

NOTAS

**CORPUS DE HABLA INFANTIL DE CARACAS 1996:
UNA HISTORIA DE PEQUEÑOS HABLADORES A PROPÓSITO DE LOS 70
AÑOS DEL IFAB Y EL ESTUDIO DE LA ORALIDAD EN VENEZUELA***

Krístel Guirado
Universidad Central de Venezuela
Universidad de Zaragoza
kristelguirado@gmail.com

La Lingüística de Corpus viene desarrollándose en Venezuela desde finales de los años sesenta, cuando Ángel Rosenblat pasa a integrar la primera Sub-Comisión Ejecutiva del *Proyecto Norma Culta*.¹ Desde entonces, el *Estudio sociolingüístico del habla de Caracas* constituye una de las líneas de investigación más sólidas del Instituto de Filología “Andrés Bello” de la Universidad Central de Venezuela (IFAB-UCV) y ha dado origen a diversos proyectos para la recolección de corpus estratificados socialmente, que se caracterizan por estar diseñados según una rigurosa arquitectura distribucional, haciendo uso de técnicas, herramientas y métodos estadísticamente representativos.²

De esta forma, en octubre de 2017, el IFAB celebró 70 años de fundado con una tradición de investigación lingüística que se apoya en el estudio de muestras reales del uso de la lengua. No obstante, con el tiempo, no siempre es posible honrar el esfuerzo que supone la ardua gestión de recoger y sistematizar estos materiales:

No es posible publicar una relatoría de las notas de corrección, ni la enumeración de recorridos por callejones y avenidas en busca de un hablante, tampoco el inventario de etiquetas que se han borrado y reescrito o las veces que se ha vuelto sobre un fragmento para descifrar las consonantes aspiradas y las elididas, un balbuceo, un murmullo, una intención (Guirado 2014: 22).

* Agradezco profundamente a Martha Shiro por sus acertados comentarios y sugerencias; huelga decir que todos los errores son de mi entera responsabilidad. Asimismo, quisiera extender mi reconocimiento a la investigadora Yanira Pinto Hernández, quien hizo la búsqueda y escogencia del fragmento utilizado en el ejemplo (1), cf. *infra*.

1. *Proyecto para el estudio coordinado de la norma lingüística culta de las principales ciudades de Iberoamérica y de la Península Ibérica*. Este proyecto constituye la primera empresa colectiva y multinacional que se propone estudiar el español de forma “coordinada, homogénea, común, por encima de preferencias personales o de principios teóricos particulares” (Lope Blanch 1987: 167).

2. Cf. Bentivoglio, D’Introno y Sosa 1977 (*Corpus Sociolingüístico de Caracas 1977*), Rosenblat y Bentivoglio 1979 (*Corpus del habla culta de Caracas 1968-77*), Bentivoglio y Sedano 1993 (*Corpus Sociolingüístico de Caracas 1987*), Bentivoglio y Malaver 2014 (*Corpus Sociolingüístico de Caracas: PRESEEA Caracas 2004/2010*).

Con la intención de conmemorar tales empresas, se han hecho una serie de publicaciones que incluyen una memoria descriptiva y una actualización del estado de algunos de estos proyectos (cf. Guirado 2014, 2015 y 2018, D’Aquino, Guirado y Malaver 2017). En esta ocasión, la presente nota permite presentar el *Corpus de habla infantil de Caracas 1996* –en adelante CHIC-96– con el propósito de acercar al lector al proceso de configuración y elaboración de uno de los corpus más significativos de nuestra institución.

El CHIC-96 se recopiló, entre los años 1995 y 1996, en seis colegios de la capital venezolana. La creación y recolección del corpus estuvo a cargo de Martha Shiro, quien lo diseña y estructura con el propósito inicial de conformar una muestra de habla infantil estadísticamente significativa, con la cual pudiera llevar adelante estudios científicos sobre la incidencia del estrato social y la edad en las habilidades narrativas de los niños. Esta perspectiva de análisis le permitió sustentar sendos proyectos de investigación, pioneros en el área de la psicolingüística en el país.³

Al principio, se seleccionaron al azar veinte (20) niños de cada escuela –tres públicas y tres privadas– para un total de ciento veinte (120) niños, la mitad de primer grado y la otra mitad de cuarto. No se incluyeron en la muestra a los niños con necesidades especiales, bilingües o mayores al promedio de edad de la clase. Todos los niños eran hispanohablantes nacidos en Venezuela, residenciados en Caracas, hijos de padres hablantes nativos del español.

Posteriormente, debido a diferentes motivos, no todas las entrevistas pudieron concluirse según el plan establecido, razón por la cual “la muestra se redujo a ciento trece (113) niños, cuyas entrevistas pudieron completarse de acuerdo con los procedimientos descritos” (Shiro 2007: 50).

Los hablantes seleccionados se agruparon en diferentes categorías sociales: dos grupos generacionales (A: 6 a 8 años; y B: 9 a 11 años), que se corresponden con dos grados de instrucción (A: primer grado; y B: cuarto grado del ciclo básico);⁴ dos niveles socioeconómicos (alto y bajo), que a su vez equivalen a dos tipos de escuela (alto: escuela privada; y bajo: escuela pública). Este criterio de selección permitió recolectar muestras de dos grupos diastráticos diferenciados que reflejaban las marcadas distinciones sociales que existían en la población venezolana para el año de grabación:

3. La recolección de este corpus estuvo asociada a dos proyectos específicos: *La competencia narrativa de niños caraqueños en edad escolar* y *El uso de la modalidad epistémica en las narraciones infantiles*, ambos registrados y financiados por CDCH-UCV bajo los números 07.33.3737.96/07.33.3737.2002 y PI 07-00-6333-2006, respectivamente. Fraibet Aveledo, Valentina Mujica, Cleris Malaver, María Alejandra Calzadilla, Graciela Ávila, Greysi Fernández, Marcos García, Mariana Delgado, Susana González, Arelis Ortiz, Luisana Bisbe, María José Gallucci y Kristel Guirado formaron parte del equipo de investigación que apoyó la ejecución de estos proyectos.

4. Este criterio se sustenta en el hecho de que durante este período etario se dan cambios notables en la elaboración de la estructura narrativa y en el uso de la evaluación. Asimismo, debido al impacto del nivel de desarrollo de la lengua oral en la escolarización, las habilidades lectoras difieren significativamente entre ambos grupos (cf. Peterson y McCabe 1983; Westby 1985; Bamberg y Damrad-Frye 1991; Berman y Slobin 1994).

Así pues, las escuelas privadas se consideran entre las mejores del país y están ubicadas en el este de la ciudad, donde vive gente con alto poder adquisitivo; mientras que las tres escuelas públicas se encuentran en barriadas donde la mayor parte de los habitantes vive por debajo de la línea de pobreza (Marcano 2011: 223).

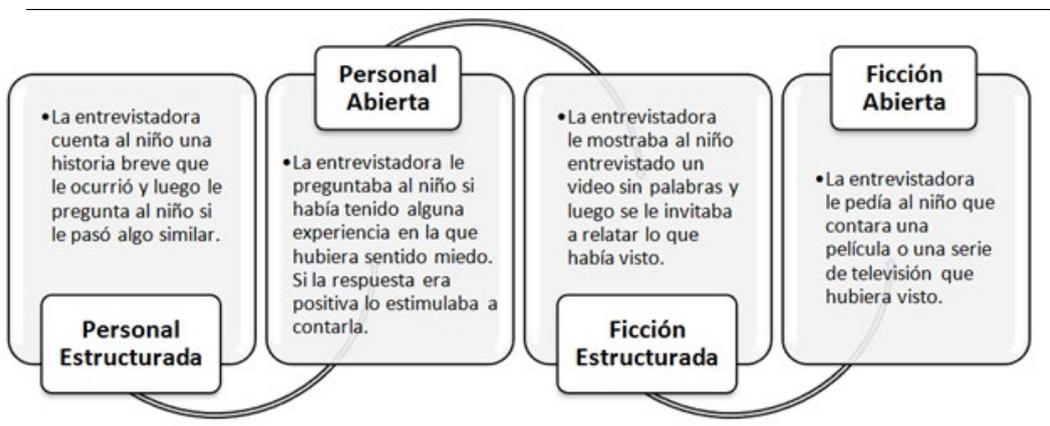
Finalmente, la muestra se clasificó según el sexo (niño y niña); respecto a esta variable, la cantidad de hablantes no resulta equitativa porque, lamentablemente, el número de niños era mayor en uno de los colegios privados como consecuencia de que este solo admitía estudiantes varones en años anteriores al de la grabación. La distribución de los hablantes del CHIC-96 se muestra en el cuadro 1:

Cuadro 1. Distribución de los hablantes del CHIC-96

Tipo de escuela Nivel socioeconómico	Grado cursado - Grupo etario / Sexo			
	GRUPO A		GRUPO B	
	1er grado - 6 a 8 años	4to grado - 9 a 11 año	niños	niñas
PRIVADA - ALTO	17	13	16	13
PÚBLICA - BAJO	14	13	12	15
TOTAL	57		56	

Estas grabaciones se realizaron pidiéndole a cada niño que participara en cuatro tareas narrativas, las cuales requerían que los hablantes produjeran dos narraciones de experiencia personal y dos narraciones de ficción según el formato que hemos diseñado en el diagrama 1:

Diagrama 1. Tipos de entrevista y procedimientos



De este procedimiento se obtuvo un total de cuatrocientas cuarenta y cuatro (444) grabaciones, cuya distribución se puede observar en el cuadro 2:

Cuadro 2. Distribución de las grabaciones del CHIC-96

		Grado cursado - Grupo etario / Escuela - Nivel socioeconómico				
		GRUPO A		GRUPO B		
		1er grado - 6 a 8 años		4to grado - 9 a 11año		
Narración	Tarea	Privada - alto	Pública - bajo	Privada - alto	Pública - bajo	TOTAL
PERSONAL	<i>estructurada</i>	29	25	29	27	110
	<i>abierta</i>	29	26	29	25	109
FICCIÓN	<i>estructurada</i>	29	27	30	27	113
	<i>abierta</i>	29	26	30	27	112
TOTAL		116	104	118	106	444

Las narraciones fueron transcritas en CHAT,⁵ para su posterior análisis con el programa CLAN.⁶ A continuación se presenta un fragmento etiquetado de una entrevista donde se incluye una narración personal (pn) estructurada (str3)⁷ y una narración ficcional abierta (fic:op):

(1) (072.RG.120.M)

@bg: pn:str3

*EXP: mira, y # pasó alguna otra vez que te dio fiebre o algún dolor que te tuvieron que llevar rápido al hospital?

*CHI: ah, que yo # a mí como siempre me da gripe [c].

*CHI: siempre me da fiebre [c].

*CHI: ah, y cuando # cuando me dio gripe [c].

*CHI: me dio fiebre [c] y me dio sarampión [c] y me tuvieron que llevar pa'l hospital [c].

*EXP: y qué pasó?

*CHI: y me dejaron hospitalizado un tiempo [c].

*EXP: y qué te hicieron?

*CHI: me # me tomaban la tensión cada ratico [c] y tenía que # todo el día vestido de rojo [c].

*EXP: rojo?

*EXP: por qué?

5. *Codes for the Human Analysis of Transcriptions* (Sokolow y Snow 1994). CHAT es un formato ideal para el análisis de interacciones de habla entre dos o más hablantes, ya que los investigadores disponen de una amplia cantidad de códigos que permiten configurar el etiquetaje tanto de categorías lingüísticas (fonológicas, morfológicas, sintácticas) como discursivas. Para una explicación más detallada de los diversos códigos de transcripción, cf. MacWhinney 1994.

6. *Computer Language Analysis*, cf. MacWhinney 1994.

7. El número 3 en str3 indica que el niño ya relató dos anécdotas personales anteriormente en el marco de la tarea estructurada.

- *CHI: no sé [c].
- *CHI: así dicen [c] que cuando uno tenga sarampión [c] que se vista de rojo [c].
- *EXP: sí?
- *EXP: está bien, y a ti te gustaba vestirse de rojo?
- *CHI: no [c].
- *EXP: no?
- @eg: pn:str3
- @bg: fic:op
- *EXP: mira, Osmel, y ver # ver televisión te gusta?
- *CHI: sí [c].
- *EXP: sí?
- *EXP: qué ves?
- *CHI: veo en la mañana [c] cuando no hay clase [c].
- *CHI: veo siempre la broma esa de # de Giros TV [c].
- *CHI: y en la tarde, empiezo a ver las comiquitas hasta las once [c].
- *CHI: después veo la noticia [c].
- *CHI: después a las dos, cuando pasaban Morelia [c], la veía [c].
- *CHI: y después ahí empezaba a hacer la tarea y ya [c].
- *EXP: ya.
- *CHI: y después empezaba a jugar [c].
- *EXP: y qué película te gusta ver?
- *CHI: en # cuando pasaban la # cuando pasaban los Power Rangers [c] yo lo veía [c].
- *EXP: mhm.
- *CHI: cuando pasaban Indiana Jones [c] también la veía [c].
- *CHI: cuando pasaban Rambo también, cuando # [c].
- *EXP: y cuál # cuál me vas a contar?
- *CHI: bueno, Terminator [c].
- *EXP: mm?
- *CHI: Terminator [c].
- *EXP: cuéntame Termineitor.
- *CHI: okey, ahá, que ellos # ellos empezaron # que ellos eran de una # de una guerra, no, del espa # ahá del espacio [c].
- *CHI: entonces uno era malo [c].
- *CHI: y otro ahí era bueno [c].
- *CHI: Arnold Schwarzenegger quería [c] matar a una mujer [c].
- *CHI: entonces el otro la estaba defendiendo [c] pa' que no la mataran [c].
- *CHI: entonces iba [c], compraba armas y broma [c].
- *CHI: entonces después casi a lo último, lo último, llegaron a # a una fábrica de # [c] donde hay bromas [c] que aplastan y broma [c].

- *CHI: entonces el # Terminator, como le habían quitado toda la # la bicha [c] y era una # un robot [c], quedó así de metal [c].
- *CHI: entonces el ro # mató al otro muchacho [c] que estaba defendiendo a la mujer [c].
- *CHI: entonces la mujer se iba [c], se escondía [c].
- *CHI: entonces llegó a una parte [c] que estaba trancado el paso [c].
- *CHI: entonces había una broma [c] que aplastaba, no [c]?
- *CHI: y ella tenía la mano así [c].
- *CHI: y cuando el robot se montó así [c] le aplastó la cabeza [c].
- *CHI: y el # y el robot se murió [c].
- *EXP: y así termina la película?
- *CHI: sí [c].
- *EXP: está bien, y qué parte te gustó más?
- *CHI: bueno, cuando # cuando mataron al robot [c].
- *EXP: sí?
- @eg: fic:op

El CHIC-96 tiene un total de ciento veintiséis mil setecientos noventa y ocho (126.798) palabras, a un promedio de mil ciento veintidós (1.122) palabras por niño (cf. Gallucci 2005: 117). Todas las entrevistas están disponibles en formato digital (audio y texto). Las transcripciones pueden utilizarse tanto en la versión etiquetada para CLAN,⁸ como en la versión sin etiquetas. Asimismo, el material de las entrevistas está incorporado a la conocida base de datos denominada CHILDES (*Child Language Exchange System*)⁹ y ha servido como base de datos para numerosos investigadores y múltiples publicaciones.

REFERENCIAS BIBLIOGRÁFICAS

- Bamberg, Michael y Robin Damrad-Frye. 1991. On the ability to provide evaluative comments. Further explorations of children's narrative competencies. *Journal of Child Language* 18. 689-710.
- Bentivoglio, Paola e Irania Malaver. 2014. *Corpus sociolingüístico Preseea Caracas 2004-2010*. CD-rom. Caracas: UCV-FHE-IFAB.
- Bentivoglio, Paola y Mercedes Sedano. 1993. Investigación sociolingüística: sus métodos aplicados a una experiencia venezolana. *Boletín de Lingüística* 8. 3-35.
- Bentivoglio, Paola, Francesco D'Introno, y Juan M. Sosa. 1977. *Proyecto de*

8. Para una descripción más amplia de las categorías de análisis etiquetadas, cf. Shiro (1996, 1998, 2000 y 2007) y Marcano (2011).

9. Véase en línea: <http://childes.psy.cmu.edu/>. En esta página web también se puede descargar gratis el manual de uso de CHAT y el programa CLAN.

- investigación para un análisis sociolingüístico del habla de Caracas*. Material mimeografiado. Caracas: Universidad Central de Venezuela.
- Berman, Ruth y Dan I. Slobin. 1994. *Relating narrative events: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- D'Aquino, Giovanna, Krístel Guirado e Irania Malaver. 2017. La pasión por las fuentes: Notas a dos proyectos. *Boletín de Lingüística* 47-48. 60-169.
- Gallucci, María José. 2005. El número de palabras: un nuevo criterio para describir tres corpus del habla de Caracas. *Boletín de Lingüística* 24. 108-121.
- Guirado, Krístel. 2014. Corpus Diacrónico del Habla de Caracas 1987/2013. *Boletín de Lingüística* 41/42: 19-44. En línea: <https://dialnet.unirioja.es/servlet/articulo?codigo=6435356> [Consulta: 1º de noviembre de 2018].
- Guirado, Krístel. 2015. *Reingeniería de Corpus* en Venezuela: una propuesta metodológica para diversificar el análisis de los corpus del español hablado en Caracas. *Letras* 92. 55-87. En línea: <http://revistas.upel.edu.ve/index.php/letras/article/view/5962/3239> [Consulta: 1º de noviembre de 2018].
- Guirado, Krístel. 2018. Habla culta de Caracas 1973-2011. Un subcorpus de propósito especial para el estudio diacrónico del habla caraqueña. *CHIMERA. Romance Corpora and Linguistic Studies* 5, 1. 111-116. En línea: <https://revistas.uam.es/index.php/chimera/article/view/9884/10304>. [Consulta: 1º de noviembre de 2018].
- Lope Blanch, Juan M. 1987. El Estudio coordinado de la norma culta de las principales ciudades de lengua española. *Actas del VII Congreso. Asociación de Lingüística y Filología de América Latina (ALFAL) I*, 163-67. Santo Domingo: Asociación de Lingüística y Filología de América Latina.
- MacWhinney, Brian. 1994. *The CHILDES Project: Tools for analyzing talk*. Hillsdale, N.J: Lawrence Erlbaum Associates.
- Marcano, Zayra. 2011. Reseña al libro *La construcción del punto de vista en los relatos orales de niños en edad escolar* de Shiro 2007. *Boletín de Lingüística* 35-36. 218-229.
- Peterson, Carole y Allyssa McCabe. 1983. *Developmental psycholinguistic*. New York: Plenum Press.
- Rosenblat, Ángel (dir.) y Paola Bentivoglio (ed.). 1979. *El habla culta de Caracas. Materiales para su estudio*. Caracas: Facultad de Humanidades y Educación, Universidad Central de Venezuela.
- Shiro, Martha. 1996. *CHILDES database. Romance Languages Corpora, Spanish-Shiro* [En línea]. Disponible en: <http://childes.psy.cmu.edu/data/Romance/Spanish/>.
- Shiro, Martha. 1998. *Los pequeños cuentacuentos. El desarrollo de las habilidades narrativas de niños en edad escolar*. Trabajo de ascenso. Caracas: Universidad Central de Venezuela.
- Shiro, Martha. 2000. Los pequeños cuentacuentos. *Cuadernos Lengua y Habla* 2. 319-337.

- Shiro, Martha. 2007. *La construcción del punto de vista en los relatos orales de niños en edad escolar: un análisis discursivo de la modalidad*. Caracas: Fondo Editorial de Humanidades, Universidad Central de Venezuela, DICORI.
- Sokolov, Jeffrey y Catherine Snow. 1994. *Handbook of research in language development using CHILDES*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Westby, Carol. 1985. Learning to talk - talking to learn: Oral-literate language differences. En Charlann S. Simon (ed.), *Communication skills and classroom success*, 181-218. San Diego, CA: College-Hill Press.