Comparaciones interlingüísticas. ¿Es posible su sistematización?

Miguel Pérez
Becario del Departamento de Lingüística
Elena Garayzábal Heinze
Profesora Asociada del Departamento de Lingüística

RESUMEN: La Lingüística así como la Psicología se interesan por los aspectos de adquisición del lenguaje y de las comparaciones lingüísticas entre niños que hablan muy diferentes lenguas. Si bien esto es un hecho, lo que es cierto es que hay que profundizar en las propiedades estructurales de las lenguas para descubrir mecanismos de carácter específico que ayuden a entender el proceso y las alteraciones que no pueden ser abarcadas desde las otras perspectivas. Sin embargo, hay una falta de sistematicidad a la hora de escoger las lenguas representativas de una muestra interlingüística. Es decir, la falta de conocimientos lingüísticos y la ausencia casi total de colaboración interdisciplinar entre lingüistas y psicólogos llevan a conclusiones estériles y poco científicas.

ABSTRACT: Language acquisition and crosslinguistic studies are part of the linguistic and psychological interest. The fact is that structural properties of languages are not taken deeply into account from the psychological point of view. There is a lack of systematicity when they select representative languages in a crosslinguistic sample. The absence of linguistic knowledge and the lack of interdisciplinary collaboration between linguists and psychologists lead to futile and non-scientific conclusions.
INTRODUCCIÓN

Las últimas décadas han visto cómo el campo de estudio de la Adquisición del Lenguaje Oral ha experimentado un significativo aumento en la cantidad e intensidad de sus aportaciones. Dentro de esta vasta área de conocimiento ha surgido una amplia serie de estudios cuyo objetivo se centra en dilucidar los procesos subyacentes a este mecanismo que es el de la aparición del lenguaje. Así, desde hipótesis distintas y dispares se han ido configurando multitud de diseños experimentales que han contribuido y contribuyen a aportar datos acerca de este proceso de adquisición, redundando todo ello en el diseño de un intricado modelo acerca de su naturaleza.

Entre las distintas dimensiones y perspectivas desde las que esta realidad se aborda, destacan, en primer lugar, las líneas de investigación que vienen a focalizar su interés en la estructura y evolución de los primeros vocabularios infantiles y su influencia en el desarrollo léxico y gramatical posterior. Los trabajos realizados desde esta perspectiva han destacado la importancia que la composición y la conformación de las primeras 50 palabras en expresión tienen para que aparezcan determinados fenómenos de especial relevancia en el período de adquisición del vocabulario y que parecen observarse en un porcentaje muy alto de niños entre los 14 y los 24 meses: la que ha sido denominada «explosión léxica». Ésta, caracterizada por un aumento súbito en la tasa de adquisición de palabras, será un importante factor de determinación de una constante «precoz» o «tardía» en el desarrollo del lenguaje, puesto que supone el inicio de un cambio cuantitativo en el proceso de adquisición del léxico que arrastrará, a su vez, subsidiarios cualitativos fundamentales en el desarrollo normal, a saber, fenómenos de gramaticalización de distintas clases de palabras, procesos de combinación que dan paso a la aparición de la sintaxis, etc. (D’Orico, Carubbi, Salerni y Calvo, 2001; Bassano, 2000; Rescorla, Mirak y Singh, 2000; Schwartz, 1988; Tamis-Lemonda, Bornstein y otros, 1998).

En segundo lugar, y teniendo como base el conjunto de investigaciones del bloque anterior, es posible encontrar multitud de tendencias e intentos explicativos que tratan de hacer frente a los mecanismos implicados en la asignación de referente por parte de los niños. Por un lado, las habilidades perceptivas han recibido mucha atención (Shi, Werker y Morgan, 1999; Jusczyk y Aslin, 1995;
Comparaciones interlingüísticas. ¿Es posible su sistematización?

Gogate y Baharick, 1998; Benasich y Spitz, 1999). Así, estos trabajos experimentales señalan la existencia de fenómenos y mecanismos asociativos y de procesamiento perceptivo inespecífico de dominio que parecen estar detrás en el proceso de adquisición de las primeras palabras. Por otro, el uso de las claves sociales también parece ser un factor clave para entender el desarrollo del lenguaje y ciertos tipos de alteraciones en determinadas poblaciones. En el establecimiento temprano de vínculos entre etiquetas lingüísticas y objetos, el entorno humano que rodea al niño y sus intenciones referenciales en la denominación de dichos objetos juegan un papel esencial. De esta manera, los fenómenos de contigüidad temporal y la covariación entre las palabras y los objetos no son suficientes para asegurar la asociación palabra-objeto (Baldwin, Markman y otros, 1996; Baron-Cohen, Baldwin y Crowson, 1997).

Junto a estos dos grandes conjuntos de teorías explicativas, en lo que concierne a la asignación de referente, aparecen también los argumentos que enfatizan las restricciones y sesgos del propio sistema lingüístico en el desarrollo de la sintaxis y la gramática en general (Naigles y Hoff-Ginsberg, 1998; O'Hara y Jhonston, 1997; Stevens y Karmiloff-Smith, 1997). Las dimensiones social y perceptiva no son suficientes para explicar el desarrollo del lenguaje y, por lo tanto, hay que profundizar en las mismas propiedades estructurales de éste para descubrir mecanismos de carácter específico que ayuden a entender el proceso y las alteraciones que no pueden ser abarcadas desde las otras perspectivas. Las claves que este sistema proporciona redundan en la asignación de significados y relaciones entre las palabras en el seno de los enunciados (propiedades referenciales y combinatorias de las palabras). El paso a la combinación de las palabras lleva al niño a enfrentarse a otros tipos de problemas no planteados hasta el momento y que van más allá del aprendizaje de las primeras palabras (restricciones semánticas y sintácticas de carácter innato). Para adquirir el lenguaje no basta con el análisis acústico de frecuencias y de los aspectos sociales; es necesario plantear un nivel estructural específico del lenguaje que ayude a entender los trastornos del desarrollo que se queden en este nivel aunque funcionen los demás.

Todas estas visiones panorámicas en torno al mismo fenómeno están llevando, como se ha dicho anteriormente, a la asunción de modelos generales que intentan explicar en qué consiste la Adquisición del Lenguaje Oral y qué factores influyen en la misma y en qué orden / jerarquía tienen lugar:
«In the real world (...) children use social, perceptual, cognitive and linguistic cues to learn new words (...). Not all the cues are utilized by children as they start to their world-learning journey. Not all cues for word learning are created equal. The weighting of the cues changes over the time» (Hirsh-Pasek, Michnick y Hollich, 2000: 145, 146).

Paralelamente a este desarrollo, otro eje de especial importancia que ha venido potenciándose al unísono en este campo de estudio es el de la perspectiva interlingüística. Dada la propia esencia del fenómeno objeto de la investigación, se hace necesaria la aplicación de trabajos con hablantes de diferentes lenguas. El primero de los bloques señalados necesita de contribuciones e investigaciones realizadas con bebés y neonatos que están desarrollando distintas lenguas maternas; sólo así se refuerzan posibles interpretaciones a favor de mecanismos y procesos generales, es decir, que se den en todos los niños y que, por tanto, no sean debidos a características tipológicas propias de su lengua particular (universalidad de los fenómenos):

«Crosslinguistic study does more than reveal uniformities of development (...). It has become clear that different types of languages pose different types of acquisition problems. One cannot study universals without exploring particulars. The world provides us with a marvelous set of ‘natural experiments’ in which children with similar endowments master languages of varying forms» (Slobin, 1985: 4, 5).

Como este mismo autor afirma, es necesario combinar la atención a los universales y particulares para poder acceder a una visión general del desarrollo del lenguaje. Sólo así se pueden señalar qué patrones de desarrollo varían de lengua a lengua y cuáles están determinados por principios comunes de un orden superior. Sin embargo, también es necesario hacer uso de esta perspectiva interlingüística en el diseño experimental de cualquiera de los trabajos del resto de las líneas de estudio señaladas más arriba: el conjunto de las lenguas y culturas del mundo proporciona multitud de datos que deben ser controlados como variables a tener en cuenta.

La trayectoria seguida por los estudios en Psicolingüística, en general, y en Adquisición del Lenguaje, en particular, durante todo el siglo XX, ha sido la de la incorporación progresiva y paulatina de estos principios multilingües. Así, mientras en las primeras décadas predominaba una visión principalmente universalista, en la que se daba por hecho la existencia de evidentes paralelos interlingüísticos basados en los principios comunes de la psicología infantil, en la
actualidad se considera de vital importancia la atención a las dos dimensiones—particular y universal—con el objetivo de poder estudiar el curso de la Adquisición del Lenguaje. Por esta razón, los objetivos de los investigadores a principios de siglo se encaminaban hacia hablantes de lenguas individuales (lenguas europeas) para descubrir principios generales, mientras hoy en día se busca la validación interlingüística de cualquier propuesta. Un ejemplo representativo de esta emergencia lo proporciona la obra de Dan Isaac Slobin (1985) «The Crosslinguistic Study of Language Acquisition»¹, cuya motivación es la de constituir y defender un marco de estudios interlingüísticos como un método en Psicolingüística del Desarrollo.

Ahora bien, una vez asumido por una gran parte de la comunidad científica la necesidad de esta nueva perspectiva, los problemas que en la actualidad aún perduran tienen que ver con la falta de una sistematicidad a la hora de escoger las lenguas representativas de una muestra interlingüística. Es decir, la falta de conocimientos lingüísticos y la ausencia casi total de colaboración interdisciplinar entre lingüistas y psicólogos está propiciando la aparición de publicaciones, en el terreno de la Adquisición del Lenguaje, que utilizan hablantes de lenguas tipológicamente muy cercanas o similares como forma de contrastación de datos para evidencias universalistas. Esto nos lleva irremediablemente a la extracción de conclusiones estériles que carecen de validez a poco que se examinen con un mínimo de rigor científico.

Este trabajo, por tanto, tratará de ofrecer un marco de apoyo que, desde los principios de la Lingüística Tipológica, sirva de base para la aplicación de investigaciones psicológicas a muestras de lenguas con una perspectiva interlingüística verdadera. Así pues, el siguiente apartado (2) afrontará el análisis de las lagunas y los vacíos principales, en torno a la validación interlingüística, de los que adolecen las iniciativas e investigaciones psicolingüísticas que en los últimos años predominan. A continuación (3), se utilizará el marco teórico de la Tipología Lingüística como fuente para el establecimiento de una propuesta de sistematización general y básica.

¹ Ver bibliografía.
ANÁLISIS DE LOS TIPOS DE LAGUNAS TEÓRICAS

La introducción de contextos multilingüísticos en los estudios de Adquisición del Lenguaje ha traído consigo toda una serie de déficits en cuanto a la rigurosidad de los conocimientos lingüísticos se refiere. Además, todavía sigue sin generalizarse, todo lo que debiera ser, la validación interlingüística, con lo que siguen preponderando las investigaciones con aspiraciones concluyentes realizadas exclusivamente en el ámbito anglosajón. Ahora bien, ¿cuál es la tipología de errores predominante cuando sí se hace uso de marcos plurilingües?

2.1. Utilización de algunos datos de lenguas de dudosa legitimidad lingüística

La apertura de determinados círculos o grupos de investigación a este nuevo panorama de pluralidad lingüística parece haber desatado la búsqueda de determinadas lenguas, de cierto calado «exótico», que sirven de comodín para el contraste con cualquier aspecto o nivel estructural de las lenguas occidentales. Sin embargo, muchas veces nos encontramos con que estas lenguas han sido colocadas en paradigmas tipológicos y gramaticales que no les corresponden, es decir, que han sido forzadas a representar imágenes falsas con el objetivo de resultar ser una fabulosa fuente de comparación con los datos que un estudio concreto presenta:

"The universality of the predominance of nouns over verbs in early vocabulary development has been challenged on the basis of data regarding the acquisition of Korean (...) and Chinese (...). The Italian language appears to be a good candidate for testing the hypothesis of the influence of input language on early vocabulary composition in so far as it has structural characteristics (it allows extensive word order variation, the subject is often omitted and verbs are often located in sentence-initial or sentence-final positions) which make it quite different from English and more similar to Korean and Chinese..." (D’Orico et al., 2001).

El trabajo de Laura D’Orico parte de un marco teórico impregnado de fuerte debate y controversia por la existencia de datos contradictorios, acerca de la validación de la hipótesis de la explosión léxica y el desarrollo y composición del vocabulario antes y después de ésta, en función del origen y la lengua materna de los sujetos con los que se han llevado a cabo las investigaciones. Por un lado,
estudios con hablantes chinos y coreanos muestran que la aparición de la explosión léxica estaría relacionada con una aceleración en la adquisición de verbos que es seguida más tarde por un aumento del porcentaje de nombres. Los estudios con hablantes ingleses, en cambio, indican un progreso totalmente contrario, es decir, de nombres a verbos. Es por esta razón por la cual la autora plantea el desempeño de un estudio con hablantes italianos, quienes permitirán contrastar, o no, los datos chinos, ya que las características tipológicas y gramaticales del italiano, según ella, se asemejan a las del chino y coreano y se oponen a las del inglés.

Parece, no obstante, que estas afirmaciones sobre la naturaleza estructural del chino (¿mandarín?) no se sostienen desde ningún ámbito de la Lingüística, puesto que el orden de palabras no marcado de la misma (SVO), y la falta de información morfológica, cargan el peso del nivel sintáctico e impiden una variación libre de lexemas y omisiones frecuentes de determinados elementos sintácticos como el sujeto 2. Un vistazo a cualquier obra o autor que desde los estudios tipológicos pase revista a estas relaciones entre los diferentes niveles lingüísticos permite contrarrestar las afirmaciones trasladadas anteriormente:

«The amount of data that nowadays is available on known languages allows us to identify certain general tendencies in the evolution of linguistic types. Broadly speaking, we can say, for instance, that languages containing short words or monosyllables [without morphemes or other grammatical particles], such as Chinese and Vietnamese, apart from processing a generally complex phonological system (see tonal differentiation in Chinese and Vietnamese), are governed by a positional syntax and rigid word order rules...» (Shibatani et al., 1995: 35)

Estas características sitúan al chino mandarín, muy al contrario de lo que D’Orico afirma, más cerca de la lengua inglesa y alejada del italiano. El coreano, en cambio, es una lengua completamente distinta y ajena a cualquier lengua de la familia síntica, a la cual pertenecen las más de 200 lenguas descritas en china (ethnologue) 3. Ésta se caracteriza por un orden de palabras distinto (SOV) y por


3 Este dato debe ser suficiente para llamar la atención sobre la necesidad de especificar de qué lengua se habla cuando se dice «chino». Se supone que la autora se está refiriendo a la lengua oficial de China, el chino mandarín.
una morfología mucho más rica que permiten una mayor libertad sintáctica y la posibilidad de la omisión del sujeto en el nivel oracional, aunque en ningún caso hacen frecuente y constante la aparición del verbo en las posiciones iniciales\(^4\).

Hechos como éstos llaman la atención sobre la necesidad del manejo de los datos lingüísticos con una mayor precisión. Lo común entre los estudios de Adquisición del Lenguaje, como suele ocurrir en todas las áreas de conocimiento, es la continua referencia bibliográfica intradisciplinar, incluso cuando se tratan aspectos de otros campos. Esto lleva irremediablemente al arrastre de estos tipos de aberraciones teóricas de autor en autor, puesto que en ningún momento se contrastan los datos con los que proceden de las fuentes de estudios especializados en las otras parcelas de investigación. Lo esperable, por tanto, sería una mayor colaboración interdisciplinar en aquellos aspectos que suponen una fuente de interés para dos o más ámbitos de estudio, como es aquí el caso de las lenguas para la Lingüística y la Psicolingüística. Así, existen numerosas y exhaustivas descripciones gramaticales, acerca de toda lengua conocida y estudiada, que desde la Lingüística Teórica y Descriptiva proporcionan una enorme cantidad de datos y evidencias en las que se pueden apoyar los estudios de Adquisición\(^5\).

2.2. Problemas en la determinación de las muestras de lenguas

Partiendo de las limitaciones metodológicas y prácticas que impiden la elección de hablantes de cualquier lengua hablada en el mundo, es necesario tener presente que la perspectiva interlingüística hace referencia a la necesaria representatividad de los grupos de lenguas que en la actualidad conocemos. Sin embargo, la realidad en los estudios que desde el campo de la Adquisición del Lenguaje Oral se ponen en marcha es muy diferente, y esto no se cumple. En muchos casos, la muestra acaba siendo de carácter «intra-europea», y las lenguas que soportan los datos sobre los que se basan las investigaciones siempre son las mismas, a saber: inglés, francés o español (D’Orico et al., 2001; Bassano, 2000; Rescorla et al., 2000; Schwartz, 1988; Tamis-Lemonda et al., 1998; Jusczy\(\)c

---

5 Para acceder a aspectos gramaticales de más de 400 lenguas, ver Campbell (1991).

Otras veces, cuando se deciden utilizar lenguas más alejadas, la elección se hace sin ningún criterio definido y estandarizado. De esta manera, nos encontramos con el ejemplo que el trabajo de Laura D’Orico et al. (2001), donde se agrupan lenguas tan dispares genética, estructural y tipológicamente como el italiano, el chino (mandarín) y el coreano, en contraposición con el inglés, que casualmente sí podría ser asociado con el chino si escogiésemos como variable crítica justamente los parámetros estructurales que los autores utilizan como eje de oposición.

En otras muchas ocasiones, sin embargo, el criterio de selección de lenguas es únicamente el de la filiación genética, esto es, se escogen lenguas de familias lingüísticas distintas. El problema entonces es que no se suele tomar en cuenta que lenguas de filiación diversa pueden coincidir en un parámetro tipológico / estructural determinado. En Rushen Shi et al. (1999), por ejemplo, encontramos que, con el fin de estudiar si los recién nacidos son sensibles a las claves perceptivas que distinguen palabras gramaticales y lexicales en el maternés, toman como punto de partida el hecho de que esta distinción es universal. Para ello, se apoyan en estudios anteriores realizados por ellos mismos en los que parecía evidenciarse estas diferencias acústicas en el análisis de las muestras de habla dirigidas a bebés con hablantes ingleses, turcos y chino-mandarines. En este caso, aunque el turco, el inglés y el chino mandarín son lenguas de familias distintas, hay que preguntarse si su contraste es relevante para los aspectos de interés: el nivel fonético-fonológico y el léxico-gramatical (¿son de tipos diferentes con respecto a estas variables?). En este caso, aunque las tres lenguas pertenecen a tipologías estructurales distintas, podrían no ser las más apropiadas para los fines que se persiguen en dicha investigación. En primer lugar, el turco y el inglés no presentan diferencias relevantes en lo que al patrón de distribución de las clases de palabras en la oración se refiere; en segundo lugar, la rica evidencia que el chino mandarín podría aportar ha sido anulada al no considerar correctamente sus características estructurales y tipológicas. La distinción entre palabra léxica y gramatical responde a un canon lingüístico de corte occidental que encuentra algunos problemas en lenguas como el chino mandarín, donde palabras de una clase pueden funcionar también y con la misma frecuencia en la otra, según el contexto (preposiciones
que funcionan como verbos). En estos casos, es obvio que la clave perceptiva asociada a una partícula será la misma cuando ésta hace las veces de palabra de clase abierta o cuando actúa como palabra de clase cerrada. Sin embargo, pudiera ser que no, pero estos aspectos no han sido considerados ni controlados en el estudio mencionado, eliminando toda posibilidad de hacer valiosa la comparación de esta lengua con las otras dos citadas. Quizá, para comprobar la universalidad de claves perceptivas que distinguen categorías sintácticas diferentes hubiera sido más interesante escoger lenguas del tipo común en Occidente (una lengua indo europea) y otros tipos de lenguas donde las categorizaciones lingüísticas europeas encuentran más problemas. Además del chino mandarín, podría ser interesante el contraste con lenguas del tipo lingüístico «incorporante», que se caracterizan porque en ellas varias raíces se funden en una sola palabra (con los correspondientes morfemas gramaticales), obteniéndose palabras que incluyen el significado de toda una oración (Moreno Cabrera, 1997). Un ejemplo de este tipo de lenguas es el inuit (esquimal), en donde la palabra *anginirnuinnanginnirnarpaat*, con partículas de clase abierta y cerrada en su interior, se puede parafrasear como «él le dijo a él que no es posible que sea más grande» (Bernárdez, 1999: 152). ¿Cuál sería aquí el patrón de percepción acústica? ¿Son válidas en este tipo de lenguas las claves acústicas que diferencian clases de palabras separadas, tal y como las concebimos desde nuestras lenguas occidentales? Habría mucho que decir al respecto.

Como se puede apreciar, la ausencia de un método claro para la comparación se une a la falta de representatividad de las lenguas seleccionadas cuando se llevan a término estudios multilingüísticos cuyo objetivo es señalar un proceso como universal. En el artículo de Rushen Shi et al., tras el aventurado uso de los hallazgos en las tres lenguas mencionadas antes como un validador interlingüístico, se pasa a la comparación de los resultados obtenidos con recién nacidos cuyas madres hablan inglés, español, francés, suizo-alemán, punjabi, cantonés, chino mandarín y tagalo. Parece haber así, por tanto, una mayor variedad lingüística, aunque es necesario un análisis detallado de cada una de estas lenguas sin perder la perspectiva general de los tipos y familias lingüísticas existentes en el mundo. La crítica más relevante a tal selección viene dada precisamente por la poca variedad que hay después de una muestra tan numerosa de lenguas: si el criterio de conformación de la muestra no ha sido una variable estructural concreta (fonológica...
Comparaciones interlingüísticas. ¿Es posible su sistematización?

o léxica), puesto que no se argumenta, la única consideración que se ha hecho, entonces, ha sido la de la presencia de muchas lenguas distintas, lo que choca con la verdadera pluralidad filogenética que bajo las ocho lenguas se esconde. Detrás de tal número de sistemas lingüísticos sólo hay tres filos: indoeuropeo, austronesia y sino-tibetano. Hay que destacar, además, que bajo las docenas de filos y lenguas aisladas (es decir, que no han sido agrupadas en ningún filo) que actualmente hay descritas (Nichols, 1992; Moreno Cabrera, 1995), pudiera ser que una muestra de tres no fuera excesivamente significativa, máxime cuando dicha muestra ha sido construida de forma que de las ocho lenguas escogidas, cuatro pertenecen a un mismo filo (indoeuropeo).

 Esto nos lleva a otra característica propia de los estudios en Adquisición del Lenguaje, los cuales, así como ha venido dándose en la misma Lingüística Tipológica durante muchas décadas, están marcados por un profundo sesgo empírico en la selección de las lenguas-muestra: «La investigación tipológica que tenemos en la actualidad está claramente viciada por un etnocentrismo indoeuropeísta. La crítica de la base empírica de la investigación en Tipología Lingüística, que supone las consideraciones sobre el muestreo lingüístico, hace evidente este viciamento de partida y pone las bases para corregirlo...» (Moreno Cabrera, 1995: 72). La propia obra de Slobin es un ejemplo de ello, puesto que, a pesar de defender la postura interlingüística como un método para los estudios en Adquisición del Lenguaje, ofrece una ejemplificación de hallazgos psicolingüísticos que no responde a ese espíritu inicial6.

 Hasta aquí se ha intentado bosquejar una «radiografía» general del panorama interlingüístico en lo que vienen siendo las tendencias de las últimas décadas en Adquisición del Lenguaje Oral. Un repaso como éste pone de manifiesto, sin duda, el hecho de que el que aún nos encontremos en un proceso de generalización y extensión de esta perspectiva de estudio, que afecta a muchas y diferentes lenguas, está condicionando la asistematicidad en su aplicación. Propuestas como

6 El conjunto de lenguas sobre las que se basan los datos fue conformado de acuerdo con la elección de diferentes filos genéticos y la subdivisión de éstos en varias subfamilias lingüísticas, pero (1) aparecen representados siete filos sin una fundamentación concisa de por qué esos y no otros; (2) en el filo indoeuropeo se consideran tres familias lingüísticas distintas (algunas hasta con cuatro representantes, como la familia romance), mientras en el resto no (sólo se proporciona una lengua de cada uno).
la que a continuación se ofrecen pretenden poner algo de orden en este caos propio de todo proceso en fase inicial. Para ello, habrán de utilizarse, necesariamente, las herramientas y constructos teóricos que disciplinas como la Lingüística Tipológica ofrecen a esta parcela de la Adquisición, haciendo efectiva una verdadera colaboración interdisciplinar.

**PROPUESTA DE SISTEMATIZACIÓN**

Los estudios de Adquisición y Desarrollo del Lenguaje se ubican necesariamente en el seno de comunidades lingüísticas concretas. Esto nos lleva directamente al terreno de las lenguas del mundo, un marco caracterizado por un eje principal que oscila entre la variabilidad de las distintas muestras existentes (6528 lenguas descritas en la actualidad⁷) y la regularidad de los fenómenos y estructuras posibles en ellas:

> «The approximately 6,000 languages spoken on this planet differ from one another in many ways. And yet there are limits to this diversity, and within the diversity it is possible to discern certain regular patterns and to formulate certain generalizations. Such regularities within cross-linguistic variation and such limits of the observable variation are the subject matter of language typology and the study of universals» (Haspelmath et al. 2001: 1).

Desde el punto de vista de la diversidad, es necesario reseñar la desigual distribución geográfica de todas estas lenguas, de manera que nos encontramos que de ese total de idiomas conocidos, el 31% se localiza en Asia (2034), el 30% en África (1995), el 21% en Oceanía (1341), el 15% en América (949) y el 3% en Europa (209) (Moreno Cabrera, 1997: 14). Vemos, pues, que nos hallamos ante una realidad compleja y variada de miles de lenguas extendidas por todo el mundo en grados diferentes de densidad⁸, lo que a su vez, y teniendo en cuenta que la mayor parte de nuestros conocimientos sobre la estructura de éstas provienen de idiomas europeos, refleja perfectamente el sesgo empírico y etnocentrísta que domina en los estudios tipológicos y que se extiende a otras muchas áreas de

⁸ Solamente en la isla de Nueva Guinea, con una población de 5 millones de habitantes, se hablan más de 800 lenguas, que representa cuatro veces las lenguas habladas en toda Europa (Bernárdez, 1999: 69).

98
conocimiento como ya se ha indicado anteriormente para el caso de la Psicolingüística. Para evitar este mismo hecho en las investigaciones que hagan uso de las tipologías lingüísticas, hay que partir de este panorama general y pluriforme y realizar muestras lingüísticos que respondan al mismo. Así, si queremos llevar a cabo un estudio de un determinado aspecto en Adquisición del Lenguaje, con hablantes de diferentes lenguas, hay que hacer lo posible para mantener esa proporción descrita antes, de forma que, por ejemplo, sólo el 3% de las lenguas usadas para la investigación sean europeas.

Sin embargo, dentro de este mismo cuadro de variabilidad y «caos» es posible encontrar patrones estables, ya que «la variedad de lenguas (...) está sujeta a generalidades y leyes que hacen que la diversidad (...) esté constreñida o delimitada de una manera precisa. No podemos encontrar cualquier tipo de diferencia lógicamente concebible entre dos o más lenguas o familias» (Moreno Cabrera, J.C., 1997: 13). De acuerdo con este marco de relaciones y similitudes entre las lenguas del mundo, se identifican tres modos fundamentales de clasificación de éstas (Greenberg, 1966): la clasificación genética, que hace referencia a la evolución de las lenguas y las agrupa según su descendencia; la clasificación tipológica, establecida en función de las características gramaticales o estructurales; y la clasificación territorial, como consecuencia del ámbito geográfico en el que se hablan. Todas ellas deben ser encuadradas dentro de los estudios de la Tipología Lingüística para ofrecer un verdadero marco de clasificación, del cual nos aprovecharemos para extraer un modelo de sistematización interlingüística válido para los objetivos de la Psicolingüística y, más en concreto, para las investigaciones en Adquisición del Lenguaje.

Pero entonces, ¿qué debemos entender por tipología lingüística? ¿Qué métodos de los utilizados por ésta pueden ser útiles para los fines que aquí perseguimos?

Sólo a partir de lo explicitado hasta el momento podemos llegar a una definición satisfactoria de lo que es la Lingüística Tipológica. Existen multitud de definiciones que responden a distintos intereses teóricos. Paolo Ramat (1987: 3,4), por ejemplo, presenta en su obra un recorrido histórico a través de las consideraciones planteadas por multitud de autores que reflejan las ideologías dominantes de las diferentes épocas en el ámbito lingüístico. En este recorrido se encuentran, entre otros: Marouzeau (1951: 234): «L’étude typologyque des langues
est celle qui définit leurs caractères en faisant abstraction de l’histoire»; Dressler (1973: 470): «In particular, linguistic typology menas generalizing typology of languages according to the similarity or dissimilarity in their linguistic structure»; o Greenberg (1974: 9): «an approach which transcends the boundaries of any particular science». El marco diseñado anteriormente, no obstante, exige una definición de la Lingüística Tipológica más amplia que incluya los tres modos de clasificación de las lenguas ya vistos. Ésta será la perspectiva que más nos interese, puesto permitirá articular un ulterior modelo de sistematización más completo y provechoso. Por esta razón, nos sumaremos a las palabras del profesor Juan Carlos Moreno Cabrera, para quien la Lingüística Tipológica y Universalista «es aquella rama de la ciencia del lenguaje que se ocupa de estudiar las semejanzas y las diferencias entre las lenguas humanas» (Moreno Cabrera, 1997: 13). Pero, además, hay que señalar la existencia de varios tipos generales de tipologías. Por un lado, las tipologías parciales; por otro, tipologías holísticas (Shibatani y Bynon, 1995: 1-23). Las primeras, que son las más practicadas en la actualidad, se caracterizan por la comparación de lenguas diferentes en función de construcciones específicas y fenómenos concretos (orden de palabras, la construcción pasiva, etc.), mientras las segundas estudian e identifican conjuntos de rasgos y propiedades que, en el interior de las lenguas, se organizan de forma jerárquica e interactiva. Obviamente, es el modo parcial el que interesa en este trabajo, puesto que se trata de poner el conocimiento tipológico y lingüístico de las lenguas y tipos de lenguas del mundo al servicio de la investigación psicolingüística: controlar una determinada variable o estructura gramatical para estudiar la universalidad de ciertos procesos cognitivos entre los hablantes de diferentes lenguas.

Ahora bien, cuando emprendemos la tarea de la comparación interlingüística, y adaptando las explicaciones de Shibatani y Bynon (1995: 89-91), tenemos que enfrentarnos ante una serie de problemas de partida: por un lado, la representatividad del rango de variación genética y geográfica en las lenguas del mundo; por otro, la selección de la variable gramatical a examinar; por último, garantizar la independencia de casos. Es decir, que el fenómeno gramatical escogido no co-ocurra en las lenguas que se comparan o sean objeto de comparación. Este hecho puede darse por dos razones, principalmente: que las lenguas estén genéticamente relacionadas y el aspecto gramatical en cuestión sea una reminiscencia de una protolengua común; o bien que las dos lenguas estén en una misma
región geográfica y el fenómeno se haya difundido de una lengua a otra (o a dos lenguas desde una tercera en común).

Para abordar todos estos inconvenientes se viene utilizando en las últimas décadas el procedimiento del «muestreo lingüístico», sucintamente evocado al principio de esta sección, al que se le deben añadir algunos complementos teóricos que permitan llevar a cabo comparaciones interlingüísticas haciendo frente a los problemas mencionados de la diversidad e independencia de casos. Uno de estos conceptos, que se verá más adelante, es el de «área ligüística».

La necesidad de realizar muestreos lingüísticos viene determinada por la imposibilidad de considerar, cuando queremos estudiar un determinado fenómeno, todas las lenguas existentes en un momento determinado. Por esta causa nos vemos obligados a seleccionar un número limitado de lenguas que serán objeto de análisis, y es aquí donde hay que combatir el sesgo empírico que, como hemos visto, suele condicionar los estudios en Psicolingüística: «no podemos suponer que una investigación tipológica es representativa si incluimos datos sobre diez o doce lenguas europeas (la mayoría indoeuropeas). Si no aplicamos un criterio más objetivo, nuestras conclusiones tipológicas estarán de partida viciadas y tendrán una base empírica muy débil» (Moreno Cabrera, 1995: 67,68).

¿En qué consiste pues el método del muestreo lingüístico? Partiendo de esa imposibilidad práctica para estudiar un fenómeno determinado en todas las lenguas existentes, hay que diseñar una muestra representativa de los diferentes tipos lingüísticos excluyendo la aparición de demasiados ejemplos de un tipo específico. Los dominios que tenemos en cuenta en una muestra de este tipo son los siguientes: universo, marco y muestra. Existen diferentes concepciones de lo que estos tres dominios son. Sin embargo, es la propuesta ofrecida por Haspelmath et al. (2001) la que nos parece más útil. Según estos autores, aunque el universo de interés son todas las lenguas humanas, el universo relevante en investigación tipológica son los tipos de lenguas establecidos en función del parámetro gramatical sobre el que se centra la propia investigación. El subconjunto de atributos que se considerará será el utilizado para establecer los tipos y las relaciones entre ellos; las lenguas relevantes para el estudio sólo serán tomadas en cuenta después de controlar las relaciones genéticas y de áreas que entre ellas se dan. El marco, por ende, debe estar conformado por la lista de lenguas existentes y descritas que entran dentro de los tipos del universo, a partir de la cual se seleccionará la
muestra de lenguas relevante. Hay varios procedimientos para seleccionar una muestra de lenguas de un marco. Veamos algunos de los que más nos interesan: conveniencia, elección experta y muestras de probabilidad. Las muestras de conveniencia comprometen a los casos / lenguas que son fácilmente accesibles para el investigador. Las muestras de elección experta reflejan el juicio de un investigador acerca de qué casos son más representativos en función del parámetro escogido. Las muestras de probabilidad incluyen un procedimiento de selección aleatoria que puede ser simple, estratificada y agrupada. En todas ellas cada caso tiene una probabilidad definida de ser seleccionada

Para llevar a cabo estos procedimientos es necesario afrontar la cuestión de las relaciones entre las lenguas. De esta forma, y dado que las lenguas pueden estar asociadas con otras mediante relaciones genéticas y/o geográficas que las hagan converger en determinados haces de rasgos estructurales, se deben hacer uso aquí de las clasificaciones genéticas que la Lingüística proporciona (distribuyendo las lenguas en familias lingüísticas), así como de la división de la superficie terrestre en regiones o áreas de influencia (hemisferios, continentes, subcontinentes regiones, etc.), para la selección última de la muestra que garantice la diversidad y la independencia de casos (ya sea mediante el procedimiento de conveniencia, el de juicio experto o el de probabilidad estratificada sobre estos estratos genéticos y geográficos). Este ámbito geográfico nos lleva, por tanto, al terreno de lo que se ha denominado en los últimos años «Tipología de Área», que es la disciplina que se encarga del estudio de patrones en la distribución por áreas de rasgos de lenguas tipológicamente relevantes (Haspelmath et al. 2001: 1456), y, más en concreto, al concepto de «área lingüística». Tomando las definiciones de diversos autores, en diferentes épocas, como punto de partida (Trubetzkoy, 1928; Emeneau, 1980; Masica, 1992; Campbell, 1994), podemos definir «área lingüística» como aquella que incluye lenguas pertenecientes a diferentes familias lingüísticas que, sin embargo, poseen una serie de rasgos gramaticales compartidos debido al contacto lingüístico y que no pueden ser explicados por relaciones genéticas. Es por esto por lo que la noción de área debe ser incorporada y tenida en cuenta en

9 Las muestras de probabilidad estratificada dividen el marco en estratos, según unas variables determinadas, a partir de los cuales se aplica el criterio de selección aleatoria; y son usadas para aumentar la diversidad y para asegurar la independencia de casos a diferencia de la selección aleatoria simple.
nuestro modelo, ya que si sólo prima el criterio de clasificación genética en el
momento de seleccionar las lenguas pertinentes del marco, se puede dar el caso de
que se escojan dos lenguas de familias lingüísticas distintas que compartan un
conjunto de rasgos debido al contacto lingüístico, haciendo peligrar el principio
de independencia de casos.

Vemos entonces, que una vez utilizado uno de los modos de clasificación de
las lenguas según Greenberg –el tipológico– para restringir el universo de interés
del trabajo, hay que hacer uso de los otros dos –genético y geográfico– para una
correcta aplicación del muestreo sobre las lenguas del marco. Sólo así se garantiza
la diversidad y la independencia de casos en la muestra final.

Ahora bien, ¿cómo llevar a la práctica estos dos criterios –genético y geográfico–
para la determinación de la muestra a partir del marco seleccionado y
cumplir con estos requisitos de diversidad e independencia de casos? Para ello se
aglutinarán las propuestas de Bell (1978), Tomlin (1986), Bybee (1985), Perkins
modelo unitario que sirva de base para la propuesta que en este trabajo se ofrece.
En primer lugar, y siguiendo a Nichols (1992: 25-33), dividiremos el mapa
distribuyéndolas entre las distintas macroáreas y áreas. Después, hay que
asegurarse de que, en función del número de lenguas que vayamos a utilizar en
la investigación, se seleccionen lenguas de diferentes familias que se sitúen en
distintas macroáreas / áreas, contribuyendo así a la maximización de la diversi-
dad y de la distancia y controlando los efectos del contacto lingüístico y de la
familia genética sobre el conjunto de lenguas tipológicamente relevantes para
el estudio.

10 La primera se compone de África, Oriente Medio, Eurasia Septentrional y Sureste de
Asia. El Nuevo Mundo se descompone en América del Norte, Mesoamérica y Sudamérica. La
macroárea pacífica se divide en Oceanía, Nueva Guinea y Australia.

11 En Nichols (1992) se puede acceder a una descripción exhaustiva de lenguas y familias
distribuidas por áreas y macroáreas.
Si nos vemos obligados a utilizar, en cambio, más de una lengua por área o macroárea (porque es imposible incluir muchas lenguas de muchas áreas distintas, por ejemplo), entonces hay que tener muy presente el mapa lingüístico general y hacer uso de la proporción equitativa defendida por Bell y Tomlin, entre otros, considerando tantas lenguas representantes de una familia según la representatividad de dicha familia en el conjunto de las lenguas del mundo\(^{12}\).

Así, es este marco teórico del muestro lingüístico el que puede ser utilizado y exportado al ámbito psicolingüístico, en general, y a los trabajos en Adquisición del Lenguaje, en particular. Aunque el objetivo que se persigue con el uso de este método desde el campo de la Lingüística Tipológica se orienta al estudio de la distribución de rasgos gramaticales en las lenguas del mundo, se puede realizar una adaptación a la parcela que nos interesa dirigiendo su acción a la comprobación de la generalidad / universalidad de un determinado fenómeno a partir del contraste de lenguas cuyas estructuras gramaticales sean pertinentes para la validación psicológica de dicho fenómeno. En una investigación sobre Adquisición del Lenguaje que quiera aplicar un estudio sobre hablantes de lenguas diferentes, en busca de evidencias universales, es fundamental partir de la variable crítica o el parámetro gramatical que interesa. De esta manera, se puede confeccionar una muestra de lenguas siguiendo los criterios destacados. Sólo así se hará posible que estos trabajos descansen sobre unos pilares lingüísticos mínimamente rigurosos y, por tanto, puedan alcanzar los objetivos diseñados\(^{13}\).

**DISCUSIÓN**

Un repaso a los trabajos y líneas de investigación en el ámbito de la Adquisición del Lenguaje en la última década pone de manifiesto el incipiente resurgimiento de la perspectiva interlingüística. Sin embargo, el uso de esta validación lingüística para la aplicación de investigaciones con propósitos

\(^{12}\) Para consultar la representatividad de una familia lingüística en el total de lenguas y familias del mundo, acudir a las referencias Bell (1978: 126) o Perkins (1989: 293-315).

\(^{13}\) Si no se utilizase un parámetro gramatical concreto porque el objetivo es estudiar la generalidad de un fenómeno particular entre los hablantes, el universo de la muestra del que partimos serán todos los tipos de lenguas existentes, a partir del cual se seguirán los mismos pasos ya mentionados. Para estos casos, no obstante, es especialmente importante el hecho de maximizar la variedad y disponer de un número alto de casos para asegurar una tendencia universal.
psicolingüísticos aún deja mucho que desear. Por un lado, es frecuente la extracción de conclusiones de carácter universalista a partir de la comprobación de trabajos experimentales en una sola lengua; por otro, y cuando se deciden llevar a cabo estudios con hablantes de diferentes lenguas con la finalidad de comprobar si los resultados son debidos a las características estructurales de la lengua materna o a mecanismos psicológicos generales, se utilizan datos lingüísticos de forma errónea o se contrastan lenguas sin ningún criterio claro y definido que corresponda con una base lingüística mínima.

La revisión de las principales lagunas que en estos aspectos se observan en el terreno de la Psicolingüística en general, y en el de la Adquisición del Lenguaje en particular, llama la atención sobre la necesidad de una mayor y más eficiente colaboración interdisciplinar entre la Lingüística y la Psicología. El ámbito de la Psicolingüística es un campo teórico de contacto entre estas dos disciplinas y, por ello, deben ponerse las bases para una mayor colaboración a través de investigaciones y estudios conjuntos en los que cada una aporte sus conocimientos específicos y necesarios para la correcta delimitación de los fenómenos y variables que entran en juego en el terreno del lenguaje humano. Este tipo de cooperaciones evitaría el uso erróneo, por parte de cada una de las áreas implicadas, de los conocimientos especializados de la otra, redundando todo ello en una práctica investigadora más exitosa y precisa.

En este trabajo, como muestra de lo que puede y debe ser una colaboración interdisciplinar, se pretende proporcionar unos fundamentos mínimos para el establecimiento de una propuesta de sistematización interlingüística en los estudios de Adquisición del Lenguaje, la cual es extrapolable, de forma general, a los estudios psicolingüísticos. Así, y partiendo del marco teórico de la Lingüística Tipológica, se han intentado adaptar algunos métodos y conceptos de este campo del estudio científico a los propósitos de la investigación psicolingüística en Adquisición del Lenguaje. El método del muestreo lingüístico, junto con la inclusión de algunas mejoras teóricas como es el caso de las aportaciones de la Lingüística de Áreas, ofrece herramientas muy útiles a este fin. De esta manera, el modelo ofrecido arranca de los tres grandes ejes/Tipos de clasificación tipológica señalados por Greenberg, los cuales se aplicarán sobre el muestreo lingüístico de forma que, partiendo de la variable crítica de interés, esto es, el parámetro gramatical que se quiere controlar para estudiar un determinado fenómeno.
psicolingüístico, se utilizan los ejes genético y geográfico para la selección final de la muestra de lenguas. Así pues, la variable tipológica se puede utilizar para determinar el universo de la muestra señalando los tipos de lenguas relevantes para un estudio particular (todos los tipos de lenguas que compartan un rasgo o que se opongan respecto éste); después, y haciendo uso de las clasificaciones de la Lingüística Genética y de Área, se escogerán las lenguas que, dentro del marco (aquellas que entran dentro de los tipos lingüísticos relevantes del universo de la muestra), y a través de procedimientos alternativos como el de «elección experta», «conveniencia» o «muestra de probabilidad», pertenezcan a familias lingüísticas diferentes y se ubiquen en áreas / macro-áreas distintas. Con este método aseguramos que se cumplan algunas de las condiciones fundamentales de toda muestra lingüística: que las lenguas seleccionadas sean diversas e independientes genética y geográficamente.

La cruda realidad a la que se enfrenta este tipo de estudios en Adquisición del Lenguaje, sin embargo, limita la participación de muchas lenguas, pues al final no siempre se dispone de hablantes de lenguas muy diversas genética y geográficamente. Aun así, la propuesta que aquí se ha delimitado sólo pretende, de momento, ofrecer un marco que hace las veces de punto de partida y de referencia al mismo tiempo, de modo que cualquier iniciativa desde las líneas psicolingüísticas de la Adquisición del Lenguaje puede ser adaptada a este modelo en función de las posibilidades y necesidades. Por ello, y ante la difícil situación en la que estas investigaciones se suelen desenvolver, en donde sólo se puede trabajar con hablantes de pocas lenguas, hay que hacer hincapié en los niveles superiores de diversificación -macroáreas y familias lingüísticas-, a partir de los cuales se escogerán los casos concretos y representativos siguiendo las máximas de diversidad e independencia de casos. Si aportamos datos de más lenguas dentro de cada área/macroárea y/o familia lingüística contrastada, por falta de acceso a más lenguas de distintas áreas o familias, tendremos que respetar, al menos, el requisito de la representatividad. Para ello, se utilizará el nivel de la subfamilia (dentro de la familia indoeuropea, por ejemplo, la subfamilia romance, germánica, etc.) para ir seleccionando más casos dentro de cada familia respondiendo al peso de ésta en la configuración lingüística general, esto es, en el mapa lingüístico mundial. De esta forma se evita el fenómeno del sesgo empírico en la determinación de la muestra.
BIBLIOGRAFÍA


*Child Language*, 25, 675-700.


Comparaciones interlingüísticas. ¿Es posible su sistematización?


Rescorla, Mirak y Singh (2000). Vocabulary growth in late talkers: Lexical development from 2;0 to 3;0. Journal of Child Language, 27, 293-311.


