Viejos nuevos datos
Un *corpus* para el estudio histórico del cambio lingüístico en una lengua indígena[1]

Frida Villavicencio
Centro de Investigaciones y Estudios Superiores en Antropología Social
CIESAS-México
Juárez 87, Col. Tlalpan
C.P. 14000 México, D.F.
frida@ciesas.edu.mx

Resumen

En este trabajo reflexiono sobre el proceso mediante el cual se constituyó un corpus *para el estudio de un fenómeno de cambio lingüístico en una lengua originaria de México, el purépecha. Las características del material (textos del siglo XVI y del siglo XX), el limitado conocimiento —o, mejor dicho, desconocimiento— del objeto de estudio (la lengua purépecha), y el punto de vista (dinámico) que guía el análisis definen la especificidad de la mirada. La reflexión sobre las sucesivas etapas de trabajo que fue necesario desarrollar y los criterios que nos guiaron hasta lograr integrar un cuerpo de datos útiles para dar cuenta de la historia de una lengua con poca tradición de escritura, aspira a hacer evidente un momento de la investigación que es compartido por otros campos de la disciplina.*

Palabras clave: lingüística histórica, gramaticalización, lenguas indígenas, purépecha, *corpus* textual, fuentes documentales

Fecha de recepción del artículo: 15 de junio de 2005
Fecha de aceptación de versión revisada: 25 de enero de 2008

[1] Me congratula poder responder a la invitación que amablemente me hiciera Teresa Carbó para colaborar en el presente volumen sobre el dato en lingüística, tema que a lo largo del tiempo ha constituido un constante punto de encuentro entre dos colegas que comparten, además de un espacio físico, la fascinación por el lenguaje. A pesar de mirar la lengua desde perspectivas distintas, en no pocas ocasiones hemos conversado sobre este tópico, primero a raíz del seminario *La construcción del dato en lingüística y sociolingüística*, que se realizó hace ya más de una década en el seno del Área de Lenguaje y Sociedad del CIESAS. Agradezco cuanto me he venido beneficiando de su experiencia, perspicacia y amor por el lenguaje en sucesivas discusiones informales que entablamos de tanto en tanto, la más reciente a raíz de la elaboración de este artículo cuya responsabilidad, cabe decir, es exclusivamente mía.

46.indd 177 10/3/09 09:03:18

Abstract

In this paper I explore the process by which a corpus *for the study of the linguistic change in a native language in Mexico, the purépecha, was constituted. The characteristics of the material (data from 16th and 20th centuries), the limited knowledge—or, shall we say, lack of it—of its object of study and the dynamic point of view that guides this analysis define the specificity of the gaze. The reflection on the following stages of the work that needed to be developed and the criteria that lead us to finally integrate a real* corpus *that were useful in order to explain the history of this native language with little written tradition aims to show a moment of the research that inevitably is shared by other fields of the discipline.*

Key words: historic linguistic, gramaticalization, native languages, purepecha, textual *corpus*, documentary sources

46.indd 178 10/3/09 09:03:18

Introducción

En lingüística histórica, el enfoque que se interesa por el cambio y los procesos de gramaticalización que han tenido lugar a lo largo del tiempo presupone la observación de la lengua durante amplios periodos. Se trata de un punto de vista que intenta dar cuenta, en algún grado, de un fenómeno cuya complejidad representa, quizás, nuestro mayor reto. Cuando se pretende observar la lengua en uso, como es el caso, un asunto medular para el análisis es el que tiene que ver con los datos; asumir el objetivo, que bien puede ser calificado de anhelo, de observar la lengua en toda su complejidad (Jakobson, [1952] 1984) implica —ante la imposibilidad de resolverlo— adoptar de manera consciente una postura frente al problema que supone contar con un cuerpo de datos empíricos que aspira a constituirse en una muestra confiable de la lengua en uso en un determinado periodo.

Uno de los primeros problemas que todo investigador interesado en el desarrollo diacrónico de una lengua enfrenta es, pues, el de cómo construir un *corpus* capaz de exhibir rasgos significativos con respecto al tema que analiza. Sin embargo, como bien lo han señalado otros autores, la cuestión del dato no es un tema exclusivo del enfoque diacrónico; en todo caso es un asunto central al trabajo lingüístico en sí mismo y, como tal, ha sido objeto de una continua reflexión teórica; ante la imposibilidad de trabajar con fenómenos de tal envergadura, la confiabilidad de la muestra a partir de la cual se realiza un análisis resulta un punto medular que se extiende a todo el campo de las disciplinas sociales (Carbó, 2001a:21).

El presente artículo tiene como finalidad mostrar, de manera detallada, el proceso que siguió la constitución de un *corpus* adecuado para el estudio histórico de un fenómeno de cambio lingüístico en una lengua con poca tradición de escritura. Parto de una experiencia concreta: el *corpus* en cuestión sirvió de base para realizar una investigación sobre el proceso de cambio que se observa en el sistema de casos del purépecha[2] en el periodo comprendido entre el siglo XVI y el

[2] El purépecha o tarasco es una lengua indígena del occidente de México, que actualmente es hablada por poco más de 100 000 personas en la región noroccidental del estado de Michoacán. Hasta ahora no se ha podido establecer con seguridad su filiación lingüística; no es seguro que forme parte de alguna de las familias lingüísticas indoamericanas y tampoco se le puede considerar como una lengua típicamente mesoamericana, por lo que se le percibe como una isla en Mesoamérica. La incierta filiación lingüística del purépecha ha sido un tema recurrente en los estudios de esta lengua; sin embargo, hasta ahora poco se ha podido avanzar en su esclarecimiento (Villavicencio, 2002).

46.indd 179 10/3/09 09:03:18

siglo XX.[3] La reflexión que surge en torno a dicho proceso permite destacar las similitudes y especificidades que este ejercicio particular guarda con otros procesos paralelos en el campo de los estudios lingüísticos (como el análisis del discurso, la pragmática o los estudios sobre variación). Me guía el afán de contribuir, en la medida de lo posible, a la discusión teórica y metodológica en torno al proceso de constitución de un *corpus* para trabajo lingüístico desde un enfoque que pretende sustentar su análisis en una muestra representativa de la lengua en uso.

En primer lugar considero las causas que me llevaron a optar por una propuesta metodológica comúnmente asociada al estudio de lenguas de larga tradición escrita —como lo son las indoeuropeas— para dar cuenta de una lengua que se considera básicamente de tradición oral, como lo son la mayoría de las lenguas indígenas de México. En segundo lugar describo las etapas mediante las cuales se integró el acervo documental que sirvió de base para la investigación: en un primer momento, la recopilación de los textos y, posteriormente, su clasificación. En tercer lugar presento los criterios que permitieron establecer los estados de lengua que se reconocieron en este estudio, así como los que guiaron la selección de los materiales que finalmente constituyeron el *corpus*. Por último describo, de manera breve, el tratamiento que se dio a los materiales tanto en la fase de transcripción como en el análisis lingüístico propiamente dicho.

La propuesta metodológica

La investigación sobre la *Estructura y cambio del sistema de casos en el purépecha. Del siglo XVI al siglo XX* (Villavicencio, 2002, 2006a) se inscribió en el marco de los

[3] Dicho trabajo, con el título de *Estructura y cambio del sistema de casos en el purépecha. Del siglo XVI al siglo XX*, fue desarrollado como tesis doctoral bajo dirección de la Dra. Concepción Company y el Dr. Thomas Smith-Stark y se presentó en el Centro de Estudios Lingüísticos y Literarios de El Colegio de México en enero de 2002. Debo a la Dra. Company haberme introducido al campo del estudio de los fenómenos de cambio lingüístico desde el enfoque de la gramaticalización y haber apoyado, con la amplia visión que la caracteriza, un proyecto que se apartaba de los cánones establecidos por la tradición en el estudio de lenguas indoeuropeas. El Dr. Smith-Stark ha sido un interlocutor constante, con el cual he discutido puntos específicos de la estructura de esta lengua; su amplia experiencia en el campo de las lenguas mesoamericanas y su interés manifiesto por los estudios de corte filológico sobre lenguas indígenas de México ha sido inapreciable en todo momento. La versión revisada de este trabajo fue publicada en 2006 con el título de *P'orhépecha kaso sïrátahenkwa. Desarrollo del sistema de casos del purépecha*.

46.indd 180 10/3/09 09:03:18

estudios de gramaticalización que se interesan en explicar cómo las formas y construcciones gramaticales emergen, se usan y llegan a conformar la lengua (Hopper y Traugott, 1993). Este enfoque tiene sus antecedentes en las ideas de Etienne Bonnot de Condillac quien, a finales del siglo XVIII, propone que los elementos gramaticales evolucionan a partir de lexemas (Lehmann, 1982), y en los trabajos de filología comparativa desarrollados durante el siglo XIX por autores como Franz Bopp, Georg von der Gabelentz, Wilhelm von Humboldt, August Wilhelm von Schlegel, William Dwight Whitney y Franz Wüllner, entre otros (Heine, Caudi y Hünnemeyer 1991; Hopper y Traugott, 1993). Durante los dos primeros tercios del siglo XX, el concepto de gramaticalización se fue desarrollando paulatinamente gracias al trabajo de insignes lingüistas como Benveniste (1966), Jakobson (1959), Kuryłowicz (1965), Malkiel (1967), Meillet (1912), Sapir (1921) y Žirmunskij (1966), que continuaron en buena medida la tradición de los neogramáticos. A Meillet se le reconoce como el fundador de los modernos estudios de gramaticalización aunque la definición más ampliamente aceptada proviene de Kuryłowicz.

A partir de la década de los años setenta del siglo pasado, se reconoce el potencial que ofrece la gramaticalización como un parámetro explicativo tanto para los estudios de corte diacrónico como para los de corte sincrónico. En esa época se empieza a tomar en cuenta la relación que existe entre dominios cognitivos como el espacio, el tiempo y el modo, así como el papel que tienen procesos creativos tales como la metáfora y la metonimia en la estructura del lenguaje (Heine, Claudi y Hünnemeyer, 1991). En los estudios sobre gramaticalización que se han desarrollado en los últimos años se pueden reconocer diferentes tópicos, aproximaciones y posiciones (Bybee, 1985; Company, 1991; Givón, 1971; Heine y Reh, 1984; Hopper, 1987; Lehmann, 1982; Li, 1975; Mithun, 1991; Nichols, 1986; Sweetser, 1988; Timberlake, 1977; Traugott, 1985 y Wischer y Diewald, 2002).

Si bien en las últimas décadas se ha venido acumulado evidencia de lenguas no europeas (Heine y Reh, 1984; Heine y Kuteva, 2002), un buen porcentaje del estudio sobre gramaticalización se basa en el análisis de lenguas indoeuropeas. Uno de los factores que explican este hecho es la naturaleza de los datos con los cuales se trabaja; al igual que el método comparativo utilizado por los neogramáticos, los modernos estudios de gramaticalización de corte diacrónico centran su análisis en los textos, únicos testimonios de antiguos estados de las lenguas. De esta manera, las lenguas que cuentan con una larga tradición no sólo de escritura, sino también de crítica textual y constitución de *corpora* como lo son, entre otras,

griego, latín, inglés, alemán, francés, italiano y español, han captado la atención de los estudiosos de la gramaticalización. Sin embargo, aun en este grupo, el panorama es heterogéneo; la disponibilidad de *corpora* adecuados para estudios de gramaticalización es diferente en cada una de estas lenguas (Company, 2001).[4]

Al inscribirnos en el marco de los estudios de gramaticalización, la primera decisión metodológica que se tomó atendió, entonces, a la naturaleza del *corpus*. El tipo de trabajo que se pretendía llevar a cabo se basa en el análisis y la comparación de sucesivos estados de lengua, para lo cual se requería contar con muestras documentadas de la lengua en estudio, en este caso el purépecha, en cada una de las épocas que se consideraron significativas para el análisis. De esta manera, la investigación puso a prueba una propuesta novedosa en el campo de los estudios históricos de las lenguas indígenas de México: el análisis de un *corpus* constituido por datos documentados escritos en purépecha.

Tomar esta decisión implicó apartarse de algunas de las metodologías más comunes en los estudios de lenguas indígenas: el método histórico comparativo y la elicitación directa con un hablante. El primer método parte de la comparación de elementos cognados que se encuentran actualmente en varias lenguas de una misma familia, para proponer protoformas que hipotéticamente estuvieron presentes en una lengua antecesora común. Hasta ahora, éste ha sido uno de los métodos más frecuentemente utilizados en estudios de corte diacrónico (Villavicencio, 2001: 435). Cabe hacer notar que, al ser el purépecha una isla lingüística en el mapa mesoamericano, sería muy difícil llevar a cabo un estudio de reconstrucción histórica que no fuera estrictamente interno.

El segundo método se basa en obtener oraciones, frases y paradigmas de la lengua que se consideran pertinentes para el análisis con ayuda de un hablante; el lingüista va preguntando los ítems de un cuestionario previamente elaborado, que el hablante va pronunciando en el idioma nativo. Si bien en los últimos años se ha ponderado la importancia que tiene trabajar con fragmentos de "producción espontánea" —cualquiera que sea el significado dado a este concepto—, la elicitación directa con un hablante sigue siendo uno de los métodos más ampliamente

[4] Company ha señalado en reiteradas ocasiones la falta de *corpora* adecuados para el estudio histórico de las diferentes variedades del español. Ella misma ha iniciado la constitución de *corpora* del español de México. Su propuesta está fundamentada en una seria reflexión teórica y metodológica sobre lo que debe ser el trabajo en lingüística histórica (Company, 1994, 2001 [en prensa]).

utilizados en muchos de los trabajos descriptivos que se realizan sobre las lenguas originarias de México, entre ellas, el purépecha.

De esta manera, la propuesta de trabajar con un *corpus* documentado constituyó una innovación teórica y metodológica en el campo de estudio de las lenguas indígenas de México. En lo que respecta al primer punto, el marco de gramaticalización ha sido un enfoque escasamente adoptado hasta ahora; algunos autores como Briseño (1997, 1998, 2000a, 2000b) y Chamoreau (2006) han apelado al marco de gramaticalización para intentar explicar fenómenos de variación que observan en la lengua actual. Si bien estos autores acuden a fuentes novohispanas para fundamentar su análisis, su interés se centra en entender un fenómeno sincrónico, por lo que no contemplan la constitución de un *corpus* documentado del tipo que aquí nos ocupa.

En lo que toca al aspecto metodológico, trabajar con documentos nos acercó a los estudios filológicos y a la tradición ecdótica que se ha desarrollado para lenguas indoeuropeas. Constatamos que el método ecdótico nos proporciona lineamientos generales para el tratamiento de los materiales (Pérez Priego, 1997; Sánchez Prieto, 1998); sin embargo, también pudimos comprobar que un trabajo de este tipo en lenguas indígenas implica consideraciones específicas y requiere de soluciones particulares (Villavicencio, 2001; Villavicencio [en prensa c]).[5]

En los últimos 30 años, el interés por los textos escritos en alguna lengua indígena de México —sobre todo aquellos que fueron producidos en la época novohispana— ha venido incrementándose. Lingüistas, historiadores y etnohistoriadores se han empezado a interesar por este tipo de documentos; sin embargo, la mayor parte del trabajo realizado en este campo se ha concentrado en lenguas como el náhuatl y el maya,[6] de las cuales se han empezado a conformar colecciones de textos novohispanos (Anderson, Berdan y Lochart, 1976; Karttunen y

[5] Se puede decir que no existe una tradición ecdótica en ninguna de las lenguas indígenas. Esta situación está en relación directa con una tardía valoración lingüística de los textos, la desigual disponibilidad de las fuentes existentes en las distintas lenguas indígenas, el poco conocimiento que todavía se tiene de algunas de ellas y la ausencia de una reflexión de corte metodológico sobre la edición crítica de este tipo de textos (Villavicencio [en prensa]).

[6] El mapa sociopolítico del Nuevo Mundo que caracterizó tanto en la época inmediatamente anterior al contacto como el periodo novohispano permite entender en buena medida el predominio histórico que han tenido —y conservan— algunas lenguas indígenas sobre otras y, con ello, la heterogeneidad en la disponibilidad que existe de documentos en las distintas lenguas indoamericanas.

Lockhart, 1976; Restall, 1995; Rojas *et al*., 1999). En su trabajo pionero, Karttunen y Lockhart (1976) han mostrado parte del potencial lingüístico de estas fuentes; con base en el desarrollo del léxico y de algunos fenómenos fonológicos que se observan en el conjunto de textos náhuatl que recopilan, establecen diferentes momentos sociolingüísticos para el periodo novohispano.

Nada paralelo existe aún para el purépecha;[7] hasta ahora, los mayores esfuerzos por catalogar, editar y reeditar textos en purépecha han provenido de historiadores y etnohistoriadores, quienes han reconocido la riqueza de información cultural que atesoran los vocabularios novohispanos y han recurrido a ellos para realizar estudios sobre aspectos específicos de la cultura y la sociedad de la época (Gringberg, 1989; Lucas, 1998; Martínez, 1997; Navarrete, 1997; Sepúlveda, 1988 y Warren, 1994).

En los estudios propiamente lingüísticos se observa un panorama similar al que se ve en otras lenguas indígenas. Por una parte, hasta ahora predominan los trabajos que toman como fuente de datos únicamente las *Artes* y los *Vocabularios* novohispanos. El aprovechamiento de otro tipo de documentos para el estudio de la lengua es aún incipiente. Por otra parte, el trabajo lingüístico que hasta ahora se ha realizado sobre las *Artes* y los *Vocabularios* novohispanos se ha centrado en estudios de corte historiográfico, cuya reflexión discurre sobre temas tales como el concepto del lenguaje que subyace a estas primeras descripciones, el modelo que sirvió de base para el trabajo gramatical y lexicográfico desarrollado en ellas, y el grado de adecuación de las distintas obras; en este tipo de acercamiento, el foco está puesto en los autores de las obras como representantes de una época determinada (Launey, 1997; Smith-Stark, 1995; Zimmermann, 1997).

Buena parte del trabajo que sobre las *Artes* y los *Vocabularios* purépechas se ha realizado es de corte historiográfico (Monzón, 1995, 1997a, 1997b; Villavicencio, 2003, 2006b [en prensa a]). Los trabajos que toman los textos novohispanos como fuentes de datos para un análisis propiamente lingüístico son los menos (Monzón, 2000; Villavicencio, 1996, 1999, 2000a, 2006 [en prensa b]).

[7] Eréndira Nansen inició en 1991 el proyecto "Antología de textos en lengua purépecha de los siglos XVI al XX", que al parecer ha quedado inconcluso. Como un producto de la investigación diacrónica que emprendí en 1996 sobre el sistema de caso del purépecha, en 2000 reuní una colección de padrenuestros purépechas como muestra de lo que idealmente debería ser una diacronía ejemplar, misma que permanece hasta ahora inédita (cf. Villavicencio, 2000).

El acervo documental

Cuando propuse como tema de investigación un estudio diacrónico de la lengua purépecha con base en datos documentados, el primer reto que enfrenté fue contar con la materia prima para realizarlo. No bastaban para ello las frases o los enunciados aislados que se citan como ejemplos en distintas gramáticas y trabajos lingüísticos ni el conjunto de vocablos registrados en los distintos diccionarios; la empresa filológica requería, además, trabajar textos que nos dieran una idea más clara de lo que podía ser la lengua en uso. La tarea inicial fue entonces conformar un acervo en el que se reunieran los diversos documentos que podían ser útiles para conformar el *corpus*. La integración de dicho acervo requirió dos etapas: una de recopilación y otra de clasificación.

La recopilación

El primer paso consistió en localizar y recopilar textos escritos en y sobre purépecha que se encontraban dispersos en distintos repositorios. No se trataba de obtener textos raros o inéditos ni joyas de coleccionista; lo que se requería era reunir textos de los que ya se tenía noticia y que estaban disponibles en instituciones, bibliotecas y archivos públicos y privados (Villavicencio, 1997). Por razones tanto de tiempo como de presupuesto, en esta ocasión la búsqueda se circunscribió al país; el trabajo en bibliotecas y archivos del extranjero quedó para futuras investigaciones.

El único criterio que guió esta primera fase del trabajo fue, precisamente, el uso de la lengua; es decir, la condición necesaria y suficiente para incluirlos en el universo documental era que estuvieran escritos en purépecha o trataran sobre esta lengua. Con este criterio se recopilaron documentos de índole y características diversas: Artes, vocabularios, diccionarios, gramáticas, testamentos, narraciones, títulos de tierra, recibos, confesionarios, catecismos, libros de texto, cartillas, periódicos, revistas, listas de topónimos, poemas, obras de teatro, pastorelas, autobiografías, entre otros. De esta manera se logró reunir alrededor de cien textos producidos desde los primeros años de la Colonia hasta nuestros días.

Lejos de lo que pudiera pensarse, la tarea que en un principio parecía casi mecánica resultó lenta y agotadora; los documentos se encuentran dispersos y el grado de disponibilidad que cada uno de ellos presenta es muy variado. Existen

46.indd 185 10/3/09 09:03:20

textos que han merecido ediciones recientes y que son relativamente accesibles a todo público, mientras otros están en fondos reservados y su consulta lleva tiempo y requiere permiso especial. Algunos otros, como los diversos materiales de lectura que se produjeron para el Proyecto Tarasco,[8] están esparcidos; nadie tuvo el cuidado de preservar un archivo completo de los mismos y ninguna biblioteca los tiene en su totalidad.[9] El Cuadro 1 muestra el acervo documental disponible; los textos recopilados se presentan ordenados cronológicamente.

Como puede observarse, existe documentación disponible a lo largo de casi cinco siglos, por lo que puede hablarse de un *continuum* documental; sin embargo, no todas las épocas se encuentran igualmente documentadas. La mayor parte de los textos que se pudieron localizar en ese momento correspondía a los siglos XVI (22 documentos) y XX (64 documentos); la documentación correspondiente a los siglos XVII (15 documentos) y XIX (9 documentos) resultaba, en comparación con los dos periodos anteriores, relativamente limitada. En este conjunto, el siglo XVIII (con un solo documento) representaba un hueco documental significativo.

La clasificación

Una vez reunido este conjunto, tan vasto como variado, el siguiente paso consistió en ordenar y clasificar cada uno de los documentos localizados. Esta tarea requirió considerar los diversos factores que enmarcan su producción. La gama era amplia: algunos fueron producidos directamente en la lengua indígena, otros, en cambio, eran traducciones de algún texto escrito originalmente en otro idioma. Tal es el caso de las distintas versiones que, en diferentes momentos, ha merecido

[8] En 1939 se llevó a cabo una campaña intensiva de alfabetización en el grupo purépecha que se conoce como el Proyecto Tarasco. En este proyecto se cuenta, por primera vez en México, con la participación de un grupo de destacados lingüistas encabezados por Mauricio Swadesh. El Proyecto Tarasco concibió la producción de material de lectura en la lengua indígena como apoyo al proceso de alfabetización. De esta manera se editaron no sólo cartillas sino también periódicos y folletos que trataban diversos temas de interés general para la población indígena.

[9] Cabe reconocer que en esta etapa conté siempre con la generosa ayuda de otros investigadores interesados en la cultura y la lengua purépecha, que compartieron conmigo ya un dato, ya un documento, ya una cita. Aunque mi deuda se extiende a muchos otros, mención especial merecen Claudin Chamoreau, Benjamín Lucas, Cristina Monzón, Fernando Nava, Carlos Paredes, Hans Roskamp, Gerardo Sánchez, Thomas Smith, Benedict Warren y el grupo del Proyecto Gilbertiano: Moisés Franco, Agustín Jacinto, Rosa Lucas y Pedro Márquez.

46.indd 186 10/3/09 09:03:20

46.indd 187

Cuadro 1. Disponibilidad documental por siglo

Siglo XX	**1900**	**1910**	**1920**	**1930**	**1940**	**1950**	**1960**	**1970**	**1980**	**1990**
	1912	1912		1939	1940	1951 1951	1960 1960	1972 1973	1980 1981	1990 1991
					1946	1951 1951	1961 1963	1973 1973	1982 1983	1991 1992
						1953 1954	1966 1969	1974 1975	1984 1985	1993 1993
						1955 1957	1969 1969	1975 1975	1985 1985	1994 1994
						1957		1976 1978	1985 1987	1995 1996
								1979 1979	1987 1987	1997 1997
									1987 1989	1997 1997
										1997 1998
Siglo XIX	**1800**	**1810**	**1820**	**1830**	**1840**	**1850**	**1860**	**1870**	**1880**	**1890**
				1834					1883 1886	1891
									1886 1887	1994
									1887	1996
Siglo XVIII	**1700**	**1710**	**1720**	**1730**	**1740**	**1750**	**1760**	**1770**	**1780**	**1790**
						1756				
Siglo XVII	**1600**	**1610**	**1620**	**1630**	**1640**	**1650**	**1660**	**1670**	**1680**	**1690**
		1610 1610	1625	1632		1650	1663	1679	1682	1697
		1610 1610							1689	
		1611 1614								
		1614								
Siglo XVI	**1500**	**1510**	**1520**	**1530**	**1540**	**1550**	**1560**	**1570**	**1580**	**1590**
				1539		1553 1558	1563 1563	1574 1574	1584	1590 1591
						1558 1559	1567 1567	1574 1574		1592
						1559 1569		1575 1578		
						1559				

10/3/09 09:03:20

el Padrenuestro Purépecha (Villavicencio, 2000b). Un buen número de ellos, especialmente aquellos de mayor antigüedad, eran de corte religioso: había confesionarios, catecismos y oraciones; otros podían considerarse textos de corte civil o no religioso: entre ellos estaban peticiones de justicia, testamentos, calendarios cívicos y cuentos. Los estilos también variaban: había narraciones, diálogos, poesía. Algunos eran ejemplos claros del lenguaje coloquial y otros podían inscribirse en un registro culto del idioma.

Aunque algunos textos eran anónimos, la mayoría registraba un autor. Las características de dichos autores también eran variables: para algunos el purépecha era la lengua materna, otros eran personas que aprendieron el idioma casi siempre con fines de evangelización (tal fue el caso de los frailes en la época novohispana y de los pastores protestantes en el siglo XX). También existía diversidad en lo que respecta a la procedencia geográfica de los escritores y a su nivel socioeconómico y escolar, factores que pueden implicar algún grado de variación dialectal.

Algunos documentos contaban con su respectiva traducción al español —aunque también los hay que están traducidos al inglés o al francés—, pero otros se encuentran escritos sólo en purépecha y están en espera de ser traducidos. La extensión de los textos también varía: existen documentos de unas cuantas líneas y hay otros, en cambio, que constituyen libros enteros. Algunos textos han sido editados en una o varias ocasiones, pero otros se encuentran aún inéditos; éste es el caso de la mayoría de los textos de procedencia judicial, que están en espera de ser paleografiados, impresos y difundidos.[10] Como se puede apreciar, la variedad de estilos, temas, autores, procedencia y diversidad del material reunido es amplia.

Una vez reunido el acervo, se procedió a catalogarlo. La ficha de catálogo consigna: el número progresivo del documento, el título del texto en cuestión, la fecha en la que éste fue realizado, el autor del mismo, la extensión aproximada en palabras y páginas del documento, el año de edición del mismo cuando éste difiere de la fecha de elaboración, el repositorio o el lugar donde se lo puede consultar,

[10] En la última década se ha visto un creciente interés por los documentos escritos en lengua purépecha, especialmente aquellos correspondientes a la época novohispana. Un grupo de investigadores del Centro de Estudios de las Tradiciones de El Colegio de Michoacán ha emprendido la edición de la amplia obra de Maturino Gilberti (siglo XVI). Benajmín Lucas, Cristina Monzón, Hans Roskamp y yo misma hemos localizado, paleografiado y editado, cada uno con su respectivo criterio, diversos textos novohispanos.

46.indd 188 10/3/09 09:03:20

el lugar de procedencia del documento (si se conoce), si existe una traducción disponible del texto, la obra de referencia donde se da noticia de él y, por último, un espacio para observaciones.

Documento	XVI. 001
Texto:	Título de tierras de Cherán Hatzicurin
Fecha:	1539
Autor:	
Traductor:	Pablo Velásquez
Extensión:	546 pal., 6 pp.
Edición:	1952
Ubicación:	AGN. Tierras, vol. 8778
Procedencia:	Tzintzuntzan
Traducción:	Pablo Velásquez
Referencia:	*Tlalocan* 3, 3 (1952) 238 – 245
Observaciones:	Probablemente fecha apócrifa. Muchos documentos de límite de tierras eran falsificados por las comunidades para reclamar sus derechos sobre la tierra. (Hans Rskamp, comunicación personal).

Documento	XVII.010
Texto:	Libro de cuentas y diezmos
Fecha:	1663-1664
Autor:	
Traductor:	
Extensión:	
Edición:	
Ubicación:	Parroquia de San Jerónimo de Aranzan Leg. 141 bis, 1610-1699, exp. 46 fs.
Procedencia:	Aranzan
Traducción:	no
Referencia:	Dr. Carlos Paredes CIESAS-UMSNH, SXVII/0013/CP
Observaciones:	

La constitución del *corpus*

Una vez determinado el universo documental disponible, el siguiente paso consistió en seleccionar, a partir de éste, el *corpus* propiamente dicho; es decir, la serie de documentos que se analizarían. Dicha selección se realizó tomando en cuenta tres criterios: la fecha de elaboración del documento, el tipo de texto del que se trataba y la disponibilidad del mismo. El primer criterio permitió establecer los estados de lengua que se reconocerían en esta investigación y, con ello, las fechas a las que deberían corresponder los documentos que se seleccionarían; los dos últimos criterios guiaron la selección de los documentos específicos que integrarían el *corpus*.

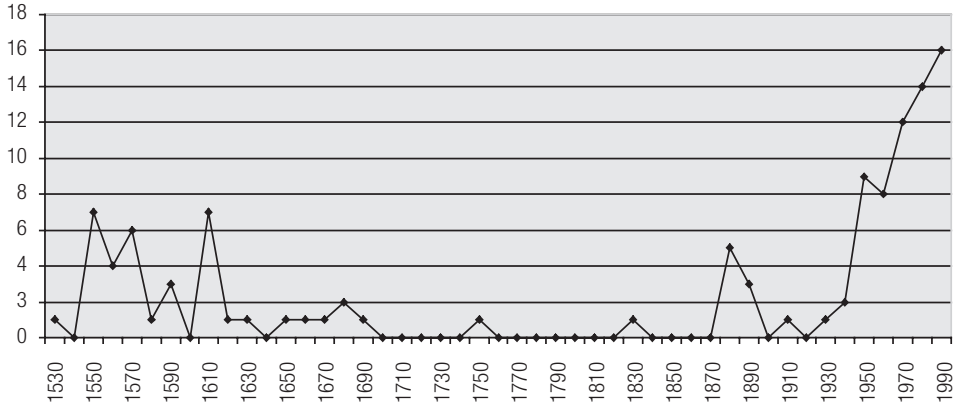
Los estados de lengua

La determinación de los estados de lengua pertinentes para el estudio es uno de los problemas medulares en el análisis. El marco adoptado tiene que ver con la naturaleza de los límites; para un enfoque que considera la gramaticalización como un *continuum* evolutivo, cualquier intento de segmentación de dicho *continumm* en unidades discretas resulta arbitrario (Heine y Reh, 1984; Hopper y Traugott, 1993); el lingüista debe recordar siempre que se trata, en todo caso, de un recurso analítico que en todo momento está a prueba. Describir el desarrollo de un proceso de gramaticalización requiere observar el fenómeno en cuestión en sucesivos estados de lengua, para constatar su evolución. La pertinencia de los estados de lengua que el lingüista, con base en su experiencia y conocimiento de la lengua, propone reconocer se confirma, o no, con el análisis mismo.

En nuestro caso, el análisis del acervo disponible permitió delimitar, en un primer momento, los periodos para los que se tenía una mayor concentración de documentos. Al establecer la densidad cronológica del universo documental se pudo apreciar que la mayor parte de los textos se agrupaba en la segunda mitad de cada siglo. Este hecho sugirió hacer cortes cada cien años, con lo cual se podía hablar de cinco periodos localizados respectivamente en la segunda mitad de los siglos XVI, XVII, XVIII, XIX y XX. Los textos producidos en la primera década del siglo XVII podían agruparse con los de la segunda mitad del siglo XVI; de la misma manera, los textos producidos en la primera década del siglo XX podían considerarse junto con los textos de la segunda mitad del siglo XIX. El Cuadro 2 muestra la

densidad cronológica del universo documental y permite apreciar aquellos periodos que se caracterizan por una mayor disponibilidad de documentación.

Cuadro 2. Densidad cronológica del universo documental



Sin embargo, no todas las etapas eran comparables cuantitativamente; como se puede apreciar en la gráfica anterior, los periodos más prolíficamente documentados fueron aquellos que se encontraban en los dos extremos del *continuum*, es decir, el siglo XVI y el siglo XX. Un análisis diacrónico ideal tendría que considerar varios estados sucesivos de la lengua en estudio, pero la escasa documentación disponible en los siglos XVII y XIX, y sobre todo la ausencia casi total de textos en el siglo XVIII, nos llevó a considerar la pertinencia de optar por un análisis contrastivo que tomara en cuenta sólo aquellos periodos que en el acervo tenían la documentación más nutrida. Se reconocieron entonces dos estados de lengua: el purépecha de la segunda mitad del siglo XVI y el purépecha correspondiente a la segunda mitad del siglo XX. Convencionalmente, para fines de la investigación, se denominó al primero PURÉPECHA NOVOHISPANO y al segundo PURÉPECHA CONTEMPORÁNEO. El Cuadro 3 representa los estados de lengua que se determinó contrastar.[11]

[11] No existe aún un estudio que proporcione los elementos necesarios para poder caracterizar algo como PURÉPECHA NOVOHISPANO o PURÉPECHA CONTEMPORÁNEO y delimite los periodos específicos que cada uno de ellos abarcaría.

Cuadro 3. Estados de lengua

El *corpus* se integró con materiales correspondientes al purépecha novohispano —segunda mitad del siglo XVI— y al purépecha contemporáneo —segunda mitad del siglo XX. El documento más antiguo está fechado en 1553 y el más reciente, en 1998. El lapso comprendido entre ellos es de más de cuatrocientos años lo que asegura una distancia suficiente para poder percibir cambios lingüísticos significativos. De esta manera, si bien el tema de estudio y el marco teórico determinaron en un principio la metodología a seguir y, con ello, la naturaleza del *corpus*, fue la disponibilidad de documentos registrada en cada uno de los periodos lo que condujo a tomar una decisión teórica sobre el tipo de trabajo analítico que se habría de seguir.

La selección de documentos

Una vez determinados los estados de lengua pertinentes para el estudio, se debía elegir, entre la amplia variedad de textos correspondientes al siglo XVI y al siglo XX, aquellos que se someterían a análisis. En esta etapa, el tipo de documento y la disponibilidad del mismo fueron los criterios elegidos para guiar la selección. De acuerdo con el primero se eligieron tres tipos de documentos: *gramaticales*, *léxicos* y *textuales,* ya que en los dos estados de lengua que se trabajarían se tenían ejemplos de todos ellos. Para los documentos de tipo textual se realizó, además, una subclasificación que permitió reconocer, *grosso modo*, dos *tipos textuales*: uno religioso y otro civil o no religioso. El criterio de disponibilidad atendió, a su vez, dos aspectos: la posibilidad que existía de acceder al documento original así como la disponibilidad de una traducción al español del documento en cuestión. El primer aspecto se observó de manera estricta; el segundo se tomó como un índice de preferencia, pero no fue excluyente.

46.indd 192 10/3/09 09:03:22

La observación estricta del primer aspecto tiene que ver con la naturaleza misma de la investigación. El cambio mediante el cual se crean nuevas categorías se caracteriza por ser progresivo y gradual; una consecuencia de lo anterior es que frecuentemente los elementos antiguos conservan su forma y su función por largos periodos de tiempo y conviven con las nuevas formas durante amplios lapsos. En este marco, el concepto de pauta y la frecuencia de uso son útiles como indicadores para determinar del grado de gramaticalización que alcanzan los nuevos elementos en los distintos periodos de la lengua (Bybee y Pagliuca, 1985); en ellos se apoyan las decisiones de análisis en los casos dudosos (Company, 2001).

Entender la gramática como un conjunto de rutinas y hábitos lingüísticos permite caracterizar estados de lengua con base en el empleo más o menos generalizado de ciertos patrones lingüísticos. Desde esta perspectiva, aquellos usos que se apartan de la pauta que se considera propia de una determinada época resultan de particular interés para el estudioso del cambio; es precisamente en estos lugares en donde suele iniciarse un fenómeno de esta naturaleza. El concepto de pauta va estrechamente ligado a la utilización de la frecuencia de uso como índice de cambio; se espera que el grado de gramaticalización que un elemento tiene en un determinado estado de su historia guarde una relación directa con la frecuencia de aparición que dicho elemento registra en ese periodo. Un alto grado de gramaticalización conlleva una distribución más general lo que, a su vez, da origen a una frecuencia de aparición más alta.

La accesibilidad al documento original, como se puede apreciar, resultaba medular para nuestros propósitos; se debía poder asegurar la fidelidad del dato. En éste, como en otros acercamientos al complejo fenómeno de la lengua en uso, el *cómo se dice* importa tanto como el *qué se dice*. En el caso de los estudios de gramaticalización, asegurar la fidelidad del dato permite que sea el lingüista el que pondere el peso específico de cada variación que se registra en un determinado documento y época, y tome una decisión cuando existe la posibilidad de análisis alternativos. Ya que en algunas ediciones y reimpresiones los criterios de edición que se han seguido no se precisan con claridad y en otras, dichos criterios no se siguen de manera consistente, se decidió integrar el *corpus* únicamente con versiones originales o ediciones facsimilares.

El segundo aspecto, referente al criterio de disponibilidad, lo constituyó la existencia o no de una versión paralela en español del documento elegido. Esto tiene que ver con un tema no menor, el de la traducción —directamente ligado

con el problema de la significación. No es éste el lugar para resolver un tema de tal envergadura, sin embargo, es pertinente reconocer que cuando se trabaja con una lengua que no es la materna —como es mi caso— el problema de la traducción adquiere una relevancia particular. La traducción desempeña siempre un papel ambivalente; por una parte, representa un apoyo de primer orden para el análisis —más básico cuanto menos familiaridad se tiene con la lengua—; por otra parte, constituye una especie de velo o lente que puede distorsionarlo.

Como todo acercamiento al lenguaje, el estudio de una lengua no materna requiere del investigador una escucha específica; ésta consiste en atender cada vez más la lengua misma y menos la traducción. Se puede decir que en este caso el reiterado trabajo con el material va desarrollando una escucha que se hace cada vez más "interna"; conforme avanza el análisis y con ello el conocimiento del idioma mismo, el lingüista está en posibilidad de dar a la traducción su justa dimensión.

En nuestro caso, contar con una traducción al español del texto elegido resultó ser un elemento importante en una etapa de la investigación en la cual el conocimiento de la estructura propia de los dos estados de lengua establecidos era, a todas luces, escaso. Aunque disponemos de estudios gramaticales del purépecha desde los primeros años de la Colonia, los estudios específicos sobre sintaxis han empezado a desarrollarse sólo a partir de la última década del siglo XX; la mayor parte de las obras precedentes presta poca atención a esta parte de la gramática, de la que se establecen sólo los lineamientos generales.

Dentro de este panorama general caracterizado por la parquedad del trabajo con las estructuras sintácticas no existían descripciones específicas sobre el sistema de casos que presenta la lengua. El tema había merecido hasta entonces muy poca atención y, sobre todo, no se había tratado desde el punto de vista diacrónico. En este contexto, la traducción al español constituía, en un primer momento, una guía importante para el análisis inicial de los textos, especialmente de aquellos correspondientes al purépecha novohispano.

De acuerdo con los criterios anteriores, para integrar el *corpus* se prefirieron aquellos documentos que correspondían a alguno de los tipos elegidos, eran accesibles ya sea en su versión original o en una edición facsimilar y para los cuales se contaba con un texto paralelo en español. El *corpus* quedó entonces integrado de la siguiente manera:[12]

[12] Una relación detallada del *corpus* puede consultarse en Villavicencio, 2002.

a) *Documentos de tipo gramatical*. Se consideraron aquellos trabajos que tratan algún aspecto de la gramática del purépecha. El conjunto de documentos gramaticales del siglo XVI incluyó las dos gramáticas que se elaboraron en este periodo. Para el siglo XX se consultaron los trabajos lingüísticos que se han producido en los últimos cincuenta años.

b) *Documentos de tipo léxico*. Incluyeron vocabularios y diccionarios bilingües. Se trabajaron todas las obras de carácter lexicográfico, tanto del siglo XVI como del siglo XX, que han llegado hasta nosotros.

c) *Documentos de tipo textual*. Se consideraron escritos tanto de corte religioso como civil. Para el siglo XVI se trabajaron dos textos religiosos y uno civil. Para el siglo XX se trabajó una antología que integra narraciones y cuentos de varios autores purepechahablantes y fragmentos del Nuevo Testamento.

El conjunto de documentos que integraron el *corpus* textual constó aproximadamente de 3,000 palabras por cada uno de los estados de lengua. La parte correspondiente a la época novohispana está conformada en su mayoría por textos de corte religioso, ya que estos fueron los más accesibles en ese momento; el único texto de corte civil que se incluyó representa aproximadamente 16% de la extensión total de datos disponibles para el siglo XVI. En la parte correspondiente a la época contemporánea, los textos religiosos representan aproximadamente 50% del total analizado para este periodo. El Cuadro 4 muestra la integración del *corpus* textual.

Cuadro 4. Integración del corpus textual

	E 1		**E 2**	
Textos religiosos	Gilberti	1,153 palabras (38%)	Lathrop	1,543 palabras (49.8%)
	Lagunas	1,378 palabras (45.4%)		
Textos civiles	AGN	507 palabras (16.6%)	Soto	1,556 palabras (50.2%)
Total		3, 038 palabras (100%)		3,099 palabras (100%)

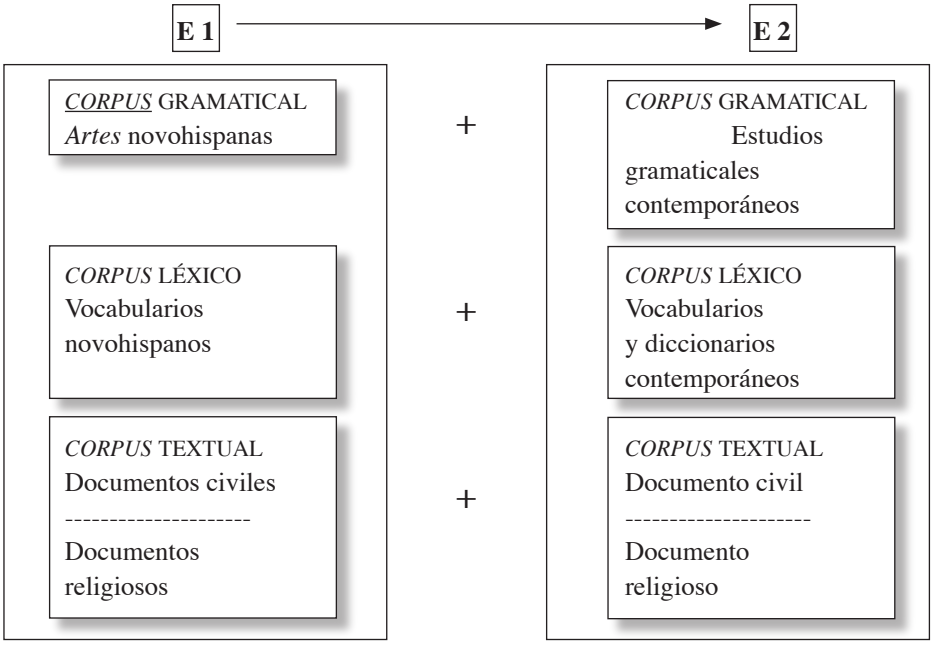
Idealmente, un *corpus* de índole diacrónica debería de poder contar con materiales totalmente comparables para cada periodo estudiado; sin embargo, en

el caso del purépecha, este ideal está aún lejos de poderse alcanzar. La principal limitación para lograrlo es la escasa documentación disponible hasta ahora, sobre todo para los periodos intermedios correspondientes a los siglos XVII, XVIII y XIX. El laborioso trabajo de análisis que implicó cada uno de los textos que lo integraron (véase la sección correspondiente al Análisis lingüístico), aunado al escaso conocimiento que hasta ahora se tiene de algunos aspectos de la gramática purépecha en general y de la época novohispana en particular, mostró que era preferible restringir el análisis a un *corpus* textual limitado pero significativo con el fin de ganar profundidad y, a partir de ahí, postular la dinámica de los fenómenos de cambio que se observan. La extensión del *corpus* textual con el que trabajé, no obstante ser modesta, mostró ser suficiente para los propósitos de la investigación; los datos mostraron recurrencias significativas que permiten sustentar el análisis.

El *corpus* en su totalidad quedó integrado con dos subconjuntos, uno por cada uno de los estados de lengua determinados, el siglo XVI y el siglo XX. Cada estado de lengua, a su vez, quedó compuesto por tres tipos de documentos: documentos de tipo gramatical, documentos de tipo léxico y documentos de tipo textual. Por último, en los documentos de tipo textual se incluyeron documentos de corte civil y documentos de corte religioso. La conformación total del corpus se puede observar en el Cuadro 5.

Para integrar el *corpus textual* se eligieron algunos fragmentos de las extensas obras religiosas. La cuestión de la elección en este último nivel tuvo que ver con el asunto de la pertinencia. En este punto, la sensibilidad del investigador, respaldada por el conocimiento profundo y puntual del material y el desarrollo de una escucha cada vez más "interna" de la lengua que se estudia, fue el instrumento básico que guió esta selección. El desarrollo de una escucha atenta y sensible al material puede concebirse como una cuestión de educación y formación del analista (Carbó, 2001b:58), o bien como un sentido práctico basado en la sagacidad, el sentido común y la amplitud de perspectivas (Company, 2001); en todo caso, creo yo, se trata de una habilidad que se adquiere con la práctica del oficio; una habilidad que, como todas, requiere tiempo. Por último, el conocimiento profundo y puntual del material que conforma el corpus y el de la lengua que se estudia tiene como principal vía de acceso un trabajo de análisis lingüístico fino.

Cuadro 5. Conformación del *corpus*



Pruebas de control

Además del *corpus* propiamente dicho se utilizó, como pruebas de control, otro conjunto de datos tanto orales como escritos. Las pruebas de control de tipo oral fueron obtenidas en elicitación directa con un hablante.[13] Para las pruebas de control de tipo escrito se utilizaron los datos obtenidos en calas de otros textos tanto de la época novohispana como del siglo XX, que no fueron considerados dentro del *corpus textual* propiamente dicho.

[13] El señor Agustín Tellez† de la comunidad de Tiríndaro, municipio de Zacapu, Michoacán.

El tratamiento de los materiales

La transcripción

En este nivel, nuevamente se aplicó el criterio de fidelidad; en la transcripción de los documentos que conforman el *corpus*, el original fue respetado fielmente. Como ya se ha dicho, para el tipo de estudio que nos propusimos todo uso lingüístico, por extraño que parezca, representa un valioso dato para el analista que tiene que tomar decisiones sobre fenómenos poco diáfanos, como suelen ser los procesos de gramaticalización. El cumplimiento de este criterio es particularmente importante para lenguas poco estandarizadas como el purépecha, en el que casi cada autor presenta una ortografía distinta. Conservar la fidelidad al texto es, por una parte, la única manera de preservar la posibilidad de estudiar una serie de fenómenos lingüísticos de los cuales hasta ahora se sabe muy poco y, por otra, las variaciones ortográficas atestiguadas en los textos también son índices para determinar la época de la que data un documento y el autor del mismo.

La transcripción de los textos del siglo XVI siguió los lineamientos más comunes del ejercicio paleográfico: se recuperaron las letras que en el texto se expresan por medio de tildes [~], así la palabra *hingũ* fue transcrita como *hingun* y la palabra *q̃* se desató como *que*. Asimismo se respetaron los usos gramaticales y la ortografía del español que se documenta en los textos.

El análisis lingüístico

En el marco de gramaticalización, el cambio lingüístico se concibe como un proceso multifactorial que típicamente incluye cierto desgaste fonológico, una mayor generalización de las marcas y un menor grado de especificidad semántica; si bien estos factores se encuentran estrechamente relacionados, el peso relativo que cada uno de ellos tiene en los distintos procesos de gramaticalización es diferente. Se espera que esta correlación se modifique en cada uno de los estados de lengua; en todo caso se trata de una cuestión de grado y no de límites absolutos. La modificación en la forma de un elemento, la función que éste desempeñe, el significado que tenga así como los contextos de aparición del mismo, el grado de independencia que manifieste y su frecuencia de aparición son índices significativos para el análisis.

46.indd 198 10/3/09 09:03:24

En este contexto, el análisis lingüístico al que son sometidos los materiales del *corpus* adquiere una relevancia especial: cuanto más fino sea éste, más y mejores elementos se tendrán para caracterizar el cambio. Con esta perspectiva, el *corpus textual* en su totalidad así como partes seleccionadas del *corpus léxico* y aquellos ejemplos provenientes de los diversos estudios gramaticales que se consultaron fueron sometidos a un análisis fino. Éste implicó un trabajo meticuloso que siguió las normas de la tradición existente en lenguas indígenas y que procedió en niveles reconociendo párrafos, oraciones, sintagmas y morfemas.

En los documentos que integran el *corpus textual*, dividí cada texto en oraciones que fueron numeradas de manera consecutiva para facilitar su localización. El análisis se realizó en un formato interlineado que considera cuatro niveles. En la primera línea se copia el texto tal como aparece en el original; para este nivel se utilizan cursivas. En la segunda línea se realiza el análisis morfémico de cada palabra; para ello se utiliza ortografía práctica, lo que permite trabajar en este nivel con una misma convención ortográfica en los dos estados de lengua. En la tercera línea se da la glosa correspondiente a cada uno de los morfemas reconocidos. Por último, la cuarta línea corresponde a la traducción al español de la oración analizada.

Respeté la traducción del autor del documento cuando ésta existía. Cuando el texto carecía de traducción o cuando la que se daba era muy libre, proporcioné entre corchetes [ ] mi propia traducción. En la traducción proporcionada por el autor encierro entre paréntesis ( ) aquellos elementos que no se derivan automáticamente del análisis morfémico, y entre paréntesis agudos < > aquellos elementos que, aunque ocurren en el análisis morfémico, no aparecen en la traducción que da el autor. El formato completo queda de la siguiente manera:

SIGLO XVI

ka ementahco *varuquevati*
ka eménda =k'o uarhú –ke -ua -ti
Y tiempo.de.aguas =únicamente pescar –PAS –FUT –3IND
'[y únicamente en tiempo de aguas se pescará]'

ambequet cuhche thirequa atangaca
ambáketi =kuch'e t'irékua a -tá -nha -ka
bueno =1PL comida comer-TR–PAS-1/2IND

‘luego fuimos recebidos con muy buena comida’
[se nos dio de comer buena comida]

SIGLO XX

imanijtunha sontku andanguni tata janikua
imá –ni =t’u =nha sóntku andángu –ni tatá janíkua
DEM-OBJ=también=REF pronto alcanzar –INF REV lluvia
‘fue alcanzado prontamente por el camino por “tata Janikua”
‘[(se dice que) también a ése pronto lo alcanzó el padre lluvia]’

Elisabeti ku Väpka *Mariani*
Elisabeti kurhá -a -p -ka maría–ni
Elisabet escuchar-vientre-PRT-SUBJ María-OBJ
‘Cuando oyó Elisabet la salutación de María’
[elisabet escuchó <en su vientre> a María]

Como puede observarse, cada texto requirió de un laborioso trabajo de análisis, que si bien consumió buena parte del tiempo de la investigación, constituyó el acceso analítico básico al fenómeno que se estudiaba, toda vez que el sistema casual transcurre predominantemente en el nivel sintáctico-morfológico. Por otro lado, el análisis lingüístico fino de textos completos y fragmentos discursivos amplios de obras extensas permitió observar, además del sistema de casos, una serie de fenómenos gramaticales interesantes tanto desde el punto de vista sincrónico como diacrónico; en este aspecto constituye una excelente base para futuras investigaciones.

Conclusiones

He intentado mostrar la complejidad de un proceso que pocas veces se hace explícito de manera detallada en cada una de sus etapas y niveles: la conformación de un *corpus* para trabajo lingüístico, en este caso específico, para el estudio histórico de un fenómeno de cambio lingüístico en una lengua con poca tradición de escritura, como lo son la mayoría de las lenguas indígenas de México. El ejercicio tuvo como finalidad hacer evidente un momento de la investigación que, con sus particularidades, es compartido por otros campos de la disciplina. El

proceso mediante el cual el analista estructura una serie de decisiones que toma con base en criterios de diferente índole y nivel, y que le permite convertir un determinado material documentado en dato pertinente para el cambio lingüístico, es semejante a otros enfoques que se interesan, si bien desde otro punto de vista, en dar cuenta de la complejidad de la lengua en uso.

La especificidad de la mirada, en este caso, fue el producto de tres factores: el tipo de evidencia con la cual se trabaja, es decir, la manifestación material del fenómeno (textos escritos en lengua indígena), el grado de conocimiento —o desconocimiento— del objeto de estudio (la lengua purépecha y, de manera específica, el sistema de casos de dicha lengua), y el punto de vista desde el cual se mira el objeto (el marco de gramaticalización y la noción de cambio lingüístico). Los criterios adoptados en las diversas etapas del proceso de constitución del *corpus* atendieron a esta especificidad

Cabe destacar que en éste, como en otros acercamientos que pretenden trabajar con una muestra de la lengua en uso, el problema de la fidelidad al dato resultó central a la discusión. En gramaticalización, la fidelidad al dato tiene que ver con la posibilidad del analista de acceder a la forma exacta que se documenta en los textos, y permitirle con ello ponderar el papel que desempeña cierto uso lingüístico en un estado de lengua determinado. El tema de la fidelidad al dato fue considerado en dos niveles: en la selección de los materiales y en el tratamiento de los mismos. En la selección de los materiales, la necesidad de asegurar el acceso del analista al documento original llevó a seleccionar únicamente documentos disponibles en su versión original o en ediciones facsimilares. En cuanto al tratamiento de los materiales, garantizar la fidelidad del dato implicó aplicar un análisis lingüístico fino a todo el *corpus* textual, así como a partes seleccionadas del *corpus* léxico y gramatical.

Por último, el trabajo con una lengua que no es la materna puso en relieve un problema no menor: el del papel de la traducción y su carácter ambivalente. En cierto sentido, la cuestión de la fidelidad al dato es un problema paralelo al de la traducción y emerge en los mismos niveles. En el momento de seleccionar los materiales se prefirieron aquellos textos que contaban con una traducción al español. El momento del análisis fino, en cambio, demandó del investigador ser especialmente cuidadoso y asegurarse, en la medida de lo posible, de estar trabajando con la estructura misma de la lengua y no con los reflejos que de ella se desprenden en una traducción. Cabe decir, para finalizar esta disertación —aunque de

ninguna manera la discusión—, que el trabajo con una lengua que no es la materna requiere del analista una escucha especial que conforme avanza la investigación se hace cada vez más interna; el investigador aprende gradualmente a escuchar la lengua indígena misma y no los ecos que ella produce en las traducciones.

Bibliografía

ANDERSON, A. J., F. BERDAN & J. LOCKHART (1976). *Beyond the codices*. Berkeley/Los Angeles: Berkeley University Press.

BENVENISTE, É., (1978). Las transformaciones de las categorías lingüísticas. En *Problemas de lingüística general*, pp. 130-140, vol. 2. México: Siglo XXI.

BRISEÑO, F. (2000a). La gramaticalización del verbo "terminar": Ts'o'okol en maya yucateco. *Lingüística Mexicana*, 1 (1): 79-91.

——— (2000b). Gramaticalización del verbo "ir" en las lenguas mayas de la rama yukateka. En M. del C. Morúa Leyva y G. López Cruz (eds.). *V Encuentro internacional de lingüística en el Noroeste. Memorias*, pp. 85-102, tomo 1, vol. 1. Hermosillo: UniSon.

——— (1998). La gramaticalización del verbo *bin* (ir) en el maya yucateco. En Z. Estrada Fernández, M. Figueroa Esteva, G. López Cruz y A. Acosta Félix (eds.). *IV Encuentro de lingüística en el Noroeste. Memorias*, pp. 69-93, tomo 1, vol. 1, Hermosillo: UniSon.

———(1997). *De gramaticalización y degramaticalización. Dos procesos en el maya yucateco actual*. (Tesis de maestría, ENAH).

BYBEE, J. L. (1985). *Morphology. A study of the relation between meaning and form*. Amsterdam/ Philadelphia: John Benjamins.

BYBEE J. & W. PAGLIUCA (1985). Cross linguistic comparison and the development of grammatical meaning. En J. Fisiak (ed.) *Historical semantics, historical word formation. Trends in Linguistics, Studies and Monographs, 29*.

CARBÓ, T. (2001a). El cuerpo herido o la constitución del *corpus* en análisis del discurso. *Escritos. Revista del Centro de Ciencias del Lenguaje*, 23: 17.

——— (2001b). Tocar el lenguaje con la mano: experiencias del método. *ALED Revista Latinoamericana de Estudios del Discurso*, 1 (1): 43-67.

CLAUDI, U. & H. BERNARD (1986). On the metaphorical base of grammar. *Studies in Language*, 10: 297-335.

COMPANY, C. (2001). Para una historia del español americano. La edición crítica de documentos coloniales de interés lingüístico. En L. Funes y J. L. Moure (eds.). *Studia in honorem Germán Orduna*, pp. 207-225. Alcalá de Henares: Universidad de Alcalá de Henares.

——— (1991). *La frase sustantiva en el español medieval. Cuatro cambios sintácticos*. México: IIF, UNAM (Publicaciones del Centro de Lingüística Hispánica, 32).

CHAMOREAU, C. (2006). En busca de un verbo "ser" en purépecha. Cadena de gramatización y gramaticalización en cadena. En R. M. Ortiz (ed.). *VIII Encuentro internacional de lingüística en el Noroeste. Memorias*, pp. 65-84, tomo 2, Hermosillo: UniSon.

GIVÓN, T. (1971). Historical syntax and synchronic morphology. An archaeologist's field trip. *Chicago Linguistic Society*, pp. 394-415.

GRINGBERG, D. M. KRASNOPOLSKY DE (1989). Técnicas metalúrgicas tarascas. *Ciencia y Desarrollo*, 15 (89): 37-52.

HEINE, B., U. C. & F. HÜNNEMEYER (1991). *Grammaticalization. A conceptual framework*. Chicago/London: University of Chicago Press.

HEINE, B. & T. KUTEVA (2002). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

HEINE, B. & M. REH (1984). *Grammaticalization and reanalysis in African languages*. Hamburgo: Helmut Buske.

HOPPER, P. (1987). Emergent grammar. *Berkeley Linguistics Society*, 13: 39-57.

HOPPER, P. J. & E. C. TRAUGOTT (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

JAKOBSON, R. (1984). El lenguaje común de antropólogos y lingüistas. En *Ensayos de lingüística general*, pp. 13-33. Barcelona: Ariel (Letras e ideas).

——— (1971). Boas's view of grammatical meaning. *Word and language. Selected writings*, pp. 489-496, vol. 2. La Haya/París: Mouton.

KARTTUNNEN, F. & J. LOCKHART (1976) *Nahuatl in the middle years. Language contact phenomena in texts of the Colonial Period*. Berkeley: University of California Press.

KURYŁOWICZ, J. (1965). The evolution of grammatical categories. *Diógenes*, 55: 55-71.

LAUNEY, M. (1997). La elaboración de los conceptos de la diátesis en las primeras gramáticas del náhuatl. En K. Zimmermann (ed.). *La descripción de las lenguas amerindias en la época colonial*, pp. 21-41. Frankfurt/Main/Madrid: Vervuert-Iberoamericana.

LEHMANN, C. (1982). *Thoughts on grammaticalization. A programmatic sketch*. Vol. 1 (Arbeiten des Kölner Universalien – Projekyd 48). Köln: Universität zu Köln. Institut für Sprachwissenschaft.

LI, C. (ed.) (1975). *Word order and word order change*. Austin/Londres: University of Texas Press.

LUCAS JUÁREZ, B. (1998). Vocabulario p'urhépecha - español con terminología constructiva y afín usada en la época colonial. En C. Paredes Martínez (coord.). *Arquitectura y espacio social*

entre los purépechas durante la época colonial, pp. 371-384. Morelia: Universidad Michoacana de San Nicolás de Hidalgo/CIESAS/Universidad de Keiko, Japón.

MALKIEL, Y. (1967). Multiple versus simple causation in linguistic change. *To honor Roman Jakobson*, pp. 1128-1246, vol. 2. La Haya/París: Mouton.

MARTÍNEZ BARACS, R. (1997). El vocabulario en lengua de Mechuacan (1559) de Fray Maturino Gilberti como fuente de información histórica". En C. Paredes Martínez (coord.). *Lengua y etnohistoria purépecha. Homenaje a Benedict Warren*, pp. 67-162. Morelia: Universidad Michoacana de San Nicolás de Hidalgo/CIESAS (Encuentros, 2).

MEILLET, A. [1912], (1948). L'evolution des formes grammaticales. *Linguistique historique et linguistique générale*, pp. 130-149, vol. 1. París: Edouard Champion.

MITHUN, M. (1991). The role of motivation in the emergence of grammatical categories: the grammaticization of subjects. En E. Traugott y B. Heine (eds.). *Approaches to grammaticalization*, pp. 159-184. Amsterdam: John Benjamins.

MONZÓN, C. (2002). El clítico topicalizador =ni en el tarasco del siglo XVI. En Z. Estrada Fernández y R. M. Ortiz (eds.). *VI Encuentro internacional de lingüística en el Noroeste*, tomo 1, pp. 291-311. Hermosillo: UniSon.

——— (1997a). La influencia de Nebrija en la gramática phurhépecha de Gilberti. En I. Guzmán Betrancourt y E. Nansen (eds.). *Memoria del coloquio La obra de Antonio de Nebrija y su recepción en la Nueva España. Quince estudios nebricenses (1942-1992)*, pp. 107-119. México: INAH.

——— (1997b). Terminología y análisis de la estructura morfológica en el *Arte en Lengua Michoacana* de fray Juan Baptista de Lagunas (siglo XVI). En K. Zimmermann (ed.). *La descripción de las lenguas amerindias en la época colonial*, pp. 133-145. Frankfurt/Main/Madrid: Vervuert-Iberoamericana.

——— (1995). La morfología de las primeras *Artes* de lenguas amerindias. *Amerindia*, 19/20: 253-261.

NAVARRETE PELLICER, S. (1997). La tecnología agrícola tarasca del siglo XVI. En C. Paredes Martínez (coord.). *Historia y sociedad. Ensayos del seminario de Historia colonial de Michoacán*, pp. 74-223. Morelia: Instituto de Investigaciones Históricas, Universidad Michoacana de San Nicolás de Hidalgo/CIESAS (Encuentros, 3).

NICHOLS, J. (1986). Head-marking and dependent-marking grammar. *Language,* 62 (1): 56-119.

PÉREZ PRIEGO, Á. M. (1997). *La edición de textos*. Madrid: Síntesis.

RESTALL, M. (1995). *Life and death in Maya community. The Ixil Testaments of the 1760s*. Lancaster, CA: Labyrinthos.

ROJAS RABIELA, T., E. L. REA LÓPEZ & C. MEDINA LIMA (1999). *Vidas y bienes olvidados. Testamentos indígenas novohispanos*. 3 tomos. México: SEP/CONACYT/CIESAS.

SÁNCHEZ-PRIETO, P. (1998). *Cómo editar los textos medievales. Criterios para su presentación gráfica*. Madrid: Arco Libros.

SAPIR, E. (1980). *El lenguaje. Introducción al estudio del habla*. México: FCE.

SEPÚLVEDA, M. T. (1988). *La medicina entre los purhépechas prehispánicos*. México: IIH, UNAM.

SMITH-STARK, T. (1995). Apuntes sobre lexicografía novohispana [manuscrito]

SWEETSER, E. E. (1988). Grammaticalization and semantic bleaching. *Berkeley Linguistic Society*, 14: 389-405.

TIMBERLAKE, A. (1977). Reanalysis and actualization in syntactic change. En C. Li (ed.). *Mechanisms of syntactic change*, pp. 141-177. Austin: University of Texas Press.

TRAUGOTT, E. C. (1985). On regularity in semantic change. *Journal of Literary Semantics*, 14 (3): 155-173.

VILLAVICENCIO, F. [en prensa a]. Crónica de una tradición misionera. Lexicografía michoacana en el siglo XVI. En A. Acosta y Z. Estrada (coords.). *Estudios sobre lexicografía*. Hermosillo: Unison.

——— [en prensa b]. Léxico purépecha en tiempos de la Ilustración. Un acercamiento a los apartados léxicos del Cathesismo de Botello [1756]. En O. Zwartjes y R. Arzápalo. *Lingüística misionera*. Amsterdam: John Benjamins.

——— [en prensa c]. De la paleografía a la edición crítica. ¿Una ecdótica para lenguas indígenas? En C. Company, L. Godinas y B. Clark (coords.). *Ecdótica novohispana*. México: UNAM/El Colegio de México.

——— (2006a). *P'orhépecha kaso sïratahenkwa. Desarrollo del sistema de casos del purépecha*. México: CIESAS.

——— (2006b). El Dictionarito en lengua de Michuacan [1574] ¿Primer calepino novohispano? En P. Máynez y M. R. Dosal (eds.). *V Encuentro internacional de lingüística en Acatlán*, pp. 357-381. México: Facultad de Estudios Superiores Acatlán, UNAM.

——— (2003). El Catecismo de Joseph Zepherino Botello, un eslabón en la historia del purépecha. En Joseph Zepherino Botello [1756]. *Cathecismo breve en lengua tarasca y recopilación de algunos verbos los mas commuines para el uso de la misma lengua*, pp. 25-48. Tomo 1: edición facsimilar, tomo 2: análisis. Estudio introductorio por Frida Villavicencio, prefacio biográfico y paleografía del texto por Benedict Warren. Morelia: Fimax [Fuentes de la Lengua Tarasca o Purépecha VII].

——— (2002). *Estructura y cambio del sistema de casos del purépecha. Del siglo XVI al siglo XX*. (Tesis de doctorado, El Colegio de México).

——— (2001). Rescate documental para investigación diacrónica en lenguas indígenas. En B. Clark de Lara y F. Curiel Defossé. *Filología mexicana*, pp. 435-466. México: UNAM.

——— (2000a). La gramaticalización del acusativo purépecha. En Z. Estrada e I. Barreras (coords.). *V Encuentro internacional de lingüística en el Noroeste*. Tomo 1, pp. 273-296. Hermosillo: UniSon.

——— (2000b). El Padrenuestro purépecha. Una diacronía ejemplar. Ponencia presentada en el *III Coloquio de lingüística en la ENAH*. México. ENAH.

——— (1999). Palabras nuevas para conceptos nuevos. Un asomo a la neología de la lengua de Michoacán. En B. Skinfill y A. Carrillo (coords.). *Estudios michoacanos 8*, pp. 257-289. Zamora: El Colegio de Michoacán.

——— (1997). Fuentes documentales para el estudio del purépecha. Ponencia presentada en las *IV Jornadas gilbertianas*. Zamora: El Colegio de Michoacán.

——— (1996). Un caso en diacronía, la gramaticalización del genitivo purépecha. *Dimensión Antropológica*, 3 (6): 61-97.

WARREN, B. (1994). La incorporación de conceptos europeos en la lengua de Michoacán. Ponencia presentada en la *XXIII Mesa redonda de la Sociedad Mexicana de Antropología*.

WISCHER, I. & G. DIEWALD (eds.). (2002). *New reflections on grammaticalization*. Amsterdam/ Filadelfia: John Benjamins.

ZIMMERMANN, K. (ed.). (1997). *La descripción de las lenguas amerindias en la época colonial*. Frankfurt/Main/Madrid: Vervuert-Iberoamericana.

ŽIRMUNSKIJ, V. M. (1966). The word and its boundaires. *Linguistics*, 27: 65-91.