

## EN TORNO A LOS METODOS DE LA SOCIOLINGUISTICA

Jeanina Umaña Aguiar

### ABSTRACT

An analytic description of the most commonly used methods for gathering and analysing sociolinguistic data.

Los métodos sociolingüísticos contemporáneos se basan en una combinación orgánica de los procedimientos para la investigación seguidos por la sociolingüística propiamente dicha y los usados en otras ramas de la lingüística, así como en la sociología, la etnografía, y la psicología social. Si bien la sociolingüística toma en calidad de préstamo diversos métodos de otras disciplinas, estos se modifican sustancialmente en razón de la naturaleza específica del objeto de estudio.

Es posible dividir los métodos sociolingüísticos en los usados para obtener datos y los utilizados para el análisis respectivo. En el primer grupo se incluyen diferentes métodos de observación, entrevista, experimentación, etc., que se pueden considerar, en su totalidad, como técnicas para la observación de campo. El segundo grupo abarca los procedimientos usados por los sociolingüistas para analizar la información obtenida en el transcurso del trabajo de campo. La selección de un tipo de método de cada grupo depende lógicamente de la naturaleza de la investigación, del tipo de material que se va a analizar, de la hipótesis de investigación propuesta, y de otros factores. A continuación se presenta una descripción analítica de ambos grupos de métodos.

### 1. Métodos para la obtención de datos sociolingüísticos.

El problema de la obtención de datos primarios es uno de los más complejos en la investigación sociolingüística. La dificultad estriba en que cualquier estudio sobre el papel determinante que juegan los factores sociales en el lenguaje requiere que el investigador tome en cuenta un número considerable de variables estratificacionales y situacionales

diferentes entre sí en cuanto a la naturaleza de las relaciones que revelan, tanto en lo relativo al lenguaje como con otras variables.

William Labov (1972b:208) enfoca claramente el problema al describir uno de los aspectos más complejos del trabajo de campo como *la paradoja del observador*: si bien el objetivo de la investigación es descubrir cómo hablan las personas cuando no son objeto de observación sistemática, a fin de obtener esa información es necesario hacer observación sistemática. El hecho es que esa observación en sí origina una situación del habla que afecta el registro estilístico, hace que los informantes traten de usar formas "correctas" del lenguaje y, por lo tanto, estimula una orientación hacia el habla estándar mucho mayor que la que generalmente se da en la comunicación informal cotidiana.

### El cuestionario

Los cuestionarios constituyen uno de los métodos comunes a la sociología y a la sociolingüística para la obtención de datos. Un cuestionario consiste en una serie de preguntas conectadas de manera lógica entre sí y con el problema central de la investigación. Idealmente, cada pregunta del cuestionario es en sí una hipótesis o parte de una hipótesis.

Cuando se usa un cuestionario surgen serios problemas en el análisis sociolingüístico en relación con la validez de la información obtenida por este método. Estos problemas se han discutido durante mucho tiempo en sociología y en psicología social, y adquieren un carácter central en la sociolingüística. De hecho, cuando se usa un cuestionario, la paradoja del observador realmente adquiere relieve, ya que aumenta la posibilidad de que el in-

formante se ajuste a las normas y expectativas del interrogador. Para evitar este peligro, o al menos para reducirlo al mínimo, se les debe dar especial atención a las técnicas de interrogación, así como a la formulación y la secuencia de las preguntas. En realidad, ningún cuestionario puede sustituir por completo la observación de la actividad verbal de los informantes en un ambiente lo más cercano posible al natural.

Cabe preguntarse, entonces, por qué continúan siendo los cuestionarios una de las herramientas más frecuentemente usadas en las encuestas sociolingüísticas. La ventaja de este método es indudablemente el carácter masivo con que se puede aplicar. Lo cierto es que la observación directa de informantes individuales o de grupos pequeños no siempre es suficientemente representativa. Algunas veces es difícil eliminar lo que es típico únicamente de un informante o de un grupo en particular, para poder luego distinguir sin ambigüedades entre la conducta normativa y la desviación individual de la norma.

Un ejemplo interesante del uso de cuestionarios es la encuesta masiva de hablantes de ruso estándar hecha por el Instituto de Lengua Rusa de la Academia de Ciencias de la URSS (Svejcer, 1986; 149 y sig.). Para probar la hipótesis de que el uso de las variantes lingüísticas depende de las características sociales del hablante, se investigaron variantes del ruso estándar que difieren en la connotación estilística (académica, neutral y espontánea), así como en su relativa frecuencia de uso.

La selección de una variante lingüística se relacionó con los siguientes rasgos sociales: edad, educación y lugar donde se obtuvo, estatus social, lugar en donde se vivió la infancia, lugar de mayor período de residencia, regularidad con que se escucha radio o se ve televisión, estatus social de los padres, etc. El método usado corresponde a lo que en sociología se conoce como una encuesta por muestreo mediante cuestionario sin contacto directo. Sin embargo, debido al tipo específico de metas sociolingüísticas, el cuestionario usado difiere sustancialmente de los sociológicos tanto en cuanto a la naturaleza y tipo de preguntas como en el uso de una estrategia más diversa. La principal diferencia se debe a la necesidad de obtener información sobre el habla del informante, más que sobre cualquier concepto sustantivo.

La encuesta se basó en un cuestionario sociológico que incluía los rasgos sociales antes mencionados, y otro lingüístico, es decir, una serie de preguntas sobre las variantes lingüísticas relativas a un

cierto nivel de estructura del lenguaje. El contenido de las preguntas de la parte lingüística se centró en hechos del lenguaje. Ese tipo de pregunta garantiza el carácter objetivo de los datos mejor que las preguntas que elicitán la opinión del informante con base en categorías definidas de manera muy difusa (del tipo, por ejemplo, "¿Tiene usted un buen dominio de la lengua X?"). Las preguntas de selección múltiple, como "Seleccione la palabra más adecuada para...", fueron las más comunes.

Para garantizar la estabilidad de las respuestas, las preguntas relativas a un mismo fenómeno a menudo se duplicaron, planteándolas de manera diferente. Por ejemplo: "La palabra A, ¿la pronuncia usted en forma X o Y?" y más adelante: "Subraye la palabra en que pronuncia X. Si no hay diferencia de pronunciación, subraye ambas".

Finalmente, se hizo un uso amplio de varias técnicas dirigidas a alejar la atención de los informantes del objetivo real del experimento y, en consecuencia, de una orientación deliberada hacia el habla "correcta". Por ejemplo, en el cuestionario morfológico el informante tenía que llenar los espacios en blanco en el texto que eran de interés para el investigador, así como también muchos en los que no había interés. En cuanto a la variación en la acentuación, se pedía marcar el acento no sólo en las palabras claves sino en todas las de la oración. Según los investigadores, se pudo eliminar en un alto grado el peligro psicológico inherente en la habilidad para controlar nuestra propia habla, así como en la orientación hacia las formas de prestigio, mediante la estrategia flexible de la encuesta por cuestionario y, en especial, mediante el sistema de control cruzado y de distractores, en tanto que el carácter masivo de las respuestas compensa la variación causada por las diferencias psicológicas entre los informantes.

### La entrevista

Por otra parte, el objetivo de un estudio muchas veces requiere la observación sistemática de los informantes y el uso de grabaciones. Una conversación directa entre el investigador y el informante, es decir, una entrevista, puede ser una de las formas de dicha observación. Sin embargo, la técnica de estas observaciones en sociolingüística tiene un cierto número de rasgos característicos determinados por el objetivo de la investigación. Después de todo, tanto los cuestionarios como las entrevistas en sociología están dirigidas a la recopilación de in-

formación sustantiva relativa a los fenómenos sociales, en tanto que en sociolingüística se usan técnicas de investigación similares principalmente para obtener la información objetiva que caracterice el habla de los informantes. Esto explica algunas de las diferencias básicas entre estos métodos en ambas disciplinas. Por ejemplo, desde el punto de vista de un sociólogo, usualmente sería bastante ineficaz tratar de lograr una atmósfera durante el curso de una entrevista, ya que se considera necesario mantener una cierta distancia entre el entrevistado y el entrevistador. Por el contrario, en la entrevista sociolingüística es necesario crear una atmósfera de total relajamiento a fin de eliminar la tensión natural que crea el micrófono en muchos hablantes.

La elaboración de la entrevista y las tácticas por seguir requieren un grado de atención considerable. Por ejemplo, el programa de Labov (1966) estaba diseñado de tal modo que le permitiera al entrevistado revelar las diferencias entre los estilos contextuales, desde el espontáneo hasta el estrictamente formal. A fin de estimular el habla emotiva y distraer la atención del informante para que no se orientara hacia el estándar de prestigio, el investigador le pedía que narrara un episodio de su vida en el que hubiera estado en verdadero peligro. También se partía de la premisa de que era posible elicitarse muestras del habla informal espontánea al grabar las observaciones del informante a una tercera persona (cónyuge o hijos, por ejemplo), su conversación con el entrevistador fuera de la entrevista formal, el folklore infantil (rimas para contar, para retar a otros niños), etc. Luego venían, en orden de aproximación al registro formal y de orientación creciente hacia el estándar de prestigio, una entrevista, cuyo objetivo ("un estudio del habla del informante") se le anunciaba al entrevistado, la lectura de un texto que incluía contrastes fonológicos de diagnóstico, la lectura de una lista de palabras y, finalmente, la lectura de pares mínimos.

Otros investigadores usaron una técnica más simplificada para entrevistar a los informantes. Por ejemplo, Levine y Crockett, en su estudio de la /r/ posvocálica, tomaron en cuenta solamente dos estilos de pronunciación: lectura de oraciones separadas y lectura de listas de palabras (Levine y Crockett, 1967). Por otra parte, R. Shuy, W. Wolfram y W. Riley (1966) hicieron un estudio de dialectos urbanos comparando tres estilos de habla: narrativa, comentarios breves y lectura.

La observación puede estar relacionada no sólo con el aspecto objetivo de la actividad verbal del

informante, sino también con los aspectos subjetivos asociados con la orientación de los valores y con las actitudes sociales relativas a los varios códigos o variables sociolingüísticas individuales. En su estudio sobre la estratificación social del inglés en Nueva York, W. Labor (1966) anota que dicha estratificación existe en dos dimensiones: la diferenciación social del lenguaje y la evaluación social de la variación. Una porción considerable del estudio la dedica a los aspectos evaluativos subjetivos de la variación social del habla y, a fin de revelar esos aspectos subjetivos, diseñó varios experimentos. Atención especial merecen los intentos para dar evidencia de la habilidad de los informantes para relacionar ciertos rasgos lingüísticos con variables sociales y étnicas, y los esfuerzos para determinar los estereotipos sociolingüísticos que subyacen este proceso. Con este fin se aplicaron pruebas de identificación socioétnica: al escuchar muestras grabadas del habla "típica" que incluían marcadores socioétnicos, los entrevistados tenían que determinar a qué grupo étnico y social pertenecían los hablantes. (Ver también Brown, 1969).

En otros estudios, a los entrevistados se les pidió que escucharan la grabación de un cierto número de hablantes, cada uno de los cuales hablaba en diferentes lenguas o dialectos. El objetivo, claro está, era poner a prueba las actitudes sociales de los entrevistados con respecto a los códigos en cuestión. Al expresar ciertos juicios de valor en relación con las voces en las grabaciones, los entrevistados automáticamente extendían hacia los hablantes sus actitudes en torno al habla que escuchaban (Lambert, 1967, como estudio clásico de este tipo).

Las actitudes hacia variables sociolingüísticas estables también se pueden evidenciar con ayuda de la llamada "prueba de autoevaluación", para la cual se les pide a los entrevistados que escuchen varias muestras de habla socialmente estigmatizada y que indiquen cuál muestra se parece más a su propia habla. Los resultados de las pruebas indican que, por lo general, los informantes no escogen el habla que se semeja a la propia sino la que corresponde en un mayor grado a su noción de lo que es un estándar de prestigio (Švejc, 1986).

Siguiendo una línea de investigación similar, Zemskaja (1973), les pidió a los informantes que llenaban un cuestionario que indicaran cuál alternativa creían que usaba la gente en el transcurso de una conversación informal. Luego les pidió que escucharan las mismas formas alternativas grabadas en un tono informal y que contestaran la misma

pregunta. La mayoría de los informantes rechazaron las frases estigmatizadas por escrito pero las aceptaron en su versión oral. Es más, en muchos cuestionarios en los que el entrevistado rechazaba las formas socialmente marcadas, los entrevistadores anotaban comentarios como "pero la usa", "las frases fueron obtenidas precisamente de este informante", etc.

#### La observación participativa.

La observación de la actividad verbal de los informantes debe eliminar, o al menos reducir a un mínimo, la influencia que sobre su conducta verbal pueda ejercer el observador. A este respecto resulta de considerable interés la observación participativa, en la que el observador actúa no como un entrevistador sino como uno de los participantes directos en el evento de la comunicación. Este método resulta sumamente eficaz para el estudio de patrones condicionados de conducta verbal en grupos pequeños, cuyos miembros están unidos por relaciones informales. Este es el método que usó, por ejemplo, J.J. Gumperz (1966) para estudiar la conmutación de código determinada por el tópico, en casos de diglosia. En vez de entrevistar a los informantes, el investigador se integró en sus círculos de conversación. La observación era controlada ya que se mantuvieron constantes todas las variables que determinan la conducta verbal excepto una: el tema de la conversación, que variaba periódicamente a fin de revelar el efecto que tiene en la selección de variantes socialmente marcadas.

Métodos similares usó W. Labov para estudiar la conducta verbal de grupos de adolescentes en los ghettos negros de Nueva York. Si bien es cierto que a Labov le ayudó el hecho de que el trabajo de campo lo dirigió un asistente oriundo de Harlem y que dominaba perfectamente el dialecto local y que, además, estaba totalmente familiarizado con el entorno psicológico y sociológico de los grupos de adolescentes, otro factor tuvo una importancia decisiva: la estructura de las relaciones informales dentro del grupo. El hecho de que un grupo informal se reuniera con sus propios líderes en un lugar corriente (los llamados "centros de atención diurna") para participar en un evento comunicativo, ayudaba a producir una atmósfera relajada y aparentemente demostró ser más fuerte que el efecto negativo de un micrófono o la presencia de una persona ajena al grupo. El habla de cada adolescente se grabó en una pista separada gracias al uso de micrófonos individuales, y el uso de la grabación

en video hizo posible registrar no sólo el habla sino también los signos paralingüísticos (expresiones faciales y gestos) (Labov, 1972).

De todo lo anterior se desprende que los métodos en alguna medida se complementan unos con otros. Por lo tanto, el uso de una combinación de varios métodos para recolectar datos sociolingüísticos aumenta la validez del análisis, pero la combinación de métodos siempre estará en función del objeto de estudio, así como de la profundidad y extensión del trabajo de campo por efectuar.

Entre los procedimientos para la obtención de datos citados, es posible distinguir, por una parte, los que se relacionan con la observación controlada y, por otra, los relacionados con la observación no controlada. En el primer grupo se incluyen todos los casos en que el observador controla el progreso del experimento y puede alterar las variables que afectan el fenómeno observado. El segundo grupo cubre todos los casos en los que la comunicación verbal se desarrolla frente al observador en una situación natural de la vida real. En este segundo caso no se presenta la paradoja del observador. Sin embargo, las posibilidades de las observaciones no controladas son claramente limitadas, por cuanto el investigador en esos casos no tiene posibilidad de dirigir el experimento.

Un ejemplo clásico de "observación anónima" de la conducta verbal se encuentra en Labov (1966). Con la ayuda de un micrófono escondido, el investigador grabó el habla de adolescentes que jugaban en la calle, de personas en un café, etc. El punto débil de tales observaciones, desde el punto de vista de las metas básicas de los estudios correlacionales, es que en este caso el investigador no tiene datos precisos para caracterizar socialmente a los hablantes, y se ve obligado a establecer estas características de manera aproximada y con base en las impresiones visuales principalmente. A pesar de su desventaja, las observaciones de la conducta verbal de los informantes en un contexto natural hacen posible determinar en qué grado la situación de entrevista puede afectar la validez de los datos.

Dependiendo de su objetivo, la observación puede catalogarse, además, como exploratoria, de control, o básica. Las observaciones exploratorias preceden al experimento básico y se usan para establecer la hipótesis inicial y el programa de investigación de campo. Las de control se usan para probar los datos experimentales y permitir un cierto grado de distorsión. Por ejemplo, en el estudio de Labov antes citado, la investigación central, basada

en la observación controlada, fue precedida de entrevistas exploratorias con dependientes de tres tiendas de departamentos de Nueva York. En el curso de estas breves entrevistas, el entrevistador se hacía pasar por cliente y hacía preguntas cuyas respuestas requerían el uso de ciertas variables fonológicas. Este estudio exploratorio hizo posible la formulación de una hipótesis de trabajo relativa a la correlación entre las características sociales de los informantes y los valores en los índices obtenidos para las variables fonológicas. En cuanto a las observaciones de control, tienen como fin eliminar el efecto distorsionador de factores externos en los resultados del experimento.

Recientemente se le ha dado un amplio uso a la observación directa de conversaciones espontáneas con la subsecuente interpretación de su contenido mediante la ayuda de informantes. Así, en uno de sus estudios, J.J. Gumperz (1966) grabó las conversaciones entre dos chicanos. La conversación era en inglés con cambios periódicos al español. El texto de la conversación se dividió en episodios centrados en un tema, y los episodios se dividieron en pasajes. Luego, el investigador y cada hablante trataron de analizar el lugar que ocupaba cada pasaje en el discurso para determinar el "significado social" del cambio de código. Se usó el método de sustitución; es decir, la frase que contenía el cambio de código se sustituía por otra en la otra lengua a fin de aclarar el papel funcional y semántico del cambio de código.

## 2. Métodos de análisis de datos sociolingüísticos.

Para procesar los datos obtenidos en los estudios de campo, ciertas variedades de análisis correlacional se usan muy frecuentemente. Como es sabido, este análisis se utiliza para estudiar interrelaciones entre variables independientes y dependientes. Típicas en sociolingüística son las correlaciones en las que ciertos parámetros sociales, estratificacionales y situacionales, son variables independientes, en tanto que los fenómenos lingüísticos son variables dependientes.

La atención del investigador se centra no sólo en los casos de dependencia total (funcional) entre las variables, sino también en aquellos de dependencia parcial.

El análisis correlacional como método de estadística matemática no se debe confundir con el llamado enfoque "correlacional" a los fenómenos lingüísticos. Este último se basa en la necesidad de

evidenciar la covariación de rasgos lingüísticos y sociales, y el problema de las relaciones causales entre los fenómenos en realidad se excluye de toda consideración. A la vez, el análisis correlacional como rama de la estadística matemática, se caracteriza no sólo por el interés de establecer interrelaciones entre ciertos índices, sino también por tratar de encontrar la causa de las relaciones; es decir, las dependencias causales.

Un método muy efectivo de análisis correlacional para procesar datos sociolingüísticos es el que usó Krysin (1974) en su estudio del ruso, basado en una encuesta masiva y en el que empleó el siguiente procedimiento. Las dependencias se describieron según la distribución cuantitativa de las variantes independientemente para cada corte transversal del grupo social (edad, escolaridad, estatus social, características territoriales, etc.). Este procedimiento requería el uso de datos tabulares, gráficos de dependencias y algunos criterios matemático-estadísticos. Así, los rasgos sociales se describieron en términos de su jerarquía y efecto en la distribución de las variantes lingüísticas.

Se supone que entre más amplio sea el rango de un rasgo social, mayor será el número de variantes lingüísticas para las que es relevante. El efecto de un rasgo es directamente proporcional a la magnitud de las diferencias de frecuencia de las variantes lingüísticas (es decir, en términos de la relación del número de personas que prefieren esta variante con el total de la muestra de este grupo social) en grupos definidos según ese rasgo.

Un rango de magnitud (D) se mide en términos de las relaciones de un número de rasgos de las variantes, cuyo uso depende de las características sociales del informante, con el número total de rasgos variantes de un tipo o de un nivel de estructura lingüística dados. La magnitud del efecto (F) se mide en términos de la relación del rango de la variación media, que caracteriza la distribución cuantitativa de los rasgos variantes en grupos sociales, con la frecuencia media del rasgo. Se calcula mediante la fórmula:

$$F_y = \frac{\bar{R}_x}{\bar{P}_x}$$

en donde  $F_y$  es el efecto de un rasgo social (por ejemplo, la edad) en el uso de las variantes,  $\bar{R}_x$  es el rango de la variación media que caracteriza la distribución social de los lexemas en los que se realiza el rasgo  $x$ , y  $\bar{P}_x$  es la frecuencia media del fenómeno (Krysin, 1974; 36-37).

Entre otras cosas, la encuesta reveló que, de acuerdo al rango de fenómenos fonéticos, los rasgos sociales se jerarquizan en el orden (1) condición territorial, (2) estatus social, (3) edad, (4) influencia familiar, (5) educación, (6) influencia de la radio y que, de acuerdo en el efecto que tienen en los fenómenos fonéticos, se ordenan así: (1) condición territorial, (2) edad, (3) estatus social, (4) educación, (5) influencia de la radio.

La encuesta no se limita a establecer ciertos patrones cuantitativos. Su comentario sociolingüístico incluye como regla una evaluación del carácter causal de las relaciones entre los rasgos sociales y las variantes lingüísticas asociadas. Así, al apuntar hacia un amplio rango de fenómenos fonéticos cubiertos por el rasgo de territorialidad, así como hacia el efecto notable de este rasgo, la encuesta indica que esta relación se debe muy probablemente a la formación de variedades locales de pronunciación estándar bajo la influencia del entorno del dialecto (Krysin, 1974; 116-117).

Un intento por establecer relaciones causales con base en datos cuantitativos lo constituye el estudio antes mencionado de L. Levine y H. Crockett, quienes analizaron la variedad del habla socialmente determinada en una de las comunidades de Piedmont, Carolina del Norte, EEUU. El objeto directo de sus observaciones era el uso de /r/ en posición posvocálica. Para cuantificar los datos usaron la variable R, la relación entre el número total de /r/ y el número potencialmente posible de apariciones, multiplicado por 100. Una vez clasificados los informantes según su índice de R, obtuvieron la distribución de los valores de esta variable entre 0 y 100, tanto para la lectura de oraciones (un estilo que se aproxima al habla espontánea) como para la lectura de palabras aisladas (un estilo más cercano al habla formal). La distribución de estos índices revela la ausencia de una distribución normal en torno a valores medios. El análisis de los datos obtenidos tanto para la lectura de oraciones como para la de palabras separadas demostró que la probabilidad de observar esa frecuencia en la distribución normal de la población era mucho menor a 1 en 1000. En otras palabras, de ninguna manera se puede caracterizar la comunidad lingüística como promedio si usamos los valores promedio de R. Por el contrario, los datos parecen sugerir la existencia de dos normas polarizadas: la de valor alto de R y la de valor bajo de R.

Para un análisis adicional de la relación entre las situaciones sociales y la conducta verbal de los informantes, éstos se agruparon en percentiles de

acuerdo con los índices de R. Los autores analizaron la relación entre el valor de R de cada percentil y las diferentes variables sociodemográficas en las situaciones examinadas, y tabularon los resultados. Así obtuvieron correlaciones múltiples, basadas en las interrelaciones entre tres filas de variables: sociodemográficas, situacionales y lingüísticas.

El análisis mostró que el grupo de informantes con educación universitaria incluía un número desproporcionadamente grande de personas con valores máximos y mínimos de R. Un fenómeno similar se presentó al examinar la relación entre los valores de R de los percentiles y la categoría "ocupación". El predominio del grupo superior (intelectuales) se vio en los percentiles con valores máximos y mínimos de R. Esto indicaba una relación curvilínea entre el estatus social, medido según los parámetros antes mencionados, y el índice de R, en donde un nivel educativo alto y el prestigio ocupacional se relacionan tanto con los índices máximos como con los mínimos de R. De aquí la conclusión de que ambas pronunciaciones (con y sin /r/ posvocálica) están sancionadas por las normas sociales de la comunidad lingüística. Entre los informantes de más alto estatus se observa una polarización drástica de las normas de pronunciación.

Al analizar las causas de este fenómeno, los autores proponen una hipótesis relativa a la existencia de dos *grupos de referencia*, orientados ya sea hacia la norma nacional con /r/ posvocálica o hacia la normal sin /r/ de los estados sureños. Estos grupos se constituyen en ejemplo para otros e influyen en la conducta verbal de éstos. Una confirmación final de esta hipótesis requeriría, según Levine y Crockett, el uso de datos adicionales.

Esto parece convincente. Debe tenerse en mente, sin embargo, que difícilmente se puede igualar la influencia de los llamados *grupos de referencia* con el impacto de la norma nacional estadounidense por medio de canales tales como el sistema educativo y los medios de comunicación masivos. Las limitaciones de un enfoque metodológico orientado exclusivamente hacia la teoría de grupos pequeños se hacen sentir en los procedimientos analíticos (Labov, 1971).

Se ha intentado combinar los métodos cuantitativos de análisis con los métodos de la gramática generativa. Por ejemplo, W. Labov introdujo el concepto de *regla variable*, que se supone combina elementos de la gramática generativa y del modelo probabilístico que subyace al análisis estadístico de la conducta verbal (Labov, 1972b; 230-232). Al hacerlo, Labov desarrolla el concepto de *regla op-*

tativa, propuesto por Chomsky y Halle, y lo aplica a la variabilidad socialmente condicionada del habla. La regla se formaliza con base en la notación de Chomsky y Halle, excepto por la inclusión en paréntesis de las variables. Por ejemplo, la regla variable para la vocalización de la /r/ retrofleja en una comunidad del East Side de Nueva York es:

$$[ + \text{cen} ] \rightarrow ( [ - \text{cons} ] ) / [ - \text{cons} ] \text{ \_\_\_\_\_\_ } \sim V$$

y dice que el segmento central /r/ variablemente pierde su carácter consonántico después de una vocal o de una semivocal, a menos que le siga inmediatamente una vocal. El carácter probabilístico de este patrón se refleja en la magnitud, que denota el porcentaje de aplicaciones reales de la regla en relación con sus aplicaciones posibles. Esta magnitud se determina según la fórmula

$$= 1 - K_o,$$

en donde  $K_o$ , incluyendo los entornos que afectan la aplicación de esta regla, es una función de la clase socioeconómica (CSE) y del estilo del habla:

$$K_o = f(\text{CSE, estilo}) = a(\text{CSE}) + b(\text{estilo}) + c$$

Un cuadro similar se observa en la variación de otros elementos de los sistemas fonológicos y morfológicos que están marcados socialmente.

Posteriormente, H. J. Cedergren y D. Sankoff (1974) analizaron el modelo de Labov desde un punto de vista probabilístico y propusieron métodos estadísticos basados en el concepto original de Labov. Formularon tres modelos probabilísticos: el sumatorio, el multiplicativo de no aplicación, y el multiplicativo de aplicación. Según el modelo sumatorio, en el entorno que incluye los rasgos a, j, ... k, la probabilidad de que una regla optativa se aplique es

$$(1) \quad p = p_o + p_i + p_j + \dots p_k,$$

en donde  $p_o$  es un parámetro inicial constante, y  $p_i, p_j, \dots p_k$  son los efectos de los rasgos i, j, ...k, respectivamente. Cada rasgo puede ser una condición fonológica o sintáctica en particular o puede representar las tendencias de un hablante en particular, o de un grupo social, o de una variable social.

El modelo multiplicativo de no aplicación se resume en la fórmula.

$$(2) \quad (1-p) = (1-p_o) \times (1-p_i) \times \dots (1-p_k),$$

en donde se requiere que  $p_o, p_i, \dots p_k$  sean todas probabilidades con un valor entre 0 y 1.

Según el modelo multiplicativo de aplicación de una regla variable,

$$(3) \quad p = p_o \times p_i \times p_j \times \dots p_k,$$

en donde es necesario que  $p_o, p_i, p_j, \dots p_k$  sean valores de probabilidades.

Para P. Rousseau y D. Sankoff (1978; 62) el problema, que se convertía en una de las mayores desventajas asociadas con el uso de reglas variables, era cómo decidir cuál modelo escoger para un grupo de datos dado, por cuanto no existía ninguna prueba lo suficientemente confiable para hacer la selección automáticamente.

A fin de corregir las deficiencias de las reglas variables (programación *ad hoc*, problemas con la prueba de ajuste y la selección de modelos), se reexaminó el problema en su totalidad y se propuso un único modelo para reemplazar a los tres anteriores:

$$\frac{p}{1-p} = \frac{p_o}{1-p_o} \times \frac{p_i}{1-p_i} \times \frac{p_j}{1-p_j} \times \dots \frac{p_k}{1-p_k}$$

Este modelo tiene varias ventajas. A diferencia de (2) y (3), es simétrico con respecto a las probabilidades de aplicación y a las probabilidades de no aplicación de una regla y, al igual que (2) y (3), los valores de  $p_i, \dots p_k$ , son probabilidades que conducen a interpretaciones imposibles de lograr en el modelo sumatorio (1). Además, resuelve los problemas de ajuste y selección, ya que cuando la  $p_i$  tiene valores cercanos a 1, se reduce al modelo (2); cuando los valores son cercanos a 0 se reduce al modelo (3), y se comporta como el modelo sumatorio cuando todas las  $p_i$  tienen valores intermedios (cerca de 0.5).

Este modelo, conocido como el *modelo logístico*, tiene ventajas incuestionables. Aun así, su interpretación probabilística conllevaba algunos problemas y, aún cuando sus autores propusieron un procedimiento para enfrentarlos, las interpretaciones propuestas no han demostrado ser útiles en la práctica. Se ha desarrollado un programa para implementar un algoritmo para calcular los parámetros del modelo logístico, pero muchos de los problemas originales existen todavía, y el trabajo en torno a las reglas variables está lejos de ser consi-

derado resuelto.

Sin poner en duda el valor de los comentarios de Labov en relación con el mecanismo de la influencia de los factores sociales en la selección de las variantes lingüísticas, no se puede pasar por alto la vulnerabilidad de la concepción teórica que subyace las reglas variables. Baste con anotar que los proponentes de la teoría de reglas variables confunden claramente dos aspectos de la comunicación verbal: el subjetivo y el objetivo. Después de todo, las llamadas *reglas variables* se derivan de datos cuantitativos objetivos que caracterizan la actividad verbal de los miembros de una comunidad lingüística. Sin embargo, se les identifica con la *competencia lingüística*, es decir, con una cierta cantidad de conocimiento relativo al aspecto subjetivo de la comunicación. Pero las características estadísticas objetivas de la actividad verbal no son de ninguna manera idénticas al conocimiento subjetivo.

En el grado en que los determinantes sociales de la actividad verbal sean un conjunto de variables, y que parte de cada una contribuya para que se haga una selección específica, en tanto que el resto la restringe, apenas si podremos hablar de la total predictibilidad de la escogencia en cada caso particular. Así, G. Sankoff (1972) estaba en lo cierto al anotar que el poder de predicción de las "reglas sociolingüísticas", incluyendo las reglas variables, es muy limitado, y que estas reglas son básicamente interpretativas.

Otra falla de las reglas variables fue señalada por Paul Kay (1978), quien anotó que ninguno de los modelos de reglas variables propuestos da cuenta de la interacción entre hablante y entorno lingüístico. En otras palabras, en la teoría de una gramática de la comunidad, en la cual se basan las reglas variables, se postula que los entornos lingüísticos son compartidos por todos los miembros de la comunidad, lo cual, como lo han demostrado los estudios sincrónicos, no es siempre el caso. Por ende, Kay llega a la conclusión de que la utilidad teórica de las reglas variables para modelar la heterogeneidad normal de las comunidades lingüísticas parece ser limitada, por cuanto la heterogeneidad a menudo conlleva la interacción de los entornos lingüísticos y los sociales. No obstante, las reglas variables constituyen un valioso instrumento estadístico para analizar la fuerza diferenciada de las variables lingüísticas y sociales que determinan la conducta lingüística en sectores homogéneos y estables de las comunidades lingüísticas, aun cuando no constituyan un modelo adecuado de las comu-

nidades en las que la heterogeneidad sea parte de un proceso de cambio lingüístico en acción.

También se ha tratado de dar cuenta de la heterogeneidad de las comunidades lingüísticas y del cambio lingüístico mediante el uso de las *escalas de implicación*. Uno de los primeros en usarlas fue C. J. Bailey (1969), quien señaló la existencia de relaciones de implicación subyacentes a la selección de variables lingüísticas. Su modelo se basaba en la teoría de las ondas y refleja la difusión de una innovación lingüística a lo largo de un continuum del habla. R.T. Bell (1976) usó una escala de implicación relativamente simplificada para ilustrar algunos rasgos del inglés en el noreste de Inglaterra: (1) el uso de "happen" [ˈappən], en vez de "perhaps" o "maybe", y (2) el uso de /u/ en vez de /ʌ/ en palabras como *but*. La siguiente escala ilustra el uso de estos rasgos en tres hablantes, A, B, y C:

	(1)	(2)
A	+	+
B	-	+
C	-	-

Con base en la selección de variantes, se puede predecir que A es un hablante de la variedad local, en tanto que C es aparentemente un hablante estándar. Por otra parte B, quien rechaza la variante dialectal local extrema (1) pero retiene la regional (2), es un hablante de una variedad regional modificada del inglés estándar.

Las ventajas del modelo de implicaciones residen en su capacidad para reflejar el cambio lingüístico en marcha, para cerrar la brecha entre los modelos sincrónicos y los diacrónicos y para revelar lazos implicacionales no sólo dentro del mismo nivel estructural sino de un nivel a otro.

Se ha tratado de combinar el enfoque implicacional con los modelos generativos. Esto se hace construyendo escalas de implicación en las que se presenta el continuum dialectal como una jerarquía de isolectos, cada uno de los cuales difiere del adyacente por el cambio de una regla de transformación únicamente. No hay diferencias cuantitativas ni cualitativas fundamentales entre los isolectos. Los hablantes de un isolecto son los que usan el mismo conjunto de reglas, generando así las mismas estructuras patentes indicadas en la línea horizontal de la escala. El hecho de que algunos isolec-

tos tengan muchos más hablantes que otros parece ser irrelevante, ya que no se puede considerar el número de hablantes como un rasgo distintivo del sistema. Algunos de los isoclectos son, sin lugar a dudas, más estables que otros, pero este hecho no parece tener importancia para los autores de gramáticas basadas en escalas de implicación. Su meta es explicar las reglas para la generación de isoclectos teóricamente equivalentes, presuntamente en el orden cronológico de su formación, como resultado de la difusión de una onda de cambio lingüístico.

Lo que se propone, en otras palabras, es un modelo transformacional de variación lingüística basado en la premisa de que la variación sincrónica refleja un cambio diacrónico y que, por lo tanto, cada segmento sincrónico se caracteriza por el hecho de que un cambio en una regla se ha extendido a sólo parte de los hablantes.

En las escalas de implicación se representan muchos isoclectos con 1 o 2 hablantes. Los autores en realidad hacen caso omiso del concepto tradicional de comunidad lingüística. Así, uno de los proponentes del enfoque, D. Bickerton (1973), admite que al mismo apenas si se le puede considerar informativo, desde el punto de vista sociológico.

Es difícil estar de acuerdo con los críticos de las reglas variables que tratan de presentar las escalas de implicación como una alternativa a los modelos de Labov. No podemos sino concordar con Labov (1973b), quien cree que al negar la existencia de dialectos (usados por comunidades relativamente homogéneas), al subestimar el papel que juegan los factores sociales en la limitación de la libertad en cuanto a la variación de las unidades lingüísticas, y (añado yo) al aumentar el número y extensión de las reglas de la gramática generativa, los proponentes de las escalas de implicación en realidad reviven el enfoque tradicional de la lingüística estadounidense que, al basarse en el estudio de individuos en vez de comunidades, se orienta hacia la idea de "idiolectos libres".

Como es bien sabido, los estudios descriptivistas tradicionales subestiman la variabilidad y la heterogeneidad de la estructura lingüística (Umaña, 1982). Los modelos de implicación, por otra parte, tienden a sobreenfatizar estos aspectos del lenguaje. Paradójicamente, estas dos tendencias que parecen ser diametralmente opuestas, tienden a ignorar de igual forma la naturaleza social del habla.

Por otra parte, las escalas implicacionales pueden tener un propósito muy útil si no se les considera como alternativas frente a otros métodos sino como complementarios. Parece que en cierto mo-

do las escalas de implicación y las reglas variables se complementan y que un enfoque corrige parte de las deficiencias del otro. El modelo de reglas variables se basa en un concepto más o menos estático de una comunidad relativamente homogénea que usa una sola gramática, en tanto que la escala de implicaciones se basa en un enfoque más dinámico, que enfatiza la heterogeneidad y la variación. Desafortunadamente, todavía no está claro hasta qué punto son mutuamente convertibles estos dos modelos. P. Kay (1978; 81) cita algunos datos según los cuales algunos estudios implicacionales (por ejemplo, los datos de Bickerton sobre los complementizadores en el habla criolla guayanesa) se pueden ajustar bien por medio del modelo logístico. Sin embargo, indica que las escalas de implicación con número elevado de columnas y filas no son susceptibles de análisis por medio de reglas variables, dado el número limitado de factores que éstas aceptan.

Por otra parte, pareciera que algunos de los modelos anteriores de variación lingüística que generalmente se excluyen de la literatura reciente, en especial la estadounidense, sobre teoría de la variación, bien se pueden aplicar al estudio de una estructura lingüística socialmente determinada. Uno de ellos es el modelo de *base común* (Hockett, 1959), que ve el continuum dialectal como un conjunto de microsistemas con una base común que se intersectan. Basado en este modelo existe la posibilidad de un enfoque que haga posible el contraste de subsistemas comunicativos (por ejemplo, un dialecto social con otro, o un dialecto social con una lengua estándar) no como objetos externos entre sí, sino como componentes del mismo microsistema, unidos por una sola base común.

Podemos examinar algunos aspectos de este método con especial referencia a las diferencias léxico-semánticas socialmente determinadas, entre elementos de dialectos sociales y el vocabulario estándar del inglés. Un análisis contrastivo de tales diferencias se puede dividir en dos etapas. En primer lugar, es necesario examinar los principales tipos de oposiciones, aislados al comparar rubros léxicos específicos. Esto hará posible el paso a la etapa siguiente del análisis, es decir, tratar de descubrir la forma en que las diferencias entre los elementos individuales del sistema léxico-semántico afectan otros elementos unidos con los primeros por relaciones intrasistémicas.

Para determinar los principales tipos de diferencias que se evidencian en el nivel léxico-semántico, es necesario tomar en cuenta el plano en el

que las unidades en cuestión revelan sus diferencias, así como el plano de los rasgos comunes que forman las bases de su comparación (el plano de la expresión o el plano del contenido). Los rubros que son idénticos en el plano de la expresión pero que difieren en el del contenido se consideran *divergentes*, en tanto que los rubros que difieren entre sí en el plano de la expresión pero no en el del contenido se catalogan como *análogos*.

En la llamada jerga general (*slang*) del inglés estadounidense, *the man* es "un policía", mientras que en el dialecto etnosocial de los negros (el "inglés negro"), significa "hombre blanco" o "jefe"; en la jerga de los drogadictos significa "narcotraficante" y en la jerga de los músicos es un "director de orquesta". Es decir, el término constituye una correlación múltiple de divergentes léxico-semánticos, cada uno de los cuales es una variante léxico-semántica (semema) de la misma palabra.

Como ejemplo de un grupo de analogías podemos citar la forma estándar *policeman* y sus equivalentes de jerga general y social como *pig*, *big John*, *cop*, *fuzz*, *the Man*, *blue boy*, todos con el significado "policía". Con este significado, *pig* es a la vez parte de la jerga de los jóvenes y del argot del bajo mundo; *big John* lo usan los drogadictos; *blue boy* y *cop* son elementos de la jerga general, mientras que *fuzz* es un elemento que actualmente se usa entre los *declassé*.

Como es de conocimiento general para los lingüistas, el término de oposición al cual se le adscriben ciertos rasgos es un término marcado. Tomando en cuenta que un rasgo social subyace las oposiciones anteriores, podemos hablar de oposiciones que incluyen uno, dos, o varios términos socialmente marcados.

El concepto de rasgo social requiere aclaración. Podemos hablar de dos tipos de marcación social: la estratificacional y la situacional. Por lo tanto, en términos de variación estratificacional, los sememas "hombre blanco", "jefe", "narcotraficante" y "director de orquesta" son términos socialmente marcados para oposiciones divergentes en el tanto en que se asocian con ciertos grupos sociales o socioétnicos. Al mismo tiempo, el semema "policía" es neutro en cuanto a rasgos estratificacionales, ya que la jerga general no sirve como marcador de ningún estrato social. Los términos de oposición análoga como *cop* y *blue boy* tampoco son específicos para ningún grupo. Todos estos rubros se pueden clasificar como términos no marcados estratificacionalmente para  $\Lambda$  oposiciones divergentes y análogas.

Un mismo rubro (semema o lexema) puede no estar marcado en términos de variaciones estratificacionales, pero puede estarlo en términos de la variación situacional del lenguaje. Por ejemplo, el divergente *the Man* ("policía") y sus análogos *cop* y *blue boy*, elementos de la jerga general, se asocian con ciertas situaciones y por lo tanto, se les puede considerar como situacionalmente marcados en términos de las oposiciones correspondientes.

Es necesario distinguir entre las oposiciones marcadas unilateralmente y las marcadas bilateralmente. Por ejemplo, la oposición de los dos divergentes *pig*<sup>1</sup>, "animal doméstico" y *pig*<sup>2</sup>, "policía", se caracteriza por la marcación unilateral, ya que el único término marcado socialmente es *pig*<sup>2</sup>. Al mismo tiempo, la oposición análoga *pig/big John* está marcada bilateralmente, ya que ambos términos están marcados socialmente.

Al pasar de las oposiciones de rubros individuales a un análisis de las consecuencias estructurales (término de V. Weinreich (1954) de las variaciones léxico-semánticas, es necesario examinar, en primer término, la conjugación de oposiciones análogas y divergentes. Volvamos a la oposición divergente *the Man*<sup>1</sup> "policía" / *the Man*<sup>2</sup> "hombre blanco", dos sememas del mismo lexema. Cada término de esta oposición entre en relaciones análogas con otras unidades léxicas. Por ejemplo, *the Man*<sup>1</sup>, se ubica con *pig*, *blue boy*, *big John*, *cop* y *fuzz*, en tanto que *the Man*<sup>2</sup> es parte del grupo de analogías que incluye *whitey* y *pale face* (en el dialecto etnosocial de los negros), así como *paddy* y *patty* (en el dialecto etnosocial de negros y chicanos). Por lo tanto, en torno al lexema *the Man* surge una cadena de analogías y divergencias que incluye elementos marcados estratificacional y situacionalmente:

*the Man*

<i>the Man</i> <sup>1</sup> = <i>pig</i> , <i>blue boy</i> , <i>big John</i> , <i>cop</i> , <i>fuzz</i>	<i>the Man</i> <sup>2</sup> = <i>whitey</i> , <i>pale face</i> , <i>paddy</i> , <i>patty</i> .
--	--

En muchos casos se hace un análisis contrastivo de campos semánticos completos a fin de poder evidenciar las relaciones sistemáticas entre los léxicos de los dialectos sociales. Por ejemplo, el análisis se puede centrar en el campo semántico de la jerga general del inglés estadounidense formado con base en el eje semántico "expresión de aprobación"; *beat*, *the cat's pajamas*, *drooly*, *gas*, *George*,

*smooth, super, way out, hot, gone, cool, hip, in, far out, freaky, groovy, out of sight.* Un análisis sincrónico de dicho cambio mostraría, entre otras cosas, la coexistencia en un cierto período de tiempo de jerga especial marcadora de la asociación del hablante con grupos de diferentes edades. El hecho de que un estadounidense contemporáneo use en su habla términos como *the cat's pajamas* (de la década de 1920) o *cool* (de la jerga de los años 50), es indicador de su asociación con grupos de cierta edad. Esos elementos marcados estratificacionalmente sirven como indicadores de la variación sociodemográfica del lenguaje. A la vez, un análisis diacrónico de ese campo semántico podría suministrar datos sobre una secuencia cronológica de "subculturas" de referencia cuya jerga penetró la jerga general de manera muy intensa durante ciertos períodos. Así, el término *cool*, de la jerga de los músicos de jazz, penetró la jerga de estudiantes y jóvenes en la década de 1950 y luego la jerga general. El término *freaky* se difundió ampliamente en la década de los años 60, junto con otros rubros léxicos tomados en préstamo de la jerga de los hippies.

Este método contrastivo no pretende tener un carácter universal. El rango de su aplicación lo determina su meta: explicitar la relación entre los elementos particulares y comunes de los microsistemas, determinar su base común y la proporción de elementos diferenciadores. Tampoco puede pretenderse universalidad para ninguno de los otros métodos de análisis sociolingüístico mencionados anteriormente. Al igual que los métodos de trabajo de campo, los métodos para el procesamiento de los datos se complementan entre sí. Pareciera que los resultados más efectivos se pueden obtener mediante una combinación racional de métodos, siempre que se le preste la debida atención a la naturaleza específica del material en estudio y a los objetivos de la investigación planteados en las respectivas hipótesis de trabajo.

#### BIBLIOGRAFIA

- Bailey, C.J. 1969. "The integration of linguistic theory: internal reconstruction and comparative method in descriptive linguistics". *Working Papers in Linguistics* 2. Honolulu: University of Hawaii.
- Bailey, R.W. y J.L. Robinson (Comp.). 1973. *Varieties in present-day English*. New York: Academic Press.
- Bell, R.T. 1976. *Sociolinguistics: goals, approaches and problems*. Londres: Penguin.
- Bikerton, D. 1973. "The structure of polylectal grammars". En R.W. Shuy (Comp.)
- Bright, William. (Comp.). 1966. *Sociolinguistics*. La Haya: Mouton.
- Brown, L. 1969. *The social psychology of variations in French Canadian speech*. Tesis doctoral. Montreal: McGill University.
- Cedergren, H.J. y D. Sankoff. 1974. "Variable rules: performance as a statistical reflection of competence". *Language*, 50.
- Fishman, J.A. (Comp.). 1971. *Advances in the sociology of language. Vol I. Basic concepts, theories and problems: alternative approaches*. La Haya: Mouton.
- Gumperz, J.J. 1966. "On the ethnology of linguistic change". En W. Bright (Comp.)
- Hockett, C.F. 1959. *A course in modern linguistics*. New York: Macmillan.
- Kay, Paul. 1978. "Variable rules, community grammar and linguistic change". En D. Sankoff (Comp.)
- Krysin, L.P. 1974. *Russian according to a mass survey*. Moscú. En Svejcer (1986).
- Labov, William. 1966. *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
- \_\_\_\_\_. 1971. "The study of language in its social context". En J.A. Fishman (Comp.) (1971).
- \_\_\_\_\_. 1972a. *Language in the inner city. Studies in the Black English vernacular*. Philadelphia: The University of Pennsylvania Press.
- \_\_\_\_\_. 1972b. *Sociolinguistic Patterns*\* Philadelphia: The University of Pennsylvania Press.
- \_\_\_\_\_. 1973. "The logic of non-standar English". En R.W. Bailey y J.L. Robinson (Comp.)
- Lambert, W.E. 1967. "A social psychology of bilingualism". *The Journal of Social Issues*. Vol. 23, No.2.
- Levine, L. y H.J. Crockett, Jr. 1967. "Speech variation in a Piedmont community: the postvocalic /r/'. *Explorations in sociolinguistics. International Journal of American Linguistics*. Vol. 33, No.4.
- Pride, J.B. y H. Holmes. 1972. *Sociolinguistics. Selected readings*. Harmondsworth: Social Press.
- Rousseau, P. y D. Sankoff. 1978. "Advances in variable rule methodology". En D. Sankoff (Comp.) (1978).
- Sankoff, David. (Comp.). 1978. *Linguistic variation. Models and methods*. New York: Academic Press.

- Sankoff, G. 1972. "Language use in multilingual societies: some alternative approaches". En J.B. Pride y J. Holmes (Comp.).
- Shuy, Roger W. (Comp.). 1973. *23rd. Annual roundtable. Sociolinguistics: current trends and problems*. Washington, D.C.: Georgetown University Press.
- Shuy, R.W., W. Wolfram y W. Riley. 1966. *Field techniques in an urban language study*. Washington, D.C.: Center for Applied Linguistics.
- Spears, Richard A. 1982. *Slang and Euphemisms*. New York: Signet.
- Švejc, A.D. 1986. *Contemporary sociolinguistics*. Philadelphia: John Benjamin Publishing Company.
- Umaña, Jeanina. 1982. "El concepto de variación y el desarrollo de la lingüística". *Káñina*. Vol. VI, Nos. 1 y 2.