

«ALLÁ LLEGA A LO QUE ES EL PUEBLO DE SAN AGUSTÍN». EL CASO DE LA PERÍFRASIS INFORMATIVA CON *LO QUE ES* EN EL «CORPUS SOCIOLINGÜÍSTICO DE LA CIUDAD DE MÉXICO»*

Yolanda Lastra, Universidad Nacional Autónoma de México
y El Colegio de México, México

Pedro Martín Butragueño, El Colegio de México, México
ylastra@unam.mx, pmartin@colmex.mx

Recibido: 16/04/2015 – Aceptado: 29/04/2015

DOI: 10.17533/udea.lyl.n69a12

Resumen: Este artículo estudia la construcción *lo que es* como perífrasis informativa de relativo en 108 entrevistas del «Corpus sociolingüístico de la ciudad de México» (CSCM), en el contexto de todos los casos de *lo que* presentes en dicho corpus. Se analizan sus propiedades sintácticas y pragmáticas y se establece su distribución social. Debe destacarse la eliminabilidad de las perífrasis y el carácter de *ser* como verbo débil. Se defiende la idea de que se trata de un marcador de prominencia informativa.

Palabras clave: variación sintáctica, ciudad de México, *lo que es*, estructura de la información.

«ALLÁ LLEGA A LO QUE ES EL PUEBLO DE SAN AGUSTÍN». THE CASE OF THE INFORMATIVE PERIPHRASES WITH *LO QUE ES* IN THE «CORPUS SOCIOLINGÜÍSTICO DE LA CIUDAD DE MÉXICO»

Abstract: This paper studies the construction *lo que es* as a cleft sentence in 108 sociolinguistic interviews from the «Corpus sociolingüístico de la ciudad de México» (CSCM), in the context of all *lo que* found in the corpus. Syntactic and pragmatic properties are analysed and its social distribution is established. The removability of cleft construction and the character of *ser* as a weak verb are especially relevant. We argue that *lo que es* is a marker of informative prominence.

Keywords: syntactic variation, Mexico City, *lo que es*, information structure.

* Esta investigación se deriva del proyecto «Diversidad y variación fónica en las lenguas de México: hacia una nueva caracterización de la diversidad geolingüística» (2011-2015, Consejo Nacional de Ciencia y Tecnología de México, 127876, convocatoria CB-2009-01). Nos es muy grato que este trabajo aparezca en un volumen dedicado a la memoria del admirable profesor José Joaquín Montes Giraldo, de quien tanto hemos aprendido. Una versión previa, con la mitad de datos, se leyó en el XVII Congreso Internacional de la ALFAL, celebrado en João Pessoa (Paraíba, Brasil) entre el 14 y el 19 de julio de 2014, donde fue muy productivo el diálogo con los participantes, especialmente con el Dr. Esteban Montoro, de la Universidad de Granada. Agradecemos el apoyo de Mary Carmen Aguilar y de Jorge A. Kasep, becarios del Laboratorio de Estudios Fónicos de El Colegio de México, quienes nos ayudaron a extraer los ejemplos de las transcripciones del CSCM. Han sido también muy productivas las observaciones efectuadas por dos evaluadores anónimos.

1. Introducción

Construcciones como las de (1) parecen tener las mismas condiciones de verdad que las de (2):

- (1) a. en el Olivar viví [...] hasta los ocho años/ ya después/ ya me fui para allá [...] Chimalhuacán/ Estado de México [...] o sea sí/ o sea **lo que es la avenida Chimalhuacán**// hasta ahora <~ora> sí que/ hasta donde llega (risa)// allá llega a **lo que es el pueblo de/ San Agustín** [...] ahí es donde/ tienes tu casa (ME-232-21H-04, turno 620).¹
- b. agarré **todo lo que es la Central**// hasta entroncar con Churubusco (ME-284-12H-07, turno 1273).
- c. no tiene nada que ver/ eso sí// los libros/ no tienen nada que ver **con lo que es el/ el cuerpo humano** (ME-252-31M-05, turno 55).
- d. Tengo **lo que es** una tarjeta de <identificación> (mujer radicada en la ciudad de México, conversación espontánea).
- (2) a. Me fui para allá [...] / la avenida Chimalhuacán. Allá llega **al pueblo de San Agustín**.
- b. Agarré **toda la Central**// hasta entroncar con Churubusco.
- c. No tiene nada que ver/ eso sí// los libros/ no tienen nada que ver **con el/ el cuerpo humano**.
- d. Tengo **una tarjeta de identificación**.

Dado que la construcción *lo que es* + frase nominal y otras afines son recurrentes en los datos del «Corpus sociolingüístico de la ciudad de México» (CSCM) (y en otras variedades de español mexicano y no mexicano), este trabajo está dedicado a esbozar sus propiedades sintácticas, semánticas y pragmáticas, así como sus semejanzas y diferencias con otras estructuras mejor conocidas, y a considerar su dispersión social.

1 Se indica entre paréntesis la entrevista de donde procede el ejemplo, así como el turno de habla específico. Así, ME-232-21H-04 debe leerse como ME= «Corpus sociolingüístico de la ciudad de México» (CSCM); 232= número de entrevista en el mismo corpus; 21H= 2 por nivel medio de estudios, 1 por joven (20-34 años) y H por hombre; por fin, 04= 2004, año en que se grabó el material. Se menciona también la procedencia cuando el dato no hace parte del CSCM, y si no se dice nada es por que se trata de un ejemplo hipotético, aunque habitualmente validado por otros hablantes. Otro tipo de comentarios se introducen entre corchetes.

Nuestra hipótesis es que se trata de una *perífrasis informativa de relativo* (PIR), es decir, una perífrasis de relativo (forma) cuyo cometido (función) es revelar la existencia de un constituyente oracional o de un orientador extraoracional que estructura la información expuesta por un enunciado inserto en un discurso.² Obsérvese (3) a ese respecto:

- (3) a. yo me dedico a **lo que es la construcción/ lo que es acabados** (ME-258-11H-05, turno 12) [foco].
- b. bueno para empezar los/ **lo que son los cubanos/ y los colombianos**/ para ellos es más fácil arreglar papeles/ ellos tienen más facilidad de arreglar papeles/ que el mexicano/ el mexicano/ nos ponen muchas trabas (ME-275-22H-06, turno 260) [tema extraoracional].
- c. **lo que era el salón**/ había dos mesas de más/ fueron las dos mesas que se me quedaron vacías/ todo lo demás/ estaba lleno (ME-197-31H-01, turno 489) [tema extraoracional].
- d. **Lo que es las cebollas** se llevan muy bien con todo [tópico intraoracional].
- e. El Colonial/ también era carpa [...] que está ahí en San Juan de Letrán/ **lo que es San Juan de Letrán** (ME-300-13H-07, turno 1073) [apéndice].
- f. **Lo que es mi casa**, todos están comiendo a esa hora [orientador espacial].

En el sentido en que Dik (1997, cap. 13) se refiere a las funciones informativas, la perífrasis informativa de relativo puede incrementar el reconocimiento de un foco (3a), de un tema extraoracional (3b y 3c), de un tópico oracional (3d), de un apéndice o tema pospuesto (3e), e incluso de orientadores espaciales y temporales (3f). Foco y tópico son funciones pragmáticas intraoracionales, mientras que temas, apéndices y orientadores son extraoracionales y mantienen una relación pragmática con la oración. El tópico se asigna a la parte de la oración que refiere a la información dada (*givenness*), aunque también existen cambios de tópico o tópicos nuevos; el foco se asocia habitualmente con aquella parte de la oración que proporciona información nueva, aunque también puede definirse como la aportación relevante o como la selección entre varias opciones posibles (Krifka, 2007).

2 Coincidimos en ese sentido con Pérez Álvarez (2014, p. 86, pp. 90-92), quien también caracteriza las estructuras con *lo que es* y *lo que viene siendo* como oraciones escindidas o hendidas, en el marco de un análisis más general sobre las funciones oracionales y discursivas de *lo que*. En particular, considera que *lo que viene siendo* «se trata de una fórmula utilizada mientras se planea el discurso subsecuente» (p. 91). Nuestro análisis, centrado solo en *lo que es* (aun habiendo fichado todos los casos de *lo que*), sin negar la contribución discursiva de la construcción, apuesta más bien por los valores informativos y señala varias diferencias con las copulativas enfáticas.

Un análisis más simple puede consistir en establecer que en la PIR el consecuente³ siempre refiere al foco, aunque a veces este se encuentre incrustado en una unidad informativa de otra naturaleza, es decir, un análisis del tipo [lo que_i es [X_i]_{FOCO}]_{FINF}, donde FINF es cualquier función informativa, como las mencionadas en (3). Así, se tendría [lo que es [el salón]_{FOCO}]_{TEMA}/ había dos mesas de más..., donde el foco *el salón* forma parte de una estructura informativa mayor, que es tema extraracional con respecto a la siguiente oración, tal como propone Krifka (2007, p. 40 y ss.) a propósito de diferentes construcciones informativas. Este análisis tendría la ventaja de unificar todos los casos posibles, aunque su discusión no es central en este momento.

2. Metodología

Se han buscado todos los contextos de *lo que* en 108 entrevistas pertenecientes al módulo PRESEEA (Moreno Fernández, 1996; Cestero Mancera, 2012) del «Corpus sociolingüístico de la ciudad de México» (CSCM: Martín Butragueño y Lastra (2011, 2012 y en preparación)), distribuidas equitativamente por nivel de estudios, edad y género. La razón para extraer todos los casos y no solo los de la construcción que nos interesa es poder establecer las semejanzas y las diferencias entre las diversas estructuras. En consecuencia, se han fichado 3 671 contextos, de los cuales 3 220 pertenecen a las personas entrevistadas. La mayoría son casos de RELATIVAS SEMI-LIBRES (*me gustaría saber lo que contiene el paquete*) y de COPULATIVAS ENFÁTICAS (*eso es lo que siempre te digo*), además de algunas CONSTRUCCIONES PARCIALMENTE LEXICALIZADAS (del tipo *lo que pasa es que es una persona muy ocupada*). No han aparecido casos claros de RELATIVOS COMPLEJOS (*es muy prepotente, lo que me molesta mucho*), lo que sugiere su carácter poco coloquial. Solo una minoría del total de ejemplos fichados son casos de PIR diferenciadas de estas otras construcciones. En las relativas semilibres «se omite fonéticamente el antecedente nominal, pero aparece realizado el artículo definido» (NGLE, 2009, p. 3293). Las copulativas enfáticas son «construcciones copulativas construidas con el verbo *ser* en las que se realiza o se destaca uno de sus componentes» (*ibid.*, 40.10a, p. 3020)⁴; este carácter realizador es compartido por las PIR, pero en las páginas siguientes se irán mostrando las semejanzas y diferencias entre unas y otras. Construcciones parcialmente

3 De la misma forma que el antecedente es un correferente anafórico, el consecuente es un correferente catafórico. En [lo que_i es [X_i]_{FOCO}]_{FINF}, X es el consecuente. La correferencia se marca por medio del subíndice. El dominio es el ámbito sintáctico en que se desarrolla la correferencia estudiada (la oración de relativo para el caso, como se menciona *infra*).

4 Acerca de las copulativas enfáticas o de relieve relativas, o construcciones pseudohendidas, o perifrasis de relativo, cf. Martín Butragueño (1997), Sedano (2003, 2011, cap. 18), RAE-ASALE (2009, 40.10-11) y Bosque y Gutiérrez-Rexach (2009, cap. 11).

lexicalizadas son *lo que pasa es que*; *lo que sea de él, es serio*; y *lo que tú quieras*; *va a lo que va* y otras documentadas en la base de datos, que sin duda requieren un análisis específico, pero que se alejan claramente de las propiedades sintácticas y pragmáticas de las PIR. Por fin, los *lo que* relativos complejos presentan un antecedente externo oracional, son reemplazables por *lo cual* (así como por *hecho que*, *cosa que*) y aparecen en construcciones explicativas (*ibid.*, 44.3ñ, p. 3316).

Dejando al margen los ejemplos producidos por los entrevistadores o por otros participantes esporádicos, hemos encontrado 254 ejemplos de PIR. Los que cumplen estrictamente con las características canónicas de la construcción son 205 casos (*infra* se establecen cuáles son los casos canónicos).

Todos los contextos se han introducido en una base de datos,⁵ en cuyas filas figura cada ejemplo y en cuyas columnas la descripción de una veintena de características sintácticas, semánticas, pragmáticas y sociales. La acotación de la variación (*envelope of variation*)⁶ se lleva a cabo en seguida, al discutir las características de las construcciones consideradas y sus diferencias con otras más o menos cercanas. Se analizan por tanto las propiedades lingüísticas de la PIR, en primer término, y más adelante se esboza su distribución social.

3. Propiedades sintácticas, semánticas y pragmáticas

Se examinan en esta sección las propiedades de las PIR, algunas de las cuales son compartidas por otras construcciones de relativo, mientras que la conjunción de otras permite esbozar su carácter específico. Todo esto al hilo de lo que va ofreciendo el examen de diferentes factores en el conjunto completo de los datos disponibles, presentados en este orden: a) la invariabilidad de *lo que*; b) el carácter objetival del correferente; c) la naturaleza referencial de *lo* (que puede ser absoluto o partitivo, y referir a entidades o cantidades); d) la relación predicativa entre *lo* y el correferente; e) la explicitud y posición consecuente del correferente, así como su dominio oracional y carácter de frase nominal; f) la función informativa desempeñada por el consecuente; g) la concordancia no categórica entre el verbo débil y el consecuente; h) el carácter no extraíble del consecuente; i) la tendencia al presente en el verbo débil y la improbabilidad de que reciba prominencia prosódica; j) la existencia de PIR prototípicas y PIR periféricas; k) la modulación de la fuerza ilocutiva por medio de

5 Es decir, se introduce en la base de datos no solo el ejemplo o enunciado específico que se está considerando, sino todo el contexto anterior y posterior necesario para entender claramente el desarrollo de la estructura informativa.

6 Véase por ejemplo el capítulo que dedican al respecto Otheguy y Zentella (2012) para acotar el estudio de la expresión del sujeto pronominal en el español de Nueva York.

algunos casos de PIR; l) la naturaleza no epistémica, no evidencial y habitualmente no contrastiva de las PIR; m) el papel de la cuantificación y de la negación; y n) el carácter comúnmente aseverativo de los actos de habla en los que se instalan las PIR.

En las PIR siempre aparece el relativo *lo que*, NEUTRO y por tanto INVARIABLE; compárese *fue a lo que es la Plaza de la Constitución* (PIR) con *fue a la que es la Plaza de la Constitución* (relativa semilibre). En general, suelen referirse a OBJETOS y no a seres animados (los consecuentes animados parecen más admisibles cuando la PIR es tema, y mucho menos cuando es foco: *lo que es Juan y su familia son simpáticos* vs. *??vinieron lo que es Juan y su familia*; en todo caso, con animados se inducen lecturas inespecíficas, como en (3b)). En la construcción, el artículo neutro *lo* es REFERENCIAL O INDIVIDUATIVO (NGLE, 2009), y no un *lo* enfático, pues refiere a una entidad y encabeza una oración que expande una frase nominal, y no es cuantificador de adjetivos o adverbios. El *lo* referencial de la PIR puede ser ABSOLUTO (4a) o PARTITIVO (4b) (pero no introduce atributivas cuantificadas, como en (4c)):

- (4) a. si uno prepara al muchacho/ para que vaya a dar clases y/ hacer sus prácticas y/// creo que le va mejor/ [se va] [...] a defender mejor [...] [sí/ sí] entonces los cursos que damos nosotros/ son para nuestros estudiantes [...] y les enseñamos/ pues <~pus>/ **lo que es el oficio** (ME-249-33H-05, turno 953) [absoluto].
- b. haz de cuenta que tú entras a la vecindad/ y derecho aprecias lo que es la [capilla] [...] a un costado no hay muro/ **solamente lo que es del lado [...] de <~de:> del lado derecho**/ entonces ahí <~ahí:>/ lo que hice yo fue dibujar a Juan Diego (ME-301-11H-07, turno 384) [partitivo].
- c. Juan consideró lo difícil del curso / #Juan consideró **lo que** es difícil del curso [sí si se interpreta como relativo semilibre y verbo pleno, pero no como PIR].

Predominan las PIR referenciales absolutas (89.8%), pero no faltan diferentes casos de referenciales partitivas. No parece haber en esto muchas diferencias con cualquier otro tipo de construcciones relativas con *lo que*, donde discursivamente predomina también la referencialidad absoluta sobre la partitiva.⁷ Por otra parte, *lo que* puede hacer referencia a ENTIDADES (5a), el caso más común con mucho (97.6%), pero también a CANTIDADES (5b), exactamente igual que en otras construcciones, como la relativa semilibre:

- (5) a. servían carnes/ cortes este/ de carne/ pero/ su especialidad era la/ la carne de/ o sea el pescado/ **lo que es el pescado y el marisco** (ME-275-22H-06, turno 74) [entidad].

7 Sabemos esto porque nuestra base de datos no solo recoge las PIR, sino todas las construcciones con *lo que*.

b. [¿y un cuartillo cuánto es?] un cuartillo/ pues <~pus> así de a litro/ **lo que es un litro de maíz** (ME-007-21M-97, turno 414) [cantidad].

La relación entre el relativo y su correferente es PREDICATIVA, no anafórica (6a). Este hecho es una constante definitoria, aunque la propiedad es compartida con las copulativas enfáticas (6b):

(6) a. los chalecos// por dentro tienen/ haz de cuenta como esta mallita [...] [viene] mucha malla forrada/ entonces <~entós> lo que pasa que como la otra es ¡más gruesa!/ y es de impacto/ la/ **lo que es la cuarenta y cinco**// la especial/ el revólver que tenemos nosotros/ el treinta y ocho especial// y mm/ o sea son de impacto/ esa te pega (ME-232-21H-04, turno 352) [PIR].

b. **Eso es** lo que te quería decir / Lo que te quería decir **es eso** [copulativa enfática].

Como puede observarse en los ejemplos en que se van mencionando casos de PIR (como en (6a) y muchos otros), el correferente siempre aparece EXPRESO (salvo que se produzca alguna interrupción discursiva) y es siempre CONSECUENTE (no antecedente) del relativo. Su DOMINIO de aparición es estrictamente la propia oración de relativo (a diferencia de las copulativas enfáticas), aunque en muchos casos el referente pueda haber aparecido antes en el discurso. Ya se ha dicho que el consecuente relacionado predicativamente con el relativo puede desempeñar cualquier FUNCIÓN INFORMATIVA (o la propia PIR, si se opta por la idea de que el consecuente es siempre foco, como se apuntó *supra*), pero es claro que en la mayor parte de los casos funciona como FOCO (88.2%).

Desde el punto de vista sintáctico, el hecho de que opcionalmente el VERBO DÉBIL concuerde con el correferente sugiere que su función es la de sujeto de la oración constituida por la PIR (donde el relativo funcionaría como atributo) (7a). El consecuente constituye, pues, una FRASE NOMINAL, que puede consistir en un nombre común (8a) o propio (8b), así como en una frase propiamente dicha o en un nombre escueto (8c).⁸

La cuestión de la CONCORDANCIA amerita una consideración más detallada. En general, los hablantes sí establecen la concordancia sintáctica, con vacilaciones cuando el consecuente es plural por coordinación (como ocurre en general en otros casos): *fui a lo que es el Zócalo y La Villa* / *fui a lo que son el Zócalo y La Villa*. Más interesante es que se pueda decir *me comí lo que son unos tacos* o *me comí*

8 En ese sentido, las PIR orientadoras, como en (3f), constituyen un problema. Hay varias soluciones posibles, que van más allá de lo que podemos discutir ahora: a) establecer que la PIR no se construye solo sobre una FN sujeto; b) dejar aparte de las PIR las posibles construcciones con valor orientador; o c) suponer que el sujeto es nulo y que el material visible es un circunstancial suyo.

lo que es unos tacos, pero no tengo lo que son dos hijos y quizá más difícilmente *??tengo lo que es dos hijos*, quizá por la más clara divisibilidad del referente *hijos* que del referente *tacos*.

Los casos pertinentes en términos variables son los consecuentes plurales (43 casos), que entablan la concordancia con el verbo en singular en 23 ocasiones, es decir, en 53.5% de los ejemplos pertinentes. Requeriría un análisis más detallado establecer los factores que inducen en la falta de concordancia (como en (7b) y (7c)), sea que se trate de a) factores generales en relación con la concordancia misma; b) de un proceso de desgramaticalización de la construcción (cf. Company (2003), Torres Cacoullós (2011)), donde tiende a fijarse la estructura *lo que es*; o c) de una refuncionalización: así, *las plantas es mi pasión* no parece posible con *las plantas* como sujeto, pero sí como tópic.

- (7) a. la gente que vive/ en las este// **lo que son los límites del país**/ al ver una persona que va/ con una/ mochilita/ porque no puedes llevar gran cosa/ tampoco/ con una mochilita/ con una muda de ropa/ pues <~pus> lógicamente/ saben que vas a la aventura/ te vas a aventurar a cruzar [a Estados Unidos] (ME- ME-275-22H-06, turno 152) [verbo en plural, consecuente en plural], vs. *la gente vive en lo que es los límites del país*.
- b. en la mesa ponen toda la ofrenda naranjas/ todo **lo que es/ tamales chocolate**/ que todo eso lo hacen ellos (ME-296-12M-07, turno 660) [verbo en singular, consecuente en plural], vs. *ponen todo lo que son tamales, chocolate...*
- c. me gustó mucho **lo que es/ las plantas** (ME-055-32M-99, turno 30).
- (8) a. hay que saber/ vivir la calle/ eh/ sin quemarte las manos/ sin que seas un drogadicto <~drogadito>/ o co- un/ un delincuente/ vivir/ **lo que es una vida**/ pacífica y tranquila (ME-314-13H-07, turno 88) [nombre común].
- b. ahí cenamos/ en/ un restaurante/ de la torre del Niágara que es giratorio// o sea <~sea> vas cenando y// va girando y// ves toda la// ves las este/ las cataratas// eh// Toronto/ colinda/ con/ bueno/ el Niágara/ con **lo que es/ Nueva York** (ME-250-31M-05, turno 240) [nombre propio].
- c. [ay <~oy>] pues <~pus> una lana// <lo est->/ era lo que estaba haciendo/ fácil eh/ llegué como a unos/ ciento diez mil pesos [...] **lo que es salón/ vestido/ invitaciones**/ [no no no] (ME-197-31H-01, turnos 477-479) [FN escueta].

Dos propiedades muy importantes del correferente son a) que NO ES EXTRAÍBLE (9) y b) que en muchas ocasiones (37.4% exactamente) refiere a un ESPACIO o a un TIEMPO (10):

- (9) a. Juan está en lo que es el Zócalo.
b. *Es el Zócalo en lo que está Juan [como PIR no hay posibilidad de extracción].
- (10) a. me subo al metro y hasta Pantitlán [...] sí/ pues <~ps> se va casi **todo lo que es Circuito Interior** (ME-155-31H-01, turno 473) [espacial].
b. y tenía funciones/ todo **lo que era el mes de septiembre**/ como todos los grupos de todos lados/ en el mes de septiembre/ que es el mes patrio (ME-259-32M-05, turno 65) [temporal].

La imposibilidad de extraer el consecuente diferencia con claridad las PIR de las copulativas enfáticas, donde sí es posible la extracción (*lo que dijo Juan es eso; es eso lo que dijo Juan; eso es lo que dijo Juan*; (9a) y (9b)). En cuanto a su vinculación con lecturas orientadoras del tiempo y el espacio, así como del tema, estas podrían estar en la base de su difusión a través de ciertos registros, como se comenta más adelante. Como estereotipo, los ejemplos que de inmediato vienen a la mente de las personas a las que hemos consultado son casi siempre de referencia espacial.

Solo el verbo *ser* como verbo débil parece construir PIR; sin embargo, considérese (11a). En el ejemplo, *tener que ver con* parece funcionar exactamente igual que una PIR: la perífrasis es prescindible, el consecuente no personal no es extraíble, se respeta y potencia su función informativa, se ofrece una lectura absoluta de entidad, etc. Es probable que la idea de PIR pueda extenderse a otras construcciones predicativas, siempre que se emplee un verbo débil y el consecuente sea una FN, características que no cumple la relativa semilibre (11b), que no significa lo mismo que *Juan solo estudia lo importante*, aun cuando guarda ciertas analogías con las PIR. Por otra parte, el propio verbo *ser* con valor existencial cancela la estructura perifrástica y da lugar a lecturas de relativo semilibre, como en (11c). Convendría estudiar en detalle si estos casos con otros verbos son menciones metalingüísticas. Considérense *estoy en el Zócalo; estoy en lo que se llama/llaman el Zócalo; estoy en lo que le dicen el Zócalo* y los ejemplos (11d) con *llamar* y (11e) con *referirse*. En cualquier caso, estos datos no se han incluido en los recuentos efectuados *supra* acerca de las PIR.

- (11) a. Se dedicará a **lo que tiene que ver con** la mejora del metro.
b. Juan solo estudia **lo que considera importante** [relativa semilibre].
c. ahí en Caracoles todo eso/ te cobran/ así este/ no traen taxímetro/ yo he tomado taxis de ahí [...] y este/ y ya me subo/ y <~y:> busco/ «oye/ ¿y cómo me vas cobrar (sic)?»/ «no pues <~pus> es que/ cobro tanto así es»/ entonces

<~entós> ya veo yo/ si si/ sí es **lo que [es]** [...] adelante/ si no/ «¿sabes qué?/ mira mejor/ espero otro no te preocupes» (ME-284-12H-07, turno 1132).

d. a mí me gusta mucho <~mucho:> hacer mueble pero **lo que le llaman** el mueble clásico [...] el de estilo (ME-144-23H-01, turno 206).

e. pues <~pus> tuve una familia mm eh <~eh:>/ pues <~pus> con grandes problemas porque imagínate/ ¡mi abuela sufrió eh <~eh:> mm// todo **lo que se refiere** a las guerras!/ este internas (ME-137-32M-01, turno 94).

Cuatro pruebas que subrayan el carácter débil del verbo *ser* en las PIR son a) el carácter (casi siempre) ELIMINABLE de la perífrasis con *lo que* (ejemplos de (12)); b) la tendencia a aparecer en TIEMPO PRESENTE: (13a) muestra dos casos en copretérito, pero que podrían haber estado en presente (13b);⁹ c) la posibilidad de cancelar la CONCORDANCIA con su sujeto sintáctico ((14^a) y (14b) y ejemplos de (7), y compárense con (14c)); y d) la IMPROBABILIDAD DE ASIGNAR UNA PROMINENCIA PROSÓDICA sobre el verbo cuando forma parte de una PIR.

(12) a. de ahí nos trasladamos a **lo que es** la <~la:>/ colonia <~colonia:>/ del Gas (ME-283-23M-06, turno 64).

b. De ahí nos trasladamos a la colonia del Gas.

(13) a. se llevaron todo lo que era de valor desde luego **lo/ que eran jaboncitos** cositas así baratas pues <~pus> eso no se lo llevaron [...] pero **todo lo que era la latería** [...] una televisión/ nuevecita que tenía yo ahí mi Hitachi/ este/ tenía treinta mil pesos en efectivo [...] entonces total que me/ me partieron (ME-281-23H-06, turno 749).

b. Lo que **son** los jaboncitos, eso no se lo llevaron. Todo lo que **es** la latería, eso sí se lo llevaron.

(14) a. Los vendedores se ponen en **lo que son las calles del Centro**.

b. Los vendedores se ponen en **lo que es las calles del Centro**.

c. Las calles del centro son/*es para los vendedores.

Con respecto a la propiedad d), compárense *me gusta lo que son las mesas* y *me gusta lo que SON las mesas* (con verbo existencial, pleno, sin o con prominencia prosódica) con *me gusta lo que son las mesas* como PIR, forzosamente sin prominencia (#*me gusta lo que SON las mesas*).

9 Cuando el verbo no aparece en presente, y especialmente cuando lo hace en copretérito, desata lecturas contrastivas (*infra*).

Se ha venido señalando que una de las principales características de las PIR es que casi siempre son prescindibles, como se ha ejemplificado en (1), (2) y (12). Vamos a llamar PIR CANÓNICAS o PROTOTÍPICAS a las que cumplen con los criterios de eliminabilidad sin cambio en las condiciones de verdad, aparición del verbo *ser* en presente y falta de prominencia prosódica en la cópula. Las PIR canónicas tienen una dimensión pragmática, pues señalan que el constituyente desempeña *alguna* función informativa que el hablante considera relevante. La PIR sería, en suma, una estructura para marcar la relevancia informativa. Ya se ha mencionado que una posibilidad económica es suponer que el consecuente siempre es foco, mientras que la perífrasis de relativo como tal sería capaz de desempeñar también otras funciones informativas.

Por otro lado, hay algunas PIR no completamente canónicas, con una dimensión semántica, pues su eliminabilidad puede desatar diferencias en las condiciones de verdad, a pesar de la proximidad estructural y de muchas semejanzas en la utilidad comunicativa. Se trata de ejemplos con copretérito, o con *ahora* estableciendo un contraste con un *antes* (véanse ejemplos en (18)), entre otros. No siempre es claro que solo la condición de copretérito produzca una diferencia semántica: considérese *me trepaba todo lo que eran / son figuritas*.

Más allá del valor pragmático principal (que recuerda al del *pues* comentador), de relevancia informativa, existe una serie de valores que de manera más particular aportan las PIR, o cuando menos algunas de ellas.

En ocasiones, la PIR tiene un valor ATENUADOR, en el sentido de disminuir la fuerza ilocutiva del acto de habla. Considérese (15a) frente a (15b):

(15) a. Quisiera pedirte el dinero **que** me debes, pues lo necesito.

b. Quisiera pedirte **lo que es** el dinero que me debes, pues lo necesito.

Parece claro que la petición de (15a) es mucho más directa que la de (15b), es decir, (15b) salvaguarda en mayor grado la imagen negativa del interlocutor. La atenuación está relacionada en ciertos casos con el hecho de que las PIR facilitan las LECTURAS INESPECÍFICAS.¹⁰ Véanse los ejemplos de (16):

10 Se ha observado que el *lo que* presente en contextos atributivos alude a cualquier tipo de argumento no animado estableciendo una relación predicativa y no de correferencia anafórica; la mención es genérica y puede identificar no solo el ejemplar sino la clase de objeto (RAE-ASALE, 2009, 14.10c): «Saca del maletín un espectacular llavero y entra en *lo que debe ser* la sacristía de la catedral» [frente a *la que debe ser*] (Vila-Matas, *Suicidios*, *id.*). En el mismo sentido, *fui a lo que es el Teatro de los Insurgentes* vs. *#fui a lo que es una obra de teatro*.

- (16) a. Es muy satisfactorio poder vivir **lo que es un parto** (*Reforma*, 29 de junio de 2014).
- b. Es muy satisfactorio poder vivir **un parto como el que tuvo Luisa** [específico].
- c. ??Es muy satisfactorio poder vivir **lo que es un parto como el que tuvo Luisa** [contradicción entre específico e inespecífico].
- d. Para Luisa, fue muy satisfactorio **poder vivir su parto** [específico].
- e. ??Para Luisa, fue muy satisfactorio **poder vivir lo que es su parto** [contradictorio].
- f. se hace un relajo// pues <~ps> para llegar al taekwondo/ ¿qué día?/ el/ jueves [...] había un tráfico/ para cruzar **lo que es Pericoapa y Galerías**// dije/ «apúrese señor/ que ya no llego a clase de seis/ (risa)/ muévale/ muévale» (ME-214-21M-02, turno 1022) [específico].
- g. Me senté en **lo que es mi mesa**.

El ejemplo (16a), tomado de la prensa a propósito de un hecho muy concreto, *añade* una lectura inespecífica, sin necesariamente impedir la específica. El caso de (16b), sin PIR, promueve una lectura específica, y prueba de ello es que (16c), donde se introduce una PIR sobre el material de (16b), tiene problemas de aceptabilidad para algunas personas, por la contradicción entre lo específico y lo inespecífico.¹¹ En el mismo sentido aparecen (16d) y (16e), siendo (16e) de todavía más dudosa aceptabilidad. Obsérvese que (16c) y (16e) se vuelven completamente aceptables en dos casos: a) si el hablante emplea la PIR para reducir la fuerza ilocutiva (atenuación), como en (15b); o b) si el hablante despragmaticaliza la PIR (16g). Por último, es importante observar que, aunque la PIR facilita lecturas inespecíficas, algunos ejemplos, como (16f), siguen otorgando lecturas específicas. Por su parte, los casos específicos permiten con facilidad lecturas atenuadas o despragmaticalizadas.¹²

Es posible hablar de MODULACIÓN DE LA FUERZA ILOCUTIVA en general, porque la perífrasis informativa parece útil para acentuar las lecturas asertivas en algunos casos, como en (17):

-
- 11 Es importante aclarar que, cuando se habla de aceptabilidad, no es solo la de los autores. Los casos hipotéticos mencionados durante el trabajo fueron presentados al menos a media docena de jueces, todos ellos hablantes naturales de la ciudad de México, con nivel medio o alto de estudios, hombres y mujeres, de tres grupos de edad: 20-34, 35-54 y 55 en adelante.
- 12 Otro aspecto en el que hay que profundizar es en la lectura como ejemplar o clase que suscitan ciertos ejemplos: *voy a comprar lo que son dos mesas* refiere a ejemplares, pero *voy a comprar lo que es dos mesas* suscita referencia de clase.

(17) a. el arma/ **lo que es** el arma/ el cañón/ trae un rallado adentro (ME-232-21H-04, turno 356).

c. Voy a **lo que es** precisamente el Zócalo.

Se ha venido estableciendo que las PIR subrayan o resaltan un constituyente que desempeña cierta aportación informativa. Algunos casos, como (17a), parecen aportar un mayor grado de asertividad (contundencia). Un punto interesante es si las PIR producen INTERPRETACIONES EXHAUSTIVAS. Dada la concepción del foco como la presencia de alternativas que son relevantes para la interpretación de expresiones lingüísticas (Krifka, 2007, p. 18), podría tratarse de una construcción que contribuya a la interpretación *exhaustiva*, es decir, que la denotación del foco sea la única que conduzca a una proposición verdadera (p. 33). Obsérvese, en ese sentido, cómo muchos de los ejemplos parecen poder combinarse (o incluso reemplazarse) con *precisamente* (al igual que las copulativas enfáticas). La interpretación exhaustiva encajaría bien asimismo con los casos de atenuación (algo así como *quisiera pedirte solo lo que es el dinero que me debes*).

Las PIR no llevan a cabo APORTACIONES EPISTÉMICAS O EVIDENCIALES. Aunque era una de nuestras expectativas iniciales, no hemos encontrado casos claros en que la PIR colabore a establecer grados de certidumbre sobre lo enunciado o a determinar la fuente o la veracidad de la fuente.¹³

En general, las PIR NO SON CONTRASTIVAS, aun siendo focales. Es decir, cuando son focos tienden a ser focos informativos.¹⁴ Suelen presentarse sin contraste, como en (18a), (18b) y (18c), pero cuando lo presentan suelen recurrir a elementos espacio-temporales que contribuyan a marcarlo, como *ahora* en (18d) y el copretérito en (18e) (es decir, los casos de PIR periféricas o no canónicas); también, la polaridad negativa en la oración principal permite establecer el contraste (18f). Esta

13 Se ha defendido que la evidencialidad desempeña un papel importante en algunas variables morfosintácticas, como las atribuidas a *de* en algunos trabajos sobre el (de)queísmo (García, 1986; Schwenter, 1999; Guirado, 2006; Bentivoglio y Sedano, 2011). A nuestro juicio, lo que encontramos en algunos de nuestros casos es modulación de la fuerza ilocutiva, como atenuación y quizá asertividad (*supra*), más allá de que no siempre sea claro cómo definir la marcación de modalidad epistémica y la marcación de evidencialidad (Torres Sánchez, 2013; Gil Burgoin, 2013).

14 Una posibilidad es que la marcación informativa establecida por la PIR sea de *verum focus* o foco de polaridad (Leonetti y Escandell Vidal, 2009). Un argumento a favor sería el carácter asertivo que suele tener la perífrasis informativa. Sin embargo, existen varios contraargumentos: a) la PIR no solo es una construcción focal, sino que introduce también otras funciones y, en general, cuando se trata de un foco, este no es contrastivo; b) la polaridad de la proposición previa tiene que ser negativa con el foco de polaridad; y c) las PIR no son en general casos de anteposición, y, cuando lo son, aparece copia gramatical, etc.

es una característica importante, pues separa a las PIR de las copulativas enfáticas, que tienden a ser contrastivas (19), aunque no siempre (como en *esa es la casa en la que viví de niño*). Es tentador pensar que la no contrastividad de las PIR, y su distanciamiento con respecto a las copulativas enfáticas (siempre focalizadoras y con tendencia a la contrastividad, frente a las PIR, asociadas a cualquier función informativa y pocas veces contrastivas), favorece la conversión de las PIR en perífrasis moduladoras de la fuerza ilocutiva. La contrastividad no es permitida en los casos más prototípicos de PIR, sin elementos que la autoricen, como en (20). La mayoría de los ejemplos con contraste, en cualquier caso, se apoyan en el cambio de tiempo verbal (21), ya en los límites de la PIR si se perdieran totalmente las características de verbo débil.¹⁵

- (18) a. me subo al metro y hasta Pantitlán [...] sí/ pues <~ps> se va casi todo **lo que es** Circuito Interior (ME-155-31H-01, turno 473) [PIR focal sin contraste].
- b. Estuve en lo que es el Zócalo [PIR focal sin contraste].
- c. Llegamos a lo que viene siendo la catedral [PIR focal sin contraste].
- d. y me fui así toda la/ orillita del/ **lo que ahora es** el pasillo de de/ de las peceras (ME-259-32M-05, turno 69) [PIR focal con contraste].
- e. nos tenía **lo que era** la mitad de terreno de plantas/ y la mitad de puros árboles frutales (ME-293-13M-07, turno 31) [PIR focal con contraste].
- f. **No voy a lo que es el Zócalo**, sino a la calle de Moneda [PIR focal con contraste en ámbito negativo].
- (19) a. Vino es lo que tomé, no agua [copulativa enfática con foco contrastivo].
- b. Lo que es esto es una pluma, no un lápiz [copulativa enfática con foco contrastivo].
- (20) #Estoy viendo lo que es el Zócalo, no la calle de Moneda [inapropiado como PIR].

15 Debe establecerse si ciertos tipos de PIR (como los mencionados en el texto) permiten lecturas contrastivas o si, en realidad, el hecho de que exista contraste en esos casos impide que sean estrictamente PIR. Nuestra solución por el momento es hablar de PIR canónicas y no canónicas, pues un argumento fuerte en contra del análisis como PIR estrictas es que la perífrasis con copretérito de (21a) y con futuro de (21b) no son eliminables sin un cambio de significado. Un problema interesante es el planteado por ejemplos como (18e), en copretérito y con contraste y, sin embargo, eliminable.

(21) a. Estoy en lo que es/era La Merced [sin contraste/con contraste].

b. Estoy en lo que es/será el nuevo museo [sin contraste/con contraste].

Cierto número de casos presentan un CUANTIFICADOR, casi siempre *todo*, como se aprecia en (22).

(22) a. entonces yo sola me fui/ caminando/ **todo lo que es la Lagunilla** (ME-293-13M-07, turno 179).

b. le gusta el fútbol soccer <~sóquer> [a mi hijo]/ de hecho/ lo estamos metiendo/ pero ahorita <~oíta>/ lo sacamos por/ necesidad (risa)/ pero este/ le gusta/ vaya le gusta también/ **todo lo que** es realmente el deporte/ le gusta/ porque lo he visto (ME-279-22H-06, turno 1585).

c. Están los puestos de comida y luego **todo lo que es** el mercado (hablante varón joven de la ciudad de México en conversación espontánea).

Todo cuantifica a la PIR completa, pues concuerda con ella y no con el correferente: *echaron cohetes por toda la plaza y echaron cohetes por todo lo que es la plaza*. Además, contribuye a la lectura como cantidad, antes que como entidad, del relativo: *crucé todo lo que es Miramontes* con el sentido de ‘crucé todo cuanto es Miramontes’. La presencia de *todo* como cuantificador neutro en las PIR podría ser un argumento a favor del carácter inespecífico de estas.¹⁶ Obsérvese (23), sin embargo:

(23) a. ya tiene [un rato <~rato:>] [...] no <~no:> pues <~pus> un año un año y medio [más o] [...] menos/ de la droga-/ de la <~la:> o sea [**lo que es la droga** ¿no?] [...] lo que sí es del <~del:>/ de la l- ahora <~ora> sí que de la cerveza pues <~pus> de vez en cuando me echo unas/ unas cervezas (ME-304-11H-07, turno 414) [sin *todo*, pero con lectura inespecífica tanto del correferente como de la PIR].

b. aquí **lo que es San Agustín** se encharca/ se <~se:~>/ ¿cómo te diré?/ se inunda ¿no? (ME-304-11H-07, turno 414) [lectura específica del correferente y de la perífrasis].

c. aquí **todo lo que es San Agustín** se encharca [siguen siendo específicos ambos].

Que favorezca lecturas inespecíficas no quiere decir que las obligue, como puede verse en (23c), donde tanto el consecuente como la PIR cuantificada se pueden leer de forma específica.

16 Compárese *me fui todo Tlalpan* con *me fui por toda la calzada de Tlalpan*.

La POLARIDAD NEGATIVA es posible en la oración principal, tal como ocurre al comparar (24a) y (24b) y como se ve en (24f). Es posible que la negación desate lecturas contrastivas, como se comentó *supra* y como se ve en (24d). Por otra parte, la negación dentro de la construcción relativa es incompatible con la lectura como PIR, como se ve en (24c) y (24e).¹⁷

- (24) a. El platillo trae lo que es queso y papa.
 b. El platillo **no** trae lo que es queso y papa.
 c. #El platillo trae lo que **no** es queso y papa [sí como relativa semilibre, no como PIR].
 d. ??**No** estoy en lo que es el Zócalo [parece necesitar contraste].
 e. #Estoy en lo que **no** es el Zócalo.
 f. desde luego lo/ que eran jaboncitos cositas así baratas pues <~pus> eso **no** se lo llevaron (ME-281-23H-06, turno 747).

Casi siempre, el acto o el evento de habla al que se suma la PIR es de carácter ASEVERATIVO, como en (25a), y solo muy esporádicamente se integra en otros, como el directivo de (25b).

- (25) a. y finalmente/ yo fui haciéndome cargo poco a poco de **lo que son** los/ los tianguis (ME-277-22H-06, turno 123) [aseverativo, narración].
 b. [me gustaba] venir porque/ cuando veníamos a/ a/ a visitar// ves **lo que eran** las vías/ to-/ ¿todavía/ viste las vías [tú?] [...] atrás (ME-279-22H-06, turno 750) [directivo].

No parece que se trate de un hecho casual, aun considerando que el material aseverativo sea abundante en general en las entrevistas del CSCM. Piénsese en la frecuencia con la que aparecen referencias espaciales y temporales en las PIR, aspecto que se liga fácilmente con discursos descriptivos y con actos de habla aseverativos.

4. Dispersión social de la construcción

La perífrasis informativa de relativo no es exclusiva de México ni mucho menos de la ciudad de México, como puede observarse en (26):

17 Un problema práctico es la escasez en el CSCM de casos en que esté involucrada la negación: los pocos en que aparece esta se encuentra en la oración principal, nunca en la propia PIR. Por ello acudimos a datos de introspección.

(26) a. Estoy en **lo que es** Las Nueve Esquinas (Guadalajara).¹⁸

b. Estamos en **lo que es** la plaza Mayor de Madrid (Madrid, reportaje en TVE Internacional).

c. Me puede dar **lo que es** su nombre, por favor (conversación en un centro de llamadas, julio de 2014, con valor atenuador).

En (26a) puede verse un ejemplo de Guadalajara, Jalisco; (26b) ofrece un dato español. Por otra parte, (26c) presenta un dato a distancia, en un registro telefónico.

Asimismo considérense los casos de (27):

(27) a. **Lo que es yo**, me levanto temprano (quizá general).

b. No; **pues lo que es miedo** no tiene vd. (Emilio Rabasa, 1871, CORDE).

c. Llegas a **lo que viene siendo** el mercado de Portales.

A reserva de un análisis detallado, llama la atención que ciertas construcciones con valor temático, como (27a), son generales en todos los lugares y épocas; obsérvese que el consecuente es animado, como en otras PIR de tema. Podría considerarse si las PIR se expanden desde los temas a otras funciones informativas. (27b) sugiere un tema inanimado en un ejemplo de 1871. El caso (27c) es especialmente interesante, pues la variante *lo que viene siendo* se ha convertido en un estereotipo, es decir, se ha adquirido una relativa consciencia social acerca de su uso.¹⁹ Aunque es posible oír esta particular PIR en boca de ciertas personas, su uso no es en realidad tan abundante: en las 108 entrevistas examinadas no hemos encontrado casos, aunque sí los hemos oído en situaciones cotidianas. Los hablantes más conscientes censuran incluso las construcciones con el verbo en presente simple, lo que sugiere un grado creciente de percepción social.

La tabla 1 presenta la dispersión social de la construcción en el CSCM en los datos prototípicos y en los más afines a las características prototípicas, tal como emerge en los datos de los 108 informantes, según nivel de estudios, género y edad.

18 Agradecemos el ejemplo (26a) a Leonor Orozco; en su opinión, la construcción se oiría en Guadalajara en boca de personas con pocos estudios. También le agradecemos el ejemplo (26c).

19 Se trata de un estereotipo negativo atribuido a personas de bajo nivel socioeducativo. «Presidente de la RAE se suicida ante tanto uso de la frase *lo que viene siendo*», anuncia de manera jocosa *Eldeforma.com*, parodia de periódico electrónico: <http://eldeforma.com/2014/02/27/presidente-rae-se-suicida-lo-que-viene-siendo/>

		F	total <i>lo que</i>	% F/total	% F/254	% esperado	(o-e)×100/e
Estudios	Bajo	77	984	7.8	30.3	30.6	-0.8
	Medio	86	1124	7.7	33.9	34.9	-3.0
	Alto	91	1112	8.2	35.8	34.5	+3.7
Género	Hombres	200	1784	11.2	78.7	55.4	+42.1
	Mujeres	54	1436	3.8	21.3	44.6	-52.3
Edad	Jóvenes (20-34)	122	1148	10.6	48.0	35.7	+34.7
	Adultos (35-54)	60	1310	4.6	23.6	40.7	-41.9
	Mayores (55-)	72	762	9.4	28.3	23.7	+19.8

Tabla 1. Distribución de todas las PIR. N= 254 casos, de un total de 3 220 *lo que*, en 108 entrevistas del CSCM²⁰

La tercera columna por la izquierda en la tabla menciona la frecuencia absoluta (**F**), o número total de casos, de PIR en nuestra base de datos según cada uno de los factores (es decir, cada uno de los subgrupos por nivel de estudios, género y edad). La cuarta columna (**total *lo que***) muestra el número total de casos de cualquier tipo de *lo que* en los informantes (se dejan también a un lado los ejemplos de los entrevistadores y de cualquier otro participante). La quinta compara porcentualmente estas dos columnas previas (**% F/total**). La sexta de las columnas anota el porcentaje de casos de cada factor respecto al total de 254 datos en cada grupo de factores (**% F/254**). Por su parte, la séptima columna establece el número de casos de PIR que debería haberse esperado en relación con el total de datos de cualquier *lo que* según cada factor (este dato presente en la cuarta columna) (**% esperado**). La octava, por último, valora las diferencias entre los porcentajes de casos observados (*o*) y los esperados (*e*), multiplicando las diferencias entre ambas magnitudes por cien y dividiendo el resultado entre los casos esperados (**((o-e)×100/e)**). Esta última columna permite valorar descriptivamente la importancia de las diferencias entre factores y, de esta manera, clasificar de una manera simple e inmediata el efecto de cada factor en cada grupo de factores independientes.

20 Los cálculos de las tablas 1 y 2 se han efectuado con Excel. El índice que permite contrastar lo esperado y lo observado a partir de estadística descriptiva ha sido ideado por los autores del artículo.

Aunque el nivel alto de estudios otorga un saldo ligeramente favorable al uso de la PIR (+3.7), las diferencias con los niveles medio (-3.0) y bajo (-0.8) son muy pequeñas; lo mismo se desprende al comparar los porcentajes de empleo: 8.2, 7.7 y 7.8, respectivamente. Las diferencias entre hombres y mujeres, en cambio, son bastante más marcadas, pues se pasa de un 11.2% en los primeros a apenas un 3.8% en las segundas, lo que se confirma con puntuaciones de +42.1 y -52.3 al confrontar los valores observados con los esperados. En cuanto a la edad, se producen también diferencias llamativas, con 10.6% en los jóvenes, 4.6% en los adultos y 9.4% en los mayores, así como con los correspondientes índices de +34.7, -41.9 y +19.8.

Si bien la tabla 1 permite hacerse una idea general de la distribución de los datos analizados, para extraer consecuencias sociolingüísticas es necesario referirse a las tablas 2, 3 y 4, en las que se consideran solo las 205 PIR estrictamente variables, correspondientes a los casos canónicos de valor pragmático. La tabla 2 ofrece un resumen de los datos análogo al de la tabla 1, y las tablas 3 y 4 presentan sendos análisis de regresión logística escalonada, diferentes entre sí por tener en cuenta tres factores de edad en la tabla 3 y solamente dos en la tabla 4.

		F	total <i>lo que</i>	% F/total	% F/205	% esperado	(o-e)×100/e
Estudios	Bajo	64	984	6.5	31.2	30.6	+2.2
	Medio	67	1124	6.0	32.7	34.9	-6.4
	Alto	74	1112	6.7	36.1	34.5	+4.5
Género	Hombres	164	1784	9.2	80.0	55.4	+44.4
	Mujeres	41	1436	2.9	20.0	44.6	-55.2
Edad	Jóvenes (20-34)	114	1148	9.9	55.6	35.7	+56.0
	Adultos (35-54)	42	1310	3.2	20.5	40.7	-49.6
	Mayores (55-)	49	762	6.4	23.9	23.7	+1.0

Tabla 2. Distribución de solo las PIR canónicas. N= 205 casos, de un total de 3 220 *lo que*, en 108 entrevistas del CSCM

El panorama general de las PIR canónicas en la tabla 2 es semejante al de la tabla 1 con todas las PIR, aunque existe una diferencia importante. Por un lado, se confirma la importancia del género y la edad, y las pequeñas diferencias en cuanto al nivel de estudios. Si bien las personas de instrucción media se perfilan con más

claridad como los que menos emplean las PIR (-6.4 al comparar los datos observados con los esperados), las diferencias con los niveles alto y bajo siguen sin ser muy abultadas. A su vez, el género sigue siendo la variable de comportamiento más claro, pues los hombres concentran el 80% de los casos (con +44.4 en la comparación entre observado y esperado), frente al 20% (y -55.2) de las mujeres. Por último, la edad confirma el papel pujante de los jóvenes, que ostentan el 55.6% de los casos (+56.0), y el alejamiento de los adultos (20.5% y -49.6). En cuanto a los mayores, constituyen precisamente la principal diferencia entre esta tabla 2 y la 1: si bien produjeron el 23.9% de las PIR canónicas, sus resultados se encuentran muy cerca de lo esperado (+1.0), lo que hace sospechar que no tienen un papel ni a favor ni en contra.

Conviene confirmar estos datos descriptivos con un análisis de regresión logística escalonada, que se llevó a cabo con Goldvarb Lion (Sankoff, Tagliamonte, y Smith, 2012) y se presenta en las tablas 3 y 4.

Factores	F	%	P
Género			
Hombres	164	9.2	0.625
Mujeres	41	2.9	0.346
Rango= 28			
Edad			
Jóvenes	114	9.9	0.635
Mayores	49	6.4	0.522
Adultos	42	3.2	0.369
Rango= 27			
Estudios			
Alto	74	6.7	n.s.
Bajo	64	6.5	
Medio	67	6.0	

Input= 0.050, razón de verosimilitud= -714.712, sign.= 0.000

Tabla 3. Probabilidad logística de solo las PIR canónicas. N= 205 casos, de un total de 3 220 *lo que*, en 108 entrevistas del CSCM²¹

- 21 Como es claro para el lector familiarizado con el uso de los programas VARBRUL, el grupo de factores dependiente toma como factor de aplicación el hecho de ser PIR, y como valor de no aplicación todos los demás casos de *lo que*. Los grupos de factores independientes, referidos al género, la edad y el nivel de estudios, fueron seleccionados siempre como significativos en la regresión logística escalonada llevada a cabo en las tablas 3 y 4; la razón de verosimilitud y la significación del modelo se anotan al pie de cada tabla. La diferencia entre ambas tablas es que en la tabla 3 la edad se desglosa en tres factores, y en la tabla 4 en

El cálculo confirma la significación del género y la edad, y la no significación del nivel de estudios. El grupo de factores más importante es el del género, con un rango de 28 y seleccionado en primer lugar durante el procesamiento estadístico: los hombres alcanzan una probabilidad de 0.625, por 0.346 de las mujeres. Por su parte, la edad tiene 27 de rango, y, en los factores del grupo, los jóvenes alcanzan la probabilidad más alta, de 0.635, mientras que los adultos muestran la más baja, de 0.369. Las personas mayores, con 0.522, favorecen muy ligeramente y quedan cerca de la neutralidad (significada por la probabilidad de 0.500).

Dada la complicación del patrón por edades, que deja a los mayores en posición intermedia y con una puntuación poco decantada, se optó por simplificar el grupo de factores, reduciéndolo a solamente dos, los jóvenes frente a los no jóvenes, tal como se presenta en la tabla 4.

Factores	F	%	P
Género			
Hombres	164	9.2	0.628
Mujeres	41	2.9	0.342
Edad			
Jóvenes	114	9.9	0.626
No jóvenes	91	4.4	0.429
Estudios			
Alto	74	6.7	n.s.
Bajo	64	6.5	
Medio	67	6.0	

Input= 0.052, razón de verosimilitud= -718.832, sign.= 0.000

Tabla 4. Probabilidad logística de solo las PIR canónicas. N= 205 casos, de un total de 3 220 *lo que*, en 108 entrevistas del CSCM, con la edad reducida a dos factores

solo dos factores. Los grupos de factores independientes se jerarquizan por orden de importancia de arriba hacia abajo en las dos tablas, según el rango alcanzado (el rango es la diferencia entre la probabilidad o peso logístico mayor y la menor, según la serie de factores de cada grupo de factores, también jerarquizada según si se favorece o no la aplicación del factor dependiente en juego, PIR para el caso). En las tablas 3 y 4, «F» es frecuencia absoluta y «P» la probabilidad logística escalonada. Una probabilidad superior a 0.500 favorece el factor de aplicación, y una probabilidad inferior a 0.500 no lo favorece.

La edad sigue siendo el segundo grupo de factores más importante, ahora con rango 20, pero de manera tal que los jóvenes favorecen las PIR canónicas ($p=0.626$) y los no jóvenes no las favorecen ($p=0.429$). Aunque este cálculo simplifica los datos, permite apreciar dónde se encuentra la principal brecha por edades.²²

Los datos sugieren en conjunto un patrón patrocinado por los hombres y los jóvenes. Hará falta más investigación en tiempo real y aparente para poder establecer si se trata de un cambio en curso, teniendo en cuenta las redes sociales, las ocupaciones y los eventos comunicativos específicos, pues llama la atención la elevada presencia de las PIR en algunas circunstancias. Así, es especialmente frecuente en ciertos casos, como en las crónicas radiofónicas de reporteros desplazados al lugar de la noticia (28a), o en la descripción de la carta por parte de un mesero (28b), o en una agencia de viajes de la ciudad de México (28c):

- (28) a. Estamos en **lo que es** la plaza de Tlatelolco.
- b. Las enchiladas suizas traen **lo que es** queso y pollo.
- c. Le entrego **lo que viene siendo** el boleto de **lo que es** su viaje (febrero de 2015).

Los casos (28a) y (28b) tienen en común la descripción de piezas de información fuera de la vista del oyente. Ambos, junto con (28c), orientan a la audiencia al tiempo que se asegura la exactitud y confiabilidad de la fuente. Y, sobre todo, se garantiza la exhaustividad de la información comunicada. Tales utilidades discursivas pueden haber alentado la expansión de las PIR a otros contextos y registros.

5. Conclusiones

En suma, lo que hemos encontrado hasta ahora es que puede defenderse la existencia de una construcción específica, la perífrasis informativa de relativo, que resalta constituyentes que desempeñan ciertas funciones informativas, diferente a otras construcciones de relativo mucho mejor descritas en las gramáticas, como las elaboradas con relativos semilibres y relativos complejos, así como distinta también de algunas construcciones semilexicalizadas y de las copulativas enfáticas. Se mencionó también la posibilidad de un análisis del tipo [lo que_i es [X_i]_{FOCO}]_{FINP} que sería más económico y que encajaría con las visiones al uso de la estructura informativa. La PIR como perífrasis informativa se planteó desde el primer apartado, y

22 Ambos cálculos, los presentados en las tablas 3 y 4, son válidos. La diferencia consiste en el grado de abstracción que se preste a la edad.

en la descripción de la metodología (segundo apartado) se mencionaron los otros tipos de construcciones relativas.

La acotación de la variación (*envelope of variation*) se llevó a cabo en el tercer apartado del artículo, al tiempo que se establecían las propiedades sintácticas, semánticas y pragmáticas de la construcción estudiada. Entre los rasgos lingüísticos de las PIR canónicas se encuentran su eliminabilidad sin alterar las condiciones de verdad, el empleo obligatorio del relativo neutro, la presencia de un verbo débil (*ser*) que aparece preferentemente en presente y singular de tercera persona, un consecuente muchas veces ligado con el tiempo y el espacio, entre otras propiedades que deberán ser exploradas con mayor detalle.

Por último, en términos sociales (cuarto apartado), una vez acotada la variación, resultó que las PIR aparecen en el CSCM mucho más en hombres que en mujeres y más en personas jóvenes. Estos últimos datos sugieren explorar la posibilidad de que se trate de un cambio en curso, que avanza por el camino de la difusión de las formas menos marcadas y la estigmatización de las más marcadas (como *lo que viene siendo*). Para ello, será necesario comparar los datos obtenidos con los de otras variedades geográficas, con datos procedentes de otros cortes temporales y, en especial, con materiales de diferentes registros (por ejemplo, reportes periodísticos o interacciones con empleados de oficinas, restaurantes, etc.).

Referencias bibliográficas

1. Bentivoglio, P., y Sedano, M. (2011). Morphosyntactic variation in Spanish-speaking Latin America. En M. Díaz-Campos (Ed.), *The Handbook of Hispanic Sociolinguistics* (pp. 168-186). Oxford: Wiley - Blackwell.
2. Bosque, I., y Gutiérrez-Rexach, J. (2009). *Fundamentos de sintaxis formal*. Madrid: Akal.
3. Cestero Mancera, A. M. (2012). El proyecto para el estudio sociolingüístico del español de España y América (PRESEEA). *Español Actual*, 98, 227-234.
4. Company, C. (2003). La gramaticalización en la historia del español. En *Gramaticalización y cambio sintáctico en la historia del español* (pp. 1-62). México: UNAM – UAM – El Colegio de México.
5. Dik, S. C. (1997). *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Berlín – Nueva York: Mouton de Gruyter.
6. García, É. (1986). El fenómeno (de)queísmo desde una perspectiva dinámica de la lengua. En *Actas del Segundo Congreso Internacional sobre el español de América* (pp. 46-65). México: Universidad Nacional Autónoma de México.

7. Gil Burgoin, C. I. (2013). Una revisión de tres clíticos adverbiales en purépecha y su relación con las categorías de evidencialidad y modalidad epistémica. *Cuadernos de Lingüística de El Colegio de México*, 1, 159-205.
8. Guirado, K. (2006). Deixis proposicional en el habla de Caracas: un análisis cuantitativo del (de)queísmo. *Boletín de Lingüística*, 26, 130-156.
9. Krifka, M. (2007). Basic notions of information structure. *Interdisciplinary Studies on Information Structure*, 6, 13-55.
10. Martín Butragueño, P. (1997). El papel de los factores sociales en el orden de palabras en español. En R. Barriga y P. Martín (Eds.), *Varia lingüística y literaria. I: Lingüística* (pp. 511-532). México: El Colegio de México.
11. Martín Butragueño, P., y Lastra, Y. (Coords.). (2011). *Corpus sociolingüístico de la ciudad de México. I: Hablantes de instrucción superior*. México: El Colegio de México.
12. Martín Butragueño, P., y Lastra, Y. (Coords.). (2012). *Corpus sociolingüístico de la ciudad de México. II: Hablantes de instrucción media*. México: El Colegio de México.
13. Martín Butragueño, P., y Lastra, Y. (Coords.). (En preparación). *Corpus sociolingüístico de la ciudad de México. III: Hablantes de instrucción baja*. México: El Colegio de México.
14. Moreno Fernández, F. (1996). Metodología del “Proyecto para el estudio sociolingüístico del Español de España y de América” (PRESEEA). *Lingüística*, 8, 257-287.
15. [NGLE]: Real Academia Española y Asociación de Academias de la Lengua Española. (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
16. Otheguy, R., y Zentella, A. C. (2012). *Spanish in New York: Language Contact, Dialectal Leveling, and Structural Continuity*. Oxford: Oxford University Press.
17. Pérez Álvarez, B. E. (2014). Funciones oracionales y discursivas de la construcción *lo que*. En S. Bogard (Ed.), *Del léxico al discurso. La construcción gramatical del sentido en español* (pp. 71-98). México: El Colegio de México.
18. Sankoff, D., Tagliamonte, S. A., y Smith, E. (2012). Goldvarb Lion: A multivariate analysis application for Macintosh [*Software de aplicación*]. Recuperado de <http://individual.utoronto.ca/tagliamonte/goldvarb.html>
19. Schwenter, S. (1999). Evidentiality in Spanish morphosyntax: A reanalysis of (de)queísmo. En M. J. Serrano (Ed.), *Estudios de variación sintáctica* (pp. 65-87). Madrid: Iberoamericana.
20. Sedano, M. (2003). Más sobre las pseudohendidas y sobre las cláusulas con verbo *ser* focalizador en dos corpus del español hablado en Caracas. *Revista Internacional de Lingüística Iberoamericana*, 1, 175-204.

21. Sedano, M. (2011). *Manual de gramática del español, con especial referencia al español de Venezuela*. Caracas: Universidad Central de Venezuela.
22. Torres Cacoullos, R. (2011). Variation and grammaticalization. En M. Díaz-Campos (Ed.), *The Handbook of Hispanic Sociolinguistics* (pp. 148-167). Oxford: Wiley - Blackwell.
23. Torres Sánchez, N. (2013). *La evidencialidad en las lenguas indígenas americanas. Un enfoque areotipológico*. México: Instituto Nacional de Antropología e Historia.