

Aplicaciones dialectales del programa VARBRUL 2S

Javier Medina López
Universidad de La Laguna

Abstract

After the uprising of Sociolinguistics in USA, some new methods of analysing were developed in opposition to what Generative Linguistics had considered of little significance: that is, the new discipline would prove that free variations did not exist, but that it is possible to account for the process by which it is governed through the study of variation and supported by an important statistical apparatus.

In our paper we aim to pointing out some of the studies carried out in the Canary Islands, Spain, using statistical package VARBRUL 2S, as a means to analyse -from different perspectives- some variational processes. We also indicate, in a very simple way, which are the steps that form VARBRUL 2S.

I. La dialectología tradicional y la sociolingüística moderna

Considerando lo que ha sido la evolución lingüística en los últimos años, se ha podido advertir que los límites de determinadas disciplinas, como la dialectología tradicional por un lado y la moderna sociolingüística por otro, quedan desde el punto de vista teórico bien delimitados, pero cuando pasamos al plano de la puesta en práctica de los mecanismos que utilizan una y otra disciplina, las confluencias van siendo cada vez más frecuentes¹. Así, por ejemplo, no es ya extraño observar el aprovechamiento metodológico que ofrece la sociolingüística para poner en práctica descripciones que corresponden al plano dialectal². El problema fundamental, creemos, reside en el aspecto «multidimensional» que proporciona un acto comunicativo y la complejidad que conlleva el análisis -desde cualquier perspectiva- del mismo. En este sentido, coincidimos con las afirmaciones de Francisco Moreno cuando señala que la dificultad de analizar cualquier acto comunicativo exige acogerse a disciplinas distintas que se preocupan por cada una de esas dimensiones (Moreno Fernández 1990:205). De este modo, los procesos de variación lingüística pueden ser objeto de análisis dentro de la geografía lingüística, de la lingüística histórica, de la sociolingüística o de la pragmática, y cada una propondrá un determinado modelo de análisis, de acuerdo con sus fundamentos teóricos y exigencias empíricas. Moreno Fernández llega a hablar de una «Lingüística del habla» que daría cobertura teórica y metodológica a la situación que planteamos: la existencia de múltiples modelos descriptivos válidos según la posición teórica que se adopte: la geográfica, la social, la histórica o, en el caso de la pragmática, la comunicativa (Moreno Fernández 1990:206).

II. Aplicaciones del programa VARBRUL

La investigación sociolingüística ha aportado a la lingüística -junto a otras disciplinas conexas- el componente probabilístico carente en otras ramas relacionadas

con el estudio de cualquier lengua, amén de otras peculiaridades. La incorporación de programas estadísticos a los estudios lingüísticos es una circunstancia hasta cierto punto novedosa, tal y como la concebimos hoy en día, dentro de los límites de la variación sociolingüística³.

En el caso concreto del español hablado en las Islas Canarias, donde como sabemos la dialectología tradicional de corte estructural ha contado con importantes exponentes y donde ha dado numerosos y ricos estudios, la investigación lingüística está cobrando en los últimos años unos derroteros enfocados hacia la órbita sociolingüística. Consideramos el análisis dialectal sólo en lo que tiene de acotación diatópica y como variante que presenta una serie de diferencias lingüísticas frente a lo que se entiende por norma estándar de una lengua dada. Así pues, diferenciaremos entre la variante dialectal canaria por un lado (aceptada en la lingüística tradicional como «dialecto canario») y la norma estándar peninsular, o norma académica por otro (con el estatus de lengua oficial).

Dadas estas premisas teóricas, podemos hablar de aplicaciones dialectales de un determinado programa probabilístico. Aunque cierto es que en un contexto más amplio, y dentro de una corriente lingüística que impera en parte del área hispánica y anglosajona, la utilización metodológica de estas técnicas corresponde *sensu strictu* más bien a la sociolingüística.

III. Breve historia del programa VARBRUL

Una vez que William Labov publicó sus investigaciones pioneras sobre la regla variable (Labov 1969 y 1966), en las que practicó un análisis cuantitativo de las reglas que regían la variación lingüística, la eclosión de métodos y programas ha sido un hecho constante en estos últimos años. A partir de ese momento, se crearon encuentros y reuniones de investigadores, lo que dio como fruto los coloquios de la NWAVE (New Ways of Analyzing Variation in English), a partir de 1972.

Algunos años antes, en 1969, Labov propuso un modelo «aditivo», basado en el método estadístico de la varianza en su estudio sobre la elisión de la cópula y del auxiliar *be* en el inglés de la ciudad de New York, según el cual se trataría de analizar cuantitativamente el número de veces que tiene aplicación una regla en cada contexto. De esta forma, la probabilidad de aparición de una regla (p) consiste en sumar (de ahí el modelo aditivo) los distintos valores ($p_a, p_b, p_c \dots p_n$) de los contextos. La formulación del modelo aditivo es la siguiente:

$$p = p_0 + p_a + p_b + p_c \dots + p_n$$

Con posterioridad, Cedergren y Sankoff desarrollaron un modelo matemático para las reglas variables, mucho más complejo llamado *modelo multiplicativo*, que presupone que cada variable debe presentarse como independiente desde el punto de vista estadístico. De esta manera, el modelo multiplicativo da cuenta de la probabilidad de que la regla (p_0) se aplique en el contexto más favorable:

$$p + p_0 \times p_a \times \dots \times p_n$$

por el contrario, también señala las posibilidades de no aplicación de una regla. La probabilidad de que un elemento aparezca es igual a la unidad menos la probabilidad de que el elemento no aparezca. De esta forma, la probabilidad de que varios elementos independientes no aparezcan es el producto de las probabilidades independientes de no aparición:

$$(1-p) + (1 - p_0) (1 - P_a)... (1 - p_n)$$

donde (p_0) es la probabilidad de que la regla se aplique en el contexto menos favorable. Como se ve, el modelo multiplicativo nos muestra matemáticamente las probabilidades de que una regla se cumpla mediante su aplicación o su no aplicación de la misma (Moreno Fenández 1988:122, López Morales 1981).

Con la puesta en marcha de los modelos multiplicativos se crearon programas informáticos que dieron cuenta con un alto grado de fiabilidad de las probabilidades de cumplimiento de las reglas variables. De esta forma surgió el programa VARBRUL (VARiABle RULE 'regla variable'). VARBRUL fue presentado en un artículo por David Sankoff y Pascale Rousseau (1978) dentro de un volumen dedicado a la variación lingüística a finales de los años setenta (Sankoff 1978). La formulación matemática de este nuevo modelo logístico es básicamente igual a la anterior:

$$\frac{P}{1-p} = \frac{P_0}{1-P_0} \times \frac{P_i}{1-p_i} \times \frac{P_j}{1-p_j} \times \dots$$

Como se observa, la probabilidad de la regla viene dada por la multiplicación de una serie de complejos integrados por (P_0), el *input* probabilístico común a todos los contextos, y (P_i), (P_j), etc., que representan la probabilidad con que contribuyen a la regla en cuestión cada uno de los factores (i, j, etc.), divididos por la mitad menos cada uno de dichos elementos.

En definitiva, el modelo VARBRUL “no sólo permite la estimación de valores probabilísticos de cada rasgo contextual propuestos inicialmente por el analista, sino que incorpora un análisis de regresión múltiple que permite reducir el modelo inicial a sólo aquellas dimensiones que contribuyen de manera significativa, en el sentido estadístico, a explicar la variabilidad de los datos” (Cedergren, 1983:153-154).

Tal y como indica López Morales: “[Los programas multiplicativos]... tienen la ventaja de que incorporan explícitamente un cálculo de verosimilitud logarítmica siempre que analizan un conjunto de efectos. Gracias a esto, el programa va comparando los resultados para llevar a cabo un análisis de regresión múltiple con una selección escalonada de factores” (López Morales 1990:85), de tal forma que sólo quedan seleccionadas aquellas dimensiones que contribuyen de manera significativa, en el sentido estadístico, a explicar la variabilidad de los datos. Lo que quiere decir que la interpretación de las codificaciones hechas por el investigador e introducidas en el programa de manera aleatoria, serán sometidas a un proceso de selección en el cual los valores de aplicación superiores a (.05) favorecerán la regla, mientras que los inferiores a esa cifra la limitan; el valor (.05) no afecta la probabilidad de aplicación de la regla⁴.

IV. Descripción de VARBRUL 2S

VARBRUL ha tenido desde sus comienzos varias versiones (VARBRUL 2, 2S y 3). En nuestro caso, expondremos brevemente los pasos que componen la versión 2S.

A) El punto de partida fundamental es la codificación inicial que debe llevar a cabo cada investigador antes de introducir los datos en el programa. Estos ficheros reciben el nombre de *files and file formats*, y son el depósito donde se almacena toda la información que luego se procesará. Deben aparecer tanto las variables dependientes como las independientes.

B) El segundo paso está constituido por lo que se denomina *factor specification file*, o «especificación de factores». Mediante una sintaxis concreta, se le indica al programa cómo debe interpretar los datos introducidos en el fichero anterior.

C) Mediante el *checktok* se detectan los errores posibles en la codificación inicial. Ofrece también el número total de ejemplos con el que se ha trabajado.

D) Desde el punto de vista estadístico, el *countup* da cifras de las ocurrencias y calcula los porcentajes de aplicación y no aplicación de cada factor o grupo de factores. Es la primera información importante desde el punto de vista cuantitativo del programa.

E) En el *condition file for makecell* se especifican los grupos de factores que deben tenerse en cuenta para construir las celdas que servirán de *input* para el *ivarb*. De igual forma, en este momento, el programa permite la realización de cambios, reagrupaciones o estructuraciones que el investigador considere necesarias, siempre que no hubiesen sido explicitadas en el *factor specification file*.

F) *Makecell*. Este archivo crea las «celdas» de las que partirán los diferentes pasos de los que dispone el programa: *ivarb*, *tvarb*, *mvarb* y *crossstab*. Es de suma importancia este momento, pues se trata del *input* que determinará las reglas variables y la crostabulación (*cross-tabulation*) de los distintos grupos de factores. Cuenta, igualmente, todas las apariciones y calcula los porcentajes de aplicación para cada factor. El *makecell* no puede crearse bajo las siguientes condiciones:

- a) La existencia de *knockouts*.
- b) La presencia de un único grupo de factores.
- c) Grupos de factores sin factores.
- d) Exceso de celdas. En este caso aparecerá la señal *memory full*.

G) La opción del *ivarb* va a determinar la regla variable. El *ivarb* es usado cuando la variable dependiente es binomial, es decir, es válido para estudios en los que se contemplan sólo dos rasgos que mantienen oposición. Por ejemplo, *tú/usted*, *perfecto simple/perfecto compuesto*, *thou/you...* Pero, además, el programa ofrece el *tvarb*, para una variable dependiente triple; mientras que el *mvarb* permite el empleo de cuatro o cinco variables dependientes. Para cada caso, las condiciones exigidas por el *ivarb* son las siguientes:

- a) máximo número de variables dependientes: 2.
- b) máximo número de variables independientes: 20.
- c) mínimo número de factores por variable independiente: 2.
- d) máximo número de factores por variable independiente: 30.

- e) máximo número de factores en total en las variables independientes: 49 en ordenadores IBM PC, y 100 en los VAX.
- f) máximo número de celdas: 1000 en IBM PC y 2500 en VAX.

Una vez que se ha llegado a este punto, el *ivarb program* ofrece información referida al número de celdas construidas, número de factores, las probabilidades de aparición de una variable en relación con cada uno de los factores. Asimismo, muestra al final la *chi-square total* y la *chi-square/cell*.

Para la realización del *ivarb*, *mvarb* y *tvarb* debe disponerse en cualquier ordenador de un coprocesador matemático capaz de realizar las operaciones del programa.

H) El *crosstab* calcula de manera bidimensional los cruces del fichero de celdas elaborado por el *makecell*. De esta forma, conjuga los distintos factores y grupos de factores en función de una de las dos variables dependientes. Ofrece los valores también de la *chi-square for totals* y el valor de la *significance* de cada cruce.

I) El propio programa da la opción de un *step up and down*, mediante al cual selecciona -después de un lento proceso de análisis- cuáles son los factores que realmente tienen relevancia estadística. En este momento -opcional por parte del investigador- hay que observar que no siempre la información que suministra el *step up and down* coincide con nuestros propios intereses. Antes de eliminar cualquier factor o grupo de factores debe meditar su eficacia.

V. Modelos de análisis en el contexto dialectal canario

Hasta el momento, las investigaciones realizadas en el ámbito del español de Canarias en las que se ha aplicado el programa VARBRUL son, desde el punto de vista cuantitativo, más bien escasas. Sin embargo, sí existen en la actualidad algunos trabajos -en curso de realización- que cuentan con la incorporación probabilística que ofrece este programa. En la muestra que a continuación exponemos, puede advertirse que se trata de parcelas bien distintas: un trabajo dedicado a la fonética insular (Samper Padilla 1990), uno a la morfología verbal (Herrera Santana & Medina López 1991) y otro a lo que podemos llamar el análisis de los actos de habla coloquiales (Medina López 1992), en especial las «formas de tratamiento».

I. En el caso del estudio fonético, se trata de una investigación llevada a cabo en la ciudad de Las Palmas de Gran Canaria, en la que se describen las reglas de cinco variables fonológicas: *-s/*, *-r/*, *-l/*, *-n/* y *-d/*. Para el cálculo probabilístico se utilizó el primer modelo logístico, creado por Sankoff, a través del programa VARBRUL 2. A partir de un corpus de 28.003 realizaciones de *-s/* (distribuidas en cuatro clases de variantes) se llegan a establecer las reglas lingüísticas y sociales que condicionan cada una de las realizaciones. Los datos del *ivarb* (probabilísticos) indican cuáles son los factores sociales que favorecen o no la regla de aspiración de *-s/*. Así, por ejemplo, tenemos⁵:

Cuadro 1				
CONTRIBUCION DE LOS FACTORES SOCIALES A LA REGLA DE ASPIRACION DE -/S/				
Sexo	Edad		Nivel Sociocultural	
M. (.47)	I	(.61)	1	(.31)
F. (.56)	II	(.45)	2	(.45)
	III	(49)	3	(.49)
			4	(.63)

(Fuente: Samper Padilla 1990:137)

De la interpretación del Cuadro 1 podemos señalar que, dentro de los factores sociales considerados (sexo, edad y nivel sociocultural), son las mujeres (F) (.56) quienes más favorecen la aspiración de -/s/, junto con los hablantes de la primera generación (.61) y los del nivel sociocultural bajo (.63). En los demás casos, como se observa, las cifras presentan coeficientes inferiores al (.50), por lo que no afectan a la realización de la regla de aspiración.

Por el contrario, en los factores sociales que propician la «regla de elisión de -/s/» observamos los siguientes coeficientes:

Cuadro 2				
CONTRIBUCION DE LOS FACTORES SOCIALES A LA REGLA DE ELISION DE -/S/				
Sexo	Edad		Nivel Sociocultural	
M. (.54)	I	(.53)	1	(.29)
F. (.52)	II	(.52)	2	(.39)
	III	(54)	3	(.53)
			4	(.61)

(Fuente: Samper Padilla 1990:143)

Las cifras de las probabilidades apuntan hacia la poca variación existente entre los dos primeros factores sociales (sexo y edad de los hablantes), pues las diferencias entre mujeres y hombres son mínimas, al igual que ocurre con los coeficientes dados para la edad. Sin embargo, en el caso del nivel sociocultural encontramos que quienes más tienden de forma clara a la elisión de -/s/ son los hablantes de la cuarta generación (.61), seguido de los sujetos de la tercera (.53). Las generaciones primera y segunda se muestran indiferentes a la regla de elisión, ya que obtienen resultados inferiores al (.50).

II. La variación morfológica ha recibido menos atención por parte de los lingüistas. Quizá debido a la dificultad que conlleva analizar parcelas concretas de la gramática, sometidas a menores cambios estructurales que los que se producen en el campo fonético, donde se encuentran mayores ventajas al analizar variables discretas, fácilmente segmentables y muy repetidas en el discurso, a lo que habría que añadir que la variación alofónica no implica cambio alguno de significado referencial, factor fundamental para dar por válida una auténtica variación lingüística (López Morales 1989:91, Silva-Corvalán 1989:97-190).

En el español canario, ha sido tradicional la creencia de que las formas simples del perfecto (*canté*) se emplean en contextos en los que el castellano normativo prefiere la compuesta (*he cantado*). Si bien estadísticamente, en efecto, parece haber una preferencia por el uso de la forma simple, no obstante los valores atribuidos a la forma compuesta no difieren, en general, de los de la norma castellana, según hemos tenido oportunidad de analizar (Herrera Santana & Medina López 1991 y 1994).

A partir de una muestra de 24 sujetos pertenecientes al área metropolitana de Santa Cruz de Tenerife-La Laguna, recogimos un total de 1.520 casos de formas verbales de perfecto, de las cuales 1.278 (84%) corresponden a la forma simple y 242 (16%) a la compuesta. Utilizamos el programa VARBRUL 2S, para el cual seleccionamos como variable dependiente las formas *perfecto simple/perfecto compuesto*, y un total de 16 grupos de factores (considerados como variables independientes) entre los que se encontraban tanto las variables lingüísticas como las extralingüísticas. Para el análisis contrastivo de los datos, ofrecemos a continuación los porcentajes según la variable «sexo»:

Cuadro 3		
USO DEL PERFECTO SIMPLE/COMPUESTO SEGUN LA VARIABLE SEXO		
	PERFECTO SIMPLE	PERFECTO COMPUESTO
Hombres	89.0%	11.0%
Mujeres	77.0%	23.0%

(Fuente: Herrera Santana y Medina López 1991:236)

La información relativa a la crostabulación (cruce de datos según diferentes factores) es muy rentable lingüísticamente, puesto que permite poner en contacto un factor con todas las posibilidades de combinación de los otros factores contemplados. Al final de cada «cruce» se ofrecen los valores de chi-cuadrada total (*Chi-square for totals*) y la significación (*significance*). Como ejemplo, veamos el cruce entre los siguientes grupos de factores:

(A): aquellas oraciones que contienen elementos temporales objetivos de tres tipos, según la codificación inicial: a) adverbios/complementos temporales que abarcan el momento de habla (**M**) (*hoy, estos días, ahora...*); adverbios/complementos temporales que excluyen el momento de habla (**E**) (*ayer,*

hace un año...); adverbios/complementos temporales que indican duración o repetición (**D**) (*siempre, nunca, a veces...*).

(**B**): oraciones negativas (**N**) y afirmativas (**A**).

Cuadro 4		
%	Negativas	Afirmativas
D	29%	74%
	71%	26%
E	100%	97%
	0%	3%
M	0%	41%
	100%	59%

Chi-square for totals = 1.296

Significance = .261

(Fuente: Herrera Santana y Medina López, 1994)

Como interpretación de estos datos, se deduce que en enunciados negativos, que contengan elementos temporales que incluyan el momento de habla o que indiquen duración o repetición (**D**) se emplea mayoritariamente la forma compuesta (71%): “Pues *no* se le *ha ido todavía*”; mientras que la simple -en el mismo contexto- es mayoritaria cuando aparece en oraciones afirmativas (74%): “Y más tarde terminé [...] en el Güimar, donde *permanecí dos años*”. Por su parte, con elementos que incluyen el momento de habla también (**M**) en enunciados afirmativos, se prefiere la forma compuesta (59%): “*Este invierno lo hemos pasado* muy bien”.

III. Por último, presentaremos algunos resultados de la aplicación de VARBRUL 2S al estudio de las formas de tratamiento en una comunidad rural de la isla de Tenerife (Buenavista del Norte). Ciertamente, la aplicación de metodologías sociolingüísticas a zonas rurales no está libre de problemas de orden teórico y metodológico, sobre todo cuando se tienen como puntos de referencia más inmediatos los trabajos e investigaciones llevados a cabo en grandes zonas urbanas, tanto en el área anglosajona como en la hispana.

La información pronominal nos la ofreció un cuestionario elaborado *ad hoc*, en el que intentamos analizar el comportamiento de las formas de segunda persona del singular de la variable dependiente *tú/usted*, y seleccionamos para ello un total de diez variables independientes, siempre teniendo en cuenta el aspecto diádico de estas formas. Además, partimos de la base de que en cualquier interacción lingüística la existencia de un emisor y un receptor es fundamental, a la vez que consideramos los aspectos extralingüísticos que funcionan de forma clara y rotunda en este tipo de relaciones tanto coloquiales como formales. La diferencia de estatus social o de edad, por ejemplo, condiciona en muchas ocasiones la presencia de una forma de tratamiento determinada.

Se recogió un total de 7.084 casos pronominales (3.349 corresponden al pronombre *tú* y 3.735 a la forma *usted*) distribuidos en cinco grandes áreas o ámbitos en los que se desarrollaba la interacción lingüística: ámbito familiar, laboral, escolar, social I y social II. La aplicación de VARBRUL 2S resultó muy apropiada para nuestros propósitos, ya que nos ofrecía la ventaja de saber con qué factores dentro de cada variable era más proclive a aparecer una forma en un contexto determinado, y cuáles eran los que no adquirirían una relevancia estadística. Nuestro trabajo contó sólo con variables sociales, toda vez que las lingüísticas propiamente estaban fuera del alcance de nuestro objetivo, y además son las que presentan menos interés en el tema concreto de las formas de trato pronominales. Los contextos literarios o lingüísticos son más fáciles de analizar cuando contamos con un material procedente de lengua escrita, tal y como realizó Turell Juliá en su estudio sobre los pronombres de tratamiento de segunda persona (*thou* y *you*) en el teatro de Shakespeare⁶.

Los cálculos probabilísticos pronominales ofrecen datos ilustrativos de los contextos que favorecen a cada forma. Téngase en cuenta que cuando se trabaja con variables binomiales (como es el caso de *tú* y *usted*), las probabilidades significativas para una variable no lo son para la otra. Así puede comprobarse en los siguientes datos pertenecientes al «ámbito escolar» (Ver cuadro 5):

En el Cuadro 5 se exponen las probabilidades de cada pronombre en relación con las variables independientes consideradas en el estudio. Así, *tú* está unido a las *formas vocativas* del tipo *nombre propio*, *diminutivos nominales*, *hipocorísticos* y derivados del mismo (.96), convirtiéndose casi en regla categórica. En general, todas las generaciones emplean *tú* en mayor o menor medida, con la excepción de la primera, que es la única que claramente favorece el uso de *usted* (.75), circunstancia fácilmente explicable al existir un mayor distanciamiento en relación con la edad del receptor.

Es destacable también la progresión ascendente hacia los usos de *tú* cuando se tiene en cuenta el *nivel de estudios* de los sujetos. Así, por ejemplo, los analfabetos serían los que menos favorecen la aparición de este pronombre (.31) y como contrapartida, por tanto, son los que más emplearían el *usted* (.69), seguido de los de enseñanza básica a favor de *tú* (.64), enseñanza media con un coeficiente de tuteo del (.88) y la enseñanza superior que presenta el mayor índice probabilístico a favor del pronombre informal (.96).

Interesantes resultan los datos extraídos de la variable *estatus del receptor*. El tratamiento al jefe laboral suele ser *usted* (.82). Por el contrario, al compañero de trabajo se le trata de *tú* con un coeficiente de (.84). Poco favorable al *tú* se presenta la probabilidad cuando el tratamiento va dirigido al cliente (.40) y más alta si se habla con una mujer de la limpieza (.61).

En lo que se refiere al *origen paterno/materno*, hay que señalar que favorecen la presencia de *tú* el origen rural tanto del padre (.57) como de la madre (.56), mientras que, por el contrario, el origen urbano de ambos propicia la presencia de *usted* en cifras altas, tanto en el caso del padre (.65) como en el de la madre (.66).

El factor [+ confianza] es el que predispone a los hablantes a emplear *tú* (.61) frente a los casos en los que existiendo ésta sólo se da en un (.39), es decir, nada significativo. Por el contrario, la presencia del pronombre *usted* viene condicionada, entre otros

Cuadro 5			
Indices probabilísticos: (p < .001) TU USTED			
Vocativos:	respetuosos	(.01)	(.99)
	nominales	(.96)	(.04)
Generación:	1ª	(.25)	(.75)
	2ª	(.63)	(.37)
	3ª	(.51)	(.49)
	4ª	(.68)	(.32)
Estudios:	analfabetos	(.31)	(.69)
	básica	(.64)	(.36)
	media	(.88)	(.12)
	superior	(.96)	(.08)
Estatus:	jefe	(.18)	(.82)
	compañero	(.84)	(.16)
	cliente	(.40)	(.60)
	Sra.limpieza	(.61)	(.39)
Origen paterno:	rural	(.57)	(.43)
	urbano	(.35)	(.65)
Origen materno:	rural	(.56)	(.44)
	urbano	(.34)	(.66)
Grado de confianza:	+ confianza	(.61)	(.39)
	- confianza	(.39)	(.61)
Edad del receptor:	mayor	(.16)	(.84)
	joven	(.84)	(.16)

Total Chi-square= 1370.367; N° cells= 331; Chi-square/cell= 4.140

(Fuente: Medina López 1992)

factores, por una probabilidad de (.61) si no se dan las condiciones pragmático-sociales para desarrollar un grado de confianza que permita el uso de *tú*.

Determinante es, igualmente, la variable *edad del receptor*, pues cuando éste es considerado una persona *joven o de igual edad* que el emisor alcanza un coeficiente de (.84) hacia el tuteo. En el lado opuesto, si la persona que recibe el tratamiento es mayor que el que habla, *tú* tiene pocas posibilidades de aparecer (.16). Como consecuencia, si el interlocutor es mayor, la probabilidad es muy alta a favor de *usted* (.84).

VI. *A modo de conclusión*

En el presente trabajo hemos querido exponer parte de las posibilidades que ofrece el empleo del paquete probabilístico VARBRUL 2S. Para ello recurrimos al ejemplo de tres casos de análisis -verdaderamente sociolingüísticos-, aunque en el marco dialectal del español hablado en las Islas Canarias: un trabajo relacionado con la fonética, otro con la morfología verbal y, por último, uno vinculado con el análisis del discurso. El programa VARBRUL no es la solución perfecta a todos los problemas que presenta la descripción de la lengua, tanto en su vertiente oral como escrita. Será válido, precisamente, para aquellas investigaciones que se ajusten a las posibilidades reales que el programa ofrece y para las cuales fue creado, es decir, para aquellas que quieran registrar los índices de variación sociolingüística. Creemos que la utilidad de VARBRUL en estudios referentes al léxico es mucho más compleja, toda vez que este tipo de trabajos exigiría una metodología más precisa que, hasta el momento, no ha satisfecho a los especialistas. Hay que tener en cuenta que los estudios sobre variación léxica son muy escasos y difíciles por las propias características intrínsecas al léxico: problemas de sinonimia, intención comunicativa del hablante y su vinculación con el receptor, actividad pragmática, etc.

A pesar de las objeciones que se puedan hacer a este tipo de programas, creemos que en el caso de VARBRUL cuenta con una garantía de primer orden: la fecunda producción que ha tenido en círculos e investigadores de reconocido prestigio, lo que, sin duda, es una presentación eficaz para modelos probabilísticos de estas dimensiones.

Notas

1. Ciertamente es que la dialectología tradicional no ha contado con los avances y las técnicas que posee la nueva sociolingüística, circunstancia que ha favorecido a ésta en lo que al aspecto teórico y metodológico se refiere.
2. Somos conscientes de las diferencias teóricas que los términos «dialecto» y «lengua» plantean en la dialectología de corte tradicional y en la sociolingüística. Por razones expositivas, mantenemos esta clasificación.
3. Para una visión general de la incorporación del ordenador a los estudios lingüísticos véase, entre otros, Vidal Lamíquiz (1978), López Morales (1981), Moreno Fernández (1984) y Etxebarria (1989).
4. Puede obtenerse una visión general de los distintos mecanismos empleados en la sociolingüística en García Marcos (1993), en especial el capítulo III, dedicado al variacionismo.
5. Nos limitaremos a ejemplificar sólo algunos casos por cada uno de estos trabajos. Remitimos a cada obra para mayor información.
6. Dado que se trata de un trabajo que sirvió como modelo para el que llevamos a cabo en Buenavista del Norte, merece recordar del mismo que el análisis cuantitativo consistió en el examen de 1.828 unidades pronominales, de las cuales 789 (43%) son ejemplos de *thou* y 1.039 (57%) de *you*. Los factores considerados fueron los *literarios* (tipo de obra), los *sociales* (sexo, edad y estatus del hablante; sexo, edad y estatus del interlocutor); los *pragmáticos* (intención del hablante) y los *lingüísticos* (tipo de frase en la que aparece la forma, y estilo). Cfr. Turell Juliá (1987: 197-198).

Bibliografía

- Cedergren, H. (1983) «Sociolingüística», en H. López Morales (corr.) (1983) *Introducción a la lingüística actual*, Madrid: Playor, 147-165
- Etxebarria, M. (1989) «La investigación sociolingüística y el uso de ordenadores», *Lingüística Española Actual IX*, 157-164
- García Marcos, F. (1993) *Nociones de sociolingüística*, Barcelona: Ediciones Octaedro
- Herrera Santana, J. & J. Medina López (1991) «Perfecto simple/perfecto compuesto: análisis sociolingüístico», *Revista de Filología de la Universidad de La Laguna 10*, 227-239
- Herrera Santana, J. & J. Medina López (en prensa) «Sobre los usos de las formas de perfecto en el español atlántico», *Anuario de Letras*, UNAM, XXXII
- Labov, W. (1969) «Contraction, Deletion and Inherent Variability of the English copula», *Language 45*, 715-762
- Labov, W. (1966) *The Social stratification of English in New York City*, Washington D. C.: Center for Applied Linguistics
- Lamíquiz, V. (1978) «Uso del ordenador en la investigación lingüística», *Español Actual 34*, 1-5
- López Morales, H. (1989) *Sociolingüística*, Madrid: Gredos
- López Morales, H. (1981) «Estudio de la competencia sociolingüística: los modelos probabilísticos», *Revista Española de Lingüística 11/2*, 247-267
- López Morales, H. (1990) «La sociolingüística actual», en F. Moreno Fernández (ed.) (1990) *Estudios sobre variación lingüística*, Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá de Henares, 79-87
- Medina López, J. (en prensa) *Estudio sociolingüístico del tratamiento. El uso de «Tú-Usted» en una comunidad rural (Buenavista del Norte)*, Universidad de La Laguna: Secretariado de Publicaciones
- Moreno Fernández, F. (1984) «Los ordenadores en los estudios lingüísticos», *Español Actual 42*, 23-48
- Moreno Fernández, F. (1988) *Sociolingüística en EE.UU. (1975-1985)*, Málaga: Agora
- Moreno Fernández, F. (1990) *Metodología sociolingüística*, Madrid: Gredos
- Samper Padilla, J. A. (1990) *Estudio sociolingüístico del español de Las Palmas de Gran Canaria*, Las Palmas de Gran Canaria: La Caja de Canarias
- Sankoff, D. (ed.) (1978) *Linguistic Variation: Model and Methods*, New York: Academic Press
- Sankoff, D. & P. Rousseau (1978) «Advances in Variable Rule Methodology», en D. Sankoff (ed.) (1978) *Linguistic Variation: Model and Methods*, New York: Academic Press, 57-69

- Silva-Corvalán, C. (1989) *Sociolingüística. Teoría y análisis*, Madrid: Alhambra Universidad
- Turell Juliá, M. T. (1987) «La equivalencia y la diferencia en Shakespeare», en R. Portillo (ed.) (1987) *Estudios literarios ingleses (Shakespeare y el teatro de su época)*, Madrid: Cátedra, 181-202