

**Más allá de la lingüística indoeuropea:  
en busca de un nuevo marco conceptual para entender el vascuence<sup>1</sup>**

Pedro Urquijo Arregui

A finales del siglo XVIII figuras como el jesuita español Lorenzo Hervás y Panduro, el danés Rasmusch Christian Rask o los alemanes Johann Christoph Adelung y Wilhelm von Humboldt sentaron las bases del comparativismo lingüístico y, un poco después, el también alemán Franz Bopp, en el seno de la universidad de Berlín, formuló la teoría del indoeuropeo o indogermánico.

En términos generales, los conceptos principales de la lingüística indoeuropea siguen vigentes para una gran parte del mundo académico y forma parte de lo que se considera que una persona de cultura general debe conocer. La prueba es que en todos los libros de texto de enseñanza secundaria que tocan este punto, especialmente los de Lengua Castellana y los de las otras lenguas vernáculas de España, se explica dicha teoría sin ningún matiz que permita ponerla en duda, como si se tratara del proceso de fotosíntesis de las plantas o la tabla periódica de los elementos químicos.

Lo que se cuenta a nuestros alumnos de la ESO y Bachillerato viene a ser, de forma simplificada, lo siguiente:

Érase hace unos 6000 o 7000 años un pueblo de pastores que habitaba en las estepas que median entre el mar Negro y el mar Caspio, cerca de la desembocadura de los ríos Don y Volga, un territorio que en gran parte está actualmente en sur de Rusia, al este de Ucrania. Por alguna razón, esta etnia comenzó a expandirse en todas direcciones. De forma progresiva marcharon hacia el norte por todo el territorio que queda al oeste de los montes Urales, avanzaron por Asia Central, bajaron al sur de la cordillera del Cáucaso, se adentraron en la actual Irán y Afganistán e incluso Pakistán y el norte de la India, penetraron en Anatolia y siguieron por el oeste hacia los Balcanes, Europa Central y la zona del Báltico, se establecieron incluso en Escandinavia y, naturalmente, llegaron a la parte más occidental de nuestro continente. Se hace notar que no se sintieron atraídos por Mesopotamia y el resto del llamado creciente fértil de Oriente Medio, y, del litoral mediterráneo, ocuparon la vertiente europea pero desdeñaron todo el norte de África. Es por esto por lo que se les puso el nombre de indoeuropeos.

Para explicar cómo lograron establecer su poder en tan vasto territorio se aduce que, a diferencia de todos los demás pueblos, habían logrado la doma y monta del caballo y poseían avances técnicos como la rueda y el carro, utilizado este como arma de guerra a la manera de las cuadrigas romanas. Eso les habría dado una gran ventaja a la hora de plantar batalla a todos los pueblos que iban invadiendo. No se conoce de ellos, sin embargo, nada parecido a otros factores que, además

---

<sup>1</sup> Se trata de una ponencia que en la IX Jornadas llevó el título de un artículo del mismo autor publicado en 2015: «Revisión de la genealogía de las lenguas a partir del vascuence: una visión crítica de la lingüística indoeuropea». [https://www.academia.edu/9968112/Revisi%C3%B3n\\_de\\_la\\_genealog%C3%ADa\\_de\\_las\\_lenguas\\_indoeuropeas\\_a\\_partir\\_del\\_vascuence](https://www.academia.edu/9968112/Revisi%C3%B3n_de_la_genealog%C3%ADa_de_las_lenguas_indoeuropeas_a_partir_del_vascuence)

de la posesión de un ejército organizado en centurias y cohortes, permitieron a Roma convertirse en un imperio: el derecho, la ingeniería hidráulica, una red de calzadas, el desarrollo de las técnicas agrícolas, la arquitectura civil, el desarrollo urbano o una tradición literaria escrita. El único vestigio tangible de aquella supuesta civilización esteparia, luego indoeuropea, son los llamados *kurganes*, unos túmulos bajo los cuales eran sepultadas en cámaras funerarias de madera los miembros de la élite dirigente. Estos y poco más son los restos visibles del pueblo que logró domeñar a la mitad del mundo conocido entre el quinto y el tercer milenio a. de C.

Por otra parte, no está nada claro a qué se debió el formidable excedente de población experimentado en las estepas que circundan la actual ciudad de Volgogrado (Stalingrado en tiempos de la URSS; más allá, no consta que, antes de su fundación en el siglo XVI con el nombre de Tsaritsin, hubiera allí o en las inmediaciones asentamientos urbanos), ni se ha explicado de forma convincente cómo un pueblo pastoril consiguió reunir hordas tan numerosas como para imponerse a casi toda la población de un continente entero y parte de otro.

En cuanto a su lengua, no se conserva ni una sola palabra; no se sabe siquiera si conocían la escritura, pese a lo cual es éste el elemento cultural que dio pie a toda la construcción de la teoría del indoeuropeo. Las similitudes observadas en la mayor parte de las lenguas conocidas del territorio antes descrito condujeron, por la vía de un razonamiento inductivo, a la conclusión de que todas ellas tenían un origen común. A esa lengua originaria se la denominó indoeuropeo. Se conoce como protoindoeuropeo al conjunto de formas lingüísticas (señaladas con asterisco\*) reconstruidas a partir de la descendencia dejada en las lenguas adscritas dentro de la familia de lenguas indoeuropeas.

Esta es, a grandes rasgos, la teoría que me propongo poner en cuestión. Y en este empeño no estoy solo. En las últimas décadas bastantes investigadores han propuesto hipótesis muy divergentes respecto de la ortodoxia aceptada durante casi dos centurias (vigencia extrañamente duradera en cualquier disciplina sometida al escrutinio científico).

Así, por ejemplo, el británico Colin Renfrew sugirió que la generalización del indoeuropeo se produjo de la mano de la revolución neolítica iniciada en la región de Anatolia. Lo más significativo de su propuesta no fue el desplazamiento hacia el sur del foco originario, sino la afirmación de que no fue el efecto de oleadas de población invasora ni de conquistas militares, sino el resultado de una irradiación cultural.

Más audaces que Renfrew, surgen los defensores de la teoría conocida como “continuidad paleolítica”. El italiano Mario Alinei o, en nuestro país, Xaverio Ballester, de la Universidad de Valencia, son algunos de sus representantes. Según ellos, las lenguas llamadas indoeuropeas, junto con otras a las que se les ha negado tradicionalmente esa condición, forman un continuo que se remonta a épocas muy anteriores al Neolítico. En ese sentido, consideran que todas las lenguas están emparentadas, pues tienen nexo originario mucho más atrás en el tiempo. Esta teoría refuerza la posibilidad de una lengua primigenia para toda la humanidad.

En la misma línea, el norteamericano Joseph H. Greenberg alentó una revolución en la clasificación de las lenguas del mundo e introdujo el término “nostrático” para referirse a un estadio de la evolución de las lenguas del mundo muy anterior al del supuesto indoeuropeo. Y su compatriota Merritt Ruhlen, en la misma Universidad de Stanford, propone también ampliar el

contexto de la clasificación de las lenguas a todas las del mundo y, sin llegar a reivindicar la validez del relato bíblico de Babel, defiende, eso sí, la hipótesis monogenética del lenguaje humano.

Más cerca de nosotros, en la misma línea de investigación de las relaciones entre el vasco y el armenio que siguió el profesor Vahan Sarkisian, de la Universidad de Ereván, tenemos las aportaciones de Antonio Arnaiz Villena y Jorge Alonso García, que relacionan las lenguas prerromanas de la península ibérica con las caucásicas, anatolias y hasta semíticas, en un marco interpretativo mucho más amplio que el de la lingüística tradicional. También son muy estimables las agudas observaciones de José Manuel Pedrosa, de la Universidad de Alcalá de Henares, sobre la rigidez de algunos paradigmas que nublan nuestra comprensión de la manera en que se relacionan unas lenguas con otras, en referencia a la tipología defendida por la lingüística indoeuropea.

Lo que tienen en común los investigadores mencionados anteriormente es, por un lado, la ampliación en el espacio y en el tiempo de su ámbito de estudio y, consiguientemente, del número de lenguas objeto de observación; por otro lado, son muy conscientes de que el esclarecimiento de la cuestión lingüística debe hacerse desde una perspectiva interdisciplinar, con el auxilio de la antropología, la arqueología, las ciencias experimentales (muy especialmente la genética) y la matemática computacional. Junto a ellos, pero no enfrentados (presentan algunos planteamientos coincidentes), sino simplemente alineados en la lingüística indoeuropea tradicional, tenemos a autores como André Martinet, catedrático de Lingüística General en la Sorbona, presidente de la Sociedad Europea de Lingüística y conspicuo baluarte del relato de las invasiones indoeuropeas; a Francisco Rodríguez Adrados, quien fuera catedrático de la UCM y miembro de varias academias; o a Joaquín Gorrochategui, destacado discípulo de Luis Michelena, actualmente director del Instituto de Estudios Clásicos de la UPV y miembro de Euskaltzaindia. Son estos últimos, ejemplos representativos entre otros muchos, los que, en virtud del indiscutible mérito de su trabajo como investigadores, gozan de todo el prestigio. No en vano, son sus puntos de vista, lógicamente, los adoptados en los manuales que manejan nuestros alumnos de secundaria.

Un dato interesante relacionado con la elaboración del comparativismo germanista de los siglos XVIII y XIX es el papel que jugó el idioma vascuence o euskera. Nos atrevemos a decir que fue utilizado como piedra angular de la bóveda. La edificación de la teoría del indoeuropeo, cuyos precursores fueron vascoiberistas como Hervás o Humboldt, se basa en una premisa: para poder explicar la presencia en Europa y Asia central de las lenguas indoeuropeas, como resultado de la expansión de los pastores de la Estepa del Sal (Óblast de Rostov), es preciso previamente demostrar que, antes de su supuesta irrupción, había en las regiones invadidas lenguas a las que podamos llamar preindoeuropeas. En esta categoría estaban, además del vascuence, los idiomas que, necesariamente, hablarían los iberos, los tartésicos, los etruscos, los urritas-urartianos, los elamitas o los veddas. Lo que pasa es que, en aquel momento (faltaba un siglo y medio para que Manuel Gómez Moreno descifrara la escritura del ibero), la única lengua cumplidamente conocida de todas ellas era el euskera, la prueba viviente de que hubo un antes y un después de las invasiones. Si el vascuence hubiera sido una mera referencia a una lengua desaparecida, cabría dudar de que dicha lengua fuera diferente en su tipología de las llamadas indoeuropeas y, por tanto, cabría pensar que la lengua originaria de todas ellas, el protoindoeuropeo, no llegó de lejanas estepas, sino que era, lisa y llanamente, la lengua que hablaban varios miles de años atrás los pobladores autóctonos del continente europeo, de Asia Central y del subcontinente de la India.

En suma, el descubrimiento de las similitudes entre todas esas lenguas habría podido conducir a una explicación no muy diferente de la teoría de la continuidad paleolítica e incluso habría podido concluir en la postulación de una protolengua definible de forma parecida o casi igual a la del nostrático. ¿Por qué no sucedió así? Porque todos los indogermanistas, alimentados por los apologistas vascos (Astarloa, Moguel) y tomando como referencia a Larramendi y a Lorenzo Hervás, se aferraron al dogma imperante todavía hoy en día, dentro y fuera del País Vasco: el euskera no es una lengua indoeuropea, es una lengua sin parentesco alguno, es una lengua aislada. Consecuentemente, basándose en ese axioma, concluyeron que el vascuence es como el último islote no sumergido que aún aflora tras la inundación indoeuropea.

Un fenómeno de conservación como el del vascuence no produce tanta extrañeza en lugares como las montañas de Siberia oriental, las profundidades de la selva amazónica o una isla del Pacífico, pongamos Samoa o Vanuatu. En nuestro caso, sin embargo, resulta inexplicable, como desconcertante es leer en muchos libros, que, a la llegada de los pueblos indoeuropeos, el euskera busco refugio en las *inhóspitas e impenetrables* montañas del País Vasco. No creo necesario explicar que el trayecto que va de las planicies de Aquitania a las lomas de Pamplona, y de ahí a las llanuras de La Rioja y al valle del Ebro no entraña más dificultad que una excursión dominguera. En cuanto a las agrestes montañas, hay que adentrarse bastante en los Pirineos para encontrar parajes de difícil acceso. Un simple vistazo al mapa de Europa nos muestra que los territorios que conforman actualmente el País Vasco ha sido un cómodo lugar de paso para pueblos de ambas vertientes de la cordillera pirenaica. Y, en cualquier caso, ¿por qué otras lenguas preindoeuropeas no se refugiaron también en la numerosas cordilleras y parajes escarpados de los que tanto abundan en la península (y fuera de ella, en otras zonas del mismo continente) y que, en muchos casos, son bastante más inaccesibles que los amables entornos donde se encuentran los caseríos vascos?

Esta incongruencia nos invita a cuestionar la teoría y a pensar en hipótesis alternativas: una de ellas podría ser que el euskera es la lengua de un pueblo que llegó a nuestro territorio después de concluidas las invasiones indoeuropeas, como pudo ocurrir en el norte de Escandinavia con del finés, procedente de Asia. Se trata de una teoría que goza de algunos adeptos, pero, de momento, prefiero dejarla aparcada. Otra hipótesis, que me resulta más verosímil, es que esas invasiones nunca se produjeron o no, al menos, a la manera en que nos ha contado. Es posible, incluso, que el euskera haya experimentado un proceso no muy diferente al de las lenguas llamadas indoeuropeas. De hecho hay quien sostiene que el euskera no deja de ser una lengua indoeuropea. Es el caso de Florentino Castro Guisasaola, quien en 1944 publicó en la Revista de Filología Española un extenso artículo para intentar demostrarlo. No podría suscribir por completo su afirmación, puesto que niego o, al menos, pongo en duda la mayor: ¿existe realmente una familia de lenguas indoeuropeas cuya génesis es distinta del origen de las demás lenguas del mundo? Sin embargo, comparto con él la idea de que los rasgos morfosintácticos y léxicos con los que se quiere definir a las llamadas lenguas indoeuropeas también se dan de forma significativa en el euskera. Para intentar demostrarlo, apliqué la metodología del lingüista norteamericano Morris Swadesh, iniciador de la glotocronología. El estudio consistió en introducir raíces léxicas vascas en la lista de palabras clave de Swadesh para compararlas con las variantes de cada una de ellas en una docena de lenguas indoeuropeas. El resultado fue que, conforme a la tasa de retención léxica, el euskera pertenecería como mínimo a la misma macro-familia o tronco lingüístico que las lenguas indoeuropeas, dado que, dependiendo de qué lista se use como referencia y qué criterios para valorarla, el porcentaje del euskera oscila entre el 24% y el 40 %, muy por encima

del 12 % indicado por Swadesh. De hecho, si nos atenemos a las cifras de Swadesh/Yakhontov, tendríamos que concluir que el vasco actual pertenece a la misma familia lingüística que el protoindoeuropeo. Claro que muchas de esas voces del euskera actual han sido incorporadas a partir de la romanización y en tiempos modernos. Pero quienes sostienen que esas aportaciones relativamente recientes no han sido transmitidas *por herencia genética*, sino que se trata de adherencias que han ido diluyendo la genuina naturaleza del vascuence llevan, a mi juicio, demasiado lejos la analogía entre las lenguas y la filiación parental de la especie humana, pues olvidan que las promiscuas relaciones entre las lenguas se parecen más a la manera en que se mezclan las aguas fluviales y a los procesos que, lentamente, hacen que cambien el curso de sus cauces o que algunos se sequen y que otros, después de siglos de mostrarse como ramblas agostadas, recobren su caudal. En suma, las lenguas son fluidas, y nada nos impide pensar que también en épocas prehistóricas, por lo que se refiere al euskera o a cualquier otra lengua, se irían produciendo esas influencias, afluencias y confluencias. A mi entender, suponer que el euskera fue durante milenios un sistema lingüístico pétreo que no comenzó su proceso de contaminación hasta la llegada del latín, contravendría principios básicos que conocemos sobre el funcionamiento y el devenir de las lenguas.

Soy consciente de que cuestionar la definición del vascuence como lengua aislada no indoeuropea no resulta sencillo. No obstante, lo que no es tan difícil, paradójicamente, es comprobar la estrechísima relación entre el euskera y una de las lenguas más fecundas de esa gran *familia* indoeuropea. Me refiero al latín. No hay más que leer los trabajos de María Teresa Echenique Elizondo, de la Universidad de Valencia, Fernando González Ollé, de la Universidad de Navarra, o Santiago Segura Munguía, de la Universidad de Deusto. Los latinismos (repárese en que no hablamos de romanismos, procedentes del castellano o del francés, sino de términos del euskera contemporáneos de la romanización, en los albores mismos de nuestra era) son numerosísimos. Pero lo que más llama la atención es que, lejos de nombrar simplemente objetos culturales recientemente asumidos por influencia de la nueva civilización romana, una enorme cantidad de ellos se refieren a realidades digamos eternas, es decir, que ya formaban parte de la cotidianeidad de los vascos y de cualquier pueblo, sin necesidad de que un pueblo conquistador se las señalara.

Por si esto fuera poco, las concomitancias entre el euskera y el latín se dan también en el plano morfológico y aun en el sintáctico. Hace doce años, elaboré para la Consejería de Educación del Gobierno Vasco un método de aprendizaje del latín mediante las categorías de la gramática vasca (*Latina euskaraz ikastelo ariketak*); es posible encontrarlo en [www.eimakatalogoa.eus](http://www.eimakatalogoa.eus). Con él, de modo colateral, pretendía justificar la pertinencia de la lingüística vasco-románica. Ahora bien, sería un error limitar el cuestionamiento de la falaz consideración del euskera como lengua aislada a la presencia del elemento latino. Lo más relevante es lo que nos está diciendo el hecho de que el euskera fuera tan receptivo ante una lengua tan nítidamente encuadrada en la genealogía de las llamadas lenguas indoeuropeas como es el latín. No es lógico que una lengua sea capaz de asimilar tantos elementos de otra con la que, según dicen, nada tiene que ver. Sin embargo, la influencia se dio de forma muy ostensible, ante lo cual concluyo que no serían lenguas tan diferentes.

Recientemente (2023), Fernando Pérez de Laborda en su reciente publicación, *El euskera en Navarra*, también tiene la osadía de cuestionar el carácter singular y único del vascuence. Él sostiene que los vascos somos descendientes de agricultores neolíticos que, hace 7000 años, llegaron a la península ibérica procedentes de los Balcanes. Me costaría suscribir por completo

el marco conceptual en el que se mueve Pérez de Laborda, pues en gran medida rema a favor de la corriente al defender determinadas ideas dominantes en la sociedad vasca actual. Sin embargo, he de admitir que su hipótesis presenta soluciones bastante racionales a cuestiones que durante demasiado tiempo han permanecido bajo el signo del enigma. En primer lugar, propone que el establecimiento en nuestros lares de una lengua procedente del este se produjo porque la población inmigrante era mucho más numerosa que la autóctona; y que ese excedente de población en la tierra de origen se produjo por efecto de la mejora de las condiciones de vida propiciada por la agricultura que surgió en el Neolítico. Y, en segundo lugar, sitúa la procedencia del protovasco o vasco arcaico en un territorio muy próximo al mar Adriático y a Etruria, es decir, en un área donde, en el quinto milenio a. de C., se hablaría alguna lengua que luego derivó en el etrusco, no muy lejos del Lacio, donde se desarrolló el latín. Sospecho, y esto es como tirarme a ciegas a una piscina sin saber si hay agua, que la lengua que llegó a Aquitania y a la zona surpirenaica en el siglo II a. de C. en boca de soldados y centuriones itálicos (por supuesto, muy diferente de la de Cicerón) fue percibida por los vascones como un idioma más familiar de lo que hemos pensado siempre. Me baso para lanzar esta osada hipótesis en la suposición de que ambas lenguas compartían un origen común, llámeselo nostrático, llámeselo ‘x’; ubíquese en el antiguo Urartu o en el Gran Valle del Rift; remóntese en el tiempo al segundo milenio a. de C. o mucho más lejos, a los primeros atisbos del Neolítico. Y esta vez no digo “familiar” en un sentido analógico, sino muy literal. La genética ha demostrado que todos los seres humanos formamos parte de la misma familia. No es nada inverosímil que los pocos miles de *homo sapiens sapiens* que salieron de África hace 100.000 años hablaran la misma lengua (lo increíble sería que, siendo tan pocos, hablaran más de una). Por tanto, tampoco debería asombrarnos encontrar similitudes entre todas las lenguas del mundo. La mayor dificultad de probar tales asertos es la imposibilidad de que un lingüista comparativista, por muy capaz que sea su mente, conozca suficientemente miles de lenguas. Mi esperanza es que ese obstáculo se vea superado gracias a un instrumento del que se viene hablando mucho últimamente: la IA (inteligencia artificial). De hecho, ya existe desde hace tiempo la lingüística computacional. Hasta ahora, con medios mucho más limitados, ha dado buenos resultados en la lingüística forense. Cuando el caudal lingüístico universal sea volcado en la base de datos apropiada y sea convenientemente procesado por la IA, es muy posible que los muertos hablen.

## BIBLIOGRAFÍA

- Almagro-Gorbea M. y G. Ruiz Zapatero, editores. *Paleontología de la península ibérica*. Complutum, 2-3, 1992.
- Arnaiz Villena, Antonio y Jorge Alonso García. *Caucásicos, turcos, mesopotámicos y vascos*. Universidad Complutense, 2001 (2ª edición: Liber Factory, 2013)
- Ballester Gómez, Xaverio. “La *adfinitas* de las lenguas aquitana e ibérica”. *Palaeohispanica* 1, (2001)
- Ballester Gómez, Xaverio. *Las primeras palabras de la humanidad*. Tilde, 2002
- Ballester Gómez, Xaverio. «Sobre el origen de las lenguas indoeuropeas prerromanas en la península ibérica». *Arse* 32-33, 1998-1999
- Ballester Gómez, Xaverio. «Indoeuropeización en el Paleolítico: una réplica». *Estudios Románicos*», 2004, Vol. 26, p. 217-232, <https://raco.cat/index.php/Estudis/article/view/237787>.
- Ballester Gómez, Xaverio. «Lingüística indoeuropea y simulación románica». *E.L.E.A.* 19. 2020
- Castro-Guisasola, Florentino. «El enigma del Vasconce ante las lenguas indoeuropeas». *Revista de Filología Española*, anexo XXX, 1944
- Cavalli-Sforza, Luca y Francesco Cavalli-Sforza. *¿Quiénes somos? Historia de la diversidad humana*. Crítica, 1999
- García Quintela, Marco V. *El árbol de Guernica. Memoria indoeuropea de los ritos vascos de soberanía*. Abada, 2013
- González Ollé, Fernando. «Navarra, Romania *emersa* y ¿Romania *submersa*?». *Aemilianense* 1 (San Millán de la Cogolla), 2004
- Martinet, André. *De las estepas a los océanos*. Gredos, 1997
- Pedrosa, José Manuel. «¿Lenguas y mitos indoeuropeos? ¿Indoeuroafricanos? ¿Paleolíticos? ¿Neolíticos? ». *Culturas populares* 5, 2007. <http://www.culturaspopulares.org/textos5/articulos/pedrosa.pdf>
- Pérez de Laborda, Fernando. *El euskera en Navarra. Síntesis de una historia*. Mintzoa, 2023
- Roberts, Edward A. y Bárbara Pastor. *Diccionario etimológico indoeuropeo de la lengua española*. Alianza, 2007
- Rodríguez Adrados, Francisco. *Historia de las lenguas europeas*, Gredos, 2008
- Santos, J., A. Emborujó y E. Ortiz de Urbina. «Reconstrucción paleogeográfica de autrigones, caristios y várdulos». *Complutum*, 2-3, 1992.
- Segura Munguía, Santiago y Juan M. Etxebarria Ayesta. *Del latín al euskera. Latinetik euskarara*. Universidad de Deusto, 1996
- Villar Liébana, Francisco. *Los indoeuropeos y los orígenes de Europa*. Gredos, 1996