

\*\*\*pruebas no corregidas\*\*\*

Torres Cacoullos, Rena. (2011). Bajar y subir con on sin SE: las construcciones en la variación morfosintáctica. *Variación variable (I Seminario Internacional de Variación Sociolingüística)*, María José Serrano (ed.), 127-152. Almería: Editorial Círculo Rojo.

**BAJAR Y SUBIR CON O SIN SE: LAS CONSTRUCCIONES  
EN LA VARIACIÓN MORFOSINTÁCTICA**

Rena Torres Cacoullos  
*The Pennsylvania State University*

**1. La construcción como unidad del análisis lingüístico**

La propuesta principal de la teoría de la gramática basada en el uso es que el conocimiento gramatical de los hablantes está basado en su experiencia lingüística, a saber, la frecuencia y los contextos de uso de las formas lingüísticas (Bybee 2010, Croft y Cruse 2004: 291-327; *inter alia*), a diferencia del postulado de que el uso de la lengua (‘actuación’) y el conocimiento lingüístico (‘competencia’) son independientes (Chomsky 1965).<sup>\*</sup> Mientras que en la perspectiva formalista de la sintaxis autónoma las explicaciones para la estructura lingüística se centran en la forma, en la perspectiva funcionalista las estructuras gramaticales se convencionalizan a partir de patrones discursivos (*cf.* Travis y Torres Cacoullos 2011). Se han cuestionado presuposiciones formalistas como la división de la sintaxis y el léxico en componentes separados y las reglas con base en rasgos abstractos. En cambio,

desde una perspectiva funcionalista, las construcciones gramaticales combinan información tanto morfosintáctica como semántico-pragmática (y fonológica) y las generalizaciones gramaticales se derivan “de la repetición de [...] eventos locales” en el discurso (Bybee 2006: 714, *cf.* Du Bois 1987, Goldberg 2006, Hopper 1998, Hopper & Thompson 1980, 1984, Langacker 2008, *inter alia*).

Las construcciones son emparejamientos directos de forma y de función, que pueden incluir posiciones fijas y posiciones abiertas en su estructura secuencial. Así, abarcan desde locuciones fijas con piezas léxicas particulares, por ejemplo: “Lo que pasa es que...”, hasta estructuras más esquemáticas (productivas), como “Lo que VERBO SER que” (Curnow y Travis 2009). Los argumentos a favor de la construcción como unidad de análisis lingüístico han venido acumulándose a partir de observaciones sobre la idiomatidad (Fillmore, Kay y O’Connor 1988), la interacción entre sintaxis y léxico (Langacker 2008), la variación interlingüística (Croft 2001), la adquisición y el aprendizaje (Goldberg 2006, Tomasello 2003) y el cambio lingüístico (Bybee y Torres Cacoullós 2009).

Un ejemplo del papel de las construcciones particulares en la formación de categorías y estructuras más generales es el de las combinaciones de adjetivos con los verbos *quedarse*, *ponerse*, *volverse* y otros que indican cambio de estado o devenir. Los intentos de establecer generalizaciones sobre rasgos abstractos, tales como la duración del estado o el grado de implicación del sujeto, no han podido dar cuenta de los adjetivos que ocurren con cada uno de estos verbos. Bybee y Eddington (2006) proponen que las elecciones de combinaciones de verbo y adjetivo están determinadas por categorizaciones locales basadas en ejemplares. Es decir, las categorías de adjetivos con uno u otro verbo de devenir están basadas en agrupaciones de adjetivos semánticamente relacionados con voces particulares que coaparecen frecuentemente

con este verbo. Por ejemplo, con la combinación de alta frecuencia *quedarse solo* están agrupados *soltera*, *aislado* y otros adjetivos semánticamente relacionados, mientras que *parado* y *tieso* están agrupados con otra combinación frecuente, *quedarse inmóvil*.

En una prueba de juicios de aceptabilidad, Bybee y Edgington (2006) encontraron que las combinaciones de verbo con adjetivo de alta frecuencia en corpus orales y escritos se juzgaban más aceptables que las de baja frecuencia textual. Además, entre las combinaciones de baja frecuencia, resultaron más aceptables las que están semánticamente relacionadas con una frase frecuente (por ejemplo, *quedarse parado*, afín a *quedarse inmóvil*) que las que no guardan una relación tal (por ejemplo, *quedarse orgullo-sísimo*). Por lo tanto, las combinaciones más frecuentes de *QUEDARSE* ADJETIVO forman la base para diferentes grupos de adjetivos por medio de la comparación analógica. Los resultados de este experimento sugieren que los miembros centrales de alta frecuencia de una categoría atraen a miembros semejantes, de manera que la formación de las categorías está basada en ejemplares locales.

En la teoría de la gramática basada en el uso, interactúan lo léxicamente particular y lo gramaticalmente general, que se ubican en un continuo, frente a la dicotomía formalista entre léxico y sintaxis. Evidencias diacrónicas de la interacción entre lo particular y lo general son las que proporciona la evolución de la construcción progresiva *ESTAR* VERBO-ndo, si se presta atención a los casos particulares de alta frecuencia como *ESTAR hablando*. En textos del español antiguo, la frase *ESTAR hablando* se adelanta en los cambios que se observan en la gramaticalización de esta construcción, tanto los cambios semánticos (desgaste del valor locativo originario, que se mide por la coaparición de expresiones locativas) como los estructurales (la fusión de elementos, que se mide por la proporción de apariciones sin elementos intercalados

entre *estar* y el gerundio, o con pronombres proclíticos antepuestos a *estar*). De acuerdo con Bybee y Torres Cacoullós (2009), al manifestar en las etapas iniciales un grado avanzado de desgaste semántico y de unidad estructural, *ESTAR hablando* contribuye a la gramaticalización de la construcción general al degradar el estatus léxico independiente léxico del verbo auxiliar emergente. Por otra parte, *ESTAR hablando* puede servir como el centro de la categoría de los verbos enunciativos o de lengua (*alabar, decir, demandar, explicar, gritar, murmurar, razonar, rogar*), siendo esta la subclase semántica más grande en la construcción *ESTAR VERBONDO* (Torres Cacoullós 2000:120-132). Al atraer más voces semánticamente relacionadas, el caso particular contribuye a la productividad de la construcción progresiva general.

En resumen, las construcciones particulares contribuyen a la gramaticalización (la creación de nuevas construcciones productivas) en la diacronía, como en el caso de *ESTAR hablando* y el progresivo, y ocupan un lugar central en las categorías sincrónicas, como en el caso de *QUEDARSE solo* y *QUEDARSE + ADJETIVO*. En este capítulo se ejemplifica cómo las construcciones particulares contribuyen al perfil de la variación morfosintáctica y cómo pueden incluirse en los estudios variacionistas.

## 2. La teoría de la variación

La teoría de la variación, así como la teoría de la gramática basada en el uso, se ha fundamentado en la primacía del uso de la lengua como objeto de estudio. La preocupación del variacionismo es la estructura gramatical *en el discurso* y las relaciones asimétricas entre forma y función que se dan en el discurso (Sankoff 1988:141). El análisis de la variación comienza por observar la alternancia entre formas diferentes que sirven funciones discursivas similares en una comunidad (Labov 1969, Sankoff & Thibault 1981, cf. Torres Cacoullós 2010). Sin embargo, se ha

sostenido que las que parecieran ser variantes sociales pueden no ser semánticamente equivalentes y que por lo tanto no deben ser analizadas como variables lingüísticas (Lavandera 1978, *inter alia*). Por ejemplo, García (1985:198) pregunta: “¿puede de hecho haber ‘maneras distintas de decir lo mismo’ entre unidades cuya función es precisamente permitir a los hablantes decir cosas distintas?”.

Nuestra hipótesis de trabajo es la *hipótesis de la neutralización en el discurso* (Sankoff 1988). Si bien es cierto que casi siempre se pueden encontrar contextos en los que dos formas distintas cumplen, a su vez, funciones diversas, también es cierto que no todas las distinciones de función son pertinentes en cada aparición de una de estas formas, ni en las intenciones de los hablantes ni en las inferencias de los interlocutores. Por eso, *las distinciones gramaticales entre diferentes formas pueden llegar a neutralizarse en el discurso espontáneo*; es éste, además, el mecanismo discursivo fundamental del cambio gramatical (Sankoff 1988:153). Como no podemos acceder a las intenciones de los hablantes que determinaron la elección de una forma u otra, se trata precisamente de una hipótesis, como lo es forzosamente la postura contraria, de que en el empleo de una forma en vez de otra siempre entran en juego, por sutiles que sean, alguna o algunas diferencias de significado (Sankoff 1988:153-154, cf. Aijón Oliva, en este volumen).

En todo caso, una presunta distinción semántico-pragmática entre dos formas no impide que se defina un contexto variable para aplicar el principio de la responsabilidad (Labov 1972:72) y realizar el análisis multivariante como herramienta heurística para establecer los contextos de uso de dichas formas (Poplack y Tagliamonte 1996:78). A diferencia del proceder todavía común en los estudios lingüísticos de atribuir valores semánticos o comunicativos a las formas a base de ejemplos convenientes, se emplea una argumentación cuantitativa. Los pa-

trones de coaparición con elementos contextuales proporcionan una manera replicable de definir las diferencias entre dos construcciones (Aaron y Torres Cacoullós 2006, Kapatsinski 2009, Poplack y Tagliamonte 1996, *inter alia*).

### 3. El problema lingüístico: el uso variable de *SE* con verbos de movimiento

El pronombre clítico *SE* con verbos intransitivos de movimiento se ha interpretado como marcador de contraexpectativas, que expresa el punto de vista del hablante sobre eventos que se desarrollan contra las expectativas normales (Maldonado 1999:390-394). De acuerdo con este valor, se podría afirmar que se emplea *SE* en (1a) para señalar lo inusitado de acceder a un edificio de gobierno escalando por el balcón (de hecho, se trata de una protesta) y en (2a) por considerarse contrario a las normas sociales el hecho de cenar sin la pareja en la noche de bodas.

(1) como protesta eh-lo que hemos hecho ha sido subirnos tres al-al balcón del Ministerio de Justicia. (a) *SE han subido* arriba. Hemos venido dos disfrazados con un mono, hemos extendido una escalera y los - hemos extendido la escalera y (b) *han subido* tres arriba. Han desplegado una pancarta y a la media hora o por ahí pues han llegado los guardias jurados y la Guardia Civil y los ha sacado a - a palos prácticamente (COREC, CCON013F)

(2) mi mujer no tenía ganas de cenar, cuando llegamos; pero yo dije, pues, mira pues yo sí, así que (a) *ME bajo* a cenar solo y (b) *bajé* a cenar ¡je, je! al comedor, a pesar de recién casado (Madrid, 232)

Maldonado (1999) propone que con *SE*, que es marcador de voz media, las “construcciones absolutas” se vuelven “construc-

ciones energéticas”.<sup>1</sup> Una consecuencia del bajo nivel de distinguibilidad entre participantes y la consiguiente reducción del alcance de la predicación en la voz media es una función focal, en la que *SE* focaliza el cambio de estado del experimentador del evento de movimiento (Maldonado 1999:353-398, *cf.* Kemmer 1993:157). Al focalizar el cambio, las conceptualizaciones absolutas adquieren dinamicidad, se vuelven energéticas y llegan a comprender eventos que se desarrollan contra los deseos o las expectativas normales del hablante. El camino trazado en la evolución histórica de las construcciones energéticas es: absoluto > focalizado > inesperado (Maldonado 1999:390).

El uso de la forma marcada para expresar lo no esperado puede considerarse, por lo tanto, un caso de *pragmaticalización* (Maldonado 1999:437) o de subjetivización, en los términos de Traugott (*cf.* Traugott y Dasher 2002). También pragmático-subjetivo sería el valor propuesto por García (1975:158-161) de expresar un grado mayor de implicación del sujeto, al concentrar la atención en el participante focalizado mediante la doble mención (en el pronombre clítico, además del sufijo verbal de persona).

Otros valores que se atribuyen a *SE* con los verbos intransitivos de movimiento son aspectuales: telicidad, puntualidad, perfectividad. Por ejemplo, en (3a) el marcador medio se usa para una acción perfectiva (morfología verbal de Pretérito) y tética (el locativo copresente *hasta el último piso* indica el término inherente de la acción), mientras que en (3b), sin *SE*, se trata de una situación imperfectiva y atética.

(3a) Y un día, por curiosidad, yo *ME subí* hasta el último piso (Madrid, 385)

(3b) veía unas escaleras como de caracol muy raras y entonces *empezaba a subir*, cuando llegaba al cuarto piso me caía (COREC, CCON032A)



El problema que nos ocupa es el uso variable. A pesar de lo que al principio parece ser una diferenciación clara entre los valores de la forma marcada y la no marcada (ya sea como inferencias del contexto o como polisemias convencionalizadas), saltan a la vista ejemplos en los que la forma marcada (1a)-(2a) y la no marcada (1b)-(2b) aparecen en contextos casi idénticos. Dada esta polivalencia de las relaciones entre forma y función, es idóneo el método variacionista, que permite descubrir patrones globales de uso. La estructura (gramatical y social) de la variación se infiere de la distribución de las formas.

#### 4. El método variacionista: datos y contexto variable

Al aplicar el método variacionista al problema de la alternancia entre *bajar-subir* y *bajarse-subirse*, operacionalizamos las hipótesis acerca del valor de *SE* como factores en un análisis multivariante para someterlas a prueba en un corpus de producción lingüística natural. Los datos fueron extraídos de los corpus del Habla de la ciudad de México (Lope Blanch 1971), del Habla popular de México (Lope Blanch 1976), del corpus del habla de la ciudad de Madrid (Esgueva y Cantarero 1981) y del Corpus de Referencia de la Lengua Española Contemporánea (COREC) (Marcos Marín 1992), que en total contienen unas 700 000 palabras. Se extrajeron además, datos de textos publicados entre 1950 y 2000 por medio del CORDE (RAE): dieciocho obras de teatro procedentes de España, cuatro obras de teatro y seis novelas mexicanas (véase Torres Cacoulllos y Schwenter 2008 para una presentación completa del estudio).

Lo que distingue al método variacionista de otras aproximaciones al estudio de la morfosintaxis que emplean argumentación cuantitativa es el proceso de circunscripción del *contexto variable*. Se define el contexto variable para aplicar el principio de la responsabilidad (*principle of accountability*): dar cuenta no solamente de todas las apariciones de una variante sino también

de las no apariciones (Labov 1972:72, 2004). Es decir, se catalogan los casos de la forma de interés (*SE*) y además los casos de la (presumible) alternativa morfosintáctica (no aparición de *SE*).

Nos ceñimos al principio de la responsabilidad porque es engañoso concluir que una forma ‘marca’ cierto(s) valor(es) sin probar si otras formas también están asociadas a este valor (Poplack y Tagliamonte 1996; Kapatsinki 2009:168). Puede que una forma ocurra habitualmente en un contexto dado, pero esto no equivale a que mantenga una asociación privativa con este contexto si en el mismo aparecen otras formas mucho más frecuentes. Por ejemplo, Poplack y Tagliamonte (1996:84) mostraron que, mientras que uno de los exponentes de tiempo pasado en el pidgin inglés nigeriano (*bin*) se usa como expresión de anterioridad en el 90% de sus apariciones, tiene solo una probabilidad del 5% de aparecer en tal contexto.

Por lo tanto, se empieza por extraer del corpus ambas formas de manera exhaustiva; a saber, toda aparición de *bajar* y de *subir*, con o sin *SE*. Tras una primera categorización de todos los casos, se excluyen los contextos bien definidos que ocurren con suficiente frecuencia como para certificar la ausencia de variación. Se excluyeron un total de 359 casos en los que no apareció la forma marcada: usos transitivos (con objeto copresente) p.ej., *bajar la cabeza* (N = 193); la expresión *subir y bajar* (en diferentes formas finitas), p.ej., *las callecitas [...] suben y bajan así* (N = 95), los casos de aumento de precio, p. ej., *¿Que suben las cosas? la gente pide más sueldos* (N = 27) y el *se* impersonal, p.ej., *una iglesia [...] en donde se puede subir a... por un elevador* (N=11).

Quedaron, así, 646 casos en el contexto variable. A este conjunto cerrado se puede proceder a aplicar los axiomas de la teoría de la probabilidad. El análisis cuantitativo sistemático de las alternancias repetidas, p.ej., *(se) han subido arriba* (1a-1b), *(me)*

*bajé a cenar* (2a-2b), nos permitirá descubrir patrones de coaparición de las variantes con elementos contextuales.

La frecuencia relativa global de las formas marcadas es el 21% (136/646), o sea, aproximadamente un quinto de los datos en el contexto variable. Frente a la hipótesis nula de que la variación es aleatoria o libre, la hipótesis variacionista es que la variación está estructurada, es decir, que ciertos rasgos particulares del contexto lingüístico y social influyen en la elección de variantes. La hipótesis de trabajo es la siguiente:

*Si las formas marcadas subirse-bajarse indican contraexpectativas, deben coaparecer más frecuentemente (que la forma correspondiente no marcada) con elementos contextuales que indiquen contraexpectativas o que sean compatibles con tal interpretación.*

El *condicionamiento lingüístico* se entiende como el conjunto de afirmaciones probabilísticas acerca de los contextos lingüísticos que difieren en las frecuencias relativas de las formas. Predecimos que en ciertos contextos lingüísticos (los factores del análisis multivariante) habrá un incremento (un efecto favorecedor) o una disminución (efecto desfavorecedor) en la frecuencia relativa de la forma marcada. Los factores condicionantes forman grupos de factores (variables independientes).

## **5. El método variacionista: hipótesis y codificación de los datos**

La siguiente tarea es definir los grupos de factores, conjuntos de subcontextos que se presume pueden diferir de manera estadísticamente significativa en la frecuencia de las variantes. En Torres Cacoullós y Schwenter (2008) consideramos la copresencia de una preposición locativa, el tiempo verbal, la persona gra-

matical, el tipo de cláusula, la pieza léxica (*bajar* o *subir*) y, como variables independientes extralingüísticas, el dialecto (peninsular o mexicano) y la modalidad (oral o escrita). Un último factor, que se relaciona con la noción de la construcción, es la referencia a un vehículo. Hemos tenido en cuenta estos grupos de factores para codificar (categorizar) cada uno de los 646 casos.

**5.1. Copresencia de preposición locativa.** La hipótesis que se pretende poner a prueba es que *SE* focaliza el cambio de estado del experimentador del evento de movimiento (Maldonado 1999). Relacionada con esta se halla la hipótesis de que *SE* contribuye aspecto télico o puntual, es decir, implica el punto de origen o el término del movimiento (Clements 2006). Operacionalizamos esta hipótesis con el grupo de factores *Copresencia de preposición locativa*. La predicción es que las frases preposicionales con *a*, *de*, *en* (4a, 4b, 4c) deben propiciar la elección de *SE*. En cambio, *por* (5a), que indica “localización espacial contingente, no inicial ni final” (García Miguel 1995:203), o la ausencia de una preposición locativa (5b) debe desfavorecer la forma marcada.

(4a) *SE habían subido* a la caja, donde estaba la carga del camión (COREC, BNOT004A)

(4b) *SE baja* de la cama y dice: “Uno, dos, tres, cuatro” (COREC, ALUD023A)

(4c) en canoa sí *SE subió*, pero nada más de paseo (México, Habla popular, 143)

(5a) Y luego, *hemos bajado* por la, por la Costa [...] Brava hasta Barcelona. (Madrid, Habla culta, 147)

(5b) tenía una rampa interior para que *subieran* los coches (México, Habla culta, 199)

**5.2. Tiempo verbal.** La hipótesis que se pretende operacionalizar con esta variable independiente es que *SE* sirve para señalar eventos de primer plano (Clements 2006). La predicción es que el pretérito debe propiciar el uso de *SE* (cf. García 1975:269, nota 43), porque el aspecto perfectivo está asociado al primer plano, que proporciona los puntos principales del discurso (Hopfer y Thompson 1980).

**5.3. Persona gramatical.** Si *SE* expresa un mayor grado de implicación del hablante (cf. García 1975), el uso de *SE* debe verse favorecido con sujetos de la primera y la segunda persona (6a) más que con sujetos de la tercera persona (6b) o con referentes no específicos. El razonamiento es el siguiente: la primera (y segunda) persona, como todo elemento deíctico, “exhiben subjetividad” porque sus significados se basan en el punto de vista del hablante (Traugott y Dasher 2002:22, y las referencias incluidas, Serrano y Aijón Oliva 2011, Serrano, en este volumen). Por otra parte, la perspectiva del hablante debería ser menos palpable cuando se está hablando de terceras personas o referentes no específicos, que están más distantes del hablante (Aaron y Torres Cacoullos 2006).

(6) luego (a) **NOS bajamos** a tomar un chato, y (b) él **bajó** con nosotros. (México, Habla culta, 180)

**5.4. La referencia a un vehículo.** En una primera categorización de los casos nos dimos cuenta de que *bajar* y *subir* se usan con una frecuencia apreciable, 19% (125/646), para hablar de salir de un vehículo o entrar en él. El medio de transporte puede ser automóvil, autobús, metro o tren, avión u otro (7). Con base en nuestras observaciones iniciales al codificar los datos, predcimos que tales contextos favorecerán la elección de *SE*.

(7a) **Y bajo** del coche *descalza* (México, Habla popular, 59)

(7b)ME tengo que **subir** en ése. En ese camión no me gusta andar (México, Habla popular, 45)

Los datos codificados de acuerdo con estos factores fueron sometidos al análisis multivariante de “regla variable” (que, a pesar de su nombre, no implica reglas) con el programa *Goldvarb X* (Sankoff, Tagliamonte y Smith 2005). Este tipo de análisis permite modelar las regularidades en los corpus, al revelar los grupos de factores que en conjunto dan cuenta de la mayor parte de la variación, de una manera estadísticamente significativa. Se asigna a cada factor una probabilidad o peso entre 0 y 1, de tal manera que en los resultados presentados en el Cuadro 1 los factores con valores más próximos a 1 propician más la aparición de la forma marcada.

## 6. Resultados del análisis multivariante: factores que propician el uso de *SE*

Cuadro 1 Factores que contribuyen a la elección de formas marcadas *bajarse – subirse* (N = 646, Input = .17 (21% se))

		% SE	N
<b>Medio</b>			
	Oral	.71	34
	Escrito	.34	12
<b>Referencia a vehículo</b>			
	‘Entrar a/salir de vehículo’	.78	50
	Otros usos	.42	14
<b>Preposición locativa</b>			
	<i>a, de, en</i>	.67	33
	Otras o ninguna	.43	16
<b>Persona gramatical</b>			
	1a y 2a persona	.58	31
	3a y sujeto no específico	.46	16
<b>Tiempo</b>			
	Pasado	.55	27
	No pasado	.47	19

Factores no significativos: Tipo de cláusula, Dialecto, Lexema.

\* Los N en la tercera columna no suman a 646 en todos los grupos por factores excluidos (por haber poco casos y no combinarse con otros factores) (Torres Cacoullos y Schwenker 2008:1462, cuadro 3, nota).

De acuerdo con los resultados del análisis multivariante, contribuyen un efecto significativo a la elección de la forma marcada con *SE* cinco de los ocho grupos de factores considerados. No significativos resultaron los grupos *Tipo de cláusula* (principal o subordinada), *Pieza léxica* (*bajar* o *subir*) y *Dialecto o variedad* (peninsular o mexicano). En cuanto a la modalidad, es más probable que se use *SE* en textos orales (.71) que en textos escritos (.34) (para resultados similares, véase Aijón Oliva, en este volumen). De hecho, *bajarse* y *subirse* son tres veces más frecuentes en los datos orales (34% (92/275)) que entre los casos procedentes de obras de teatro o novelas (12% (44/371)). Si cruzamos la modalidad con el dialecto o variedad, en la modalidad escrita no se observa diferencia dialectal, mientras que en los datos orales son más frecuentes las formas marcadas en los corpus mexicanos (39% (58/148)) que en los peninsulares (27% (34/127)) ( $\chi^2 = 4.73$ ,  $p = 0.03$ ).

El efecto favorecedor de un locativo con *a*, *de* o *en* (.67) (véase ejemplos (4), *supra*) proporciona apoyo empírico para el papel de la focalización del cambio de estado o para consideraciones de telicidad-puntalidad en la elección de *SE* con *bajar* y *subir*, de acuerdo con la hipótesis que expusimos. Es decir, aunque siguen siendo más frecuentes las formas no marcadas, la frecuencia relativa de *SE* en este contexto (33%) es superior al promedio (la frecuencia relativa global de 21%).

Los resultados para el tiempo verbal requieren algo más de explicación, porque la distinción pertinente resultó no ser aspectual, entre el pretérito (porcentaje *SE* = 28% (34/120)) y el imperfecto (26% (23/88)), sino temporal: es más probable que aparezca *SE* en contextos pasados (combinando los casos del pretérito y del imperfecto) (.55) que en contextos no pasados (.47). De hecho, la frecuencia relativa de las formas marcadas en el presente de indicativo es apenas del 16% (37/234). Nuestra interpretación del resultado es la siguiente: el presente se usa principalmente en con-

textos estativos y habituales (Torres Cacoullós 2000:182) y también es la forma que se emplea para las afirmaciones generales sobre sujetos no específicos (por ejemplo, *Los gatos tienen cuatro patas*); el efecto desfavorecedor del presente de indicativo puede deberse a la predecibilidad de las situaciones habituales. En cambio, se podría pensar que es más probable que se indiquen las contra-expectativas para un evento pasado, que es un evento real y concreto (*cf.* Aaron y Torres Cacoullós 2006:619-620).

El efecto de la persona gramatical proporciona evidencias de un grado mayor de implicación del hablante con *SE* que sin *SE*. Los sujetos de primera y segunda persona propician la aparición de *SE* (.58) frente a los de tercera persona y los no específicos (.46), de acuerdo con la hipótesis. De hecho, incluso si consideramos por separado los sujetos de tercera persona humanos y específicos, sigue siendo más alto el porcentaje de *SE* en la primera y la segunda persona (31% (64/204) frente a 17% (40/229)).

Consideremos ahora los correlatos con la forma del sujeto de primera o segunda persona, que puede ser no expreso o pronominal. El porcentaje de *SE* es más elevado con pronombre de sujeto expresado (50% (16/32)) que con sujeto no expresado (28% (38/136)) ( $\chi^2$  5.779,  $p = .02$ ). Se ha sostenido que los pronombres de sujeto de primera y de segunda persona proporcionan *peso pragmático* (Davidson 1996) o, más generalmente, que la expresión de un sujeto pronominal subraya icónicamente a su referente como *protagonista* del evento (Serrano y Aijón Oliva 2011:147), como tal vez en el ejemplo (8). De acuerdo con esta interpretación, la frecuencia relativamente elevada de *SE* con los pronombres de sujeto de primera y segunda persona puede tomarse como una medida del mayor implicación del sujeto de *subirse* y *bajarse*.

(8)tuvo que poner unas bombillas en un techo muy alto y no *SE* subía. *ME* tuve que subir yo, lo cual, demuestra lo de la mujer. (Madrid, 189)



En síntesis: los resultados del análisis multivariante han demostrado diferencias no categóricas pero cuantitativamente patentes, al revelar los patrones de coaparición de las variantes con elementos contextuales. Concentrándonos en el condicionamiento lingüístico, resultaron ser factores que condicionan la alternancia entre las formas marcadas *bajarse-subirse* y las no marcadas *bajar-subir* la copresencia de una preposición locativa, el tiempo verbal y la persona gramatical. Se favorecen las formas marcadas cuando coaparecen locativos con *a, de, en*; cuando el evento se ha realizado en tiempo pasado; y cuando el sujeto es el hablante o el interlocutor (primera o segunda persona). Estos contextos son compatibles con la focalización del cambio de estado del experimentador del evento de movimiento y con la expresión de las (contra) expectativas y la implicación del hablante (Maldonado 1999, García 1975).

### 7. Tendencias de construcciones particulares con o sin *SE*

Aunque los datos indican una asociación de *SE* a la focalización, las contraexpectativas y la implicación del hablante, no parece justificado proponer una construcción energética o de contraexpectativas (Maldonado 1999) en un nivel de generalidad donde la única especificación formal sería la presencia de *SE* con *bajar* o *subir* u otro verbo intransitivo de movimiento. Esto es porque dicha asociación es una tendencia más que un valor semántico-pragmático-discursivo global, ya que la forma marcada con *SE* es minoritaria en los contextos (factores) que definimos para el análisis multivariante. Sin embargo, hay evidencias para construcciones particulares con fuertes tendencias de aparición o ausencia de *SE*.

El elemento contextual que más favorece la elección de *SE*, con un peso probabilístico de .78 (Cuadro 1, *supra*), es la referencia a la entrada (*subir*) o la salida (*bajar*) de un vehículo. La

frecuencia relativa de *SE* en este contexto alcanza el 50% (62/125); en los datos orales llega globalmente al 75% (45/60) y en los corpus del habla mexicana, al 83% (38/46). Esta fuerte tendencia parece incluso anular las consideraciones de contraexpectativas: en (9a) se usa *SE* para la salida del vehículo a pesar de tratarse de un evento normal; en (9b) el sujeto es no específico y el aspecto es habitual; en (9c) el sujeto es no específico y el evento es rutinario.

(9a) Yo me ba/ yo *ME bajo* antes. Normalmente, vamos.  
(COREC, CCON001)

(9b) la gente *SE baja* de los coches y compra la droga  
(COREC, ACON026A)

(9c) cada vez que un ciudadano de Madrid *SE sube* en el autobús, en cercanías o en el Metro, eh - el Estado le da 20 pesetas  
(COREC, APOL023A)

Aunque se haya considerado la referencia a un vehículo un grupo de factores en el análisis multivariante, esto no obliga a considerar este contexto como una variable independiente, que existe independientemente de (o previamente a) la elección de *SE*. Se puede considerar que la referencia a la entrada (*subir*) o la salida (*bajar*) de un vehículo y el uso de *SE* forman parte de una construcción única, cuyos elementos se encuentran asociados simultáneamente en la memoria o la representación mental del hablante (cf. Kapatsinski 2009:161).

Cuando se usa *subir-bajar* para hablar de la entrada o la salida de un vehículo, no solamente usan los hablantes *SE* con una frecuencia elevada, sino que también usan frecuentemente una frase preposicional locativa con *a*, *de* o *en*: el porcentaje de los casos de *subir-bajar* de un vehículo con una de estas tres prepo-

siciones es el 49% (61/125); en cambio, fuera del contexto de la referencia a un vehículo, estas preposiciones coaparecen con *subir-bajar* solo en el 21% (111/521) de los casos. Además, el efecto favorecedor de tales preposiciones (véase el cuadro 1) se aplica a los casos de *subir-bajar* fuera del contexto de la referencia a un vehículo, en los que el porcentaje de *SE* es más elevado en la presencia de *a*, *de*, *en* que en su ausencia (22% (24/111) frente a 12% (50/410),  $c^2 = 6.369$ ,  $p = .01$ ); en cambio, en el contexto de la referencia a un vehículo, el porcentaje de *SE* permanece relativamente elevado con o sin la presencia de *a*, *de*, *en* (52% (32/61) frente a 47% (30/64)).

La construcción que emerge de los datos se puede representar como aparece en la figura 1. En su estructura secuencial, las posiciones fijas de la construcción son (una forma de) *subir* o *bajar*; variablemente, pero con fuerte tendencia a aparecer, (una forma correspondiente de *SE*); y variablemente, pero también con tendencia a aparecer, *a*, *de* o *en*. La posición abierta es la de la frase nominal en la clase semántica de un vehículo o medio de transporte.

Figura 1: Construcción particular: *subir-bajar* de un vehículo

<i>subir</i>		<i>a / en</i>
	+ <i>SE</i> (50%-83%) +	(49%) + ‘Vehículo’
<i>bajar</i>		<i>de / en</i>

Un examen minucioso de los datos revela varias secuencias convencionalizadas de palabras, con fuertes tendencias a aparecer con o sin *SE*. Hacia el polo de las locuciones fijas con piezas léxicas particulares encontramos la expresión idiomática, con entrada propia en el *Diccionario* de la RAE, *subirse a la cabeza* (N=5), que además se registra ya en el *Diccionario de Autorida-*

*des* (1726-1739). En los datos peninsulares aparece también *subirse a la parra* (N=3). Estas expresiones aparecen con *SE*.

(10a) Esto como no *SE sube a la cabeza*. (COREC, CCON019A)

(10b) se han puesto en un plan ya de *subirse a la parra* todo el mundo (COREC, CCON037B)

Otras construcciones se ubican más hacia el polo productivo, al tener estructuras más esquemáticas. Las formas de la unidad *subir y bajar*, con varios tipos de sujeto, aparecen sin *SE* (0/95) (ejemplo 11). Tampoco aparece *SE* en las referencias a aumentos o rebajas monetarios, por ejemplo, para referirse a precios o tarifas (0/29) (ejemplo 12).

(11) *suben y bajan* los elevadores (CORDE, Azuela, El tamaño del infierno)

las callecitas [...] *suben y bajan* así (México, Habla culta, 48)

(12) ya *subieron* mucho los precios de... de camión (México, Habla popular, 422)

Es que *han bajado*. [...] hay una reducción de tarifas (COREC, CCON013G)

Cuando se refiere a ascender o descender (varios tipos de) escaleras, es casi categórica la ausencia de *SE* (3/66). Esta construcción particular parece anular el factor aspectual, según el cual la telicidad propicia la elección de *SE*. Por ejemplo, en (13a) se bajan todas las escaleras, una situación con término inherente, sin *SE*; en (13b) se suben los peldaños hasta llegar a un término también sin *SE*.

(13a) Total que que *bajé*, bueno como pude, las escaleras. (COREC, CCON037A)

(13b) Rodrigo *subió* los peldaños de azulejo hasta su pieza interior (CORDE, Fuentes, *La región más transparente*)

En contraste, *SE* aparece en los nueve casos de *subir* en contextos de trepar a un árbol. En (14a) se trata de la subida en sentido figurado de un sujeto inanimado, ‘los recuerdos’, que se comparan con una hiedra, mientras que en (14b) el sujeto es no específico en una afirmación general. Es decir, la construcción particular *SUBIRSE a* (un) ÁRBOL puede invalidar tendencias globales, como el efecto desfavorecedor de los sujetos no animados y los sujetos no específicos. Nótese que ‘subirse a un árbol’ no es más télico o puntual que ‘subir escaleras’. Tal vez las construcciones, con *SE* la de los árboles y sin *SE* la de las escaleras, se originan en considerar el ‘subirse a un árbol’ menos rutinario y más consumidor de energía que el ‘subir escaleras’.

(14a) todos los recuerdos se trastocan, *SE suben* como una hiedra al tronco macizo de un mezquite (CORDE, Azuela, El tamaño del infierno)

(14b) Las higueras yo creo que son los árboles donde más fácil *SE sube* uno. (COREC, CCON034A)

Para resumir: los datos revelan varias construcciones, desde locuciones fijas con piezas léxicas particulares hasta estructuras más esquemáticas (productivas), con tendencias fuertes a aparecer con o sin *SE*. Además, el análisis cuantitativo de la variación entre las formas marcadas y las formas no marcadas muestra el efecto importante de la construcción particular que se refiere a vehículos.

## 8. Conclusiones

Los resultados que hemos analizado se suman a los de otros estudios variacionistas que sugieren un papel importante para las piezas léxicas y las frases frecuentes en la variación morfosintáctica. Por ejemplo, el trabajo de Poplack (1992) sobre la variación en el francés canadiense demuestra que un verbo, *falloir*, ejerce un efecto enorme sobre la elección de formas en contextos subjuntivos. Asimismo, estudios del inglés afro-americano muestran que las formas no marcadas de tiempo pasado son en gran medida casos del verbo *come* (Poplack y Tagliamonte 2001:134-136). Nuestro trabajo sobre el complementarizador *that* en el inglés canadiense confirma que gran parte de su omisión variable se puede atribuir a combinaciones frecuentes de sujeto y verbo como *I think* y *I guess* (Torres Cacoullous y Walker 2009).

Estos resultados tienen sentido en la teoría de la gramática basada en el uso, según la cual la gramática incluye secuencias convencionalizadas de palabras (Bybee 2006:713) o “fragmentos reciclables” (*reusable fragments*) (Thompson 2002:141); la base para el conocimiento gramatical de los hablantes es su experiencia lingüística, y el conocimiento gramatical incluye los contextos de uso de las formas lingüísticas.

Al complementar el análisis multivariante con la consideración de secuencias particulares de palabras, hemos mostrado cómo las construcciones particulares contribuyen a las distribuciones observadas. A la vez, por encima de las construcciones particulares hay patrones generales: aun cuando se excluyen los casos de *subir-bajar* en la construcción que se refiere a vehículos, resultan significativos y con la misma dirección del efecto las variables independientes de la modalidad (oral o escrita) y de la copresencia de una preposición locativa (que operacionaliza la hipótesis de que *SE* focaliza el cambio de estado o, en términos aspectuales, que *SE* expresa telicidad y puntualidad). Concluimos

que las construcciones particulares interactúan con los patrones generales, contribuyendo a la estructura de la variación. Se espera que los estudios variacionistas venideros consideren la contribución de las construcciones particulares al perfil de la variación morfosintáctica.

## REFERENCIAS BIBLIOGRÁFICAS

Aaron, J. E. y R. Torres Cacoullós. 2006. “*Te sales: ¿puede cuantificarse la subjetivización?*”, en R. M. Ortiz Ciscomani (ed.) *Octavo Encuentro Internacional de Lingüística en el Noroeste*, 2º Tomo. Hermosillo, Sonora: UniSon, 43-63.

Bybee, J. 2006. “From usage to grammar: The mind’s response to repetition”, *Language* 82: 711-733.

Bybee, J. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Bybee, J. y D. Eddington. 2006. “A usage-based approach to Spanish verbs of ‘becoming’”, *Language* 82: 323-355.

Bybee, J. y R. Torres Cacoullós. 2009. “The role of prefabs in grammaticization: How the particular and the general interact in language change”, en R. L. Corrigan, E. Moravcsik, H. Ouali y K. Wheatley (eds.) *Formulaic language*, vol. 1: *Distribution and historical change*. Amsterdam: John Benjamins, 187-217.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Clements, J. C. 2006. “Transitivity and Spanish non-anaphoric *se*”, en J. C. Clements y J. Yoon (eds.) *Functional approaches to Spanish syntax: Lexical semantics, discourse and transitivity*. New York: Palgrave MacMillan, 236-264.

Curnow, T.J. y C. E. Travis. 2009. “Clause-combining in Spanish cleft constructions”, trabajo presentado en el 11º congreso de la *International Pragmatics Association (IPrA)*, Melbourne University, Melbourne.

Croft, W. 2001. *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, W. y D. A. Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.

Davidson, B. 1996. “‘Pragmatic weight’ and Spanish subject pronouns: the pragmatic and discourse uses of ‘tu’ and ‘yo’ in spoken Madrid Spanish”, *Journal of Pragmatics* 26: 543-565.

Du Bois, J. W. 1987. “The discourse basis of ergativity”, *Language* 63: 805-855.

Fillmore, C. J., P. Kay y M. C. O’Connor. 1988. “Regularity and idiomaticity in grammatical constructions: The case of *let alone*”, *Language* 64: 501-538.

Esgueva, E. y M. Cantarero (eds.) 1981. *El habla de la ciudad de Madrid: materiales para su estudio*. Madrid: CSIC.

García, E. C. 1975. *The role of theory in linguistic analysis: the Spanish pronoun system*. Amsterdam: North Holland.

García, E. C. 1985. “Shifting variation”, *Lingua* 67: 189-224.

García Miguel, J.M. 1995. *Transitividad y complementación preposicional en español*. Santiago de Compostela: Universidad de Santiago de Compostela.

Goldberg, A. E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Hopper, P. J. 1998. “Emergent grammar”, en M. Tomasello (ed.) *The new psychology of language: Cognitive and functional approaches to language structure*, vol. 1. Mahwah, NJ: Lawrence Erlbaum Associates, 155-175.

Hopper, P. J. y S. A. Thompson. 1980. “Transitivity in grammar and discourse”, *Language* 56: 251-299.

Hopper, P. J. y S. A. Thompson. 1984. “The discourse basis for lexical categories in universal grammar”, *Language* 60: 703-752.



Kapatsinski, V. 2009. “Adversative conjunction choice in Russian (*no, da, odnako*): Semantic and syntactic influences on lexical selection”, *Language Variation and Change* 21: 1-17.

Kemmer, S. 1993. *The Middle Voice*. Amsterdam/ Philadelphia: John Benjamins.

Labov, W. 1969. “Contraction, deletion, and inherent variability of the English copula”, *Language* 45: 715-762.

Labov, W. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W. 2004. “Quantitative reasoning in linguistics”, en U. Ammon, N. Dittmar, K.J. Mattheier y P. Trudgill (eds.) *Sociolinguistics / Soziolinguistik: An international handbook of the science of language and society*, vol. 1, 2a ed. Berlin: Mouton de Gruyter, 6-22.

Langacker, R. W. 2008. *Cognitive Grammar: A basic introduction*. New York: Oxford University Press.

Lavandera, B. 1978. “Where does the sociolinguistic variable stop?”, *Language in Society* 7: 171-182.

Lope Blanch, J. M. (ed.) 1971. *El habla de la ciudad de México: Materiales para su estudio*. Mexico: Universidad Nacional Autónoma de México.

Lope Blanch, J. M. (ed.) 1976. *El habla popular de México: Materiales para su estudio*. México: UNAM.

Maldonado, R. 1999. *A media voz: Problemas conceptuales del clítico se*. México: Universidad Autónoma de México.

Marcos Marín, F. (dir.) 1992. *Corpus de Referencia de la Lengua Española Contemporánea: Corpus Oral Peninsular (COREC)*. [www.lllf.uam.es/~fmarcos/informes/corpus/corpus-six.html](http://www.lllf.uam.es/~fmarcos/informes/corpus/corpus-six.html).

Poplack, S. 1992. “The inherent variability of the French subjunctive”, en Christiane Laeufer y Terrell A. Morgan (eds.) *Theoretical analyses in Romance linguistics*. Amsterdam: John Benjamins, 235-263.

Poplack, S. y S. Tagliamonte. 1996. “Nothing in context: Variation, grammaticalization and past time marking in Nigerian Pidgin English”, en P. Baker y A. Suya (eds.) *Changing Meanings, Changing Functions. Papers Relating to Grammaticalization in Contact Languages*. Westminister, UK: University Press, 71-94.

Poplack, S. y S. Tagliamonte. 2001. *African American English in the diaspora*. Malden, MA / Oxford: Blackwell Publishers.

Real Academia Española. 1726-1739/1990. *Diccionario de Autoridades*. Madrid: Gredos.

Real Academia Española. 2001. *Diccionario de la lengua española*, 22ª edición, <http://www.rae.es/rae.html>.

Real Academia Española. *Corpus Diacrónico del Español* (CORDE), <http://corpus.rae.es/cordenet.html>.

Sankoff, D. 1988. “Sociolinguistics and syntactic variation”, en F. Newmeyer (ed.) *Linguistics: The Cambridge survey* (Vol. 4, *Language: The socio-cultural context*). Cambridge: Cambridge University Press, 140-161.

Sankoff, D. y P. Thibault. 1981. “Weak complementarity: tense and aspect in Montreal French”, en B. B. Johns y D. R. Strong (eds.) *Syntactic change. Natural Language Studies* 25, 205-216.

Sankoff, D., S. Tagliamonte y E. Smith. 2005. “GOLD-VARB X: A multivariate analysis application for Macintosh and Windows”. Department of Linguistics, University of Toronto y Department of Mathematics, University of Ottawa. [http://individual.utoronto.ca/tagliamonte/Goldvarb/GV\\_index.htm](http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm).

Serrano, M. J. y M. A. Aijón Oliva. 2011. “Syntactic variation and communicative style”, *Language Sciences* 33: 138-153.

Thompson, S. A. 2002. “‘Object Complements’ and conversation: Towards a realistic account”, *Studies in Language* 26: 125-163.

Tomasello, M. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Torres Cacoulllos, R. 2000. *Grammaticization, synchronic variation, and language contact: a study of Spanish progressive -ndo constructions*. Amsterdam: John Benjamins.

Torres Cacoulllos, R. 2010. “El estudio de la variación morfosintáctica: volver a la «complementariedad débil» por los canales de gramaticalización”, en P. Martín (ed.) *Primer coloquio de cambio y variación lingüística: Realismo en el análisis de corpus orales*. México: El Colegio de México, 391-410.

Torres Cacoulllos, R. y J. A. Walker. 2009. “On the persistence of grammar in discourse: A variationist study of *that*”, *Linguistics* 47: 1-43.

Torres Cacoulllos, R. y S. A. Schwenter. 2008. “Constructions and pragmatics: Variable middle marking in Spanish *subir(se)* ‘go up’ and *bajar(se)* ‘go down’”, *Journal of Pragmatics* 40: 1455-1477.

Traugott, E. C. y R. B. Dasher. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.

Travis, C. E. y R. Torres Cacoulllos. 2011. “Discourse syntax”, en J. I. Hualde, A. Olarrea y E. O’Rourke (eds.) *Handbook of Hispanic linguistics*. Malden, MA / Oxford: Blackwell.

<sup>1</sup> Se considera que en la voz media (*se perdió*), el iniciador y el término son la misma entidad, pero la distancia conceptual entre estos participantes es menor que en la reflexiva (*se peinó (a sí mismo)*) (Kemmer 1993).