ocasional, ni se ha puesto de relieve su interés para aspectos básicos del quehacer diacrónico como la reconstrucción, la etimología o la comparación de los estados más antiguos reconstruibles de la lengua con sus correspondientes de otras lenguas y protolenguas (v. Lakarra 1998a). Debería ser completamente innecesario repetir aquí la importancia central de la *Fonética histórica vasca* de Michelena y otros trabajos del autor en estos ya 60 últimos años en lo que toca a las labores diacrónicas señaladas. Pues bien, aunque la *FHV* siga siendo absolutamente estándar e insuperada en los aspectos relevantes,<sup>7</sup> y a

<sup>6</sup> En realidad, dejando para mejor ocasión aquellas que contienen diptongos (v. Lakarra en prep.-6), los cuales merecen una atención monográfica.

<sup>7</sup> Creemos haber explicado en Lakarra 2009b de manera más ortodoxa casi un par de docenas de haches “adventicias” en inicial ( $*h_3 > h_1$  y  $*h_2 > h_1$ ). Sobre el tema de la /h/ en general son de la mayor importancia y constituyen un claro avance los trabajos de Igartua 2001, 2006, 2008, 2009.

día de hoy no contemos con ningún diccionario etimológico remotamente comparable a la recopilación de sus etimologías realizada por Arbelaiz 1978, creemos que es factible aplicar ahí una crítica similar a la que Benveniste hiciera en 1935 a la labor indoeuropeística anterior. En efecto, como señalaba respecto a la práctica etimológica y reconstructora IE, “Ce qu'on a enseigné jusqu'ici de la nature et des modalités de la racine est au vrai, un assemblage hétéroclite de notions empiriques, de recettes provisoires, de formes archaïques et récentes, le tout d'une irrégularité et d'une complication qui défient l'ordonnance” (147-8). Benveniste encontraba que se combinaban raíces monosilábicas y disilábicas, raíces “bilíteras”, “trilíteras”, “cuadrilíteras” y “pentalíteras”, con vocal inicial o final, en grado largo o en grado cero, con diptongo largo o breve, con sufijo o alargamiento, etc. Y de ahí su conclusión:

On serait en peine de justifier et même d'énumérer complètement tous les types de racines qui sont attribués à l'indoeuropéen. Il y a ici un abus de mots qui trahit une doctrine indécise. On n'obtient pas de l'indo-européen en additionnant les diverses formes indo-européennes d'un thème verbal ni en projetant dans la préhistoire des particularités d'un état de langue historique. Il faut essayer, par de larges comparaisons, de retrouver le système initial sous sa forme la plus simple, puis de voir quels principes en modifient l'économie (Benveniste 1935, 147-148).

Tampoco parecía verosímil dar por coetáneos en nuestro caso, como integrantes del léxico PV morfemas y raíces como *\*anai-(k)ide* (> *ahaide* ‘pariente’), *\*ahaid(e)-goa* (> *aha(i)koa* ‘parentesco’), *\*anal-gabe* (> *ahalge* ‘vergüenza’), *\*anan(t)z-* / *\*enantz-* (> *aha(n)(t)z-* ‘olvidar’), *\*ba-ahaldadi(e)n* (> *baldin* ‘sí’), *\*(a)bari ba(ga)ko egun* (> *bariaku* ‘viernes’), *\*baR* (> *barren* ‘debajo’), *\*egu-gaitz* (> *ekaitz* ‘tormenta’), *\*gara-un* (> *garaun* ‘sien’), *\*gau-haur* (> *gaur* ‘hoy’, ‘esta noche’), *\*(h)oba-ro* (> *haboro* ‘demasiado’), etc., —incluso eliminando los préstamos que pudieran hallarse— dadas sus multiformes estructuras en inicial, medial o final (v. Lakarra 2008b).

A pesar del interesante precedente de Uhlenbeck<sup>8</sup> y lo que parece tan reseñable, el perfecto conocimiento y profunda estima que Michelena mostró

---

<sup>8</sup> Cf. “(...) je crois déjà rendre service aux études basque[s] en tâchant de ramener les nombreux radicaux nominaux et les racines verbales de la langue basque restés sans analyse à quelques types déterminés, qui doivent, selon toute probabilité, être considérées comme pyrénéens occidentaux anciens. Je laisse à de plus compétents le soin de confronter les types pyrénéens occidentaux anciens établis par moi avec des types caucasiens ou autres. Pour le moment, il m'est indifférent de savoir si tel élément lexical —mot ou racine— possède ou non une parenté caucasienne plausible. Si pareil élément lexical correspond à un type ancien, il devra passer pour pyrénéen occidental ancien, sauf la possibilité qu'un élément non originellement pyrénéen occidental ancien d'importation ultérieure —mot ou racine— se soit rallié secondairement à un type pyrénéen occidental ancien. *Mais cette question est sans importance pour mes investigations, puisque je recherche non pas des étymologies, mais des types.* Des éléments d'origine étrangère secondairement assimilés à d'anciens types indigènes contribuent, eux aussi à affirmer l'existence de types anciens” (Uhlenbeck 1947, 567).

siempre por Benveniste y por su teoría de la raíz,<sup>9</sup> lo cierto es que la forma canónica no representó para él —ni para el resto de vascólogos— un tema de investigación prioritario y mucho menos la base de su labor diacrónica. Las razones metodológicas o teóricas que invitaban y siguen invitando a cambiar de ruta son evidentes, incluso sin necesidad de aludir a paralelos ilustres en la tradición IE, urálica, semítica o austronesia.<sup>10</sup>

En 1995 propusimos que la raíz en PV (al menos en PV antiguo) era no bisilábica como habían supuesto Michelena 1977b, 1979, Uhlenbeck 1947 y la mayor parte de los tratadistas —pero no Azkue 1923-25—, sino monosilábica, de estructura CVC estricta en los lexemas.<sup>11</sup> Tal propuesta permitía explicar generalizaciones como \*-T en sílabas y raíces antiguas o \*-VC, \*-CV en raíces monosilábicas o \*-TVT en bisílabas sin coda, restricciones anteriormente desatendidas o desconocidas, difíciles de abordar sin algún tipo de teoría sobre la forma canónica de los morfemas. Así, p.ej., es posible entender la inexistencia en el léxico patrimonial de voces simples bisilábicas con sílaba abierta en ambas (tipo de combinación sumamente natural y escasamente marcada) frente a la presencia de múltiples préstamos

---

<sup>9</sup> Cf. “Ejemplo típico de las actividades de reconstrucción proseguidas en el escalón más alto, conocido o restituido, que aspiran no sólo a penetrar en el pasado, sino también a comprender mejor las relaciones que entrelazaban los términos —cuyo número es siempre indeterminado en las protolenguas— son las investigaciones acerca de la “forma canónica” de morfemas o raíces; su configuración muestra curiosas limitaciones en semítico común que delatan las afinidades de los fonemas en el sistema, o en indoeuropeo, etc. De aquí teorías generales, como la de Benveniste sobre la constitución de las raíces indoeuropeas” (Michelena 1963, 40).

<sup>10</sup> Así, p.ej., ya señalé al tratar de cierto trabajo de Vennemann (cf. Lakarra 1996, Gorrochategui & Lakarra 2001) que la discusión sobre la posición del acento antiguo entre Martinet (1ª sílaba) y Michelena (en la 2ª) estaba viciada desde un principio por la falta —común a ambos (y no sólo a ellos, claro)— de una teoría de la raíz y de la formación de palabras en PV. I.e., Martinet se sentía obligado a situar el acento en la 1ª sílaba para lograr el deseado efecto demarcativo que, a su vez, derivara la clara diferencia de comportamiento en la evolución de las oclusivas en las diferentes sílabas; sabemos, sin embargo, que no hay que ir muy lejos para encontrar acentos demarcativos fuera de la 1ª posición. Por otro lado, Michelena, aún acertando al colocar el acento PV en 2ª, no podía recoger la interesante intuición de Martinet sobre las oclusivas, ni basar en ella su hipótesis sino en la distribución de la aspiración: dado que ésta se da históricamente sólo antes del acento, este último debía estar al menos en la 2ª sílaba para que fuera compatible con la aspiración en 1ª y 2ª posición. Si hubieran sido conocidos ambos límites de los morfemas, no sólo el izquierdo, —y, sobre todo, si hubiera sido conocida la estructura interna de los bisílabos como PREF + RAIZ MONOS.— es claro que el acento hubiera sido impensable en inicial; por otra parte, el acento en final hubiera conseguido automáticamente los objetivos mencionados de uno y de otro, así como otros (la relación entre PREF/SUF, posición del verbo, etc.) que la tipología holística liga básicamente a la posición del acento; cf. Donegan & Stampe 1984, 2004, Lakarra 2005a, 2006a y aquí en §3.

<sup>11</sup> Naturalmente, la cuestión es muy diferente en morfemas gramaticales y, sobre todo, en fonosimbolismos en los cuales puede pensarse en una especie de imagen especular de la misma.

con tal estructura básica en los bisílabos, por no hablar de los numerosos fonosimbolismos; v. Lakarra 2004b y Apéndice.

Los frutos iniciales de la teoría, presentados ya en anteriores ocasiones en este foro (v. Gorrochategui & Lakarra 1996, 2001) fueron los esperables en un acercamiento como el resumido: a) hallazgo de nuevas raíces monosilábicas y ampliación de redes y familias léxicas con raíces monosilábicas ya conocidas; b) nuevos elementos de la antigua gramática como eran los prefijos que aparecían en el SN y la reduplicación; además (c) la FC nos sirvió para evaluar o diagnosticar comparaciones de la lengua vasca con otras (v. Lakarra 1998b, etc.). Poco a poco la reconstrucción basada en el monosilabismo radical nos ha llevado a abordar o plantear como de necesario estudio para el futuro tres cuestiones, en buena manera interrelacionadas cuales son (1) la historia de los modelos radicales, (2) las consecuencias tipológicas de la teoría de la raíz monosilábica y (3) la periodización y cronología de la prehistoria de la lengua.

### **3. FORMA CANÓNICA Y CAMBIO TIPOLÓGICO<sup>12</sup>**

Entre las posibilidades que abría la teoría de la raíz monosilábica, mencioné ya en Lakarra 2002a la que puede resultar más interesante para las labores reconstructivas, i.e. la necesidad de postular para el PV más antiguo una tipología muy diferente (sin SOV, ni aglutinación, ni ergatividad, ni flexión verbal inextricable) a la del vascuence histórico y algunas posibilidades que muestra tal teoría para el estudio de la deriva lingüística posterior.<sup>13</sup> Creo haber mostrado (cf. Lakarra 2005a, 2006a) que hay razones internas a la lengua para postular un profundo cambio en la fonología de los dos últimos milenios —no necesariamente en la dirección de las lenguas vecinas—<sup>14</sup> y, lo que parece más relevante, buena parte de esos cambios podrían estar relacionados con la estructura y cambios de la morfosintaxis de la lengua durante ese periodo. En concreto, se trataría de la adquisición por la lengua de ciertos rasgos fonológicos (CV, vocales nasales, abundancia de consonantes oclusivas, aumento de oclusivas sordas en inicial, vocales

---

<sup>12</sup> Además de Lakarra 2005a y 2006a, pueden verse Lakarra 2007a y 2007b, más varios trabajos en preparación.

<sup>13</sup> Para Trask 1998 el vascuence ha mantenido en lo fundamental durante siglos y milenios su estructura morfológica y sintáctica libre de toda influencia ajena y similar a la PV; no ha ocurrido así, en cambio, en lo que toca a la fonología de la lengua. Junto a esto, dado que no hay alusión alguna a posibles relaciones entre estructura (y cambio de estructura) fonológica y estructuras (y cambio de estructuras) morfosintácticas, parece que las razones de ese “gigantesco y paciente” cambio de la primera han de ser propias e independientes de otros módulos, sean internas o no a la lengua. Trask es partidario de achacarlo a la influencia de las lenguas vecinas, las cuales habrían tenido más éxito en la fonología y en el léxico por su carácter más superficial frente a la morfología y a la sintaxis.

<sup>14</sup> Véase ahora Reid 2003 para un interesante caso de creación en un par de generaciones de una concordancia verbal compleja en una lengua australiana (en muy desigual “contacto” con el inglés, naturalmente, pero sin que un cambio tan relevante pueda achacársele a este).

iniciales antes inexistentes, armonía vocálica hacia la izquierda...) solidarios en alguna manera con otros nuevos rasgos morfosintácticos incluidos habitualmente por los investigadores en la definición de lengua aglutinante.<sup>15</sup>

En su brillante análisis de la evolución, prácticamente inversa en múltiples aspectos de su fonología, morfología y sintaxis de las familias munda y mon-khmer, a partir de la común protolengua austroasiática, Donegan y Stampe 1983, 2004, etc. han mostrado que, en lo fundamental, desde un punto de vista tipológico ambas familias constituyen bloques compactos y opuestos a cualquier nivel. Tras rechazar el recurso a explicaciones sustratísticas basadas en supuestas influencias divergentes alegadas con anterioridad para explicar el origen de tales diferencias, han hecho notar que los órdenes básicos OV (de las lenguas munda) y VO (de las mon-khmer y de la protolengua austroasiática) dependen a su vez del acento de frase inicial o final respectivo y que, en términos generales, es el modelo rítmico de palabra y frase el “plan interno” que gobierna la estructura holística de las lenguas:<sup>16</sup>

	Munda	Mon-khmer
Phrase Accent: <sup>17</sup>	Falling (initial)	Rising (Final)
Word Order:	Variable - SOV, AN, Postpositional	Rigid – SVO, NA, Prepositional
Syntax:	Case, Verb Agreement	Analytic

<sup>15</sup> Probablemente una investigación renovada de varios aspectos cruciales del cambio fonológico en vascuence —una especie de “relectura tipológica” de la *FHV*— podría llevarnos a una historia fonológica más trabada y a una explicación más profunda, además de a una tipología diacrónica más completa de la lengua.

<sup>16</sup> Cf. “Munda languages not only reversed the accentuation and syntax of Austroasiatic: they also moved from analytic structure with only derivational morphology, to synthetic or even polysynthetic structure with rich inflectional morphologies which include subject and verb incorporation (or at least agreement), case marking, and a far greater freedom of word order. A predominantly morphological grammar like this seems to be typical of dependent-head order (...) If morphologization and head-last order were not clearly associated in the world’s languages, the rich flowering of morphology in each of the Munda languages, once they had adopted head-last order, would certainly make the association clear (...). This indirect association of head-first order with rising word accent and head-last order with falling word accent is the default in most of the world’s languages, and it is in turn responsible for the association of head-first order with prefixing and head-last order with suffixing, as first pointed out by Greenberg” (Donegan 1993, 4-5).

<sup>17</sup> “We further claimed that the complete reversal of structure that occurred in Munda began with a single change from rising to falling accent” (Donegan 1993, 2); cf. “Reversals of rhythm and word order, as Indo-European and in Munda, are not very common in the world’s languages. Niger-Congo (Givon 1975) had a progressive shift like Indo-European —falling to rising and head-last to head-first. Tibeto-Burman, given the prefixing character of proto-Sino-Tibetan (Benedict 1972), may have had a regressive shift like Munda —rising to falling and head-first to head-last. A regressive shift entails the construction of an inflectional system (section 3.3), and surely takes far longer than a progressive shift. Judging from the time depth of the far less complete reversal of type in Indo-European, Munda must have a time depth of several millenia” (Donegan & Stampe 2004, 16).

*Forma canónica y cambios en la forma canónica de la lengua vasca...*

Word Canon:	Trochaic, Dactylic	Iambic, Monosyllabic
Morphology:	Agglutinative, Suffixing, Polysynthetic	Fusional, Prefixing or Isolating
Timing:	Isosyllabic, Isomoric	Isoaccentual
Syllable Canon:	(C)V(C)	(C)V- or (C)(C)'V(C)(C)
Consonantism:	Stable, Geminate Clusters	Shifting, Tonogenetic, Non-Geminate Clusters
Tone/Register:	Level Tone (Korku only)	Contour Tones/Register
Vocalism:	Stable, Monophthongal, Harmonic	Shifting, Diphthongal, Reductive

(Donegan & Stampe 1983, 337; cf. 2004, 3, 16)

Es interesante ver, p.ej., que los sufijos de las lenguas munda no remontan ni siquiera al protomunda (no ya al protoaustrasiático) sino que son innovaciones independientes de las diferentes lenguas de la familia; sí pueden hallarse en munda prefijos fosilizados —que en algún caso se han convertido en sufijos, compartidos con las lenguas mon-khmer—,<sup>18</sup> y que remontan a la antiquísima protolengua austrasiática. Igualmente, los lingüistas han reconstruido para épocas anteriores de las lenguas munda y para el protomunda ricos inventarios vocálicos que, si bien no se han conservado en ellas, han dejado rastros de haber sido eliminadas de maneras diferentes en cada una, mientras eran conservados y ampliados por las mon-khmer en sus sílabas plenas (las segundas) gracias a su acento ascendente. Las munda —dicen Donegan y Stampe—, como el dravídico, urálico, las denominadas lenguas altaicas y el *vascuence*, han desarrollado, gracias a su acento descendente de palabra y de frase, distintos sufijos casuales y postposiciones.

Las lenguas mon-khmer (como la mayor parte de las del SE asiático y el protoaustrasiático) son isoacentuales (“stress-timing”), lo que supone una polarización entre sílabas acentuadas o “mayores” y las inacentuadas o “menores”, estas últimas con vocales reducidas, cuando no eliminadas, como en vietnamita o en mong; las lenguas munda (“mora-timing”), en cambio, erradicaron la distinción entre unas sílabas y otras por igualación de la antiguamente existente entre largas y breves, generalmente promocionando la vocal reducida de la sílaba menor, armonizándola con la acentuada, plena, de la 2ª sílaba.<sup>19</sup> En mon-khmer no existe, en cambio, la armonía vocálica, propia de lenguas de acento descendente, como en indoiranio, dravídico, urálico, etc., pero sí existen múltiples vocales centralizadas (reconstruidas también para etapas anteriores de las lenguas munda) e incontables escisiones y diptongos en las sílabas acentuadas, fenómenos que han llevado

<sup>18</sup> Para algún caso potencialmente similar en nuestra lengua, véase Lakarra 2006a, con bibliografía, y aquí más abajo §5 y ss.

<sup>19</sup> Véase François 1999 para fenómenos paralelos en lenguas austronesias de Nueva Caledonia.

a que en mon-khmer los fonemas vocálicos —entre 3 y 4 docenas— sean incluso más numerosos que los consonánticos.<sup>20</sup>

Siendo muy interesantes en sí los detalles de la reconstrucción y de la evolución de las varias lenguas y protolenguas implicadas, la aportación de Donegan y Stampe es fundamental en otras dos direcciones. Por una parte, constituye un fuerte argumento en la construcción de la tan deseada tipología holística (cf. Plank 1998), que enmarca e interrelaciona los diversos aspectos de cada módulo de la lengua, al menos los más relevantes de la fonología, morfología y sintaxis de la misma: sin ir más lejos, muchos de los rasgos tipológicos y de los cambios experimentados en ellos por el vascuence mencionados más abajo son fácilmente subsumibles en la explicación de los hechos mundas y alcanzan ahora una lógica interna que antes era más difusa o más débil (cf. las conclusiones de Post 2006 sobre las tani). Por otro lado, desde el punto de vista del reconstructor, es evidente que a mayor trabazón (pricipiada) de la fonología, morfología y sintaxis, nuestra labor se hace necesariamente menos anecdótica, anárquica y, tal vez, “personal”, pero, a cambio, la seguridad y la profundidad explicativa de las propuestas ganan claramente. Finalmente, uno diría que la vía trazada por Donegan y Stampe encierra para el reconstructor del PV tantos paralelismos en la evolución y estructura de las lenguas austroasiáticas de datos, cambios, análisis aplicables *prima facie* y evidentes posibilidades de trascender de una manera articulada y principiada el PV clásico,<sup>21</sup> que invitan a desarrollar las consecuencias derivadas de tal modelo en nuestro trabajo actual y futuro.

En Lakarra 2005a, 2006a creemos haber dado una serie de razones que animan a investigar la (pre)historia de un tipo lingüístico (el descrito sumariamente como “canónico”) que ha sido visto por más de un autor como pancrónico o eterno, quizás babélico o anterior. La búsqueda de escenarios compatibles, coherentes o coetáneos de los reconstruidos para las diversas fases de la evolución de la forma de los morfemas —principal, pero no exclusivamente, de la forma canónica radical—, estudiada desde 1995 nos llevó en dos direcciones complementarias: por una parte nos animó a observar en otras lenguas desarrollos más o menos análogos o de interés en

---

<sup>20</sup> Tampoco la estructura silábica y las consonantes se han librado de la evolución divergente de ambas familias: mientras que en munda se han conservado todas las oclusivas (tanto sordas como sonoras) de los antiguos grupos de la sílaba inicial (débil) al desarrollarse una vocal plena en ellas, en las lenguas mon-khmer, sin esta evolución, han sufrido forticiones (sordas > aspiradas, sonoras > sordas) cuando no se han disimilado entre ellas e incluso con las vocales adyacentes.

<sup>21</sup> Si combinamos la evolución del vocalismo en munda (incipiente armonía vocálica y reducción del inventario) con lo que hallamos en ciertas lenguas africanas (cf. Williamson 2004), no resultaría en absoluto descabellado suponer que en PVant. no existía la armonía pero que con casi total certeza había un mayor número de vocales. De manera independiente, parece haber pruebas (relativamente numerosas etimologías) de que las sonantes fortes fueron anteriormente grupos consonánticos, opción que Michelena no descartaba en *FHV*, si bien le era imposible darla por demostrada en el estado de la reconstrucción y de la etimología de hace medio siglo.

las respectivas teorías de los sistemas radicales y de sus evoluciones. Por otra, hemos ido notando y recogiendo supuestas o reales irregularidades y características morfosintácticas y fonológicas que no concuerdan con el tipo canónico o cuyos análisis pueden, tal vez, iluminar etapas anteriores de la lengua, la cual como diacronistas, no teníamos por qué asumir que hubiera sido siempre tal y como la vemos o creemos ver en la actualidad, y esto, incluso, por lo que respecta a rasgos de cierta relevancia estructural.

Entre las razones aludidas, reunidas en trabajos anteriores y que nos hacen pensar en la existencia de una deriva que cambió profundamente la estructura de la lengua vasca de aislante a aglutinante, —las mismas que nos llevan a postular un PV antiguo muy diferente al estándar—, se hallan algunas como éstas:

- 1) la posición, a la derecha del N (como en las lenguas con V inicial), y escasez como clase diferenciada del adjetivo,
- 2) el orden antiguo —a la derecha del N— de las oraciones relativas y de los genitivos locativos,
- 3) la escasez en términos comparativos y el carácter claramente tardío de postposiciones y sufijos,
- 4) la evidencia de la enorme antigüedad del monosilabismo sobre el bisilabismo en la forma canónica de la raíz,
- 5) el hallazgo de algunos prefijos nominales (*\*la-*, *\*sa-*, *\*gi-*), antes desconocidos y de otros nuevos prefijos verbales que se añaden a los ya conocidos,
- 6) la constatación de la tendencia general moderna a la sílaba abierta frente a la antigua raíz con sílaba cerrada y el desarrollo (medieval) de las vocales nasales,
- 7) el desarrollo de las vocales iniciales y del inventario consonántico frente al mantenimiento general del vocálico, tendencias relacionadas con el incremento de la aglutinación, como, p.ej. en dravídico,
- 8) la armonía vocálica hacia la izquierda, coherente con la antigüedad de los prefijos y la inexistencia antigua de los sufijos,
- 9) el carácter tardío del grupo verbal conjugado,
- 10) el origen —en la posposición *-ga* de los casos locativos animados— de la marca de ergativo,
- 11) la abundancia de adjetivos procedentes de antiguos verbos estativos como en dravídico y al contrario que en tibeto-birmano,
- 12) la posible prominencia antigua de las categorías modales sobre el tiempo y el aspecto en el V, como sucede en tibeto-birmano y al contrario que en dravídico,
- 13) la progresiva especialización o “especificación” por adición de marcas de los casos locativos, anteriormente más generales, acercándose al dravídico y alejándose del tibeto-birmano,

- 14) una nueva reconstrucción (en sílaba final) del acento para el PV más antiguo.<sup>22</sup>

Naturalmente, no todos los argumentos presentados, las distintas evoluciones y características discutidas, tendrán el mismo valor; en realidad, la investigación sobre la mayor parte de ellos es casi inexistente, por lo que es probable que en un futuro reciban explicaciones diferentes o contrapuestas a las adelantadas por nosotros. Incluso aquellos argumentos que mantengan su valor habrán de recibir precisiones importantes y, con seguridad, no todos ellos serán independientes entre sí. Es lo esperable si no buscamos anécdotas sincrónicas o diacrónicas más o menos abundantes sino fundamentos de una teoría general sobre la evolución de la lengua vasca, teoría que difícilmente puede ser específica, en lo relevante, de aquélla. Es decir, es un acercamiento holístico y comparado el único que puede ayudarnos a plantear una nueva reconstrucción del PV premicheleniano y, además, a dar una explicación principiada y coherente de los diversos rasgos de ese sistema y de los cambios propuestos para llegar de lo que se reconstruya al estado lingüístico realmente presente en las épocas históricas y protohistóricas documentables o a reconstrucciones correspondientes a épocas prehistóricas más recientes y mejor conocidas (véase al final de §9).

#### 4. ETIMOLOGÍA FORMAL

Entre las virtualidades de la teoría, y no entre las menos importantes, se halla la posibilidad —no sólo la necesidad, dado el corpus de la lengua— de poder desarrollar un tipo de etimología que, a falta de otro término, hemos denominado “formal” y que puede ayudarnos a elaborar un esbozo de prehistoria del léxico patrimonial vasco. Fueron precisamente razones estructurales como la inexistencia de oclusivas finales o la imposibilidad de combinar dos TT en un bisílabo monomorfémico las que nos han llevado a emitir la hipótesis del monosilabismo PV antiguo y no, p.ej., creencias u opiniones sobre la semántica o el origen de las voces bisilábicas. Es después y no antes cuando se han postulado tales raíces monosilábicas y otros morfemas, siguiendo los procesos de triangulación habituales en la

---

<sup>22</sup> Cf. Post 2007b sobre el tani: “As is argued in more detail in Post (2007:§2), the historical morphological and phonological facts taken together suggest a scenario in which a previously isolating, analytical language with a basically morphosyllabic typological profile became increasingly synthetic and agglutinating in consort with a rhythmic shift to a trochaic pattern.

The suggestion made here, however, is that prosody has not only driven developments in some aspects of Tani grammatical organization (such as the shift from monosyllabic, simplex root to disyllabic, complex lexeme as the basic lexical unit, and the fusion of free sequences of simplex functional morphemes into complex functional words), it has in effect stayed one step ahead of grammatical organization by creating word-level units which are subject to reanalysis by speakers — and which the grammar may eventually adjust to accommodate” (25-26).

reconstrucción (cf., p.ej., Zorc 1991): *gi-zen* ‘gordo (de la carne)’ : *gi-bel* ‘hígado’ ‘parte trasera’: : *ze-zen* ‘toro’ : \**da-dar*, etc., de donde \**bel* (ya conocido de *orbel* ‘hojarasca’, *harbel* ‘pizarra’, etc.), \**gi*, \**zen*, etc.

La precisión formal ha de ser máxima en este proceso, por cuanto que no contamos prácticamente más que con ella;<sup>23</sup> i.e., la ayuda que la semántica nos pueda dar es escasa. Así, no tenemos ninguna seguridad de que *ere* en los nombres *And-ere* y *Ere-xonis* sea el mismo morfema a pesar de que los segmentos coincidan plenamente: de hecho, además de ser dudoso que en aquitano pudiera aceptarse un análisis con un morfema o alomorfo en *-d*, (¡y en PV dos raíces con V- y V-/-V!) a falta de más pruebas, nos hallamos ante dos homófonos situados en 1er y 2º elemento de la palabra, lo cual no da derecho a creer que la cuestión —la identificación de los morfemas— esté cerrada y, con ella, tampoco la del préstamo vasco → céltico (v. Lakarra 2005b, Lakarra & de Bernardo 2009). Así, frente a (Lakarra 2002b), pensamos ahora que si *gizon* ‘hombre’, *giharre* ‘carne magra’, *gizen* ‘gordo de la carne’, *gibel* ‘hígado’, y *sakon* ‘profundo’, *sabel* ‘vientre’, *samur* ‘tierno’, *samin* ‘dolor fuerte’, etc., han de analizarse como \**gi-zon*, \**gi-harr(e)*, \**gi-zen*, \**gi-bel*, y \**sa-kon*, \**sa-bel*, \**sa-bur*, \**sa-bin*, etc., tales formaciones no pueden ser tomadas como compuestos de dos raíces \**gi-*, \**sa-* combinadas con los respectivos segundos miembros de los vocablos. Esto es así pues \**sa-* y \**gi-* no cumplen con la estructura mínima necesaria para todo lexema radical PV (CVC), ni tienen la autonomía propia de otras raíces —no se documentan ni en solitario ni a la derecha de ninguna otra raíz conocida y, por tanto, sólo pueden ser prefijos o antiguas preposiciones, a no ser que aceptemos caídas de consonantes finales inexistentes con anterioridad a la Edad Media y, en todo caso, posteriores al aquitano, donde no hay nada de eso (cf. Michelena 1954 y Gorrochategui 1984). En otras palabras, nos las habemos con prefijos a no ser que hagamos trampa cambiando las definiciones de “raíz” y “prefijo” o, alternativamente, ignorando la cronología de los cambios fonéticos estándares.<sup>24</sup>

Hemos recogido en las tablas del Apéndice los principales datos resultantes de nuestra labor hacia 2004:<sup>25</sup> las raíces potenciales de cada modelo radical, las documentadas, las de etimología desconocida, las que han superado el criterio fonotáctico y geográfico, todo ello expresado en cifras absolutas y en porcentajes; por otro lado, las tomadas en préstamo y las onomatopeyas y sus porcentajes. Tanto entonces como ahora, la primera conclusión es que son muy escasas las raíces bisilábicas documentadas en

<sup>23</sup> Véanse los preliminares metodológicos de Gorrochategui 1984.

<sup>24</sup> Por ello, creo que deberíamos considerar mi conclusión actual y rechazar la anterior, por mucho que esto nos anime a —o quizás nos exija— cambiar en buena medida (cf. Lakarra 2005a y 2006a) la idea más tradicional y difundida sobre la tipología de la lengua vasca para las etapas más antiguas. Véase § anterior.

<sup>25</sup> Para avances posteriores veanse diversos trabajos en preparación, particularmente Lakarra en prep.-3 y aquí §§6-9.

general y de entre éstas las menos las que podemos clasificar como de etimología desconocida, i.e, las potenciales candidatas a PV. Creemos poder afirmar, además, que —si bien todos los modelos radicales bisilábicos están muy lejos de los monosilábicos— también entre aquellos cabe establecer diferencias y corresponden, sin duda, a épocas diferentes. Entre los modelos radicales más recientes se halla claramente cualquier combinación de una C con /f/ y /m/: magras documentaciones y porcentajes, préstamos abundantes, onomatopeyas y fósiles (que superen los filtros establecidos) escasos.<sup>26</sup> Encontramos algo similar para el modelo \*\*T<sub>1</sub>V<sub>1</sub>T<sub>2</sub>V<sub>2</sub> que estuvo en el origen de toda esta investigación (cf. Lakarra 1995); lo que antes meramente intuíamos —que este modelo no tiene posibilidad alguna de pertenecer al PV— podemos ahora afirmarlo a partir de indicios de muy diversa procedencia.

En 2004 dimos cuenta del análisis del complejo esquema radical CVCV, con múltiples divergencias internas en función de las variables utilizadas; en pocas palabras, tenemos más fonosimbolismos y fósiles y menos préstamos que en los modelos radicales estudiados anteriormente. Dentro de todos estos modelos quizás sean aquellos en los que está implicado el fonema /h/ los que menos préstamos —pero también menos fonosimbolismos— muestran. Los modelos recogidos bajo los esquemas CVCCVC, CVCCV, CVCVC (y sus correspondientes sin C-) son potencialmente 5, 10 ó 20 veces más extensos que los anteriores, pero las cifras de raíces realmente documentadas van muy parejas y dejo de lado por ahora la mayor o menor transparencia de las de cada cual. La existencia de modelos radicales así en vascuence moderno o contemporáneo<sup>27</sup> es indudable, pero ni mucho menos para el PV.

Antes de efectuar otros análisis, es conveniente efectuar una rápida comparación con los monosílabos: claramente estos se hallan lejos de los modelos radicales bisílabos sin coda por lo que se refiere a los porcentajes de raíces documentadas sobre las potenciales, pero, sobre todo, en la abundancia de aquellas que superan los filtros y criterios geográficos y fonotácticos, en la escasez de préstamos o en la existencia de los fonosimbolismos. Las conclusiones y vías de investigación que se nos antojan más relevantes y prometedoras son las siguientes: no hay pruebas de ningún cambio generalizado bisílabo > monosílabo antiguo; además de más pruebas sobre la restricción silábica ya conocida \*\*T y otras como \*\*T<sub>1</sub>V<sub>1</sub>T<sub>2</sub>V<sub>2</sub>, \*\*bisílabo (y \*\*polisílabo), \*\*V- y \*\*-V hechas notar en Lakarra 1995, donde aportamos claras pruebas de que las mismas pueden extenderse a \*\*CVCV. Igualmente, si CV y CVCC no pueden ser aceptados como raíces PV, la primera habrá de ser prefijo en las estructuras CV-CVC y que -C (ambas en caso de ser analizables, claro) habrá de ser tomado como sufijo en vocablos de estructura CVCC; en caso contrario, si CV- o -C fueran

---

<sup>26</sup> Léase lo mismo para los modelos sin coda alguna o con coda única en la sílaba de la izquierda.

<sup>27</sup> Cf. la cronología de Lakarra 1997a y ahora Lakarra 2007c.

inanalizables, podríamos estar ante indicios de préstamo (cf. Michelena 1963).

En la medida en que hemos ido precisando el modelo de reconstrucción existente en cada momento, los resultados etimológicos han acompañado al cambio de estructura propuesto. Así, el cambio de CVRST a CVRS y de ésta a CVC-C como forma canónica de la raíz, ha determinado nuevas etimologías para vocablos antes no tratados —por ser innecesario o infructuoso, aparentemente, cualquier intento en este sentido— como *hortz* ‘colmillo’ o *bortz* (> *bost*) ‘cinco’; en efecto, el resto (-tz) adquiere la condición de sufijo sólo cuando quedan explicadas o justificadas las bases *hor* y *\*bor* y hallamos nuevas formaciones en -tz (*bel-tz* ‘negro’ y *har-tz* ‘oso’, p.e.) en las cuales se aprecia alguna similitud semántica o funcional (como elementos formadores de adjetivos) con las anteriores. Igualmente, la imposibilidad de **\*\*CVCV** nos lleva a explicar como compuestos o derivados, pero en todo caso bimorfemáticos, no sólo términos evidentes como *bada* ‘sí/si es’ o *badu* ‘sí/si tiene’ sino también otros como *begi* ‘ojo’, además de ayudar en grado sumo a la clasificación como préstamo de *bide* ‘camino’ como hiciera antaño Michelena. Igualmente, el trisilabismo de *egia* ‘verdad’, *\*ardano* ‘vino’, *\*gaztana* ‘queso’, imposible de reducir hasta el momento a bases monosilábicas o meramente bisilábicas, hace que junto a las anteriores, la hipótesis del préstamo haya de ser tenida en cuenta.<sup>28</sup>

La investigación de la forma canónica de la raíz y de su evolución nos lleva a estudiar la posibilidad, sus requisitos y consecuencias, de plantear una modalidad de etimología basada precisamente en tal análisis. Son criterios como los siguientes —quizás junto a otros todavía por describir o desarrollar— algunos de los que la nueva teoría nos proporciona y que pueden permitir discernir de manera principada entre los tipos radicales y raíces antiguos y modernos:

- (a) la proporción entre raíces atestiguadas y el conjunto de las raíces posibles en cada modelo radical,
- (b) la proporción existente entre aquellas sin etimología conocida y el conjunto de las documentadas,
- (c) las condiciones fonotácticas cumplidas o incumplidas en cada fase de la historia y de la prehistoria de la lengua,
- (d) el criterio de la distribución dialectal histórica suficiente o escasa de las raíces sin etimología conocida,

---

<sup>28</sup> Incluso la indiferenciación categorial, señalada por Michelena (cf. 1964, 56) como rasgo antiguo o exclusivo de las capas del léxico vasco más arcaicas (cf. *hotz* ‘frío, el frío, friamente’, *ilum* ‘oscuro, oscuridad, oscuramente’, etc.), puede suponer un indicio relevante para discriminar entre estratos léxicos de diferente antigüedad. Con todo —o “además”, según se mire—, es pertinente señalar (cosa que no llegó a hacer Michelena), que esta indiferenciación categorial se da tanto en monosílabos como en bisílabos; además de los citados, añádase *bero* ‘calor, etc.’, *argi* ‘luz, etc.’ o *gose* ‘hambre, etc.’ por un lado y *huts* ‘vacío, etc.’ por otro. Tratándose de monosílabos y de los bisílabos aparentemente más antiguos, la duración del fenómeno pudo ser considerable. sobre *\*ardano* véase §7.

- (e) la abundancia o escasez de onomatopeyas entre las raíces documentadas en cada modelo radical y
- (f) la abundancia o escasez de préstamos en los mismos.

Siendo esto así, y si en un futuro más o menos próximo pudiéramos constatar que se trata de problemas relevantes y de vías de resolución fructíferas, habríamos conseguido motivar un paradigma más profundo que el estándar micheleniano del que disponíamos hasta ahora para la reconstrucción y etimología del PV y para la explicación de la evolución de la historia y, sobre todo, de la prehistoria de la lengua.

Hemos clasificado las voces correspondientes a cada modelo en cinco apartados: (1) préstamos no problemáticos, (2) variantes tardías, (3) compuestos y derivados, (4) onomatopeyas y fonosimbolismos y (5) de etimología desconocida. Posteriormente, se les han aplicado a los resultados iniciales del grupo (5) —que iba ya menguando a favor de los restantes, particularmente de (1), (3) y (4)— dos filtros cuales son los fonotácticos y el geográfico.<sup>29</sup> Parece claro que, siendo las voces incluidas en (5) las únicas que pueden aspirar a ser protovascas, dicho conjunto es muy limitado y menguante fuera de CVC. Aquí, en cambio, la proporción entre “fósiles” —i.e., miembros de (5) que superan los filtros fonotácticos y geográficos— respecto al total de raíces documentadas o, incluso, de las potenciales de dicho modelo es muy elevado comparativamente (15% frente a cifras en bisílabos por debajo del 3% y, sobre todo del 1%, con múltiples ceros absolutos). Pero lo más relevante, se me antoja, es que las cifras absolutas y relativas de CVC sólo pueden crecer (quizás hasta un 25-30%) con el concurso de las nuevas raíces halladas en la labor reconstructiva y no utilizadas aquí, mientras que en los bisílabos, por las mismas razones, no es previsible que dejen de seguir menguando sus ya escasas cifras.<sup>30</sup>

---

<sup>29</sup> Véanse las cifras y tablas de Lakarra 2004b, extractadas en el Apéndice. Tales tablas habrán de ser revisadas, muy “a la baja” por lo que toca a cualquier modelo bisilábico, en futuros trabajos como se ve simplemente por lo mencionado en varias ocasiones en torno a la detección de nuevos préstamos. Naturalmente, también los más de 500 nuevos términos explicados —algunas veces por procesos fonológicos y morfológicos desconocidos o poco estudiados— reducen considerablemente el % de varios modelos bisilábicos y las ya escasas posibilidades de su candidatura a pertenecer al léxico PV.

<sup>30</sup> En trabajos recientes o en preparación (v. Lakarra 2007b o Lakarra en prep.-1) hemos mostrado que nuevos análisis de antiguas raíces que en un primer momento carecían de etimología y conseguían superar los filtros habituales, cambian significativamente los resultados finales (número y % de “fósiles”) de bastantes modelos radicales, haciendo aún más despreciables los escasos restos de los candidatos bisilábicos a pertenecer a etapas anteriores a la fragmentación dialectal posterior al Vasco Común (v. § 9).

Así, si descontamos de los fósiles *barre* ‘risa, reir’ en TVRV, *herri* ‘pueblo’ en hVRV, *gazte* ‘joven’ en gVCCV, *argi* ‘luz, luminoso’ en aCCV, *belar* ‘hierba’, *bihar* ‘mañana’, *bider* ‘vez’ y *bizar* ‘barba’ en bVCVC y *labur* ‘corto’, *lagun* ‘compañero’, *lakar* ‘áspero’, *legatz* ‘merluza’, *lehen* ‘antes, primero’, y *lizar* ‘fresno’ en lVCVC (v. Lakarra en prep.-1) tendríamos que esos modelos radicales pasan de tener, respectivamente, un 0,66% al 0,49%, de 0,47% al 0,31%, del 0,18% al 0,09%, del 1,85% al 1,38%, del 0,32% al 0,14% y del 0,61%

## 5. CAMBIOS EN LA FORMA CANÓNICA DE OTROS IDIOMAS

Es sabido (cf. Lakarra 2006a) que las lenguas y familias de lenguas aglutinantes más conocidas —urálica, túrcica, mongola, japonesa, etc.) tienen una estructura radical casi exclusivamente bisilábica. Así, p.ej., en urálico (v. Bákro-Nagy 1992) vemos que casi el 100% de los morfemas reconstruidos son bisílabos CV(C)CV, siendo además las escasas excepciones no lexemas sino partículas gramaticales, las cuales muestran como en otras familias (cf. Gamkrelidze-Ivanov 1995 para el IE) una estructura fonotáctica totalmente diferente, si no una imagen especular, de las raíces de sustantivos y verbos.

Aunque se haya defendido lo contrario, el caso del vascuence queda muy lejos del urálico y del resto de lenguas aglutinantes por lo que toca al bisilabismo, tanto más según nos vamos retro trayendo a las etapas reconstruibles más antiguas de la lengua. Por una parte, en vascuence histórico tenemos docenas de términos como *gar* ‘llama’, *gor* ‘sordo’, *hor* ‘can’, *hots* ‘ruido’, *hotz* ‘frío’, *sal* ‘vender’, *sar* ‘entrar’, *sor* ‘nacer’, *zur* ‘madera’, etc.<sup>31</sup> atestiguados en el conjunto del territorio de habla vasca; por otra, aunque sean mayoritarios los bisílabos, es fácil encontrar entre ellos múltiples préstamos y compuestos y derivados como puede verse en §9 y en las tablas del Apéndice. Finalmente, buena parte de los bisílabos (ibidem) están muy lejos de poder superar cualquier filtro geográfico, por lo que su antigüedad es más que dudosa.

Si nos retrotraemos al PV reconstruido, incluso a sus etapas más recientes (i.e., al paradigma estándar micheleniano) es evidente que el monosilabismo radical hubo de ser no sólo mayoritario sino prácticamente exclusivo por lo visto reiteradamente en trabajos anteriores. La seguridad con que podemos afirmar esto es tanto mayor cuanto que somos conscientes de que habríamos de añadir a los “radicales libres”, documentados históricamente de forma autónoma, aquellos otros —no menos numerosos, incluso limitándonos a los más seguros— provenientes de la reconstrucción: i.e., *\*bel* ‘negro’, *\*dats* (cf. *adats* ‘melena’), *\*den* ‘terminar’ (cf. *atseden* ‘morir’, *eteten* ‘rompiendo’, *edeki* ‘quitar’), *\*der* (cf. *eder* ‘hermoso’), *\*din* ‘(de)venir’, *\*dol* (cf. *odol* ‘sangre’), *\*don* ‘poner, sacar’ (cf. *idoki* ‘sacar, quitar’), *\*dul* ‘salir’ (cf. *ilki, jaulki*), *\*dun* ‘haber’, *\*han* ‘animal’ (cf. *ahuntz*

---

al 0,31%. También otras estructuras bisilábicas aparentemente importantes muegan significativamente: nVCV de 0,18% a 0,06% (resta de *nafar* ‘navarro’, *nabar* ‘polícromo’, y *negar* ‘lágrima, lloro’), hVCVC de 0,37% a 0,18% (resta de *haran* ‘valle’, *hegatz* ‘ala’, *hezur* ‘hueso’, *hoben* ‘falta’), aCVC de 3,52% a 2,11% (supresión de *agir* ‘manifiesto’, *agor* ‘estéril, improductivo’, *akats* ‘fallo’, *alor* ‘sembrado’, *arrotz* ‘extranjero’) u oCVC de 1,17% a 0,93%.

<sup>31</sup> Por no recurrir a estructuras CDC como *haur* ‘niño’, *laur* ‘cuatro’ o *gain* ‘encima’; ahora bien, parecen ser más abundantes los problemas (cf. *sei* ‘seis’ o *gau* ‘noche’ sin coda) en estos que en aquellos. Creo por ello que todos los monosílabos con diptongo han de ser explicados, bien como préstamos, bien como derivados, a partir de bisílabos anteriores que han sufrido contracción; v. Lakarra en prep.-1. No parece necesario, por tanto, postular diptongo alguno para el PV (al menos para el PV antiguo).

‘cabra’ < ‘animal (de cuernos como) la hiedra’), \**nol* (cf. *ohol* ‘tabla’), \**zen* (cf. *zezen* ‘toro’), etc.

Curiosamente, no es en las lenguas y familias aglutinantes que disponen de historia mejor conocida donde encontramos más estudios sobre un tipo de cambio en la forma canónica [monosilábico] > [bisilábico] como el que parece haber experimentado la lengua vasca en algún momento —a lo largo de varios milenios según Donegan y Stampe—, en todo caso, en una *longue durée*” todavía por especificar. Si bien sería del mayor interés conocer la prehistoria del bisilabismo urálico, japonés, mongol o túrcico, no es menos cierto que diferentes estudios sobre el tema en otras lenguas y familias pueden servirnos de fuente de inspiración nada despreciable; me refiero a lo que podemos saber sobre el chino y el *tani*.<sup>32</sup>

*Chino*. Probablemente sea ésta la lengua que más bibliografía haya generado sobre el tema: se ha discutido la existencia o no en época antigua de bisílabos junto a los siempre mayoritarios monosílabos y las posibles razones para la creación de bisílabos que van desde las sociolingüísticas (cuando no metafísicas) de su necesidad para expresar nuevos conceptos,<sup>33</sup> hasta todo tipo de argumentos fonéticos, morfológicos, sintácticos y semánticos.

Dai 1990 defiende que diversos tests sintácticos y morfológicos demuestran que las formas ligadas bisilábicas de algunos verbos en chino moderno constituyen morfologizaciones de verbos independientes en chino antiguo. Tal morfologización no procedería sólo de cambios fonéticos —pérdida de oclusivas finales y reducción de tonos léxicos contrastivos— sino también de pérdidas semánticas y generalización de la nueva forma canónica. El proceso de morfologización partiría de una frase bisilábica coordinada que se desarrollaría como compuesto —con ambos morfemas aún libres— y tendería a la afijación de uno de sus miembros o de ambos.

Feng 1998 ha extendido el análisis del origen del bisilabismo al conjunto del léxico. Según él, es durante la dinastía Han cuando se

---

<sup>32</sup> No hay necesidad de repetir aquí lo que puede verse sobre el austroasiático y el *motlav* en §3.

En realidad, el cambio monosílabo > bisílabo no es desconocido sino todo lo contrario en las lenguas de Australia. Tenemos “Most Australian languages have no monosyllabic words at all (outside interjections)” en palabras de Dixon (2002, 553), quien señala un esquema general CVC(CV)(C). El mismo autor añade que ciertas lenguas tienen raíces ligadas monosílabas —particularmente verbales— pero que cualquier palabra flexionada tiene al menos dos sílabas. Al menos desde comienzos de los 70, se supone que el bisilabismo generalizado actual proviene de una situación anterior donde el monosilabismo sería mucho más abundante y que iría evolucionando por diversas razones entre las que Dixon (1980, 173-4) cita la preferencia del ritmo “sílabo acentuada — sílabo no acentuada” y el reanálisis de “raíz monosilábica + inflexión monosilábica > raíz bisilábica”, con supresión del linde morfémico en la mitad de la forma disilábica.

<sup>33</sup> Necesidad cubierta tanto por préstamos como por nuevos compuestos y derivados; v. Masini 1993.

documentan masivamente las palabras compuestas bisilábicas en chino a causa de la transformación de la estructura del pie prosódico en bisílabo, dado que el monosilabismo anterior no pudo soportar un pie bimoraico tras la caída de las consonantes finales y la simplificación de los grupos consonánticos. A diferencia de la explicación teleológica tradicional, que ve la creación de compuestos como un modo de esquivar los homófonos derivados de los cambios fonéticos, Feng enriquece el papel de la fonología prosódica a la hora de explicar el fenómeno. Por otro lado, arguye que las palabras compuestas no son estructuradas únicamente desde el punto de vista sintáctico —anteriormente solían conocerse como “palabras sintácticas”— distinguiendo para el chino clásico entre palabras, i.e., compuestos basados en la lexicalización o en reglas de cambio de categoría, y frases, i.e., palabras prosódicas generadas por el uso combinado frecuente:

Note that compound words in Classical Chinese are syntactic words because they historically originated from disyllabic phrases. Compound words are prosodic words also, because they are lexicalized PrWds. This entails that not every phrase can develop into a compound, but only those which meet the prosodic requirements. Neither can any foot be identified as being a compound, but only those that represent an independent syntactic unit, i.e., a phrase. By prosody, only phrases that fit the description of being one foot are eligible to become compounds. By syntax, only feet that represent independent phrases are qualified to be compounds (Feng 1998: 238-9).

Y, en conclusión:

Given all the analyses above, the origin of compounding can now be described as follows: the phonological change of Old Chinese resulted in a disyllabic foot, the disyllabic foot, in turn, resulted in disyllabic PrWds, disyllabic PrWds are formed by two-syllable phrases given the monosyllabic property of the language, and the two-syllable phrases are idiomatized in usage, becoming Idiomatized PrWds. When Idiomatized PrWds are lexicalized, they become an  $X^0$  level category item, i.e., a compound word in the lexicon (...) (Feng 1998: 238-9).<sup>34</sup>

*Tani*. Mark Post 2006, 2007b ha mostrado que en tani (un grupo tibeto-birmaniano situado al norte de Bangladesh, entre el Tibet y Birmania) ha experimentado un proceso de disilabificación en el que las formas monosilábicas del proto-tani se han lexicalizado como compuestos.<sup>35</sup>

---

<sup>34</sup> El análisis de Feng ha sido aplicado al caso vasco por Martínez Areta 2003 [2008]; como se muestra en §9, creo, sin embargo, que son necesarias muchas precisiones sobre el motor del cambio, su extensión y, particularmente, su cronología, la cual no me resulta evidente que sea pre-PV, ni siquiera pre-aquitana como concluye Martínez Areta.

<sup>35</sup> Véase la cita de §4. Esto, incidentalmente, ha llevado a una división moderna entre adjetivos y nombres, por un lado, y verbos por otro, frente a la tipológicamente habitual Adj-V / N de las lenguas del área. Reproducimos a continuación dos párrafos de Post 2007b correspondientes, respectivamente, a su abstract y al final de sus conclusiones:

The resulting “mismatch” between “phonological words” and “grammatical words” in Galo is argued to be theoretically non-trivial, in that its existence is capable of explaining a variety of otherwise seemingly disparate facts in the synchronic and diachronic organization of Galo grammar (...) Although there might be said to exist a very generalized functional

Mientras que el prototani, reconstruido anteriormente por Sun era una lengua “morfosilábica”<sup>36</sup> como tantas otras del Sudeste Asiático Continental o el chino (cf. Sagart 1999), las lenguas tani son sintéticas y aglutinantes en varios grados. La deriva tipológica de su morfología ha reestructurado el antiguo lexicon —básicamente monosilábico— llevándolo a un disilabismo generalizado, derivado de las anteriores “colocaciones” (raíces con prefijos y compuestos),

Post señala que mientras en tani moderno casi todas las raíces (monosilábicas) son ligadas y necesitan formar compuestos o añadir afijos para funcionar como palabras gramaticales, no era así en prototani, donde las raíces eran potencialmente libres.<sup>37</sup> Como consecuencia, en tani moderno los polisílabos son lexemas o morfemas gramaticales casi siempre fácilmente analizables como antiguas colocaciones de morfemas monosilábicos y, en general, la morfología es fácilmente reconstruible y escasamente gramaticalizada a partir de raíces léxicas; por fin, su estructura monosilábica anterior viene apoyada también por el tono léxico y una estructura segmental y silábica simples.

Por lo que toca a los aspectos fonológicos de la transformación monosilábico → bisilábico, es interesante ver que Post señala la conversión de la palabra en unidad prosódica básica en lugar de la sílaba como anteriormente y, sobre todo,

(...) while early sound changes introduced opacity between languages, they preserved the phonological integrity of monosyllabic roots within a given language. However, later changes began to introduce opacity among instantiations of the same root within a given

---

pressure towards “unification” of “phonological words” and “grammatical words”, such a pressure would not be expressible as a formal constraint on language grammar (1)

(...) in Galo at least, the driving force behind the diachronic innovation of new word structures may be primarily prosodic, and that the grammar may subsequently restructure itself to accommodate the now-prevailing whord shapes (27).

<sup>36</sup> Como explica Post 2006, 43, el término fue creado por Tymothy Light dentro de la teoría de la tonogénesis, aunque no parece haber tenido gran éxito. Habitual entre las lenguas del Sudeste Asiático Continental, una lengua morfosilábica prototípica exhibe una unidad fundamental sílaba = morfema = palabra. En ellas hallamos muy poco material morfológico y muchas veces las funciones gramaticales se codifican por variantes de las palabras léxicas. Así mismo, su gramática es fuertemente analítica y aislante, con evidentes muestras de serialización verbal. Frecuentemente se hallan tonos léxicos y la estructura segmental y silábica es simple.

<sup>37</sup> Post 2006, 43 argumenta que si bien en los últimos 40-50 años ha habido un enorme incremento de la influencia de lenguas indoeuropeas como el hindi, el asamese, el nepalí y el inglés y que “numerous lexical items, grammatical morphemes and grammatical constructions have been borrowed”, hay pocos indicios de influencias anteriores y que, en todo caso, la deriva morfológica observada en tani difícilmente puede achacarse al contacto lingüístico.

Algo similar puede decirse sobre el caso vasco; sería interesante saber, con todo, el alcance de esos “numerous lexical items” y la proporción de bisílabos en ellos para una posible comparación con lo que veremos en los siguientes apartados y, sobre todo, en §9.

language, which reflected an increase in the basic status of complex polysyllables over simplex monosyllables in the Tani lexicon (Post 2006: 45).<sup>38</sup>

## 6. INDICIOS Y EVIDENCIAS DEL CAMBIO DE FORMA CANÓNICA<sup>39</sup>

a) Uno de los principales factores conducentes al cambio de forma canónica reside en las metátesis consonánticas:

- .g-n > n-g: \*(la)grima > \*girma > \*girma > \*nirga > \*nigar ‘lloro’, ‘lágrima’
- .d-p > p-d: turpe > \*durpe > \*burde > urde ‘cerdo’, ‘sucio’
- .r-b > b-r: herba > erbar > berar > belar/bedar ‘hierba’
- .d-b > b-d: timor > ... \*dirbur > \*birdur > bildur ‘miedo’,<sup>40</sup>
- .d-z > z-d: \*edazun > \*ezadun > eza.un > ezagun ‘conocido’
- .d-g > g-d: timbal > \*dilbal > \*dibal > \*di.al > \*dikal > gidal ‘pierna’
- .p-k > k-p: bokadu > \*bokau > \*poku > kopu ‘bocado’

b) Aunque aparentemente en menor número, hallamos también metátesis vocálicas:

- .a-au > au-a: \*ha(t)s-la(b)ur > \*haslaur > \*hasnaur > hausnar ‘rumiar’
- .e-au > eu-a: \*e-radun-egi > \*erauhegi > \*herauegi > heuraegi > heuragi/heuregi ‘abundante’
- .e-oi > o-ei: \*e-don-i > \*edohi > hedoi > hodei ‘nube’, ‘tormenta’

c)  $*h_3/*h_2 > h_1$

Entre las metátesis merece especial atención (cf. Lakarra 2009b) el adelantamiento a posición inicial de /h/ intervocálica surgida en onset de segunda o tercera sílaba:

- hagin < \*ha.in < \*ahiun < \*ahinu < caninu ‘diente’
- harea < \*areha < arena
- herio < \*eriho < \*e-lin-o ‘muerte’
- heuskara < \*ehuskara < \*e-nus-kara ‘lengua vasca’

<sup>38</sup> Post 2006, 58 señala que hay muchas posibilidades de que los cerca de 100 sufijos verbales tanis deriven de fenómenos de serialización verbal. No me resisto a traer aquí la última frase de su trabajo: “In some modern Tani languages, adjectives may be moving closer to verbal status again, particularly in those languages which lack a copula which is not homophonous with the pan-Tani imperfective suffix *-duun* (< PTS *\*dun* ‘sit’). But that’s another story”. Téngase en cuenta que son muchos, y jerárquicamente importantes en la escala de Dixon 1977, los adjetivos vascos que terminan con la *-i* propia de los participios; además, en Lakarra 2008b se relaciona el *da-* imperfectivo de Trask 1977 con la raíz *\*dar* de *jarri* ‘sit’, gramaticalización ampliamente extendida en las lenguas del mundo.

<sup>39</sup> Aquí no puedo sino remitirme a futuros estudios (sobre todo a Lakarra en prep.1 y en prep.-5) donde deberán buscarse tratamientos detallados de las distintas derivaciones (p.ej., pasos intermedios que aquí no explícito por falta de espacio); sin embargo, no es posible procrastinar el enorme interés que para la historia del léxico en general y de la evolución de la forma canónica en particular podría tener la elaboración de una cronología relativa de los cambios implicados, rigurosa y detallada. Parece, p.ej., que *-rm-* > *-rn-* es anterior tanto a “/r/ a la derecha” como a *-n-* > *-h-* (véanse ejemplos en el texto); también *w-* > *ø-* ha de ser muy antiguo —quizás una de las primeras transformaciones sufridas por voces latinas al pasar a nuestro idioma—, anterior probablemente al cambio de *w-* > *b-* en protoromance.

<sup>40</sup> Cf. *debekatu* < *bedekatu* < *impedicatu* ‘prohibido’.

*hezur* < \**ehazur* < \**enazur* ‘hueso’  
*hibai* < \**ibahi* < \**u(r)bani* ‘río’  
*hidoi* < \**idohi* < \**u(r)doni* ‘estanque’  
*hodei* < \**edohi* < \**e-don-i* ‘tormenta’, ‘nube’  
*hoge* < \**ogehi* < \*(b)o(r)gehi < \**bor-gen-i* ‘veinte’

d) *r* a la derecha.

Mientras que el desplazamiento de *-n-* intervocálica a la derecha —previa nasalización de vocales y consonantización posterior— está bien descrito en Michelena 1977a,<sup>41</sup> no ocurre otro tanto con múltiples *erres* presentes en sílabas iniciales y mediales (sean CCV, sean CVC) que desembocan en la final:

*praesepe* > \**barzepe* > ... > *bazter* ‘rincón’  
*craindre* > *aiher* ‘propenso a, apesadumbrado’  
*arlo* > *alhor* ‘terreno sembrado’  
*ristra* > *istar* ‘corva’, ‘pierna’  
*tremere* > \**derme(tu)* > \**lernetu* > *lehertu* ‘agitar’, ‘explotar’  
\**nirga* > *nigar* ‘lágrima’, ‘lloro’  
\**hatz-zarpa* > *atzapar* ‘garra’

e) De manera similar con *-n* en coda:

*deffendere* > \**bende(tu)* > \**beden(du)* > ... > *geben(du)* ‘prohibido’  
*hoben* < *offende(re)* a través de \**obeden* > \**obehe*n, etc. ‘falta’

f) Contra las consonantes homorgánicas.

Aunque Michelena no insistiera en el interés del fenómeno —ni en la interconexión entre el comportamiento entre sonantes y sibilantes— en la *Fonética histórica* hay suficientes datos para ver que en vascuence histórico no ha sido aceptable la combinación de consonantes homorgánicas dentro de la misma raíz. De aquí se sigue que tienen el mayor interés los fenómenos (de asimilación y disimilación) encaminados a evitar tal combinación pues nos informan indirectamente sobre la extensión de la raíz:

*sazón* > *sasoi*  
*solaz* > *solas*  
*zin-hets-i* > *sinhetsi* ‘creer’  
*so-egotzi* > *sorotsi* ‘ayudar’  
\**buru-bar* > *burar* > *bular* ‘pecho’  
\**erbar* > *berar* > *belar/bedar* ‘hierba’  
\**erardun* > *inardun* ‘dedicarse a, trabajar en’  
\**irarrosi* > *inarrosi* ‘agitar’, ‘gritar’

g) Asimilación vocálica:

(*h*)*azkar* < *hazi-kor* ‘grande, rápido’  
*zahar* < \**zan-or(?)* ‘viejo’

<sup>41</sup> Nuevos casos de desplazamiento de *ene* intervocálica a posición final son los constituidos por *finu* > *bigun* y *caninu* > *hagin*.

*zuhur* < \*zun-or(?) ‘sabio’

h) inserción de /g/:

*hagin* ‘diente’ < \*ha.in < caninu  
*hogen* ‘falta’ < \*ho.en < offende(re)  
*nagusi* ‘mayor, jefe’ < na.us-i < \*da-dun-tz-i  
*geben*(du) ‘prohibido’ < \*be.en(du) < (de)fende(re)  
*bigun* ‘suave, débil’ < \*bi.un < finu  
*igerri* ‘adivinar’ < \*eu.erri < \*edun-berri  
*jagon* ‘vigilar, cuidar’ < \*ja.on < \*e-da-don  
*ugatz* ‘ubre, teta’ < \*u.atz < \*edo(n)-hatz

i) contracción de trisílabos y polisílabos:

<i>aulki</i> ‘silla’ < * <i>abedulki</i>	<i>hogen</i> ‘falta’ < * <i>ho.en</i> < offende(re)
<i>jabe</i> ‘dueño’ < * <i>e-da-dun-e</i>	<i>bazter</i> ‘rincón’ < <i>praesepe</i>
<i>zemai</i> ‘amenaza’ < <i>menaza</i>	<i>hezur</i> ‘hueso’ < * <i>enazur</i>
<i>gurdi</i> ‘carro’ < * <i>egur-gi</i>	<i>bular</i> ‘falta’
<i>jagon</i> ‘vigilar, cuidar’ < * <i>e-da-don</i>	
<i>hagin</i> ‘diente’ < * <i>ha.in</i> < <i>caninu</i>	

j) neutralización y caída de vocales en final de primer elemento:

<i>ama</i> > ‘madre’ <i>ama-</i>	<i>ogi</i> ‘pan’ > <i>ot-</i>
<i>aza</i> ‘berza’ > <i>aza-</i>	<i>idi</i> ‘buey’ > <i>it-</i>
<i>larre</i> ‘sembrado’ > <i>larra-</i>	<i>buru</i> ‘cabeza’ > <i>bur-</i>
<i>lore</i> ‘flor’ > <i>lora-</i>	<i>errege</i> ‘rey’ > <i>erret-</i>
<i>baso</i> ‘bosque’ > <i>basa-</i>	<i>iturri</i> ‘fuente’ > <i>itur-</i>
<i>asto</i> ‘burro’ > <i>asta-</i>	<i>eliza</i> ‘iglesia’ > <i>eliza</i>

Naturalmente, la mayor parte de los fenómenos anteriores no son exclusivos de los bisílabos aunque, sin duda, son más numerosos en estos que en los polisílabos; alternatively, podemos hallar otros fenómenos que alteran el tamaño, la forma o la estructura de las voces, como la haplogía de *voluntate* > .... \**ololde* > *olde* ‘pensamiento’, cuyo producto son nuevos bisílabos radicales simples.

## 7. SOBRE ORÍGENES DE LAS EXTENSIONES VERBALES

Como cualquiera sabe, Lafon<sup>42</sup> describió una raíz verbal absolutamente polimorfa: mono y bisilábica, de una sólo vocal, con CV o VC en monosílabos, con y sin C iniciales y finales en los bisílabos, etc:

Pour avoir du verbe basque une connaissance satisfaisante, il faudra joindre à la connaissance de son système celle des rapports de forme, de sens ou d’emploi que certaines racines ont pu avoir entre elles, c’est-à-dire la connaissance de l’histoire et de la préhistoire des racines (Lafon 1943, I, 433).

<sup>42</sup> El mismo autor tiene un pequeño artículo posterior (Lafon 1950) específico sobre la raíz, pero no va más adelante y se limita a resumir aspectos de Uhlenbeck 1947 [1942] que antes desconocía.

B. *Racines simples*. Les racines simples du basque se composent d'une ou de deux syllabes. Les types suivants se rencontrent :

1° RACINES MONOSYLLABIQUES :

- a) une voyelle : *a-* “être”;
- b) consonne et voyelle : *go-* “rester”, *gi-* “être fait”;
- c) voyelle et consonne : *bil-* “marcher”, *kus-* “être vu”.

2° RACINES DISSYLLABIQUES : *oa-* “aller”; *augi-* “venir”, *aki-* “être su”, *zagu-* “être connu” (Lafon 1943, I, 421).<sup>43</sup>

Si bien como con el resto de raíces, su única esperanza de investigar el origen de las mismas era la comparación con las lenguas caucásicas, la reconstrucción interna basada en la forma canónica CVC permite no estar pendientes de tal vía explicativa, poco o nada productiva hasta el momento: en concreto, es posible “regularizar” casi todas las raíces verbales, derivándolas de antiguas CVC con uno (-*ra-* / -*da-*) o dos (-*da-ra-*) prefijos, además de la \**e-* inicial. Así, *utzi* ‘dejar’ = \**e-dutz-i* (cf. *luz-e* ‘largo’), *jasan* ‘soportar’ = \**e-da-san*, *jakin* ‘saber’ = \**e-da-kin*, *jauzi* ‘saltar’ = \**e-da-dutz-i*, *irauli* ‘volcar’ = \**e-ra-dul-i*, etc.; cf. Lakarra 2006b.

Hemos podido mostrar que, además del conocido prefijo causativo -*ra-* caduco en época histórica, existió al menos otro, \**-da-* presente en unas tres docenas de pares como *ezarri* ‘imponer’ : *jazarri* ‘atacar’, *esan* ‘decir’ : *jasan* ‘soportar’, \**edon* ‘sacar’ : *jagon* ‘vigilar’, etc., prefijo que no puede ser otro que el *da-* de las formas de presente que Trask 1977 explicó como antiguo morfema de aspecto inacabado.<sup>44</sup> Teníamos (v. Lakarra 2007a), incluso en etapas no muy alejadas de los primeros textos, combinaciones de ambos prefijos (-*da* + -*ra*) en verbos tan irregulares como *jardun* ‘dedicarse a, trabajar en’, *jarraiki* ‘seguir’, *jarrugi* ‘colocarse junto a’, *inarrosi* ‘agitar, hacer ruido’, *urgatzi* ‘ayudar’ y, aproximadamente, otra docena similar en su formación a las conocidas combinaciones RAIZ – CAUS – APLIC bantúes (cf. Good 2005):

<i>edan</i>	:	<i>ja(a)n</i>	:	<i>eradan</i>	:	* <i>ardan</i>
* <i>edin (jin)</i>	:	* <i>jadin</i>	:	<i>eriden</i>	:	<i>jarrain</i>
* <i>editsi</i>	:	<i>jaditsi</i>	:	<i>eritsi</i>	:	<i>jardetsi/ihardetsi/inardetsi</i>
* <i>edon</i>	:	<i>edadon</i>	:	<i>eradon (erro)</i>	:	* <i>arran-</i> ( <i>arrano, arrain</i> )
* <i>edun</i>	:	<i>jaun</i>	:	* <i>eradun</i>	:	<i>jardun/ihardun/inardun</i>
* <i>edul-ki (ilki)</i>	:	<i>jaulki</i>	:	<i>irauli</i>	:	<i>arraul(tza/e)</i>
		( <i>jalgi</i> )				
<i>egotzi</i>	:	_____	:	<i>eragotzi</i>	:	<i>urgatzi</i>
* <i>enon</i>	:	<i>joan</i>	:	<i>erhan/erho</i>	:	<i>eroan/eraman</i>
<i>inotsi</i>	:	_____	:	<i>erauntsi</i>	:	<i>inarrotsi</i>

<sup>43</sup> En Lafon 1951-1952 es explícito el reconocimiento de la inexistencia de un tipo vasco-caucásico comparable, p.ej. al camito-semítico de Cohen y otros.

<sup>44</sup> Al doble valor (perfectamente conocido en los estudios de gramaticalización) de \**da* —aspectual en el verbo, espacial en el SN— y sus consecuencias para la reconstrucción me he referido en Lakarra 2006b y 2008a.

<i>ekin</i>	:	<i>jakin</i>	:	<i>irakin</i>	:	<i>jarrekin</i>
<i>ifini (ebini)</i>	:	<i>jafini</i>	:	<i>irabio</i>	:	<i>arbin</i>

Por citar algún análisis etimológico relacionado con nuestra concepción de los orígenes del estema verbal y de las extensiones derivadas (*\*e-raCVC* / *\*e-daCVC* / *\*e-daraCVC*) aludiremos brevemente a *\*ardano* ‘vino’ y *entzun* ‘oir’ y *ezagutu* ‘conocer’. El primero, se trata de una reconstrucción anterior a Michelena y que sigue gozando de consenso general; cuando en su momento traté de ir más adelante, sugerí ver en él un préstamo a partir del “hispanico antiguo” *arándano* (cf. n. 28 y texto correspondiente). Aunque “sonara bien”, es claro que la semántica no favorecía en absoluto una derivación similar; ahora, al colocar *\*ardano* junto a *arraNo* ‘águila’ y *arraultza* ‘huevo’ es preferible ir en otra dirección, reconociendo en esas voces antiguas CVC verbales (cf. en nuestro caso el *\*dan* de *edan* ‘beber’ y el *\*dul* de *ilki* ‘salir’), acompañados de una serie de prefijos desusada históricamente y desconocida hasta ahora en los análisis verbales vascos.<sup>45</sup>

De *entzun* y *ezagun* (verbos sintéticos ambos) es conocido que Schuchardt defendió que eran dos préstamos románicos derivados de *intensum* y *sabut*, respectivamente. La idea no obtuvo, ni mucho menos, el consenso general dado que no se conocen otros casos de V sintéticos tomados en préstamo, así como por la innegable anterioridad de *ezagun* sobre *ezagutu*. Que yo sepa, nadie aludió a cuestiones formales, pero es precisamente su no acomodación a ninguna estructura canónica conocida (ni CVC ni CV-CVC) la que causa mayores problemas. Ahora bien, si pensamos que la *-n* de *-nzun* es un caso de repercusión como en *zentzun* (< *sensu*) —y compárese *oihartzun* ‘eco’, lit. ‘ruido en el bosque’— y, además, suponemos que en *-zagun* tenemos no una raíz pura sino un caso de prefijo + raíz (*\*da-zun*) con metátesis e inserción en intervocálica de *-g-* antihiática (*\*e-da-zun* > *\*e-za-dun* > *ezaun* > *\*ezagun*; v. §6h), obtenemos, no sólo, que ninguno de ellos es un inverosímil préstamo sino que ambos forman parte de la misma familia constituida sobre la raíz *\*-zun* con dos prefijos *ø-* / *\*da-* / *ra-*: *entzun*, *ezagun*, *erantzun* ‘responder’, como *\*edin* ‘(de)venir’ / *jaiio* ‘nacer’ / *ediren* ‘hallar’ (v. Lakarra 2009a).

A partir de *herio* ‘muerte (activa)’ u *hodei/hedoi* ‘nube’ (< *\*elino*, *\*edoni*, respectivamente) podemos pensar (v. Lakarra 2009b) que formas como (*h*)*erro* ‘raíz, ubre’ corresponden a *\*e-ra-don*, con *\*h<sub>2</sub>* > *h<sub>1</sub>* frente al *\*h<sub>3</sub>* > *h<sub>1</sub>* de los primeros; por otro lado, *-rr-* proviene como en *erran* ‘decir’ de un grupo, bien que de *-nr-* frente al *-sr-* del otro (< *\*e-ra-san*). Así pues,

<sup>45</sup> Es elocuente, más que significativo, que Michelena 1977b analice como un antiguo y opaco nominal (y por tanto el conjunto como compuesto) el *jar-* de una serie de verbos como *jardun* ‘dedicarse a’, *jardetsi* ‘discutir’, *jarein* ‘derramar’, etc., mientras que el mismo año, en la segunda edición de la *Fonética*, se explica que la *j-* de ciertos nominales como *jaun* ‘señor’ o *jabe* ‘dueño’ se debe a que provienen de un antiguo verbo en *\*e-*. Es claro, por tanto, que el “monstruo” *jar-* debía explicarse como antigua raíz verbal o, mejor, como prefijo o amalgama prefijal de las raíces que en cada caso van a la derecha.

todas o parte de las *-rr* en  $C_1$  pueden provenir no de *R* fortis inicial de raíz sino del grupo mencionado, mientras que las *-R* fortis de  $C_2$ , a su vez, no son sino el resultado de la fortición de la *-r* final anterior. Dejo los detalles para otra ocasión pero sospecho que históricamente la fonología oculta o guarda interesantes aspectos de la morfología anterior.

Sin entrar ahora en los detalles de la discusión sobre el origen de los verbos sintéticos que Schuchardt vició radicalmente con sus arbitrariedades, creo que ha quedado establecido (v. Lakarra 2006b, 2007a) no sólo que las formas conjugadas (*narraio* ‘yo le sigo’, *diharduzu* ‘tú trabajas’, *nazagu* ‘él me conoce’, etc.) son posteriores al establecimiento tras diversos cambios previos del nuevo estema radical (el fusionado con los prefijos) y no viceversa; tales formas sintéticas provienen de la aglutinación de marcas personales al verbo impersonal perifrástico o serial anterior: *nago* = *ni-da-gon*, *hago* = *hi-da-gon*, etc.<sup>46</sup>

Resulta inaceptable pretender, por tanto, que la influencia románica trajo la creación por vez primera de las diferentes perífrasis y la ruina del verbo sintético, por mucho que sea ésta a día de hoy la idea más extendida.<sup>47</sup> En realidad, parece que las perífrasis han existido siempre mientras que los sintéticos han sido en todo momento una parte mínima del verbo vasco, con un máximo de dos docenas o poco más de raíces así conjugadas, como sabemos ahora que ocurre en múltiples lenguas australianas y papúas o en algunas amerindias (cf. Garrett 2004, Pawley 2006).<sup>48</sup>

<sup>46</sup> Para el paralelismo bantú con extensiones verbales (causativo y aplicativo entre otros) que provienen de antiguos verbos seriales v. Hyman 2007a. La bibliografía sobre extensiones verbales y aplicativos en lenguas africanas y —en menor medida— “oceánicas” (del Pacífico) es prácticamente inabarcable.

<sup>47</sup> Cf. “There is a very old discussion on whether synthetic forms in Basque are older than analytical forms or vice-versa. In earlier centuries and up to the beginning of ours, researchers thought, almost unanimously, that the ancient Basque verb had only analytical (“periphrastic”) forms, simple or synthetic forms being a corruption of those [...] Today, in contrast, there is universal agreement that the handful of verbs which have synthetic forms preserve a state of affairs which must have been much more general in older times. In fact, old texts present a greater number of verbs with synthetic forms than are currently used. Nevertheless, the system of analytical forms with an auxiliary appears as perfectly consolidated even in the earliest texts. The use of analytical forms permitted the expression of a greater number of differences in tense, mood and aspect than was possible within the bounds of the synthetic conjugation (...) Contact, first with Latin and then with the Romance languages, must have at least promoted and accelerated the rapid development and expansion of the system of periphrastic forms, if it did not in fact give rise to it. It is no clear whether in Roman times the system of synthetic forms was productive [...] The problem would be how to explain that, among the huge number of verbs which were borrowed during those periods, only two, at most, entered the synthetic conjugations” (Gómez & Sainz 1995: 237-238).

<sup>48</sup> A juzgar por formas como *e-thorr-i*, *e-kharr-i*, *e-khus-i*, etc., es razonable pensar que son anteriores a la transformación *\*Th- > h-* en los primeros siglos de la Era. No habiendo verbos en *\*e-hVC* fuera de *eho* ‘moler’ (< *\*e-non*) podríamos pensar que ese tipo de formación ya estaba agotado para el s. III-IV, si bien nada impide que se desarrollara —siempre a partir de un número muy limitado de raíces— mediante los prefijos citados más arriba, ya que no con la extensión a nuevas raíces y menos con ninguna tomada en préstamo.

En Lakarra 2006b hemos revisado parte de la formación de la extrema derecha en el verbo no conjugado, en concreto de los “participios” *-ki/-gi* y *-tsi*. Mientras que el 2º es una recategorización de dos antiguas desinencias, *-tz* (cf. *hor-tz* ‘colmillo’ de *hor* ‘can’) e *-i* (cf. *e-thorr-i* ‘venir’, etc.)— sin que pueda relacionarse en absoluto con el anterior como quería Schuchardt a través de una palatalización a la románica (*-ki > -tsi*) imposible en vasco, el primero, cuya forma antigua *-gi* es el mismo sufijo “de materia” *-ki* de *txerriki* ‘carne de cerdo’, *jaki* ‘comida’, etc., nos lleva a interesarnos por la transmigración de prefijos en sufijos.<sup>49</sup> En efecto, no es sólo que tengamos *gi-* como prefijo en *gibel* ‘hígado, detrás’, *gizen* ‘gordo’, *gihar* ‘magro’, *gizon* ‘hombre’ sino que también en otros casos parece que podemos hablar del mismo fenómeno:

*de- : le-tor, le-kar	:	nai-te, lei-te	:	ekar-le, egi-le
*da- : la-bur, la-gun	:	etxe-ra, mendia-la-t	:	*du-da
*za- : za-bal	:	diru-tza, dago-z	:	

Si bien un acercamiento como el de Donegan & Stampe 1983, 2004 — resumido aquí en §3— nos llevaba a suponer que el acento en 2ª posición implicaba un orden VO (no OV), una morfología flexiva más escasa, incluso compatible con una estructura aislante tanto en el SV como en el SN —i.e., muy diferente a lo predicado en Trask 1998<sup>50</sup>—, entiendo que podemos hablar no sólo de un “cambio de dirección general” de la afijación —de izquierda a derecha— sino, incluso, del traslado de determinados morfemas

<sup>49</sup> Hyman 2008 estudia a partir de abundantes ejemplos bantúes diversas asimetrías y su interacción en el nivel de la palabra: derecha/izquierda y estema/palabra, explicando así el predominio de la sufijación sobre la prefijación o la preferencia de la “fonología anticipatoria” respecto a la “perseverativa” o la mayor ligazón a las raíces de los sufijos respecto a los prefijos (estos son excluidos del estema y del pie métrico, mientras que los primeros se integran en él).

En el trabajo hay múltiples cuestiones de interés para el futuro de la reconstrucción de la palabra, la raíz y los afijos en PV y en la prehistoria de la lengua vasca que no podemos desarrollar aquí; así, p.ej., que “once prefixes arise, being disfavored, they are more subject to loss” (p. 313), que “speakers not only process stems before affixes, but in so doing sometimes anticipate the stem-initial, skipping over a prefix, which then gets fused into the stem” (p. 316 [la cursiva es mía]) o que “prefixes join the preceding stem to form a clitic-group-like P-domain” (p. 333; cursiva de Hyman). Sobre lo último remito a Lakarra 2006b en donde se proponía —por otras razones— una migración de los prefijos a la izquierda, en la prehistoria de la flexión nominal deviniendo en sufijos de las palabras anteriores.

<sup>50</sup> Cf. “Typologically, Basque is a rather well-behaved SOV language with almost all of the textbook characteristics of such languages: verb-final order, preposed modifiers, an abundance of non-finite verb forms, a rich case system, a highly regular agglutinating morphology with few alternations, an absence of prefixes, and so on” (Trask 1998: 313).

“With its SOV word order, with its preposed complex modifiers, with its postpositions, with its periphrastic verb-forms, with its polypersonal verb agreement, with its lack of gender, of noun classes, and of verb classes, with its uniform inflection of noun phrases, and above all with its thoroughgoing ergative morphology, Basque remains today the most typologically distinct language in Europe west of the Caucasus” (Trask 1998: 323).

de su posición original a la nueva, no de mera sustitución funcional de la prefijación por la sufijación.

## 8. CAMBIOS EN LA FORMA CANÓNICA Y NUEVAS ETIMOLOGÍAS

A lo largo de estos años, particularmente en los últimos cuatro o cinco, la reconstrucción basada en la teoría de la raíz monosilábica nos ha llevado a elaborar propuestas etimológicas para más de medio millar de voces (cf. Lakarra en prep.-1). De éstas, cerca de un centenar corresponden a irregularidades y extensiones verbales antiguas que, lógicamente, se concentran en las entradas de *e-*, *i-* y *j-*, más algunas (más marcadas y desconocidas hasta ahora) en *a-* como *arraultza* ‘huevo’, *arrano* ‘águila’, *\*ardano* ‘vino’, etc. (v. Lakarra 2007a), y, aún, de otras en *h-*: *herio* ‘muerte’, *herri* ‘pueblo’, *hego* ‘ala’, *(h)erro* ‘teta de animal, raíz’, *hebain* ‘golpeado’, etc. (v. Lakarra 2009b).

Hemos de referirnos también a otro centenar de etimologías correspondientes a préstamos anteriormente no detectados que sepamos (v. Lakarra en prep.-5). Hablo de voces como *bigun* ‘blando’, *negar* ‘lágrima’, *ezain* ‘feo’, *eskatu* ‘pedir’, *erne* ‘despierto, germinar’, *orein* ‘ciervo’, *itsu* ‘ciego’, *olde* ‘voluntad’, *alu* ‘vagina’, *hezur* ‘hueso’, *ahur* ‘palma’, *oiher* ‘oscuro’, *ister* ‘corva, pierna’, *bazter* ‘rincón’, *beldur* ‘miedo’, *belau* ‘rodilla’, *hagin* ‘diente’, *bider* ‘vez’, *bizar* ‘barba’, etc. Un rápido vistazo nos muestra estructuras CVCVC, DCVC, VCCV, VCV, VCVC e incluso CVCCVC; i.e., están representados múltiples subtipos radicales que distan de lo que razonablemente podemos esperar de antiguos derivados por sufijación de la raíz CVC.<sup>51</sup> Sin embargo, casi todas esas voces tienen una extensión considerable en la geografía e historia de la lengua y pertenecen a campos semánticos como las partes del cuerpo, adjetivos, etc. que no hacen pensar en principio en préstamos; de hecho no parecen haber suscitado sospechas en ese sentido. Con todo, pienso que no corresponden a raíces PV ni a procesos de formación de palabras patrimoniales sino que, por el contrario, obtienen como tales préstamos explicaciones o derivaciones ortodoxas respecto a la doctrina reconstructiva estándar, con escasas, si alguna,<sup>52</sup> nuevas reglas o justificaciones *ad hoc*, como puede verse en §6.

Al examinar los distintos modelos de estructura radical (vide Apéndice) se hace evidente que las cantidades absolutas y los porcentajes de préstamos y de formas sin etimología conocida son distintos, incluso muy distintos los unos de los otros. Esto nos hace pensar que la cronología y origen de los mismos puede ser muy diferente. Así, mientras que apenas

<sup>51</sup> Al parecer en ainú, —v. la reconstrucción estándar de Vovin 1993—, la antigua forma canónica CVC deviene CVCV por aglutinación del sufijo posesivo y de la marca de transitivización (comunicación personal de J. A. de Alonso).

<sup>52</sup> Entre éstas cabe mencionar *\*-rm- > -rn-* y “/t/ a la derecha”, principalmente; por grande que fuera su importancia, lo es mucho mayor la combinación de otras reglas ya conocidas por Michelena. Véase al final de la n. 57.

encontramos préstamos entre las raíces CVC monosilábicas, siendo muy elevado el porcentaje (incluso el número absoluto) de las mismas sin etimología conocida, ocurre justo lo contrario en otras sin coda: 50% en *fVCV*, 60,00% en *CVfV*, 44,8% en *CVmV* o 50,00% en *mVCV*. Por lo que toca a fósiles: CVC 14,76% (que puede incrementarse verosímelmente a 25-30%), *fVCV* 0,00%, *CVfV* 0,00%, *mVCV* 0,00%, *CVmV* 0,00%, *SVRV* 1,5%, *SVTV* 1,00, *TVT* 0,00%, *TVSV* 0,50%, *TVRV* 0,62%, *SVSV* 0,5%. En estos últimos las cifras —ya claramente residuales— no pueden sino menguar; ninguna FC así existía, por tanto, en la lengua común anterior a la partición dialectal.<sup>53</sup> Incluso sin recurrir a estos últimos modelos, tan marcados, tenemos *SVRV* 30,5%, *SVTV* 35,95%, *TVT* 42,18%, *TVRV* 46,92%, etc.<sup>54</sup>

Si atendemos al conjunto de modelos sin coda (*TVT*, *TVRV*, *TVSV*, *RVS*, etc.) vemos que también aquí su carácter no-PV y tardío es evidente: abundancia de elementos en el apartado (1) de préstamos, escasez de los mismos en el (5) “de etimología desconocida”. Algo diferente parece ser el caso de modelos bisilábicos con coda a la derecha —sobre todo de (C)VCVC y algo menos de (C)VCCVC— donde es muy inferior el % de préstamos y superior el de voces de etimología desconocida.<sup>55</sup>

En modelos con vocal inicial encontramos que se da un % de voces de etimología desconocida elevado para unas cifras absolutas reducidas: *uCCVC* 7/23 (30,43%), *aCVC* 33/110 (30%), *oCV* 7/29 (24,13%), *oCCV* 8/39 (20,53%), *aCCVC* 15/90 (16,66%), *oCVC* 7/47 (14,59%), *uCCV* 6/45 (20,93%), etc.; evidentemente, cualquier identificación de alguna voz perteneciente a esos modelos radicales como préstamo o como compuesto o derivado hace que los porcentajes caigan “dramáticamente”. En otros modelos del tipo *CVCCV* —*bVCCV* 19/96 (19,79%), *zVCCV* 12/92 (13,04%), *hVCCV* 9/60 (15,00%), *gVCCV* 6/84 (7,14%)— hallamos una situación que parece idéntica (y con idénticos riesgos), si no fuera por el agravante de que aquí la relación de raíces documentadas/posibles es aún mucho más desfavorable. En un modelo aparentemente similar al último (*CVCVC*) los resultados son más favorables: *zCVC* 34/106 (32,07%), *bCVC* 25/115 (21,73%), *sCVC* 14/62 (22,58%), *hCVC* 9/31 (29,32%); no podía ser de otra manera dado que el análisis morfológico de ambos es tan dispar: mientras que *CV-CVC* se trata en origen de un derivado con

<sup>53</sup> Todos los porcentajes incluyen el grupo (2) de “variantes”, que, en puridad —pero aquí nos ha parecido innecesario—, debería eliminarse como hemos hecho en los cálculos efectuados en el siguiente §; por tanto, habría que añadir otro 10 o 15% al porcentaje de préstamos en los bisilabos; en los monosilabos, en cambio (7,89% de préstamos) apenas hay nada que añadir en ese apartado.

<sup>54</sup> *SVSV* aparece en la tabla con 0,00% de préstamos pero se trata de un dato que debe ser revisado y que en todo caso corresponde a cifras absolutas tan escasas que su testimonio deviene insignificante.

<sup>55</sup> Parece que ha de verse aquí una prueba más del monosilabismo radical CVC en PV, con raíz a la derecha en bisilabos, con prefijos, reduplicación, etc. Esto nos lleva, naturalmente, a un tipo de lengua muy diferente a la clásica; cf. §4.

prefijo y raíz a la derecha, antiquísimo (paralelo a la reduplicación y como ésta caduco ya en protovasco), CVC-CV es un modelo tardío, con raíz a la izquierda y sufijo, posible lógicamente siempre (cf. Hyman 2008) pero sólo extendido tras el desarrollo de la sufijación, que sabemos o intuimos tardía.<sup>56</sup> Pero es que, además, es fácil ver que en la extensión de este modelo radical ha tenido gran importancia la adopción de préstamos de otras lenguas: zVCCV 30/92 (32,50%), sVCCV 19/63 (30,13%), bVCCV 43/96 (44,78%), lVCCV 19/53 (35,84%), gVCCV 35/84 (41,66%), etc.; de manera muy diferente, en CVCVC, los préstamos —zVCVC 7/106 (6,60%), bVCVC 15/115 (13,04%), gVCVC 10/68 (14,70%), nVCVC 3/24 (12,50%), etc.— resultan ser bastantes menos, tanto en cifras absolutas como relativas. Por fin, es claro que en unos y otros, representantes de sistemas derivativos de muy diversa época, los respectivos momentos de su lexicalización y conversión en bisílabos habría de ser muy diferentes y distantes entre sí.

Incluso en un submodelo radical tan “genuino” (*jatorra*) como CVCCVC, que reuniría en apariencia compuestos de raíces CVC + CVC autóctonas, el análisis debe llegar hasta la triangulación estándar de CVC<sub>1</sub> y CVC<sub>2</sub>. Es un objetivo irrenunciable y no puede ser sustituido por meras presunciones si no queremos caer en burdos espejismos como *bildur* ‘miedo’ y *bazter* ‘rincón’, donde ni *\*\*bil*, ni *\*\*dur*, ni *\*\*baz*, ni *\*\*ter* están justificados formal y/o semánticamente y, por tanto, hemos de pensar más bien en sendos casos de préstamo (de *timor* y *praesepe*, respectivamente).<sup>57</sup>

Dado que la investigación del elemento latino-románico (cf. Michelena 1974) ha sido extensa e intensa<sup>58</sup> merece que por un momento recapacitemos

<sup>56</sup> Sin embargo, no podemos olvidar que las cifras finales (tras la aplicación de los distintos filtros, particularmente el geográfico) caben pocos distinguos entre los modelos radicales bisilábicos pues todos ellos quedan por debajo de cualquier umbral de validez estadística; v. n. 65.

<sup>57</sup> Algo similar ocurre en CV-CVC con *bigun* < *finu* o *nigar* < (*la*)*grima*; véase Lakarra en prep.-1, en prep.-5 y aquí §6. Para los escépticos me permito señalar que no hay regla alguna que ligue *negar* y *nigar* y permita explicar la una a partir de la otra, p.ej. con *i + r > er*, dado que ésta no es una regla a distancia, por lo que algo como *\*nirga* es imprescindible. Por supuesto, de ahí a (*la*)*crima* hay un trecho largo; quizás alguno esté dispuesto a recorrerlo, al menos en parte, teniendo en cuenta lo que Michelena advierte en la *FHV* sobre la escasez de *n-* antigua en bisílabos y polisílabos; además, la estructura CVCCV de *\*nirga* (cf. *neska* ‘muchacha’) tampoco es que haga pensar en antigüedades por sí misma; por fin, no vemos qué puedan ser *\*nir* o *\*-ga* o, yendo más adelante en la reconstrucción, *\*gir* y *\*-na*.

*Pro domo* he de decir que la primera vez que uno ve la derivación *biao* ‘siesta’ < *meridianu* (Michelena 1964) agradece mucho que todos y cada uno de los numerosos cambios y estadios intermedios estén justificados independientemente y, desde luego, que el autor de la propuesta etimológica sea ése y no otro.

<sup>58</sup> Cf. “... los préstamos, de linaje latino-románico en último término en la inmensa mayoría de los casos; queda mucho por aclarar, pero es más bien cuestión de detalle y de precisión. En otras palabras, no iremos mucho más lejos del estado actual, salvo por lo que respecta al rigor y a algún que otro descubrimiento aislado que todavía pueda hacerse. En la inmensa mayoría de los casos, esto convertiría presunciones en certezas, pero las presunciones están ya presentes” (Michelena 1970, 308).

sobre qué razón ha provocado que estos y otros compañeros (en torno al centenar, de momento) no fueran detectados con anterioridad. En mi opinión la explicación puede venir en la siguiente dirección: estas voces han sufrido una importante serie de cambios que los han alejado mucho de la FC de su lengua de origen y se han acercado a otras más comunes en el euskara del segundo milenio de la Era. Es más, diríase, incluso (v. §9), que la introducción masiva de préstamos encuadrados en ciertos modelos radicales ha contribuido decisivamente en el triunfo moderno de estos: han sido capaces de introducir por su número modelos nuevos en la lengua recipiendaria, de tal manera que hablantes y aun lingüistas las consideran anodinas, similares a las auténticamente patrimoniales e indistinguibles de ellas. Sólo un análisis minucioso de sus FFCC, sus particularidades fonotácticas y sus etimologías nos dice que la historia real ha sido muy otra.

## 9. SOBRE EL TEMPO DEL BISILABISMO

Es sabido que para Uhlenbeck o Michelena las raíces vascas, históricamente bisilábicas, lo eran también en su mayoría para etapas prehistóricas o protohistóricas anteriores.<sup>59</sup> Los datos de los análisis efectuados por nosotros respecto al casi medio centenar de modelos radicales examinados hasta el presente (v. Apéndice) para clasificar las voces en ellos incluidas según el método explicado en §5, vienen a reforzar y consolidar esa impresión por lo que a los últimos 500 años de la lengua se refiere: frente a 112 radicales monosilábicos (tras descontar variantes y compuestos o derivados [cf. n. 64]) tendríamos 1400 bisílabos. Hay, por tanto, algo más de una docena de bisílabos históricamente disponibles por cada monosílabo, si bien variarán seguramente mucho los bisílabos que el hablante o el

---

Por cierto que las consideraciones formales nunca estuvieron ausentes del todo en la labor de Michelena al respecto, como se ve en *Sobre el pasado de la lengua vasca*: “Por consideraciones formales sobre todo, uno se siente p. ej., inclinado a ver préstamos en vize. *abagadaune* ‘ocasión’ (con el suf. *-une* tratado en 4.2.), occid. ant. *apukadu* ‘sucedido’, ronc. *atizatu*, *atxezatu* ‘comulgado’, *dollar* ‘ruín, malo’, *elikatu* ‘abstenido’ y ‘alimentado’, vize. ant. *endorea* ‘el alcalde’, *errexal* ‘árbol’ en Landucci, etc., pero no se ha señalado que sepamos su origen preciso de manera solvente” (Michelena 1964, 33).

Lamento tener que decir que no todo ha sido progreso desde entonces ante obras como la de Segura y Etxebarria. Por lo demás, creo ser consciente de los riesgos que las conocidas “costumbres” lexicográficas de Azkue acarrearán respecto a la presencia y al registro de la extensión dialectal de los préstamos en su diccionario. Es claro que hay una inmensa labor por delante, incluso por lo que toca a la extensión y cronología de los préstamos, antes de elaborar cualquier cronología y periodización de la lengua, inclusive la de sus formas canónicas.

<sup>59</sup> Cf. “Tales elementos por ahora últimos [los elementos más simples a los que pueden ser retrotraídas las formas complejas] no siempre son monosilábicos; ni siquiera lo son, al contrario de lo que ocurre en caucásico del noroeste, en la mayoría de los casos. Hallamos, por lo general, bases bisilábicas. Si se prescinde de algunos temas nominales (*su* ‘fuego’, etc.) y de raíces verbales como *-gi* ‘hacer’, *-go* ‘estar, permanecer’, son los sufijos de derivación y de flexión los que suelen ser monosilábicos” (Michelena, *PT* 312).

testimonio de éste o aquel dialecto conozca o utilice: en trabajos anteriores (v. Lakarra 2008b, p.ej.) hemos mostrado que son muy escasos en comparación con los monosílabos —generales en la lengua en su inmensa mayoría— los bisílabos que consiguen alcanzar un estatus pan-dialectal significativo. Muchos no llegan a cubrir ni de lejos un sólo área dialectal, bastantes reducen su implantación a (partes de) dos dialectos contiguos, etc., de manera que son los menos los que cubren de manera continua o discontinua varios dialectos alejados entre sí y, en general, aquellos que pueden aspirar a ser considerados parte del lexicón del Vasco Común.<sup>60</sup>

Si bien Martínez Areta 2003 [2008] defiende que el bisilabismo corresponde no sólo a la etapa histórica de la lengua sino incluso a todas las posteriores al PV antiguo (i.e., al PV moderno, al aquitano y al vasco medieval),<sup>61</sup> uno tiende a pensar que la “fecha” de la conversión de la FC en bisilábica —un fenómeno que, desde luego, no ocurriría de un día para otro— es mucho más tardía. Al examinar la antigüedad relativa de los rasgos mencionados arriba en §3 como muestra del cambio tipológico (“deriva”) experimentado por la lengua, observamos que buena parte de aquellos que en otras lenguas han relacionado lingüistas y tipólogos con el acrecentamiento de la aglutinación y con el orden SOV son en vascuence (en la medida en que somos capaces de detectar o de intuir su aparición) claramente tardíos, postaquitano y, probablemente, tardoantiguos o medievales: oclusivas sordas en inicial, vocales nasales, relativas a la izquierda del N, (escasos) sufijos, prominencia aspectual o temporal en el V... Incluso algunos de los restantes rasgos —aparentemente más antiguos o

---

<sup>60</sup> Espero dar cifras más afinadas y actualizadas en un trabajo en preparación sobre la distribución dialectal de bisílabos y monosílabos y las implicaciones de la misma (Lakarra en prep.-3) aunque es difícil que varíen las conclusiones generales. Por otra parte, si bien antes y ahora he utilizado a efectos prácticos los dialectos de la clasificación de Bonaparte (mediados del XIX) no asumo en absoluto que la misma refleje el estado inmediatamente posterior a la disgregación del Vasco Común, ni tampoco que ésta pudiera ser representada como un estema de 8 o de 6 ramas hermanas. Se discuten algunas de las razones y consecuencias de este escepticismo en Lakarra & Urgell 2008; con todo, me apresuro a señalar que en todos mis trabajos hasta el presente y para no parecer excesivamente cicatero y restrictivo con el “enemigo” (o demasiado generoso con mis propias hipótesis) he tratado como “común” toda voz documentada en dos o más territorios no contiguos, incluso cuando en ellos sólo ocupaba una ínfima parte y no la totalidad de su extensión.

<sup>61</sup> Cf. “Para entender cómo surgen en el vascuence histórico las raíces disilábicas canónicas hemos de dividir la época que acabamos de describir en dos. En una primera época que podemos denominar “pre-*proto-vasco-A*” se extendió en el sintagma nominal la formación de pies yámbicos, pero necesariamente ha de tenerse en cuenta que estos eran sintagmas, i.e., combinaciones sintagmáticas y no nombres disilábicos (...) Los compuestos, por tanto, se interpretaban analíticamente, pero muchos de ellos se convirtieron en fórmulas por el uso. En una etapa posterior que podemos denominar “pre-*proto-vasco-B*”, muchos sintagmas que en “pre-*proto-vasco-A*” se habían convertido en fórmulas se lexicalizaron. Los sintagmas disilábicos formados por dos lexemas monosilábicos se convirtieron en palabras disilábicas monolexemáticas, dando origen a las raíces disilábicas, las más comunes en vascuence histórico” (78-79; la traducción es mía [J.A.L.]).

más difíciles de fechar— como las posposiciones (incluida la de ergativo), el complejo verbal, una categoría de adjetivo relativamente nutrida, aumentada con antiguos V estativos, derivados, préstamos, etc... no parecen necesariamente pre-aquitano y, por tanto, asimilables siquiera al PV reciente, i.e., al reconstruido por Michelena para los últimos siglos del I milenio a. de C. Es posible que la deriva que podemos intuir tras múltiples evoluciones esbozadas más que analizadas hasta la fecha —y seguramente en otras adicionales— se haya dado en su mayor parte, no en los periodos anteriores al PV reconstruido, sino más bien, en el que va de los testimonios aquitanos a los bajomedievales.<sup>62</sup>

Antes de alegar otras razones de peso, creo que la distribución geográfica de monosílabos y bisílabos juega totalmente en contra de la conclusión de Martínez Areta: parece manifiestamente claro que mientras que las raíces monosilábicas corresponden a un estado de lengua anterior o muy anterior a la diversificación dialectal, por el contrario los bisílabos se han formado o introducido en la lengua —o al menos se han generalizado en ésta— con posterioridad a la fragmentación dialectal, pues en otro caso su extensión posterior sería mucho mayor que la realmente documentada. Desde que Michelena lo propusiera en 1981, los vascólogos venimos aceptando los siglos V-VI d.C. como la época del Vasco Común, i.e., la correspondiente al último (por más tardío) estado de lengua reconstruible antes de producirse la fragmentación dialectal); pues bien, habríamos de concluir entonces que la FC bisilábica —¿cuál de ellas, por cierto?— es posterior al menos a esa época, si no al s. X que Michelena fijó como fecha *antequam* de la aparición de los primeros rasgos dialectales.<sup>63</sup>

A este argumento de índole geográfico cabe añadir varios otros complementarios. Por una parte, los bisílabos modernos han sido afectados por una serie de cambios a lo largo de toda su estructura (inicial, final y, sobre todo, medial), como las caídas de vocales altas en final de antiguos primeros elementos, las *-r* o *-n* en idéntica posición, etc., que son posteriores

---

<sup>62</sup> En el SN es sabido que las posposiciones son claramente tardías y aun hoy su grado de gramaticalización muy escaso; los sufijos casuales, antiguas posposiciones y otras desinencias son tardías, particularmente las correspondientes al plural, basadas en la gramaticalización del artículo que tampoco pudo ser muy anterior a los ss. X-XI; vide Manterola 2006 y 2009.

Los derivados, por fin, han tenido un desarrollo históricamente tardío (v. Lakarra 1997a, Sarasola 1997 y Bueno 2004, 2006), relacionado muy directamente con el de la lengua literaria y, en buena parte, no son conocidos hasta fechas muy recientes fuera de vocabularios como el de Pouvreau o el de Larramendi. Muchos supuestos sufijos latinos o románicos antiguos todavía eran claramente ajenos a la lengua popular en el XVIII, pues no parecen haberse dado combinaciones mixtas de raíz tomada en préstamo y sufijo autóctono (ni viceversa), por lo que no ha habido análisis ni integración de tales sufijos hasta hace pocas décadas, ni siquiera siglos.

<sup>63</sup> En el trabajo citado tres notas más arriba creemos haber mostrado la inverosimilitud de que todos los dialectos —sean estos los de Bonaparte u otros cualesquiera— sean coetáneos y que, desde luego, hayan mantenido sus actuales límites —sean los que fueren— desde la noche de los tiempos.

no ya al PV reciente o micheleniano sino, por lo menos, al aquitano. Es cierto que, carentes de documentación directa para los siglos que van del III al X —y de estudios de cronología relativa de los cambios *à la* Straka o *à la* Pensado, todo hay que decirlo—, no somos capaces de datar todavía con más precisión tales transformaciones pero, posiblemente, no sean anteriores al Vasco Común de Michelena o quizás aún más tardías.

Por otro lado, hemos reiterado que la proporción de préstamos de los bisílabos es muy elevada; veamos simplemente y por encima los casos de TVTV, TVRV, zVCCV, TVSV y mVCV, los cinco modelos radicales bisilábicos más numerosos.

Tabla 1<sup>63b</sup>

Mod.rad.	R.Pos.	Doc.	C/D	Var.	Resto	Prést.	Onom.	ED
CVC	325	152	04	36	112	12	30	70
TVTV	900	190	24	35	131	80	36	15
TVRV:	600	195	21	37	137	95	16	25
mVCV:	425	94	03	16	75	47	16	12
zVCCV:	1075	92	10	21	61	30	19	12
TVSV:	600	95	10	22	63	47	02	14

Sólo podremos valorar adecuadamente la relevancia de los datos aportados si tenemos en cuenta que estos cinco modelos radicales suponen el 33,35% de todos los bisílabos<sup>64</sup> (467 / 1400) integrados en los cuarenta y siete modelos analizados. Pues bien, el 64,28% de los nuevos radicales bisílabos son préstamos indudables<sup>65</sup> y, por si hiciera falta algo más para cerciorarnos de la escasa antigüedad de los modelos particulares que estamos examinando, recuérdese que todos ellos carecen de coda en la segunda sílaba (sólo uno la tiene en la primera) por lo que es inverosímil que la raíz haya residido nunca ahí dada la estructura (-)CVC de ésta en épocas antiguas.

En efecto, una estructura CVCV puede derivarse por sufijación de una raíz monosílaba CVC si el sufijo es -V o, también, con -CV por simplificación del grupo -CC- resultante, bien que esto último en época tardía

<sup>63b</sup> Leyenda: Mod.Rad. = Modelo radical; R.Pos. = Número de raíces potenciales del modelo; Doc. = Id de documentadas; C/D = número de compuestos y derivados atestiguados; Var. = Id de variantes; Prést. = Id de préstamos; Onom. = Id de onomatopeyas y fonosimbolismos; ED = Voces de etimología desconocida.

<sup>64</sup> Entendiendo esto aquí como la resta de (2) “variantes” y (3) “compuestos y derivados” a (1) “total de documentados”.

<sup>65</sup> Como se sabe (véase la última nota del §4, el grupo (5) va reduciéndose y, alternativamente, ampliándose los restantes, particularmente el (1); por tanto, el porcentaje de préstamos sólo puede aumentar en el grado que sea. Además, según avanzan los estudios etimológicos se hace evidente, si no lo era ya desde un comienzo, que los escasos y menguantes bisílabos sin explicación no son sino muestra de nuestra incapacidad momentánea para acabar con unos restos que, por otra parte, perdieron ya hace mucho tiempo todo interés estadístico y cualquier capacidad real de constituirse en alternativa, incluso en alternativa muy minoritaria, a la hegemonía absoluta de los monosílabos.

(medieval). Es lo que ocurre con varias de las voces que hallamos en los grupos “(3)” [= compuestos/derivados] de los modelos analizados:

TVTV: *baku, dago, dina, duka, bage, bako, bape, batu, begi, beko, bete, beti, biga, biki, biko, gogo, guti, pape, peka, peko, tegi, toka, tuka, tupa.*

TVRV: *bana, barru, bele, bere, bero, berri, bina, gari, gerri, giño, gora, gorri, gune, kola, kona, tori, dena.*

TVSV: *baso, batsu, batza, batze, batzu, beso, bizi, gatzu, gatzí.*

mVVCV: *meki, mihi.*

zVCCV: *zaldi, zantza, zazpi, zinka, zolda, zortzi, zozko, zunka.*<sup>66</sup>

Pero es que, aun aceptando la antigüedad de estas voces —lo cual en su mayoría es difícil dada su transparencia y su escasa implantación dialectal o la circunstancia de que algunas sean elementos gramaticales evidentes como *gino, kola, kona, dago, dena*, etc.—, debe señalarse que, como se ve en la tabla 1, el número de derivados y compuestos es no sólo muy inferior al de préstamos (en proporciones que van del 1/3 al 1/15) sino incluso al de fonosimbolismos, por lo que incluso esta vía de expansión del modelo radical parece más relevante que el de compuestos y derivados. Dejando análisis más pormenorizados para futuros trabajos (v. Lakarra en prep.-3), dediquemos un momento a la extensión dialectal de los préstamos presentes en los modelos radicales bisilábicos: si tomamos los tres más numerosos de entre los arriba mencionados (TVTV, TVRV, TVSV) encontramos en ellos 210 préstamos, de los cuales 132 (62,85%) no satisfacen los criterios habituales para establecer el calificativo “general” de una voz. Convendría precisar, que mientras en TVTV hablamos de un empate a 38 (50%), en TVRV las cifras son de 28 voces que superan los criterios frente a 61 que no lo hacen (32,58% ~ 67,42%) y en TVSV de 12 frente a 29 (29,24% ~ 70,76%), respectivamente. Por tanto, cabría pensar que fue TVTV el primer radical bisilábico que consiguió el estatus de forma canónica y que TVRV y TVSV lo harían sólo posteriormente, en todos los casos valiéndose más de los préstamos que de voces del resto de las clases (compuestos y derivados, fonosimbolismos y voces “simples”, i.e., sin etimología conocida hasta el momento).<sup>67</sup>

<sup>66</sup> Dejo fuera de la lista *gero, goza, kari, kera, zildi, zorti* sobre cuyo carácter de compuestos/derivados albergo ahora dudas que no tenía anteriormente.

<sup>67</sup> Siendo verosímil esta conclusión provisional, he de señalar una cuestión metodológica básica que altera (retrasa) la cronología —cualquier cronología— basada en los préstamos: mientras que en las voces patrimoniales “general” o “común” equivale casi siempre a “incluido en la protolengua inmediatamente anterior a la ruptura dialectal” —o lo pretende: ténganse en cuenta las innovaciones comunes posteriores a la ruptura dialectal—, es claro que eso mismo predicado de un préstamo es, no imposible pues todo estado de lengua tiene préstamos, pero sí bastante menos seguro: i.e., podemos imaginarnos fácilmente un escenario en el que voces de relativamente reciente introducción en la lengua logren una amplia presencia en todo el territorio o, al menos, en dos dialectos no contiguos (labortano y suletino, navarro y vizcaíno), lo cual permite dar una lectura positiva —de “antigüedad”— incluso a voces de una historia muy limitada en la lengua.

Cabe añadir un tercer tipo de argumento a favor del carácter tardío de los radicales bisilábicos. Parece razonable suponer que es sólo tras la generalización de nuevos compuestos y derivados trisílabos y cuadrisílabos a partir de bases bisilábicas cuando cabe hablar de raíces bisilábicas o al menos de una extensa FC bisilábica. Este proceso es tan claramente tardío para cualquiera que ni siquiera ha sido estudiado hasta el momento desde un punto de vista histórico,<sup>68</sup> al carecer de cualquier opción de antigüedad respecto a idénticos procesos basados en monosílabos. En todo caso, los trisílabos y, sobre todo, los cuadrisílabos resultantes son en general muy poco opacos (lo cual no es precisamente indicio de arcaísmo) y, desde luego, muestran en su estructura evidentes rasgos de la fonética postaquitana y medieval.

Sería totalmente prematuro dar aquí por cerrada la cuestión pero creo ver bastantes más indicios y posibilidades de una datación tardía —medieval más que tardoantigua— de la FC bisilábica que de otra, anterior en milenio y medio y quizás bastante más; es, además, este escenario el que se acomoda mejor a otros aspectos de la deriva de la lengua como puede ser el carácter tardío de la sufijación derivativa —y, en otra medida, de la flexiva—, la escasa gramaticalización de las posposiciones, etc.; v. Lakarra 2005a, 2006a.<sup>69</sup>

## 10. CONCLUSIÓN

Frente a lo sucedido en otras lenguas y familias de tradición diacrónica más desarrollada, la forma canónica de los morfemas (en especial de la raíz) no ha sido estudiada hasta fechas recientes (cf. Lakarra 1995ss). Tal hecho ha motivado su no utilización en la reconstrucción del protovasco y de la

---

No es el caso, claro, de *baba*, *bago* o *bake* y, seguramente, de *butzu* y de *gozo*; no me atrevería a decir lo mismo de *bila*, *bilo*, *dallu*, *dorre*, *gona* o *kisu* y, desde luego, de *pare*, *tira*, *koba*, *paga*, *thorra*, *kexa*, *petxa* o *totxo*.

<sup>68</sup> Véase Azkarate 1991 desde un punto de vista sincrónico o pancrónico.

<sup>69</sup> Creo que, en lo fundamental, puede y debe mantenerse la relación establecida por Igartua en un brillante artículo de 2001 entre la evolución de la desaspiración y la de la forma canónica de la raíz:

(...) De ahondar en esta perspectiva y desarrollarla hasta sus consecuencias lógicas, el proceso de desaspiración no sería más que un epifenómeno de la evolución gramatical que conduce a la raíz vasca del monosilabismo al polisilabismo (la raíz canónica va experimentando cambios pero las restricciones que afectan a la aspiración continúan siendo durante mucho tiempo las mismas, de modo que la regla no varía: sólo resulta posible una aspiración por cada raíz (210).

Disiento, sin embargo, con él en denominar “Protovasco II” y “Protovasco III” a las épocas en que que la lengua adquiere, respectivamente (se supone que de manera masiva), raíces bisilábicas y polisilábicas. Creo haber dado razones para situar esos cambios de la forma canónica en etapas mucho más tardías de la lengua; incidentalmente, *andere* (y seguramente *alhaba*, con su *-ba* de término de parentesco) no es el mejor ejemplo de raíz simple trisilábica que pudiera estar presente ya en aquitano; vide Lakarra 2005b y aquí al comienzo del §4.

prehistoria de la lengua, así como en la labor etimológica, en la cual hubiera sido de gran valor, dadas las características del corpus vasco y el aislamiento genético de la lengua. Así se ve, p.ej. (cf. Lakarra 2009c y varios trabajos en preparación) en diversos problemas planteados por el acento o la aspiración, o por los fenómenos de inicial de palabra, la posición más complicada para el reconstructor según Michelena 1977a.

La teoría de la raíz monosilábica en PV ha conseguido varios tipos de frutos que avalan el nuevo acercamiento: a) nuevas raíces y ampliación de familias léxicas ya conocidas; b) nuevos elementos de la antigua gramática (prefijos y reduplicaciones) y c) filtro adicional relevante para el método comparado tradicional en la comparación con otras lenguas y protolenguas (v. §§1-2).

Durante los últimos años la investigación ha transcurrido por dos vías que motivan, creemos, el estudio aquí presentado: 1) se ha mostrado la existencia de una clara deriva en los modelos radicales de la lengua, deriva que, obviamente tuvo consecuencias en el tipo general de aquella, anteriormente muy diferente a la moderna aglutinante, SOV, con abundantes sufijos en el SN y florida concordancia verbal (v. § 3) y 2) el estudio de la distribución e historia de los modelos radicales nos ha llevado a plantear la posibilidad de una etimología formal, no atomista, basada en la FC y, quizás, a la elaboración ulterior de una nueva cronología y periodización de la lengua (v. §4).

Aquí hemos querido abordar una serie de cambios sufridos por la raíz PV que hicieron que ésta fuera históricamente bisilábica de manera muy mayoritaria desde una situación anterior monosilábica estricta, inspirándonos en otros casos similares conocidos (v. §5) como el chino, el austroasiático o el tani. Junto a cambios más generales como la fosilización de afijos y compuestos, la adopción de tipos radicales a través de los préstamos y, tal vez, el desplazamiento del acento, creemos haber mostrado otros procesos más específicos como ciertas metátesis consonánticas, el movimiento a la izquierda de la aspiración y a la derecha de enes y erres, las asimilaciones vocálicas o la adopción de restricciones contra las consonantes homorgánicas; todo ello coadyuva a la disolución de antiguas lindes internas del nuevo bisílabo, a la consolidación de las lindes externas renovadas y al logro de una mayor uniformidad interna de la raíz resultante. Fenómenos como la inserción de /g/ en hiatos, los cuales hubieran podido dar lugar a monosílabos, han de ser entendidos también en el mismo sentido de la “conspiración” favorable a la nueva FC.<sup>70</sup>

---

<sup>70</sup> También, claro está, el desarrollo de la desaspiración mencionada de pasada en la n. 69 y el de la acentuación que no he tratado aquí y no sólo por razones de espacio. Habrá de recibir un tratamiento detenido en otra ocasión, si bien hay bastante bibliografía reciente como Martínez Areta 2004 [2008] o Elordieta en prensa. No veo razones para alterar lo dicho en Lakarra 1996 y más pausadamente, siguiendo a Donegan & Stampe 1983 y 2004 en Lakarra 2006a: acento en segunda (final), en la raíz, para los pies bisilábicos en época antigua, escenario apto para prefijos pero no para sufijaciones abundantes; otra cosa es para etapas más modernas y determinar cuáles

Entre las implicaciones de la teoría de la raíz monosilábica y, más específicamente, de la investigación del cambio en la FC de los morfemas cabe mencionar varias de especial relevancia en la reconstrucción diacrónica del vascuence:

- 1) la posibilidad de entender de una manera más coherente un importante número de cambios fonológicos (segmentales y prosódicos) y morfológicos aparentemente inconexos y que coadyuvan en la transformación de la forma canónica de las raíces de monosilábica en bisilábica (§6). Idealmente, tal vía de investigación podría llevarnos a establecer en un futuro una cronología y una periodización más satisfactorias de la historia y prehistoria de la lengua;
- 2) la determinación de la antigua estructura y orígenes del estema verbal vasco (absolutamente abigarrado para Lafon), mediante la reducción de la raíz polimórfica histórica al modelo radical CVC, precedida en su caso — además de por el general \*e-, de los prefijos \*-ra, \*-da y su combinación (\*-da-ra), ésta última sujeta a transformaciones importantes antes de su fosilización, como en las lenguas bantúes (v. §7);
- 3) la obtención de más de medio millar de nuevas etimologías gracias al nuevo paradigma, entre las que cabe destacar un centenar largo de derivados verbales y nominales antiguos, tiempo ha fosilizados, así como de varias decenas de nuevos préstamos antiguos latino-románicos, previamente no detectados o explicados de manera insatisfactoria en nuestra opinión (v. Lakarra en prep.-1 y en prep.-5);<sup>71</sup>
- 4) los préstamos —con casi 2/3 de las voces documentadas en los modelos radicales bisilábicos más comunes— parecen haber tenido un papel crucial en la extensión del bisilabismo (y no sólo de los modelos en -V, imposibles anteriormente en un estadio de lengua con estructura prefijal o reduplicativa en CV-CVC o (C)V-CVC), mucho más relevante que el desempeñado por los más escasos, generalmente transparentes y tardíos, compuestos y derivados (v. §§8-9);
- 5) Los datos y argumentos anteriores nos llevan a postular fechas tardías (medievales o, a lo sumo, premedievales) para la consolidación y extensión del bisilabismo como tal forma canónica en los morfemas radicales vascos (v. §9).<sup>72</sup>

---

sean éstas. Es más, no me parece que nuestros acentólogos sigan al pie de la letra aquella máxima de Benveniste que nos decía que el establecimiento de una cronología debía ser la mayor preocupación del “comparatista”, léase del lingüista histórico.

<sup>71</sup> Como se explica en §8, la casi perfecta acomodación de tales préstamos a las FFCC tradicionales o históricamente más extendidas de la lengua llegó a tal punto que fueron tomados en buena parte como patrimoniales o indistinguibles de ellas por hablantes y aun por lingüistas.

<sup>72</sup> Aquí y en todo el trabajo (en realidad, desde nuestros primeros trabajos en 1995) nos estamos refiriendo a formas canónicas mayoritarias y no a la (in)existencia más o menos marginal *como tales raíces simples* de algunos bisílabos y polisílabos patrimoniales, ni negamos la “lectura sintagmática” de los bisílabos todavía no lexicalizados (cf. Feng para el chino arcaico y Martínez Areta para nuestra lengua).

## BIBLIOGRAFÍA

- Agud & Tovar 1988: M. Agud & A. Tovar, *Materiales para un diccionario etimológico de la lengua vasca*. Anejos de ASJU, Donostia-San Sebastián, 1988-.
- Aikhenvald 2006: A. Y. Aikhenvald, “Serial verb constructions in typological perspective” in Aikhenvald & Dixon (eds.), 1-68.
- Aikhenvald & Dixon 2006: A. Y. Aikhenvald & R. M. W. Dixon (eds.), *Serial verb constructions. A cross-linguistic typology*. Oxford 2006.
- Anderson 2006: G. D. S. Anderson, *Auxiliary Verb Constructions*. Oxford 2006.
- Arbelaiz 1978: J. J. Arbelaiz, *Las etimologías vascas en la obra de Luis Michelena*. Tolosa 1978.
- Artiagoitia 1990: X. Artiagoitia, “Sobre la estructura de la sílaba en (proto)vasco y algunos fenómenos conexos”, *ASJU* 24, 1990, 327-349.
- Artiagoitia, Goenaga & Lakarra 2002: X. Artiagoitia, P. Goenaga & J. A. Lakarra (eds.), *Erramu Boneta: Festschrift for R. P. G. de Rijk*, Supplements of *ASJU* XLIV, Bilbao 2002.
- Artiagoitia & Lakarra 2008: X. Artiagoitia & J. A. Lakarra (eds.), *Gramatika Jaietan. Patxi Goenagari Omenaldia*. Anejos de *ASJU* LI, Bilbao 2008.
- Azkarate 1991: M. Azkarate, *Hitz elkartuak euskaraz*. Donostia-San Sebastián 1991.
- Azkue 1905-06: R. M<sup>a</sup> Azkue, *Diccionario vasco-español-francés*, 1905-06 [Reed. con apéndice], Bilbao 1969.
- Azkue 1923-25: R. M<sup>a</sup> Azkue, *Morfología vasca*, I-III, Bilbao 1923-25. 2<sup>a</sup> ed. 1969.
- Bakrò-Nagy 1992: M. Sz. Bakrò-Nagy, *Proto-Phonotactics. Phonotactic investigation of the PU and PFU consonant system*. *Studia Uralica* 5, Harrassowitz Verlag 1992.
- Baldi 1991 = Ph. Baldi (ed.), *Patterns of change, change of patterns. Linguistic change and reconstruction methodology*, Berlin-New York 1991.
- Baldi & Page 2006: Ph. Baldi & B. R. Page, “(reseña) Theo Vennemann. *Europa Vasconica-Europa Semitica*”, *Lingua* 117, 2006, 2183-2220.
- Benveniste 1935: E. Benveniste, *Origines de la formation des noms en indo-européen*. Paris 1935.
- Bueno 2004: A. Bueno, *Nominalizazio atzizki deitutakoen azterketa historikoa eta morfologikoa*. Trabajo de investigación de doctorado (dir. J. A. Lakarra). UPV/EHU, Vitoria-Gasteiz 2004.
- Bueno 2006: A. Bueno, “Atzizkiak aztergai autore ezberdinen begiradapean” in Lakarra & Hualde (eds.), 221-244.

---

Es posible que Downing 2006 (cf. reseña de Ussishkin 2009) traiga alguna secuela favorable para el campo; de momento sus intereses están bastante alejados de los de cualquier diacronista y no digamos de los específicamente etimológicos.

- Bybee, Perkins & Pagliuca 1994: J. Bybee, R. Perkins & W. Pagliuca, *The evolution of grammar. Tense, aspect and modality in the languages of the world*. U. of Chicago P. 1994.
- Corominas & Pascual 1980-91: J. Corominas & J. A. Pascual, 1980-91, *Diccionario crítico etimológico castellano e hispánico*. Madrid.
- Dai 1990: J. X.-L. Dai, "Historical morphologization of syntactic words: Evidence from Chinese derived verbs", *Diachronica* 7, 1990, 9-46.
- Dixon 2002: R. M. Dixon, *Australian languages*, Cambridge 2002.
- Donegan 1993: P. Donegan, "Rhythm and vocalic drift in Munda and Mon-Khmer", *Linguistics in the Tibeto-Burman Area* 16, 1993, 1-43.
- Donegan & Stampe 1983: P. Donegan & D. Stampe, "Rhythm and the holistic organization of language structure" in J. Richardson et alii (eds.), *Papers from the Parasession of phonology, morphology and syntax*. Chicago 1983, 337-353.
- Donegan & Stampe 2004: P. Donegan & D. Stampe, 2004, "Rhythm and the synthetic drift of Munda", *The Yearbook of South Asian Languages and Linguistics 2004*, de Gruyter (Berlin-NY), 3-36.
- Downing 2006: L.J. Downing, *Canonical forms in prosodic morphology*, Oxford & NY 2006.
- Duanmu 1999: S. Duanmu, «Stress and the development of disyllabic words in Chinese», *Diachronica* 16, 1999, 1-36.
- Elordieta en prensa: G. Elordieta, "Azentu zaharraz". En prensa en las Actas del XVI Congreso de la Academia de la Lengua Vasca – Euskaltzaindia (Pamplona, octubre de 2008).
- Ernout & Meillet 1979: A. Ernout & A. Meillet 1979, *Dictionnaire étymologique de la langue latine*. 4ème éd., Paris 1979.
- Feng 1998: Sh. Feng, "Prosodic structure and compound words in Classical Chinese", in Packard (ed.), 197-260.
- François 1999: A. François, "Mouvements et clonages de voyelles en motlav. Entre phonologie et morphologie", *BSL* 94, 1999, 437-486.
- Gamkrelidze & Ivanov 1984: T. V. Gamkrelidze & V. V. Ivanov, *Indo-European and the Indo-Europeans*. Trad. del ruso: Mouton de Gruyter, Berlin & NY 1995.
- Garrett 2004: A. Garrett, "The evolution of Algic verbal stem structure: new evidence from Yurok", ms., UCB. 2004.
- Gómez 1994: R. Gómez, "Euskal aditz morfologia eta hitzordena: VSO-tik SOV-ra" in J.-B. Orpustan (ed.), *La langue basque parmi les autres*. Baigorri 1994: 93-114.
- Gómez & Sainz 1995: R. Gómez & K. Sainz, "On the Origin of the Finite Forms of the Basque Verb", in Hualde, Lakarra & Trask (eds.), 235-274.
- Good 2005: J. Good, "Reconstructing morpheme order in Bantu. The case of causativization and applicativization", *Diachronica* 22, 2005, 3-57.
- Gorrochategui 1984: J. Gorrochategui, *Estudio sobre la onomástica indígena de Aquitania*, Bilbao 1984.

- Gorrochategui 1999: J. Gorrochategui, “La romanización en el País vasco: aspectos lingüísticos” in *Antiqua. VI Jornadas sobre la Antigüedad*. Donostia: *Bitarte*, 1999, 10-23.
- Gorrochategui 2001a: J. Gorrochategui, “Planteamientos de la lingüística histórica en la datación del euskara”, *XV Congreso de Estudios Vascos*, Donostia-San Sebastián 2001: 103-114.
- Gorrochategui 2001b: J. Gorrochategui, “Aitzineko euskara”. Ponencia del *XV Congreso de la Real Academia de la Lengua Vasca – Euskaltzaindia*.
- Gorrochategui 2008: J. Gorrochategui, “Euskara zaharra”. Ponencia del *XVI Congreso de la Real Academia de la Lengua Vasca – Euskaltzaindia (Pamplona, 6-10 de octubre)*. En prensa.
- Gorrochategui & Lakarra 1996: J. Gorrochategui & J. A. Lakarra, “Nuevas aportaciones a la reconstrucción del protovasco”, in Villar & Encarnação (eds.), 101-145.
- Gorrochategui & Lakarra 2001: J. Gorrochategui & J. A. Lakarra, “Comparación lingüística, filología y reconstrucción del protovasco”, in Villar & Fdez Alvarez (eds.), 407-438.
- Guiter 1989: H. Guiter, “Elementos de cronología fonética del vascuence”, *ASJU* 23, 1989, 797-800.
- Hamp 1998: E. P. Hamp, “Some draft principles for classification” in J. C. Salmons & B. D. Joseph, (eds.), *Nostratic. Sifting the evidence*. Amsterdam – Philadelphia 1998: 13-15.
- Harris 1990: A. C. Harris, “Kartvelian contacts with IE” in T. L. Markey & J. A. C. Greppin (eds.), *When worlds collide. Indo-Europeans and Pre-Indo-Europeans*. Ann Arbor 1990: 67-100.
- Hualde, Lakarra & Trask 1995: J. I. Hualde, J. A. Lakarra & R. L. Trask, (eds.), *Towards a history of Basque language*. Amsterdam & Philadelphia 1995.
- Hurch 1988: B. Hurch, “Is Basque a syllable-timed language?”, *ASJU* 22, 1988, 813-825.
- Hyman 2004: L. Hyman, “How to become a “Kwa” verb”, *Journal of West African Languages* 30, 2004, 69-88.
- Hyman 2007a: L. Hyman, “Niger-Congo verb extensions: Overview and discussion” in *Selected Proceedings of the 37<sup>th</sup> Annual Conference on African Linguistics*, ed. D. L. Payne & J. Peña, 149-163. Somerville, 2007.
- Hyman 2007b: L. Hyman, “Reconstructing the Proto-Bantu verbal unit: Internal evidence”, *SOAS Working Papers in Linguistics* 15, 2007, 201-211.
- Hyman 2008: L. Hyman, “Directional asymmetries in the morphology and phonology of words, with special reference to Bantu”, *Linguistics* 46, 2008, 309-350.
- Igartua 2001: I. Igartua, “La aspiración en vasco: ensayo tipológico y diacrónico”, *ASJU* 35/1, 2001, 185-213.

- Igartua 2006: I. Igartua, “Del origen de la aspiración como elemento morfológico en vasco”. En Lakarra & Hualde (eds.), 519-530.
- Igartua 2008: I. Igartua, “Historia abreviada de la aspiración en las lenguas circumpirenaicas”. Comunicación al *XVI Congreso de la Real Academia de la Lengua Vasca – Euskaltzaindia (Pamplona, 6-10 de octubre)*. En prensa.
- Igartua 2009: I. Igartua, 2009, “La aspiración de origen nasal en la evolución fonética del euskera: un caso de rhinoglottophilia”. En prensa en *ASJU*.
- Lafon 1943: R. Lafon, *Le système du verbe basque au XVIème siècle*, 1943. 2ª ed. Elkar, Donostia, 1980.
- Lafon 1950: R. Lafon, “La racine en basque”, *BAP* 6: 4, 1950, 303-308. Reed. en 1999, 287-292.
- Lafon 1951-52: R. Lafon, “Concordances morphologiques entre le basque et les langues caucasiennes”, *Word* 7, 1951, 227-244 y 8, 1952, 80-94.
- Lafon 1999: R. Lafon, *Vasconiana*, *Iker* 11, Euskaltzaindia. Bilbao 1999.
- Lakarra 1991: J. A. Lakarra, “(Reseña) Roman del Cerro, *El desciframiento de la lengua ibérica en ‘La ofrenda de los Pueblos’*”, *ASJU* 25:3, 1991, 1001-1004.
- Lakarra 1995: J. A. Lakarra, “Reconstructing the root in Pre-Proto-Basque” in Hualde, Lakarra & Trask (eds.), 189-206.
- Lakarra 1996: J. A. Lakarra, “Sobre el europeo antiguo y la reconstrucción del protovasco”, *ASJU* 30, 1996, 1-70.
- Lakarra 1997a: J. A. Lakarra, “Euskararen historia eta filologia arazo zahar, bide berri”, *ASJU* 31, 1997, 447-536.
- Lakarra 1997b: J. A. Lakarra, “Gogoetak aitzineuskararen birreraiketaz: konparaketa eta barnebirreraiketa”, *ASJU* 31, 1997, 537-616.
- Lakarra 1998a: J. A. Lakarra, “Hizkuntzalaritza konparatua eta aitzineuskararen erroa”, *Uztaro* 25, 1998, 47-110.
- Lakarra 1998b: J. A. Lakarra, “Gure izterlehengusuek eta guk erro bera? Gogoetak erroaz aitzinkartvelikoz eta aitzineuskaraz”. In I. Turrez et al. (eds.), *Studia Philologica in honorem Alfonso Irigoien*. Bilbao 1998: 125-150.
- Lakarra 1999: J. A. Lakarra, “Ná-De-Ná”, *Uztaro* 31, 1999, 15-84.
- Lakarra 2002a: J. A. Lakarra, “Etymologiae (proto)uasconicae LXV”, in Artiagoitia, Goenaga & Lakarra (eds.), 425-442.
- Lakarra 2002b: J. A. Lakarra, “Ez zirenez: \*\*TVTV eta haren lagunez: I. So bat erro disilabiko kodagabeez”. Ms., UPV/EHU 2002.
- Lakarra 2004a: J. A. Lakarra, “Etimología y reconstrucción en el campo vasco”. In E. Ridruejo (ed.), *Las otras lenguas de España*, Valladolid 2004, 41-116.
- Lakarra 2004b: J. A. Lakarra, “Bisílabos (¿proto?)vascos”. Incluido en Lakarra (en prensa).
- Lakarra 2005a: J. A. Lakarra, “Prolegómenos a la reconstrucción de segundo grado y al análisis del cambio tipológico en (proto)vasco”, *PalHispanica* 5, 2005, 407-470.

- Lakarra 2005b: J. A. Lakarra, «*Andere*: sobre regalos y costes en etimología y reconstrucción», in J. Alonso et al. (eds.), *Homenaje a Olga Omatos*, Vitoria-Gasteiz 2005.
- Lakarra 2006a: J. A. Lakarra, “Protovasco, munda y otros: reconstrucción interna y tipología holística diacrónica”, *Oihenart* 21, 2006, 229-322.
- Lakarra 2006b: J. A. Lakarra, “Notas sobre iniciales, cambio tipológico y prehistoria del verbo vasco”, in Lakarra & Hualde (eds.), 561-621.
- Lakarra 2007a: J. A. Lakarra, “Para la reconstrucción del verbo protovasco: irregularidades radicales y extensiones a la izquierda”. En prensa en *ASJU*.
- Lakarra 2007b: J. A. Lakarra, “Erro monosilabikoaren teoria eta aitzineuskararen berreraiketa: zenbait alderdi eta ondorio”. En prensa en J. Gorrochategui, J. A. Lakarra & B. Urgell (eds.), *Actas del 2º Congreso de la Cátedra L. Michelena*. Vitoria.
- Lakarra 2007c: J. A. Lakarra, “Protovasco: reconstrucción, cronología y periodización”. Conferencia de la Univ. de Deusto; ms. de la UPV/EHU 2007.
- Lakarra 2008a: J. A. Lakarra, “Aitzineuskararen gramatikanantz malkar eta osinetan zehar”. En Artiagoitia y Lakarra (eds.), 451-490.
- Lakarra 2008b: J. A. Lakarra, “Forma canónica, etimología y reconstrucción en el campo vasco”, *ASJU* 37, 2003, 261-391 [2008].
- Lakarra 2008c: J. A. Lakarra, “Aitzineuskara”. Ponencia del XVI Congreso de la Academia de la Lengua Vasca-Euskaltzaindia (Pamplona, octubre de 2008). En prensa.
- Lakarra 2009a: J. A. Lakarra, “Temas para un prólogo: forma canónica, tipología holística diacrónica y reconstrucción del protovasco”. En prensa en *Oihenart*.
- Lakarra 2009b: J. A. Lakarra, “ $*h_3 > h_1$ ,  $*h_2 > h_1$  eta horiei datzekien zenbait fenomenoz”. En prensa en las Actas del Coloquio de Bayona (C.N.R.S. dic. 2008).
- Lakarra 2009c: J. A. Lakarra, “Aitzineuskararen birreraiketaz: zergatik ezkerra?”. Discurso de entrada en la Real Academia de la Lengua Vasca-Euskaltzaindia (Arrazua, 30/V/2009).
- Lakarra en prensa: J. A. Lakarra, *Raíz y reconstrucción del protovasco*. Anejos de *ASJU* L, Donostia-San Sebastián, en prensa.
- Lakarra en prep.-1: J. A. Lakarra, “500 etimologías y subiendo. Materiales para un diccionario etimológico vasco”. Ms., UPV/EHU, en prep.-1.
- Lakarra en prep.-2: J. A. Lakarra, “Antiguas y nuevas iniciales en (proto)vasco”. Ms., UPV/EHU, en prep.-2.
- Lakarra en prep.-3: J. A. Lakarra, “Para un análisis histórico formal del léxico vasco: monosílabos comunes y bisílabos dialectales”. Ms., UPV/EHU en prep.-3.
- Lakarra en prep.-4: J. A. Lakarra, “V-, CV-, -VC, -C, -V, -CV: fragmentos y forma canónica en (proto)vasco”. Ms., UPV/EHU, en prep.-4.

- Lakarra en prep.-5: J. A. Lakarra, "100 nuevos (antiguos) préstamos latino-románicos: transformaciones estructurales y fonotáctica diacrónica". Ms., UPV/EHU en prep.-5.
- Lakarra en prep.-6: J. A. Lakarra, "Vocales y diptongos en bisílabos: análisis sincrónico y diacrónico". Ms., UPV/EHU en prep.-6.
- Lakarra & de Bernardo 2009: J. A. Lakarra & P. de Bernardo, "Forma canónica y etimología: de porqué *neska* puede y debe ser préstamo". Ms. UPV/EHU 2009.
- Lakarra & Hualde 2006: J. A. Lakarra & J. I. Hualde (eds.), *Studies in Basque and Historical Linguistics in Memory of R. L. Trask* (= *ASJU* XL, 1-2), Bilbao 2006.
- Manterola 2006: J. Manterola, "-a euskal artikulatu definituaren gainean zenbait ohar". In Lakarra & Hualde (eds.), 651-676.
- Manterola 2009: J. Manterola, "Is Basque an agglutinative language?". Conferencia de la Universidad de Santa Bárbara 2009.
- Martinet 1950: A. Martinet, "De la sonorisation des occlusives initiales en basque", *Word* 6, 1950, 224-233. Reed. en *Économie des changements phonétiques. Traité de phonologie diachronique*. Francke, Berne: 370-388. Trad. cast., Madrid, 1974.
- Martínez Areta 2003: M. Martínez Areta, "Konposatuak aitzineuskaraz", *ASJU* 37, 2003 [2008], 1-84.
- Martínez Areta 2004: M. Martínez Areta, "El acento protovasco", *ASJU* 38, 2004 [2008], 135-206.
- Martínez Areta 2006a: M. Martínez Areta, "Adjektiboa aitzineuskaraz". In Lakarra & Hualde (eds.), 687-722.
- Martínez Areta 2006b: M. Martínez Areta, *El consonantismo protovasco*. Tesis doctoral inédita. UPV/EHU 2006.
- Masini 1993: F. Masini, *The formation of Modern Chinese Lexicon and its evolution toward a national language: The period from 1840 to 1898*. Journal of Chinese Linguistics, Monograph Series n. 6, Berkeley 1993.
- Matisoff 1990: J. Matisoff, "On megalocomparison", *Lg* 66, 1990, 106-120.
- Melena 1985: J. L. Melena (ed.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae*. Vitoria-Gasteiz, 1985, 2 vols.
- Merlan 1979: F. Merlan, F., "On the prehistory of some Australian verbs", *Oceanic Linguistics* 18, 1979, 33-111.
- Michelena 1950: L. Michelena, "De etimología vasca". Reed. en *SHLV*, 439-444.
- Michelena 1951a: L. Michelena, "De fonética vasca. La distribución de las oclusivas aspiradas y no aspiradas". Reed. en *SHLV*, 212-219.
- Michelena 1951b: L. Michelena, "La sonorización de las oclusivas iniciales. A propósito de un importante artículo de André Martinet". Reed., *SHLV*, 203-211.
- Michelena 1954: L. Michelena, "De onomástica aquitana". Reed., *LH*, 409-445.

- Michelena 1956: L. Michelena, "La lengua vasca como medio de conocimiento histórico", *Zumarraga* 6, 1956, 49-70.
- Michelena 1957a: L. Michelena, "Las antiguas consonantes vascas". Reed., *SHLV*, 166-189.
- Michelena 1957b: L. Michelena, "Basque et roman". Reed., *SHLV*, 106-115.
- Michelena 1963: L. Michelena, *Lenguas y protolenguas*, 1963. Reed., Anejos de *ASJU* 20, Donostia-San Sebastián 1990.
- Michelena 1964: L. Michelena, *Sobre el pasado de la lengua vasca*, 1964 [Reed., *SHLV*, 1-73].
- Michelena 1970: L. Michelena, "Nombre y verbo en la etimología vasca". Reed., *PT*, 283-309.
- Michelena 1971: L. Michelena, "Toponimia, léxico y gramática". Reed., *PT*, 141-167.
- Michelena 1974: L. Michelena, "El elemento latino-románico en la lengua vasca". Reed., *PT*, 195-219.
- Michelena 1977a: L. Michelena, *Fonética histórica vasca*, 2ª ed. corregida y aumentada [1ª ed. 1961], Anejos del *ASJU* 4, Donostia-San Sebastián 1977.
- Michelena 1977b: L. Michelena, "Notas sobre compuestos verbales vascos". Reed., *PT*, 311-335.
- Michelena 1979: L. Michelena, "La langue ibère". Reed., *LH*, 341-356.
- Michelena 1981: L. Michelena, "Lengua común y dialectos vascos". Reed., *PT*, 35-55.
- Michelena 1985: L. Michelena, *Lengua e Historia*, [*LH*], Paraninfo, Madrid 1985.
- Michelena 1987a: L. Michelena, *Palabras y Textos*, [*PT*], J. Gorrochategui (ed.), Bilbao 1987.
- Michelena 1987b: L. Michelena, *Orotariko Euskal Hiztegia-Diccionario general vasco*. Bilbao 1987.
- Michelena 1988: L. Michelena, *Sobre historia de la lengua vasca*, J. A. Lakarra (ed.), Anejos de *ASJU* 10, Donostia-San Sebastián 1988, 2 vols.
- Ozanne-Rivierre 1995: F. Ozanne-Rivierre, "Structural changes in the languages of Northern New Caledonia", *Oceanic Linguistics* 34, 1995, 44-72.
- Ozanne-Rivierre & Rivierre 2004: F. Ozanne-Rivierre & J.-C. Rivierre, "Évolution des formes canoniques dans les langues de Nouvelle-Calédonie" in E. Zeitoun (ed.), *Les langues austronésiennes (= Faits de Langues 23-24)*, Paris 2004: 141-153.
- Packard 1998: J. L. Packard, (ed.), *New approaches to Chinese Word formation: morphology, phonology and the lexicon in modern and ancien Chinese*. Berlin 1998.
- Pawley 2006: A. Pawley, "Where have all the verbs gone? Remarks on the organisation of languages with small, closed verb classes". 11<sup>th</sup> Biennial Rice University Linguistic Symposium, 16-18 March 2006.

- Plank 1998: F. Plank, "The co-variation of phonology with morphology and syntax: A hopeful history", *Linguistic Typology* 2, 1998, 195-230.
- Post 2006: M. W. Post, "Compounding and the structure of the Tani lexicon", *Linguistics in the Tibeto-Burman Area* 29, 2006, 41-60.
- Post 2007a: M. W. Post, "Grammaticalization and compounding in Thai and Chinese", *Studies in Language* 31, 2007, 117-175.
- Post 2007b: M. W. Post, "The phonology and grammar of Galo "words"". Ms., Research Centre for Linguistic Typology. La Trobe University 2007.
- Potet 1995: J.-P.-G. Potet, "Tagalog monosyllabic roots", *Oceanic Linguistics* 34, 1995, 345-374.
- Reid 2003: N. Reid, "Phrasal verb to synthetic verb: recorded morphosyntactic change in Ngan'gityemerri" in N. Evans (eds.), *The Non-Pama-Nyungan languages of Northern Australia. Comparative studies of the continent's most linguistically complex region*. Pacific Linguistics 555, The Australian National University 2003: 95-123.
- de Rijk 1969: R. P. G. de Rijk, "Is Basque a SOV language?". Reed. in 1998, 13-38.
- de Rijk 1992: R. P. G. de Rijk, "'Nunc" Vasconice". Reed. in 1998, 347-376.
- de Rijk 1998: R. P. G. de Rijk, *De lingua uasconum. Selected writings*. Anejos de ASJU, Donostia-San Sebastián 1998.
- Sagart 1999: L. Sagart, *The roots of Old Chinese*. Amsterdam – Philadelphia 1999.
- Sagart & Xu 2001: L. Sagart & Sh. Xu, "History through loanwords: the loan correspondences between Hani and Chinese", *Cahiers Linguistique de l'Asie Orientale* 30, 2001, 3-54.
- Sarasola 1984-95: I. Sarasola, *Hauta-Lanerako Euskal Hiztegia*, Gipuzkoako Kutxa, Donostia 1984-1995.
- Sarasola 1997: I. Sarasola, "Euskal hitz altxorraz", *ASJU* 31, 1997, 617-642.
- Schuchardt 1906: H. Schuchardt, "Baskisch und Romanisch". Trad. de A. Goenaga ("Vascuence y romance"), *BAP* 13 (1957), 463-487, 15 (1959), 181-205, 16 (1960), 339-363.
- Schuchardt 1972: H. Schuchardt, "Sobre la formación de las flexiones de relación del verbo vasco", *BAP* 28, 1972, 217-337.
- Segura & Etxebarria 1996: S. Segura & J. M. Etxebarria, *Del latín al euskara. Latinetik euskarara*. Bilbao 1996.
- Sneddon 1993: J. N. Sneddon, "The drift towards final open syllables in Sulawesi languages", *Oceanic Linguistics* 32, 1993, 1-44.
- Tovar 1959: A. Tovar, *El euskera y sus parientes*, Madrid 1959.
- Tovar 1997: A. Tovar, *Estudios de tipología lingüística. Sobre el euskera, el español y otras lenguas del Viejo y el Nuevo Mundo*. J. Bustamante (ed.), Madrid 1997.
- Trask 1977: L. R. Trask, "Historical Syntax and Basque Verbal Morphology: Two Hypotheses", in W. Douglass et al. (eds.), *Anglo-American*

- contributions to Basque studies: Essays in honor of Jon Bilbao*. Reno 1977, 203-217.
- Trask 1985: L. R. Trask, "On the reconstruction of Pre-Basque Phonology", in Melena (ed.), II, 885-891.
- Trask 1997: L. R. Trask, *The history of Basque*. Londres 1997.
- Trask 1998: L. R. Trask, "The typological position of Basque: then and now", *Language Sciences* 20, 1998, 313-324.
- Uhlenbeck 1947: C. C. Uhlenbeck, "Les couches anciennes du vocabulaire basque", *Eusko Jakintza* 1, 1947, 543-581. Original holandés de 1942.
- Ussishkin 2009: A. Ussishkin, "reseña de Downing 2006", *Phonology* 26, 2009, 363-366.
- Vennemann 1994: Th. Vennemann, "Linguistic reconstruction in the context of European Prehistory", *TPS* 92, 1994, 215-284.
- Vennemann 2003: Th. Vennemann, *Europa Vasconica – Europa Semitica*. NY-Berlin 2003.
- Villar & Encarnação 1996: F. Villar & J. D. Encarnação (eds.), 1996, *La Hispania prerromana. VI Coloquio sobre lenguas y culturas prerromanas de la Península Ibérica*. U. Salamanca & U. Coimbra. Salamanca 1996.
- Villar & Fdez Alvarez 2001: F. Villar & M<sup>a</sup> P. Fdez Alvarez (eds.), *Religión, lengua y cultura prerromanas de Hispania*. Universidad, Salamanca 2001.
- Vovin 1993: A. Vovin, *A reconstruction of Proto-Ainu*. Brill: Leiden, etc. 1993.
- Vovin 1994: A. Vovin, "Genetic affiliation of Japanese and methodology of linguistic comparison", *Journal de la Société Finno-Ougrienne* 85, 1994, 241-256.
- Vovin 2009: A. Vovin, "Japanese, korean and other "Non-Altai" languages", *Central Asiatic Journal* 53/1, 2009, 105-147.
- Watkins 1984: C. Watkins, "L'apport d'E. Benveniste à la grammaire comparée" in G. Serbat (ed.), *E. Benveniste aujourd'hui. Actes du Colloque international du CNRS*, Peeters, Louvain 1984, I, 3-11.
- Watkins 1991: C. Watkins, "Etymologies, equations, and comparanda: types and values, and criteria for judgment". In Baldi (ed.), 289-304.
- Williamson 2004: K. Williamson, "Typical vowel systems and processes in West African Niger-Congo languages", *Journal of West African Languages* 30, 2004, pp. 128-142.
- Zorc 1991: R. D. Zorc, "The Austronesian monosyllabic root, radical or phonestheme" in Baldi (ed.), 175-194.

Joseba A. Lakarra  
Universidad del País Vasco – Euskal Herriko Unibertsitatea  
e-mail: joseba.lakarra@ehu.es

APÉNDICE<sup>73</sup>

Tabla A

Mod.rad.	R.Pos.	Doc.	%	ED	%-1	%-2	Fós.	%-1	%-2
CVC	325	<b>152</b>	46,76	70	46,05	21,53	48	31,18	14,76
TVRV	600	<b>195</b>	32,50	25	13,07	04,33	04	02,05	00,66
TVT	900	<b>190</b>	21,11	16	08,42	01,77	00	00,00	00,00
TVSV	800	<b>115</b>	14,62	24	20,86	03,00	04	03,47	00,50
bVCVC	2125	<b>115</b>	05,40	25	21,73	01,17	07	06,08	00,32
aCVC	425	<b>110</b>	25,88	33	30,00	07,76	15	13,63	03,52
zVCVC	2125	<b>106</b>	04,98	34	32,07	01,60	13	12,26	00,61
bVCCV	1075	<b>96</b>	08,93	19	19,79	01,76	03	03,12	00,27
TVSV	600	<b>95</b>	15,85	14	14,73	02,33	03	03,15	00,50
mVCV	425	<b>94</b>	22,10	12	12,76	02,82	00	00,00	00,00
zVCCV	1075	<b>92</b>	08,55	12	13,04	01,11	02	02,06	00,18
aCCVC	1075	<b>90</b>	08,37	15	16,66	01,39	01	01,11	00,09
gVCCV	1075	<b>84</b>	07,81	06	07,14	00,55	02	02,38	00,18
aCCV	215	<b>82</b>	38,13	12	14,64	05,58	04	04,87	01,85
zVCCVC	5375	<b>80</b>	01,48	10	12,50	00,01	02	02,25	00,03
bVCCVC	5375	<b>77</b>	01,43	10	12,98	00,18	02	02,59	00,03
SVTV	300	<b>75</b>	25,00	13	17,33	04,33	04	05,33	01,33
SVRV	200	<b>71</b>	35,50	05	07,04	02,50	01	01,40	00,50
RVT	300	<b>69</b>	23,00	09	13,04	03,00	01	01,44	00,03
gVCVC	2125	<b>68</b>	03,20	13	19,11	00,61	06	08,82	00,28
IVCVC	2125	<b>66</b>	03,10	33	50,00	01,50	13	19,59	00,61
gVCCVC	5375	<b>65</b>	01,20	02	03,07	00,03	00	00,00	00,00
aCV	85	<b>64</b>	75,29	04	06,66	04,70	02	03,12	02,35
sVCCV	1075	<b>63</b>	05,86	10	15,87	00,93	04	06,34	00,37
sVCVC	2125	<b>62</b>	02,91	14	22,58	00,65	03	04,83	00,14
hVCCV	1075	<b>60</b>	05,58	09	15,00	00,83	02	03,33	00,18
hVCV	425	<b>54</b>	12,47	18	33,33	04,23	02	03,70	00,47
uCVC	425	<b>54</b>	12,70	05	09,25	01,17	01	01,85	00,23
IVCCV	1075	<b>53</b>	04,93	09	16,98	00,83	00	00,00	00,00
CVmV	325	<b>49</b>	15,07	09	18,35	02,76	00	00,00	00,00
sVCCVC	5375	<b>48</b>	00,89	11	22,91	00,20	00	00,00	00,00
oCVC	425	<b>47</b>	11,05	07	14,59	01,64	05	10,63	01,17
uCV	85	<b>44</b>	51,75	05	11,36	05,83	02	04,54	02,35
uCCV	215	<b>45</b>	20,93	06	13,33	02,78	00	00,00	00,00
RVRV	200	<b>41</b>	20,50	02	04,87	01,00	00	00,00	00,00
oCCV	215	<b>39</b>	18,13	08	20,53	03,72	06	15,37	02,79
hVCVC	2125	<b>31</b>	01,45	09	29,32	00,42	08	25,80	00,37
IVCCVC	5375	<b>31</b>	00,57	06	19,35	00,11	03	09,67	00,05

<sup>73</sup> Para abreviaturas y siglas véase n. 63b; en cada caso los modelos radicales están ordenados en función de la columna en negrita.

*Forma canónica y cambios en la forma canónica de la lengua vasca...*

fVCV	425	<b>30</b>	07,05	06	20,00	01,41	00	00,00	00,00
CVhV	300	<b>29</b>	09,66	04	13,79	01,33	02	06,89	00,66
oCV	85	<b>29</b>	34,11	07	24,13	08,23	06	20,68	07,05
oCCVC	1075	<b>25</b>	02,32	01	04,00	00,09	00	00,00	00,00
nVCVC	2125	<b>24</b>	01,12	08	33,33	00,37	04	16,66	00,18
RVSV	200	<b>24</b>	12,00	00	00,00	00,00	00	00,00	00,00
hVCCVC	5375	<b>23</b>	00,42	00	00,00	00,00	00	00,00	00,00
uCCVC	1075	<b>23</b>	02,13	07	30,43	00,65	03	13,04	00,27
SVSV	200	<b>13</b>	06,50	01	07,69	00,50	01	07,69	00,50
CVfV	325	<b>10</b>	03,07	01	10,00	00,33	00	00,00	00,00
nVCCV	1075	<b>09</b>	08,40	01	11,11	00,09	00	00,00	00,00
nVCCVC	5375	<b>07</b>	00,13	00	00,00	00,00	00	00,00	00,00

**Tabla B**

Mod.rad.	R.Pos.	Doc.	%	Prést.	%-1	%-2	Onom.	%-1	%-2
CVC	325	152	46,76	<b>12</b>	07,89	03,69	30	19,73	09,23
uCCVC	1075	23	02,13	<b>00</b>	00,00	00,00	00	00,00	00,00
nVCCVC	5375	07	00,13	<b>00</b>	00,00	00,00	02	28,57	00,03
SVSV	200	13	06,50	<b>00</b>	00,00	00,00	02	15,38	01,00
oCCVC	1075	25	02,32	<b>00</b>	00,00	00,00	00	00,00	00,00
nVCCV	1075	09	08,40	<b>02</b>	22,22	00,18	00	00,00	00,00
hVCCVC	5375	23	00,42	<b>02</b>	08,69	00,03	00	00,00	00,00
nVCVC	2125	24	01,12	<b>03</b>	12,50	00,14	00	00,00	00,00
hVCVC	2125	31	01,45	<b>03</b>	09,67	00,14	00	00,00	00,00
CVhV	300	29	09,66	<b>05</b>	17,24	01,66	00	00,00	00,00
uCVC	425	54	12,70	<b>05</b>	09,25	01,17	01	01,85	00,23
zVCCVC	5375	80	01,48	<b>05</b>	06,25	00,09	08	10,00	00,14
lVCCVC	5375	31	00,57	<b>06</b>	19,35	00,11	00	00,00	00,00
CVfV	325	10	03,07	<b>06</b>	60,00	01,84	03	30,00	00,92
oCVC	425	47	11,05	<b>06</b>	12,76	01,41	01	02,12	00,23
zVCVC	2125	106	04,98	<b>07</b>	06,60	00,32	06	05,66	00,28
oCV	85	29	34,11	<b>07</b>	24,13	08,23	03	10,34	03,52
hVCV	425	54	12,47	<b>07</b>	12,96	01,64	04	09,25	00,94
uCV	85	44	51,75	<b>07</b>	15,90	08,23	08	18,18	09,41
uCCV	215	45	20,93	<b>08</b>	17,77	03,72	03	06,66	01,39
sVCCVC	5375	48	00,89	<b>09</b>	18,75	00,16	01	02,08	00,89
sVCVC	2125	62	02,91	<b>10</b>	16,12	00,47	00	00,00	00,00
oCCV	215	39	18,13	<b>10</b>	25,63	04,65	00	00,00	00,00
gVCVC	2125	68	03,20	<b>10</b>	14,70	00,47	00	00,00	00,00
gVCCVC	5375	65	01,20	<b>11</b>	16,92	00,20	05	07,69	00,09
lVCVC	2125	66	03,10	<b>12</b>	18,18	00,56	00	00,00	00,00
hVCCV	1075	60	05,58	<b>13</b>	21,66	01,20	00	00,00	00,00
RVRV	200	41	20,50	<b>14</b>	34,14	07,00	06	14,63	03,00
aCV	85	64	75,29	<b>14</b>	21,87	16,46	10	15,62	11,70

RVSV	200	24	12,00	<b>15</b>	62,50	07,50	00	00,00	00,00
aCCV	215	82	38,13	<b>15</b>	18,30	06,97	03	03,65	01,39
fVCV	425	30	07,05	<b>15</b>	50,00	03,52	04	13,33	00,91
bVCVC	2125	115	05,40	<b>15</b>	13,04	00,70	04	03,47	00,18
bVCCVC	5375	77	01,43	<b>16</b>	20,77	00,29	09	11,68	00,16
aCVC	425	110	25,85	<b>16</b>	14,54	03,76	06	05,45	01,41
sVCCV	1075	63	05,86	<b>19</b>	30,14	01,76	03	04,76	00,27
lVCCV	1075	53	04,93	<b>19</b>	35,84	01,76	01	01,88	00,09
aCCVC	1075	90	08,37	<b>20</b>	22,22	01,86	03	03,33	00,27
CVmV	325	49	15,07	<b>22</b>	44,88	06,77	07	14,24	02,15
SVRV	200	71	35,50	<b>27</b>	38,02	13,50	03	04,22	01,50
SVTV	300	75	25,00	<b>28</b>	37,33	09,33	03	04,00	01,00
zVCCV	1075	92	08,55	<b>30</b>	32,50	02,79	19	20,65	10,76
gVCCV	1075	84	07,81	<b>35</b>	41,66	03,25	00	00,00	00,00
RVTV	300	69	23,00	<b>36</b>	52,17	12,00	00	00,00	00,00
bVCCV	1075	96	08,93	<b>43</b>	44,78	04,00	02	02,08	00,18
mVCV	425	94	22,10	<b>47</b>	50,00	11,05	16	17,02	03,76
TVSV	600	95	15,85	<b>47</b>	49,47	07,83	02	02,10	00,33
TVT	900	190	21,11	<b>80</b>	42,10	08,88	36	18,94	04,00
TVRV	600	195	32,50	<b>95</b>	48,71	15,83	16	08,20	02,66

**Tabla C**

Mod.rad.	R.Pos.	Doc.	%	<b>C/D</b>	%-1	%-2	Var.	%-1	%-2
CVC	325	152	46,76	<b>04</b>	02,63	01,23	36	23,68	11,07
bVCVC	2125	115	05,40	<b>38</b>	33,04	01,78	33	28,69	01,55
hVCCV	1075	60	05,58	<b>36</b>	60,00	03,34	02	03,33	00,18
zVCVC	2125	106	04,98	<b>34</b>	32,07	01,60	25	23,58	01,17
aCCVC	1075	90	08,37	<b>32</b>	35,55	02,97	20	22,22	02,04
bVCCVC	5375	77	01,43	<b>30</b>	38,96	00,55	12	15,58	00,22
TVT	900	190	21,11	<b>24</b>	12,63	21,33	35	18,42	03,88
gVCVC	2125	68	03,20	<b>24</b>	35,29	01,12	21	30,88	00,98
gVCCVC	5375	65	01,20	<b>22</b>	33,84	00,42	25	38,46	00,46
aCCV	215	82	38,13	<b>22</b>	26,84	10,22	30	36,58	13,95
TVRV	600	195	32,50	<b>21</b>	10,76	03,50	37	18,97	06,16
zVCCVC	5375	80	01,48	<b>21</b>	26,25	00,39	36	45,00	00,66
oCCVC	1075	25	02,32	<b>21</b>	84,00	01,86	03	12,00	00,27
oCVC	425	47	11,05	<b>21</b>	44,68	04,94	12	25,52	02,82
hVCV	425	54	12,47	<b>20</b>	37,36	04,70	06	11,11	01,41
sVCVC	2125	62	02,91	<b>20</b>	32,25	00,94	17	27,41	00,80
uCVC	425	54	12,70	<b>20</b>	37,03	04,70	23	42,59	05,41
hVCCVC	5375	23	00,42	<b>19</b>	82,60	00,42	02	08,69	00,03
aCVC	425	110	25,85	<b>18</b>	16,36	04,23	37	33,63	08,70
gVCCV	1075	84	07,81	<b>17</b>	20,23	01,58	26	30,95	02,41
hVCVC	2125	31	01,45	<b>17</b>	54,83	00,80	02	06,45	00,09

*Forma canónica y cambios en la forma canónica de la lengua vasca...*

SVRV	200	71	35,50	<b>14</b>	19,71	07,00	22	30,98	11,00
lVCVC	2125	66	03,10	<b>13</b>	19,69	00,61	08	12,12	00,37
bVCCV	1075	96	08,93	<b>13</b>	13,54	01,20	43	44,78	04,00
RVTV	300	69	23,00	<b>12</b>	17,39	04,00	12	17,39	04,00
uCCV	215	45	20,93	<b>12</b>	28,57	05,58	16	38,09	07,44
oCCV	215	39	18,13	<b>12</b>	07,68	01,39	09	23,07	04,18
lVCCVC	5375	31	00,57	<b>11</b>	35,16	00,20	08	25,90	00,14
sVCCVC	5375	48	00,89	<b>11</b>	21,91	00,20	16	33,33	00,07
lVCCV	1075	53	04,93	<b>11</b>	20,75	01,02	13	24,52	01,21
SVTV	300	75	25,00	<b>10</b>	13,33	03,33	21	28,00	07,00
CVhV	300	29	09,66	<b>10</b>	03,44	03,33	10	03,44	03,33
zVCCV	1075	92	08,55	<b>10</b>	01,08	00,09	21	22,82	01,95
TVSV	600	95	15,85	<b>10</b>	10,52	01,16	22	23,15	03,66
uCCVC	1075	23	02,13	<b>08</b>	34,78	00,74	08	34,78	00,74
aCV	85	64	75,29	<b>07</b>	10,93	08,23	29	45,31	34,11
nVCVC	2125	24	01,12	<b>06</b>	25,00	00,28	07	29,16	00,32
RVRV	200	41	20,50	<b>06</b>	14,66	03,00	13	31,70	06,50
SVSV	200	13	06,50	<b>05</b>	38,46	02,50	05	38,46	02,50
oCV	85	29	34,11	<b>05</b>	17,24	05,88	07	24,13	08,23
RVSV	200	24	12,00	<b>05</b>	20,87	02,50	04	16,66	02,00
nVCCVC	5375	07	00,13	<b>04</b>	57,14	00,07	01	14,28	00,01
nVCCV	1075	09	08,40	<b>03</b>	33,33	00,27	03	33,33	00,27
mVCV	425	94	22,10	<b>03</b>	03,19	00,70	16	17,02	03,76
uCV	85	44	51,75	<b>03</b>	06,81	03,52	21	47,72	24,70
CVmV	325	49	15,07	<b>02</b>	04,08	00,61	09	18,36	02,76
fVCV	425	30	07,05	<b>00</b>	00,00	00,00	05	16,66	01,17
CVfV	325	10	03,07	<b>00</b>	00,00	00,00	00	00,00	00,00