



Literatura y Lingüística

ISSN: 0716-5811

[literaturalinguistica@ucsh.cl](mailto:literaturalinguistica@ucsh.cl)

Universidad Católica Silva Henríquez

Chile

Cifuentes Becerra, Edgardo; Salamanca Gutiérrez, Gastón Felipe  
Desarrollo y aplicación piloto de un instrumento para recoger datos del nivel fonético-fonológico  
Literatura y Lingüística, núm. 25, 2012, pp. 221-248  
Universidad Católica Silva Henríquez  
Santiago, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=35223198011>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica  
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

# Desarrollo y aplicación piloto de un instrumento para recoger datos del nivel fonético-fonológico\*

Edgardo Cifuentes Becerra\*\*

Gastón Felipe Salamanca Gutiérrez\*\*\*

## Resumen

En este artículo se presenta un instrumento para recoger datos fonéticos en investigaciones del área de las ciencias del habla. En primer término, se presentan las ventajas y desventajas de los instrumentos que se utilizan actualmente en la recolección de datos lingüísticos; luego se entregan los fundamentos teóricos en los cuales se basa el desarrollo del instrumento propuesto; posteriormente, se detallan sus características más prominentes, junto con las distintas etapas de diseño, implementación y evaluación ejecutadas; se concluye con la presentación de las características más importantes que debe tener el instrumento para su aplicación más efectiva, en sus etapas de construcción, ejecución y evaluación.

**Palabras clave:** fonética, instrumento para recogida de datos, metodología, ciencias del habla.

## Development and pilot application of an instrument to gather phonetic and phonological information

### Abstract

Throughout this article we present a phonetical data gathering instrument for researches in the area of speech sciences. We first introduce the advantages and disadvantages of instruments nowadays used in linguistics data collection; afterwards, we present the theoretical foundations in which the development of the proposed instrument is based. Subsequently, its most prominent characteristics are detailed, alongside with the design stages and the implementation and evaluation carried out. We conclude with the presentation of the most important characteristics the instrument must have in order to reach its more effective application in the stages of implementation and evaluation.

**Key words:** Phonetics, phonetical data collecting instrument, methodology, speech sciences

Recibido: 24-03-2011

Aceptado: 12-04-2011

\* Este trabajo contiene los principales tópicos de la tesis de Magister en Lingüística Aplicada del autor principal de este artículo.

\*\* Magister en Lingüística Aplicada y Doctorando en Lingüística. Universidad de Concepción. edgardocifuentes@gmail.com

\*\*\* Doctor en Lingüística. Universidad de Concepción. gaston.salamanca@gmail.com

## Introducción

Actualmente, para la lingüística –y en especial para las ciencias del habla– se hace necesario el acceso a un adecuado corpus lingüístico para llevar a cabo investigaciones en distintos tópicos. En fonética y fonología, los datos –necesarios sobre todo para la descripción efectiva de diversos fenómenos– deben cumplir ciertos requisitos para que su análisis sea válido y expedito. Algunos de estos son: suficiente cantidad de apariciones de un fenómeno, representatividad respecto a la variante de discurso que se pretende describir y fidelidad acústica de la grabación.

En relación con estas pretensiones, las técnicas de recogida de datos actualmente existentes presentan importantes problemas para la investigación en fonética y fonología, que redundan en cuestionamientos de la validez de los datos o dificultad para emprender ciertas investigaciones. Algunos de estos problemas son: dificultad de acceso al discurso espontáneo, exigencia de esfuerzos demasiado grandes para acceder a pocos datos, obtención de registros de naturaleza distinta a un enunciado hablado, el riesgo de influenciar al informante con una producción fónica distinta a la suya.

En este artículo se presenta un instrumento para recoger datos fonéticos en investigaciones en el área de las ciencias del habla, que soluciona de manera importante algunos de estos problemas. En efecto, este instrumento supera algunas falencias de las técnicas de recogida de datos usadas comúnmente en la investigación fonética: permite aislar el factor influencia de la escritura en la articulación, así como la influencia de la producción fónica del investigador; minimiza las consecuencias de un eventual estado de alerta del informante al saber que se está estudiando su producción lingüística, en una situación de laboratorio (es decir, se accede a un estilo cercano al habla espontánea); y, además, permite al investigador un corpus con alta presencia del asunto que busca investigar. En suma, se desarrolla un procedimiento que permite acceder a datos fónicos, optimizando la relación esfuerzo requerido / validez de los datos.

Para lograr estos objetivos, diseñamos, probamos y mejoramos un instrumento de asociación de estímulos mixtos (imagen y enunciado). Para recoger los datos, el instrumento se vale de tres tareas de reconocimiento y recuerdo de enunciados, en orden creciente de dificultad, que buscan desviar la atención del informante respecto del nivel fónico, a la vez que permiten asegurar que el enunciado obtenido sea léxico-sintácticamente preciso.

## 2. Ventajas y desventajas de los instrumentos de recolección de datos

### 2.1. Requerimientos de un corpus

En la práctica actual de la investigación lingüística, cobra gran importancia el uso de datos “reales”, un corpus de muestras de habla (o escritura, según el caso) obtenido sistemáticamente, con fines investigativos. Para obtener un corpus adecuado en fonética y fonología, se han ideado distintas estrategias; probablemente las más utilizadas provienen de la sociolingüística, debido al particular interés de este programa en la variación que experimenta la lengua.

Las estrategias de obtención deben garantizar que las muestras de habla obtenidas cumplan con ciertos requisitos. En primer lugar, un corpus adecuado estará compuesto de una cantidad de apariciones de un fenómeno suficiente para poder estudiarlo<sup>1</sup>; en segundo lugar, es importante que los elementos del corpus representen efectivamente la variedad lingüística que se pretende estudiar; por último, también es importante la calidad de audio de los datos grabados, para permitir su análisis posterior a través de los métodos que se estimen convenientes en la investigación.

### 2.2. Instrumentos utilizados en la investigación en fonética y fonología

Samarin (1967) recopila algunas técnicas de la lingüística de campo, destinadas, en principio, para estudios iniciales de lenguas; muchas de estas técnicas han sido aplicadas además para recopilar corpus para estudios más específicos. Los métodos recopilados por este autor son: elicitación de corpus previsto, elicitación analítica, audición oculta o audición selectiva, elicitación no traduccional de textos y completación de oraciones.

Otras formas de recogida de datos utilizadas con frecuencia en ciencias del habla son las aportadas por la sociolingüística laboviana: entrevistas breves y anónimas, observaciones no sistemáticas, audición

---

1 La cantidad mínima de apariciones dependerá de diversos factores, como la naturaleza de la investigación, el tamaño del grupo cuya variedad se estudia, de las hipótesis planteadas por el programa investigativo que enmarca la investigación, etc.

y visionado de medios de comunicación de masas, test de extracción familiar y de reacción subjetiva, y test de autoevaluación (Labov, 1983 [1972]).

Moreno Fernández (1990), López Morales (1994) y Silva-Corvalán (2001) reúnen una serie de instrumentos utilizados en investigaciones en lingüística y sociolingüística. En lo que respecta a las ciencias del habla, estos instrumentos se pueden dividir en dos grandes grupos: los destinados a obtener muestras de habla espontánea y los utilizados para recoger un corpus predeterminado, preparado *ad-hoc* para una investigación (Llisterri, 1991).

Los elementos del primer tipo de corpus “suelen obtenerse realizando grabaciones de uno o varios informantes hablando de manera más o menos libre durante un periodo de tiempo generalmente largo” (Llisterri, 1991: 68). En este grupo encontramos la *Observación* (Moreno Fernández, 1990: 92; Silva-Corvalán, 2001: 54) y la *Conversación grabada* (Labov, 1983 [1972]; Silva-Corvalán, 2001). Dentro de la primera, se distinguen la *Observación estructurada* y la *Observación participativa*; dentro de la segunda, la *Entrevista sociolingüística* y la *Grabación secreta*.

Los corpus del segundo tipo -preparados *ad-hoc*- “son normalmente leídos por informantes seleccionados en condiciones precisas y bien controladas” (Llisterri, 1991: 68); están destinados a obtener enunciados específicos, de acuerdo con una hipótesis investigativa en particular (Llisterri, 1991: 71). También pueden obtenerse mediante otros medios (entrevista, entrevista telefónica, pregunta breve anónima). Entre los instrumentos que se utilizan para constituir este tipo de corpus encontramos: cuestionarios, tests y lecturas grabadas (Labov, 1983 [1972]). Dentro de los dos primeros, se distinguen los utilizados para acceder a datos de valoración o percepción fonético/fonológica, y los utilizados para acceder a producción fónica.

### 2.3. Problemas de los métodos utilizados actualmente

Estos instrumentos, pese a estar desarrollados para suplir distintas necesidades, presentan importantes problemas para la investigación lingüística. Algunos de estos problemas son:

a) *No se accede al discurso espontáneo*. El discurso espontáneo, correspondiente al “habla cargada de excitación o de emoción” (Labov, 1983 [1972]: 124), es un objetivo deseable al buscar datos fónicos. Este estilo es de difícil acceso a la mayoría de los procedimientos: conversación individual o dirigida, cuestionarios, tests –tanto de percepción como de

elicitación a partir de estímulos, cuando estos alertan al informante sobre la observación del uso de su lengua— y las lecturas grabadas<sup>2</sup>.

b) *Se requiere de un esfuerzo demasiado grande para acceder a pocos datos.* Algunos instrumentos están diseñados para acceder a algo cercano al habla espontánea; tales son las técnicas de observación o las técnicas de conversación menos estructuradas y con más participantes. Sin embargo, estas tienen en común la gran distancia entre el esfuerzo que requieren (económico, de tiempo, etc.) y la cantidad de datos que aportan a la investigación; esta diferencia se da de modo que, a mayor libertad del discurso, menos se asegura la aparición de las estructuras que se busca investigar. Este problema se acentúa cuando la búsqueda se centra en elementos o construcciones con escasa distribución y aparición en la lengua<sup>3</sup>.

c) *Se obtiene un registro de distinta naturaleza respecto del discurso casual.* Para salvar la dificultad de acceso a realizaciones de sonidos en contextos poco frecuentes, se pueden presentar enunciados previamente contruidos para que los informantes los eliciten; una forma de conseguir esto es a través de la lectura de textos que contengan estas secuencias fónicas. Los instrumentos que hacen uso de la lectura en voz alta de textos asumen la situación de lectura como la parte formal de un continuo que va de la casualidad a la formalidad (Labov, 1983 [1972]); de esta manera, si no es fundamental para la investigación acceder al discurso espontáneo o si se necesita el dato de la situación formal, se consideran válidos los datos fónicos aportados por la lectura en voz alta. Sin embargo, se ha criticado el supuesto de que la lengua hablada y la lectura pertenezcan a un mismo parámetro, y más bien se consideran, en atención a los resultados de las investigaciones en psicolingüística, actividades cognitivas de distinta naturaleza (Llisterri, 1991: 87).

Existen, además, otras dificultades prácticas de este método. Una es que la atención prestada a la lectura puede variar según características

2 También se debe hacer notar que instrumentos como la conversación grabada no superan necesariamente el problema de que se participa en una situación comunicativa “artificial”, no espontánea ni funcional a algún propósito comunicativo distinto del metalingüístico; más bien suelen dejar la responsabilidad del logro de este acceso en manos de “la habilidad del investigador para motivar la creación de una atmósfera cómoda, amistosa y despreocupada” (Silva-Corvalán, 2001: 55).

3 A esta dificultad se suma el empobrecimiento en la calidad acústica de la grabación que suponen las técnicas de observación, la grabación secreta o la grabación de conversaciones grupales, lo cual no permite identificar fácilmente las variantes fonéticas (Hudson, 1981 [1980]: 158; Llisterri, 1991: 70).

del informante, tales como su grado de alfabetización o las valoraciones subjetivas frente a la lectura (Milroy, 1987, citado por Llisterri, 1991: 87). La segunda dice relación con las características del sistema de escritura del español, que es fundamentalmente fonológico; en este caso, la lectura introduce el peligro de que el informante se influencie por el alfabeto y produzca sonidos que, de otro modo, no aparecerían.

d) *Se corre el peligro de influenciar al informante con una pronunciación distinta de la suya.* Para conseguir que el informante produzca de forma exacta enunciados previamente definidos, se le puede pedir que lea, o bien dar indicaciones verbales de distinto tipo; un caso extremo –generalmente evitado por todos los instrumentos de recogida de datos– sería la elicitación por repetición de la producción del investigador; el problema mayor de una estrategia como esta es que se corre el riesgo de influenciar la articulación del informante con la del investigador, ofreciéndole un modelo de pronunciación (Llisterri, 1991: 90)<sup>4</sup>.

### 3. Fundamentos teóricos para el desarrollo del instrumento

#### 3.1. Función de los instrumentos en investigación

La metodología de investigación científica acerca del lenguaje suele contemplar, en los diseños destinados a estudiar el nivel fónico, una etapa importante: recoger las realizaciones de habla. Precisamente en esta etapa se inserta el instrumento que proponemos, para cumplir la función de “recolectar los datos pertinentes sobre las variables involucradas en la investigación” (Hernández-Sampieri, Fernández y Baptista, 1991: 241).

---

4 Pero no solamente al producir de manera directa los enunciados esperados se presenta esta dificultad, ya que incluso el habla del investigador en la situación de entrevista, puede predisponer al informante a utilizar un registro u otro (Hudson, 1981 [1980]: 165). Para Bell, son justamente las características del interlocutor (características personales, nivel de estilo de habla y actuación respecto a variables lingüísticas específicas) el factor más relevante en la variación de estilo; el mismo autor reúne evidencia de que si una variante es propia de un grupo, el hablante la utilizará al dirigirse a un integrante de ese grupo (Bell, 1984: 167).

## 3.2. Requerimientos del instrumento

### 3.2.1. Acceso al estilo espontáneo

De acuerdo con Silva-Corvalán (2001: 87), el estilo se puede definir como la variación correlacionada con “el contexto físico y humano en el que tiene lugar la comunicación”. Las variedades lingüísticas distinguidas por estos criterios se suelen llamar, respectivamente, dialectos y registros: el dialecto está definido por el hablante y refleja la estructura social; el registro está definido por la situación (Suriani, 2008: 33).

Halliday destaca, en el hecho de la variación, la competencia en su lengua materna que posee el hablante, que no se limita al conocimiento gramatical, sino práctico, funcional; en este entendido, la competencia funcional del lenguaje incluye un conocimiento sobre la variación según el tipo de situación (Halliday, 1982 [1978]: 46). Esta variación estilística se distingue de la postulada en sociolingüística clásica, porque no incluye solo “la forma” lingüística (ya sea sintáctica, morfológica o fónica), sino que funciona como un conjunto unitario con lo semántico<sup>5</sup>, pues el sistema lingüístico “tiene un alcance de potencial de significado específico en cuanto a la cultura y sensible en cuanto a la situación” (Halliday, 1982 [1978]: 49); es decir, los significados están determinados por factores sociales, entre ellos, la situación de comunicación.

En cuanto a los factores que determinan la variación de estilo, en su fundamentación de los métodos utilizados para acceder al discurso espontáneo, Labov (1983 [1972]) establece una correlación positiva entre la atención prestada al habla propia y la formalidad del discurso resultante<sup>6</sup>.

Halliday (1982 [1978]) considera necesario que, al estudiar el registro, no solamente se tengan en cuenta los cambios lingüísticos

5 Contrátese, por ejemplo, con la definición de estilo que considera que cada variedad (dialecto) es actualizada por distintos registros o estilos, que son modos de habla usados en distintas situaciones (Suriani, 2008: 33). En esta definición es evidente una separación entre el contenido de lo dicho y su forma, pues considera que la variación estilística se da en el plano del habla.

6 El principal argumento que le permite extraer esta conclusión proviene de algunos experimentos de Mahl (1972, citado en Labov, 1983 [1972]), en los que se enmascaraba la retroalimentación del sonido de la propia habla del informante con un ruido neutro a través de audífonos y se sacaba de su vista el entrevistador. La conclusión general que obtiene el autor es la hipótesis de que los estilos del discurso están colocados en una misma dimensión, relacionada con la atención prestada al discurso (Labov, 1983 [1972]: 141) y que “La manera más importante en que esta atención se ejercita es el autocontrol auditivo del propio discurso” (Labov, 1983 [1972]: 264).



que se producen, sino también la relación entre los tipos de factores situacionales y sus tipos de selección del sistema lingüístico asociados (Halliday, 1982 [1978]: 47). En este sentido, propone la pertinencia de tres variables que regulan el registro: campo, tenor y modo (Halliday, 1982 [1978]: 48).

Bell (1984), en tanto, cita algunas investigaciones cuyos resultados ponen en duda las conclusiones de Labov sobre la atención prestada al discurso como el factor determinante de la variación de estilo<sup>7</sup>. En su artículo “Language style as audience design”, el autor reanaliza los datos de Mahl en los que se basa Labov y concluye que la variable ‘presencia a la vista del entrevistador’ (que no formó parte de la hipótesis final de Labov), en algunos casos explica mejor la variación de estilo que la atención prestada al discurso (Bell, 1984: 149). Además, reúne evidencia de diversos estudios que muestran cómo un hablante acomoda su estilo al interlocutor (según su tipo de ocupación, si es conocido/desconocido, etc.)<sup>8</sup>. Así, plantea su hipótesis de la incidencia de la audiencia en el estilo: tal como la variación social del lenguaje se correlaciona con características sociales medibles del hablante (clase social, edad, etc.), la variación estilística se correlacionaría con atributos de los oyentes<sup>9</sup>.

Podríamos decir que, si bien los factores que inciden en la variación son numerosos, estos pueden causar un efecto en el discurso sólo en la medida que el hablante accede a ellos y que se ocupa de interpretarlos y actuar en consecuencia; es decir, los factores condicionan la variación si el hablante fija su atención en ellos y en su habla. Martín Butragueño (1993) se refiere a este proceso a través de la noción de ‘control’. En esta explicación, las modificaciones que se producen en la actuación lingüística, y que responden a alteraciones en factores como la atención puesta al habla o las características del contexto comunicativo, son ajustes que el hablante aplica a su habla, guiado por la información que le entregan los canales de retroalimentación acerca del cumplimiento o no de ciertos principios.

Sobre los diversos estilos que se pueden distinguir, Labov (1983 [1972]), de acuerdo con su criterio de la atención prestada al habla,

7 Por ejemplo, Dressler (1974) y Coupland (1980).

8 Por ejemplo, cita a Bickerton (1980); Douglas-Cowie’s (1978), Thelander (1982), Coupland (1981 - también 1980 y 1984).

9 Otros factores no personales influyen también en la variación de estilo; los más importantes de ellos son el tema y el contexto situacional (Bell, 1984: 178-182); además, está el factor “iniciativa” (si el hablante decide hablar o sólo responde) (Bell, 1984: 178-186).

los categoriza en estilos A, B, C y D. El Estilo A (coloquial, vernáculo, casual) corresponde al habla propia de situaciones informales; puede darse también en situaciones de relativa formalidad, siempre y cuando se logre un habla espontánea, es decir, “cargada de excitación o de emoción y donde las constricciones de la situación formal desaparecen” (Labov, 1983 [1972]: 124). A este estilo se accede, por ejemplo, dirigiendo el discurso del informante a un tema que evoque asuntos emocionalmente significativos; así, al relatar éste una circunstancia de peligro de muerte, “Es corriente que se sumerja de tal manera en su relato que reviva el momento crítico, y que aparezcan signos de tensión emocional” (Labov, 1983 [1972]: 133).

El Estilo B (cuidado) se define por la situación de relativa formalidad que corresponde a una entrevista con un desconocido; si bien es menos formal que una entrevista laboral y que un discurso público, no alcanza el nivel de naturalidad de la situación informal (Labov, 1983 [1972]: 116).

Los Estilos C y D siguen en orden creciente de formalidad y se definen ambos por la situación de lectura; el primero corresponde a la lectura de textos redactados en un estilo informal; el segundo, a la lectura de listas de palabras y pares mínimos (Labov, 1983 [1972]: 121-123).

Esta nomenclatura, aplicada algunas veces de forma más estricta que otras, es la utilizada frecuentemente en estudios de sociolingüística, junto con las estrategias de obtención de datos que lleva implícita (ver, por ejemplo, Moreno Fernández, 1990; López Morales, 1994; Silva-Corvalán, 2001; Rincón, 2007).

La dificultad de acceso a los estilos más casuales en situación de observación sistemática, podría motivar al investigador a utilizar los estilos más formales como material de estudio. Sin embargo, existen varias razones teóricas que hacen deseable el acceso a los estilos A y B, y que exponemos a continuación; estas razones justifican la elaboración de nuestro instrumento.

En primer lugar, los estilos C y D, metodológicamente definidos por la situación de lectura, han sido cuestionados como formas de acceso válido a la gramática de los hablantes. Se ha criticado el hecho de que la lengua hablada y la lectura en voz alta pertenezcan a un mismo parámetro, que –tal como afirma Labov– exista entre el discurso hablado y el leído sólo una diferencia de grado (Hudson, 1981 [1980]; Moreno Fernández, 1990).

En segundo lugar, si lo que se busca no es estudiar la forma estándar de una lengua (una variedad relativamente fija), sino la variación que se produce en ella, entonces es importante conocer no solo la realización de un estilo muy cuidado, sino también qué sucede en los estilos menos cuidados (formal, espontáneo); de lo contrario, este programa de estudio del lenguaje sería parcial respecto a sus propios objetivos. Este objetivo no es marginal a la lengua, sino relevante para una descripción lingüística más acabada, porque la variación –la sujeta al estilo, entre otras– también estaría codificada en la competencia lingüística de los hablantes (Medina, 2005: 117).

Por lo tanto, para el desarrollo completo de un programa variacionista (y cualquier otro paradigma investigativo que considere la variación de estilo, como parte del conocimiento lingüístico sistemático del hablante), se hace necesario describir los distintos estilos y sus reglas internas de funcionamiento.

El discurso espontáneo presenta otras características que lo hacen un objetivo deseable. La primera de ellas es su papel predictor del cambio lingüístico. En efecto, es en el discurso casual donde se puede descubrir el cambio en proceso; la corrección, que superpone una variedad sobre otra, no permite ver la sistematicidad del cambio. La variedad estándar muestra los cambios terminados, no deja ver el proceso (Medina, 2005: 118).

Otra razón que justifica los esfuerzos metodológicos conducentes a acceder a estilos más espontáneos es su interés intrínseco para el lingüista. Los hablantes, aun sin tener conciencia explícita de sus reglas, saben que en la situación formal su comunidad les exige aplicar algunos cambios a su estilo de habla.

### 3.2.2. Evitar influencia de la escritura

En español, la estrecha cercanía entre la escritura y la pronunciación, que se verifica mayormente en la situación de lectura, provoca que muchas veces los hablantes no vean las incongruencias entre la escritura y la fonología, y el uso de ciertos sonidos al leer ciertos grafemas sea considerado un “error de pronunciación”. Un caso conocido es la difundida creencia de que el grafema “v” representa un fonema distinto que el grafema “b”; esta afirmación causa que se enseñe la necesidad de pronunciar un sonido [v, ʋ] cuando se lee “v” y [b, β] cuando se lee “b”, y que algunas personas consideren apropiada esta distinción y la apliquen sobre todo en situación de lectura.

De este modo, la naturaleza del sistema de escritura del español, sumada a una tradición de creencias sobre el lenguaje y la escritura, puede propiciar que en la situación de lectura aparezcan sonidos que –incluso en el habla formal– no aparecerían.

### 3.2.3. Minimizar influencia del habla del investigador

Si el informante escucha el enunciado en el momento previo a elicitarlo, existe la posibilidad de que lo repita copiando alguna característica articulatoria. Así parece entenderlo, por ejemplo, Navarro Tomás, cuando advierte en un cuestionario lingüístico que, para lograr que los informantes produzcan ciertos enunciados específicos, el investigador debe “ingeniárselas” para provocar su enunciación, debiendo evitar la pregunta directa, con la finalidad de que la pronunciación del sujeto no se influya por la de la persona que lo está interrogando.

También el informante puede ser influenciado por el estilo del investigador. Por un lado, el registro utilizado por el investigador podrá predisponer al informante a utilizar un registro u otro (Hudson, 1981 [1980]: 165); por otro, existe evidencia que sugiere que, si al grupo de pertenencia del interlocutor le es propia una variante, el hablante utilizará esa variante al momento de dirigirse a él (Bell, 1984: 167). Y no solamente el estilo de habla del investigador influiría en el discurso del informante, sino que también, como hemos visto, sus características personales y la actitud de su grupo frente a variables lingüísticas en particular (Bell, 1984).

## 4. El instrumento propuesto

El instrumento que proponemos consiste en una sesión de tres tareas en que el informante interactúa con un computador, observando imágenes y oyendo enunciados de entrada; él debe realizar distintos procesos con esos enunciados para elicitarlos nuevamente. En todos los casos, los enunciados están grabados por hablantes de lengua extranjera, para evitar que la voz del investigador sea interpretada como “modelo” por los informantes. En la primera tarea, se le presentan series de tres imágenes, acompañadas cada una de un enunciado alusivo, consistente en un adjetivo y un sustantivo (en este orden); una vez terminada la serie, se presentan sólo las imágenes, y el informante debe recordar el enunciado que originalmente las acompañaba y repetirlo, invirtiendo el orden de las palabras. En la segunda tarea, se presenta al informante una imagen asociada a un enunciado; el enunciado está desordenado, y el

informante debe darle un orden natural. En la tercera tarea, las imágenes van asociadas a un enunciado ralentizado artificialmente; el informante debe reconocerlos y repetirlos.

Hemos explicado cómo la atención se puede constituir en un mecanismo facilitador del control sobre el registro. Así, pese a que los factores que influyen en la variación de estilo son de otra naturaleza, la atención prestada al habla es un mecanismo mediante el cual estos factores pueden afectar el estilo. Tal como en otras tareas cognitivas, factores del contexto pueden afectar la atención en una tarea, lo que afecta la actuación. En este entendido, un objetivo del instrumento que diseñemos será distraer la atención del informante sobre los factores que inciden en la manipulación del estilo, para evitar, en parte, que opere el sistema de control sobre el habla.

La posibilidad de considerar algunas muestras obtenidas con el instrumento como habla espontánea, está dada por varias características de la interacción configurada durante la aplicación del mismo. En primer lugar, la interacción incluye un ingrediente lúdico en las tareas, tendiente a distraer la atención del informante de los elementos situacionales que podrían gatillar los mecanismos de control lingüístico. Además, el instrumento propone tareas cognitivamente desafiantes, en lo que a agilidad mental y memoria se refiere, distanciándose así de la entrevista; otra diferencia capital con el mismo mecanismo es que, mientras en una entrevista sociolingüística el “objeto declarado es el lenguaje del hablante” (Labov, 1983 [1972]: 116), en nuestro instrumento las tareas inducen al informante a creer que el foco de la atención está puesto sobre su comprensión y su memoria. En efecto, la desviación de la atención sobre una tarea distinta de la sola elicitación de los enunciados pretende equiparar este discurso, en formalidad, al estilo B. De este modo, el instrumento constituye una mejora a los procedimientos de entrevista, pues recoge datos similares de forma mucho más dirigida y económica.

Esta forma de búsqueda del registro más natural está apoyada en ciertas características de la mente humana: percibimos y retenemos gran cantidad de información por poco tiempo, pero sólo seleccionamos una pequeña parte de esa información para procesarla conscientemente (Crespo, 2002: 188-189); esta selección de información relevante obedece a dos motivaciones: por las metas u objetivos fijados y por la capacidad de los mismos estímulos para capturar la atención (Crespo, 2002: 196-197). Así, la selección implica desatender de alguna forma la información que no está puesta en foco, ya sea por las metas o por su

propia capacidad de llamar nuestra atención. De este modo, la elección de un tema interesante para el informante (tal como se hace en la entrevista sociolingüística) permite que este focalice su atención en el tema, antes que en su actuación lingüística.

Sin embargo, como nuestro instrumento no consiste en una conversación, es poco probable lograr que el hablante fije su atención en el contenido del mensaje (y, por lo tanto, preste menos atención sobre su habla) a través de la misma estrategia; en lugar de esto, debemos adaptarla utilizando el mismo principio. Así, proponemos que es posible dirigir la atención del informante a través de varias características del instrumento: 1) se induce al hablante (con la naturaleza de las tareas) a pensar que se está desafiando su memoria o sus habilidades comprensivas, desviando hacia este ámbito su preocupación por el rendimiento; 2) se le hace ver al informante que se considera el factor tiempo, de manera que su habla sea lo menos planificada posible; 3) las tareas presentan un cierto nivel de dificultad, lo cual las asemeja al juego y al desafío; esto permite que el hablante se involucre emocionalmente en la interacción, a la vez que configura una situación comunicativa en donde el lenguaje no es el objetivo declarado; 4) se estimula el recuerdo de los enunciados a través de imágenes relacionadas con éstos, lo cual permitiría focalizar su atención en el nivel semántico antes que en el fónico.

Por otro lado, hemos afirmado que es necesario lograr del informante enunciados precisos previamente definidos, sin recurrir a la lectura. Para esto, se utilizará —entre otros recursos— la grabación de audio de un hablante extranjero, quien producirá una serie de enunciados esperados, al mismo tiempo que aparece una imagen alusiva; el informante deberá reconocer el enunciado, memorizarlo y elicitarlo después, aplicando algunas modificaciones de orden sintáctico. Surge aquí la siguiente pregunta: ¿es suficiente el estímulo para que el informante abstraiga de un enunciado fónicamente “no prototípico” (y que puede resultarle poco comprensible o “ambiguo”) la pronunciación que para él resulta natural?

Para responder esta interrogante, el instrumento supone un proceso de influencia semántica en la comprensión auditiva de la grabación, de modo que un enunciado poco comprensible, o ajeno al habla corriente (por ser dicho por un hablante de lengua extranjera), se reinterpreta con la fonología del informante, apoyado semánticamente por una imagen representativa del enunciado; es decir, se espera un reconocimiento de palabras por procesamiento descendente: de lo semántico a lo fónico (Vega y Cuetos, 1999). Nos apoyamos para esta suposición en la evidencia de

que “las palabras se reconocen por medio de la interacción entre sonido y conocimiento”, como demuestran estudios como el de Col y Jakimik (1980) o el de Garmes y Bond (1976), citados en Berko y Bernstein (1999: 115), en los que estímulos fónicos ambiguos (cercaños a las fronteras entre fonemas) se perciben influidos por el contenido semántico de la frase portadora, pues “los oyentes usan información semántica para ayudar a la decodificación del mensaje” (Berko y Bernstein, 1999: 115).

De este modo y con estos supuestos se pretende cautelar: 1) que el informante comprenda el enunciado, pese a las importantes diferencias con su propia articulación y, en general, con la usual para su variante de la lengua; y 2) que el informante no imite la grabación del hablante de lengua extranjera y que, en lugar de eso, reproduzca el enunciado a partir de su propio sistema fonológico.

## 5. Desarrollo del instrumento

### 5.1. Diseño del instrumento

El instrumento que diseñamos busca, primeramente, obtener un corpus de enunciados específicos, evitando la influencia del habla del investigador y de la escritura. En atención a este objetivo, el método del instrumento consiste en la presentación de estímulos visuales asociados a estímulos auditivos, para posteriormente solicitar al informante que ejecute una tarea con el enunciado y elicitarse el resultado.

La segunda necesidad que se debe cubrir es la de evitar la influencia del habla del investigador, en la producción del informante; para esto, se utilizarán estrategias y se tomarán algunas precauciones. En primer lugar, para evitar que el informante se retroalimente observando al investigador y aumente el control sobre su discurso, los enunciados de entrada para ejecutar las tareas serán entregados a través de una grabación. En segundo lugar, se tomará la precaución de que el investigador que dé las instrucciones del instrumento al informante le resulte cercano, en cuanto su edad, grupo sociocultural, etc. Por último, para evitar ofrecer al informante un modelo de pronunciación, se entregarán los enunciados de estímulo grabados en la voz de un hablante de español como lengua extranjera<sup>10</sup>.

---

10 Para que esta estrategia resulte efectiva, debe grabarse a un hablante extranjero con suficiente dominio del español como para que su habla resulte inteligible para el informante; su dominio, sin embargo, no debe ser tal que haya adquirido una pronunciación muy cercana al hablante nativo.



El uso de hablantes de español como lengua extranjera es una estrategia diseñada para evitar que el enunciado inicial sea visto como una articulación prestigiosa que el informante busque imitar (es, más bien, una producción extraña). Presumimos que la distancia entre la pronunciación extranjera y el sistema fonético-fonológico del informante obligará a éste a reinterpretar fónicamente el enunciado. La influencia de la imagen asociada y la realización de tareas relativamente complejas a base del enunciado (no la simple repetición) tributan en parte al logro de este mismo propósito.

La primera tarea consiste en la presentación de series de tres imágenes, cada una acompañada de un enunciado grabado, relacionado semánticamente; todos los enunciados son sintagmas nominales, compuestos de un sustantivo más un adjetivo antepuesto; se instruye al informante para recordar esos enunciados y repetirlos invirtiendo el orden de las palabras cuando se le muestren las imágenes nuevamente (sin sonido, esta vez). La segunda tarea consiste en la presentación de oraciones cuyos sintagmas se encuentran en desorden, acompañados de imágenes que ilustran el contenido semántico de los enunciados; se instruye al informante para que repita la oración en un orden más usual apenas la reconozca. La tercera tarea corresponde a la audición de enunciados ralentizados artificialmente, junto a una imagen que ilustre el contenido semántico de la oración; se instruye al informante para que repita el enunciado a velocidad normal cuando logre identificarlo.

## 5.2. Implementación del instrumento

Utilizamos el instrumento, en su aplicación piloto, para obtener un corpus de realizaciones fonéticas de los distintos contextos segmentales de /dʒ/; incluimos en el corpus las realizaciones de los diptongos con [j] o [ɨ] iniciales de palabra<sup>11</sup>, típicamente representadas por la grafía “hi”, que encontramos en palabras como “hielo”, “hiena”, “hierba”, etc.

Este caso fonológico reúne las condiciones que hacen problemático el proceso de obtención de datos. En primer lugar, su escasa aparición en el español no permite obtener suficientes datos con instrumentos que supongan la libertad del discurso del informante; además, la grafía “hi” representa un caso típico de ambigüedad de la relación entre fonología

11 La denominación de estos segmentos como semiconsonantes anteriores es solo una posibilidad entre otras interpretaciones; hemos usado este rótulo por considerarlo la interpretación más tradicional.



y escritura corriente: precisamente aquí se da la posibilidad de que la escritura incite una pronunciación en particular.

De entre los enunciados posibles para llenar los contextos (cuando la lengua nos proveía de ellos), decidimos seleccionar los más frecuentes y menos formales. Esta elección se inspira en nuestra interpretación –a la luz de Halliday (1982 [1978])– de la estrategia de Labov para acceder al discurso espontáneo, consultando a los informantes por discursos propios de su niñez. Ordenamos los enunciados según su frecuencia de uso, de manera que los menos frecuentes o más propios de situaciones formales quedaran agrupados hacia el final; esto nos permite sugerir al informante un incremento de la complejidad cognitiva de las tareas.

Los enunciados distractores se generaron y distribuyeron siguiendo dos objetivos: en primer lugar, evitar que el hablante se percate del foco de la investigación; en segundo lugar, crear aglomeraciones de enunciados que compartan el tema, con el objeto de desviar la atención del informante hacia el aspecto semántico de las tareas y evitar fijarla en el componente fónico.

Los enunciados de entrada fueron grabados por tres hablantes de chino. En el caso de los enunciados de la primera tarea, se grabaron directamente en el orden en que se presentarían a los informantes (adjetivo - sustantivo); los enunciados de la segunda tarea se grabaron en el orden normal de la oración y posteriormente su orden fue alterado artificialmente, a través del programa de edición de audio *Audacity*. Usando este mismo programa, ralentizamos en un 50% los enunciados de la tercera tarea (junto con bajar el tono).

### 5.3. La muestra

Se realizaron tres aplicaciones piloto, probando en cada una de ellas distintas estrategias y mejoras del procedimiento. La primera aplicación se efectuó con 10 informantes; para la segunda, debido a la especificidad de las estrategias que estaba probando, fueron suficientes 3 informantes; para la tercera aplicación, se contó con 9 informantes. La mayor parte de los sujetos eran estudiantes universitarios, y los restantes, profesionales jóvenes. Se cauteló que las ocupaciones y áreas de estudio de los informantes estuvieran alejadas del estudio o preocupación sistemática por el lenguaje o las lenguas.

## 6. Evaluación del instrumento

### 6.1. Eficiencia del instrumento

El primer objetivo de este instrumento es acceder a una serie de datos predeterminados sin necesidad de que el informante los lea, y asegurando en gran medida su aparición. La versión final del instrumento cumple este su objetivo, puesto que la pérdida de datos es mínima: para un total esperado de 279 apariciones de /dʒ/ en distintos contextos, se obtuvieron 240 datos, equivalente a un 86%.

Una aplicación piloto, antes de tomar las muestras definitivas, permitirá detectar los enunciados que se pierden más frecuentemente. Las fallas en estos casos se pueden atribuir a la claridad de la grabación del estímulo, la longitud de los enunciados o al uso de palabras poco frecuentes; mejorando estos enunciados de entrada para la aplicación definitiva del instrumento, se podrá incrementar el porcentaje de datos obtenidos.

Por otro lado, el procedimiento general de aplicación funciona en la forma esperada: los hablantes comprenden pronto las instrucciones y la sesión progresa sin mayores inconvenientes. Además, se cumple el objetivo de optimización del tiempo del investigador: en este caso, considerando las instrucciones, imprevistos y la ejecución de las tareas propiamente tales, el tiempo que tomó el instrumento en ser aplicado fue de aproximadamente media hora por cada informante.

### 6.2. Variación estilística en la muestra

En segundo lugar, nos interesa determinar a qué registro de habla se logra acceder con el instrumento y cómo se puede llegar a un registro cercano al habla espontánea. En atención a este último objetivo, en sucesivas aplicaciones pusimos a prueba distintas estrategias sugeridas por la teoría para acceder a la variedad estilística buscada; de este modo, exponemos en los subapartados siguientes la incidencia de los factores relacionados con estas estrategias en el tipo de corpus obtenido.

Para identificar el estilo obtenido con la aplicación del instrumento, comparamos la variación fonética de /s/ codal encontrada en la muestra con los datos de situación de entrevista sociolingüística provistos por tres investigaciones: Tassara (1991), Valdivieso y Magaña (1991) y Pérez (2007a). Estos datos muestran la gran sensibilidad al registro que experimenta este segmento en el español de Chile: según los resultados

de las investigaciones, el predominio de la realización [h] caracteriza al estilo espontáneo, mientras que en el habla cuidada la norma es la realización [s] (en un 62.2% de los casos, según Valdivieso y Magaña (1991), y un 49.2%, según Pérez (2007a)<sup>12</sup>.

Tabla 1. Distribución de /s/ codal en el discurso espontáneo

	[s]	[h]	Ø
Tassara, 1991	16.0%	69.5%	14.0%
Valdivieso y Magaña, 1991	10.2%	72.3%	17.5%
Pérez, 2007a	13.0%	61.9%	25.1%

La predominancia de la realización [s] en la primera aplicación del instrumento (un 67.9%), que asemeja estos datos a los propios del habla cuidada, plantea la necesidad de seguir indagando en otras estrategias que disminuyan el control de los informantes sobre su habla.

### 6.2.1. Factor tipo de tarea

Nos interesa saber si las tres tareas propuestas al informante arrojan datos equivalentes. Al observar la variación de /s/ codal notamos, en la primera muestra, que el porcentaje de [s] en la tarea C es muy superior al de las tareas A y B. Sin embargo, esta distribución no se repitió en la aplicación del segundo prototipo, por lo que no continuamos con la comparación en el tercer prototipo.

Tabla 2. Realizaciones de /s/ codal en las tareas A, B y C, muestras I y II

		[s]	[h]	Ø
Muestra I	Tarea A	55,5%	29,6%	14,8%
	Tarea B	60%	34,2%	5,8%
	Tarea C	76,7%	20,2%	3%
Muestra II	Tarea A	16,7%	66,7%	16,7%
	Tarea B	61,2%	34,8%	2,2%
	Tarea C	61,7%	30,5%	1,8%

El alto porcentaje de [h] en la tarea A de la muestra II, aunque pudiera sugerir características distintas en esta tarea, no es relevante por dos motivos: 1) no es consistente con la aplicación anterior a 10 sujetos,

<sup>12</sup> En la aplicación del primer prototipo del instrumento se comparó, además, la distribución de /d/ intervocálica con los datos de Pérez (2007b), puesto que este trabajo concluye que también /d/ se distribuye regularmente según el registro. Esta comparación sugirió una similitud entre el discurso espontáneo y la muestra de habla obtenida por el instrumento propuesto. Dado que la comparación de la distribución de /s/ codal en esa misma muestra sugería lo contrario, en las sucesivas aplicaciones se midió el estilo utilizando sólo este último parámetro.

y 2) representa un conjunto muy pequeño de realizaciones de /s/ codal (9 casos en total, considerando los 3 informantes), en el cual los porcentajes son fácilmente modificables. De este modo, podemos concluir que, en lo que respecta al registro de habla que propician, las tareas A, B y C son equivalentes.

### 6.2.2. Factor repetición

Ante la posibilidad de que la dificultad de las tareas pudiera estar incidiendo en la aparición de [s], restando dinamismo a la elocución, pusimos a prueba una variación en el guion de la toma de muestra: solicitamos al informante repitiera dos veces cada enunciado. Esperábamos que la segunda repetición fuera más segura, más fluida y, por tanto, más espontánea; sin embargo, las realizaciones de /s/ no mostraron mayor variación entre una repetición y otra.

Tabla 3. Realizaciones de /s/ codal en la primera y segunda repetición, muestras II y III

		[s]	[h]	Ø
Muestra II	Repetición 1	60,6%	32,7%	3,8%
	Repetición 2	58,5%	35,4%	3,2%
Muestra III	Repetición 1	41,2%	54,7%	4,1%
	Repetición 2	39,6%	57%	3,4%

Estos datos muestran que la segunda repetición de los enunciados es similar a la primera en su estilo, por lo que la estrategia demuestra no ser efectiva para el objetivo que fue concebida originalmente.

### 6.2.3. Factor velocidad de habla

En la primera versión del instrumento, el factor tiempo no estaba siendo controlado; solamente se comunicó a los informantes que se estaba midiendo el tiempo que tardaban en responder. El objetivo de esta instrucción era presionar a los sujetos para cumplir la tarea en menos tiempo, sin detenerse a controlar su habla. En la segunda versión se buscó reforzar esta instrucción, indicando al informante que debía cumplir las tareas en el menor tiempo posible. Sin embargo, ambos métodos para propiciar el habla más fluida sólo produjeron cambios menores, puesto que las muestras I y II se asemejan en la distribución de /s/ codal al estilo formal descrito en Valdivieso y Magaña (1991).

Tabla 4. Realizaciones de /s/ codal en muestras 1 y 2

	[s]	[h]	Ø
Lectura formal (Valdivieso y Magaña, 1991)	62.2%	36.4%	1.4%
Muestra 1	67.9%	27.3%	4.9%
Muestra 2	61%	35,5%	3,5%

La tercera versión del instrumento introdujo un mecanismo con el fin de limitar el tiempo disponible, para elicitación de los enunciados y, por tanto, controlar la velocidad de habla: se asignó un tiempo determinado (4 o 5 segundos, dependiendo de la longitud del enunciado) para que el informante elicitara 2 repeticiones de cada enunciado. Al sujeto se le informaba previamente de esta condición, y, además, la propia interfaz del instrumento iniciaba automáticamente el conteo del tiempo una vez finalizada la exposición al estímulo; el tiempo destinado para repetir los enunciados aparece representado en pantalla, a través de una animación que muestra el transcurso de los segundos; al terminar el plazo, la pantalla cambia automáticamente. Esta estrategia de control de la velocidad permitió elevar considerablemente la presencia de la realización [h] respecto a los prototipos anteriores.

Tabla 5. Realizaciones de /s/ codal en muestras I, II y III

	[s]	[h]	Ø
Muestra I	67,9%	27,3%	4,9%
Muestra II	61%	35,5%	3,5%
Muestra III	40,4%	55,8%	3,8%

Esta distribución acerca el estilo de la muestra a una variedad más espontánea. Aunque los porcentajes de [h] no son tan altos como en la situación de entrevista sociolingüística, esta muestra comparte con la entrevista la predominancia de la realización aspirada por sobre [s].

Tabla 6. Distribución de /s/ codal en muestra III, comparada con discurso espontáneo

	[s]	[h]	Ø
Tassara, 1991	16,0%	69,5%	14,0%
Valdivieso y Magaña, 1991	10,2%	72,3%	17,5%
Pérez, 2007a	13,0%	61,9%	25,1%
Muestra III	40,4%	55,8%	3,8%

Para la implementación de esta estrategia, se decidió mantener la instrucción de repetir dos veces cada enunciado. Aunque este método ya había probado no ser efectivo, por sí solo, para incrementar la velocidad de habla, se utilizó en la tercera versión, en conjunto con la limitación del tiempo de elicitación, para presionar al informante a producir un habla más espontánea. Si no se usara una repetición doble, el tiempo dispuesto para que el informante produzca el enunciado debiera ser

menor (cercano a la mitad), para evitar que quedara suficientemente holgado como para controlar su habla; sin embargo, al reducir el tiempo a la mitad, se corre el riesgo de que no alcance a terminar el enunciado; esto ocurre muchas veces con la segunda repetición en esta versión del instrumento, pero no es una pérdida relevante, pues se dispone aún de la primera repetición. Si se elicitaran los enunciados sin repetición doble, se corre el riesgo de perder más información o, por otro lado, de no incrementar lo suficiente la velocidad de habla.

#### 6.2.4. Factor ruido de enmascaramiento

La tercera muestra fue dividida en dos grupos: al grupo A se le presentó el conteo del tiempo a través de un segundero visual; al grupo B se le expuso además a un conteo de tiempo auditivo, consistente en campanadas en frecuencias altas, con un tono puro (también alto) de fondo; con esta última modificación, pretendíamos aplicar el principio de que a menor retroalimentación auditiva disminuye el control (Labov, 1983 [1972]). En la siguiente tabla se muestran los datos de /s/ codal, comparados entre estos dos grupos:

Tabla 7. Realizaciones de /s/ codal en grupos A y B, muestra III

	[s]	[h]	Ø
Grupo A	35,5%	60,1%	4,4%
Grupo B	44,1%	52,6%	3,3%

Contrario a las expectativas, disminuir la retroalimentación auditiva del sujeto mientras estaba elicitando los enunciados no permitió el acceso a un estilo menos cuidado.

#### 6.2.5. Factor idiolectal

En ninguna de las tres muestras los resultados se reparten equitativamente entre todos los informantes. En algunos informantes, el porcentaje de aparición de [h] es bastante más alto que el promedio:

Tabla 8. Realizaciones de /s/ codal en informantes con promedios más altos de [h]

	[s]	[h]	Ø
Promedio muestra I	67.9%	27.3%	4.9%
Informante 5-I	52,6%	38,6%	7%
Informante 6-I	20,7%	65,5%	13,8%
Promedio muestra II	61%	35,5%	3,5%
Informante 2-II	34,2%	55,3%	1,8%
Promedio muestra III	40,4%	55,8%	3,8%
Informante A3-III	10,8%	77,1%	12%
Informante B1-III	9,6%	77,7%	12,8%

En varios de estos informantes, la distribución se corresponde aproximadamente con la observable en el discurso espontáneo (6-I, A3-III, B1-III). Dado que las circunstancias, tareas e instrucciones son las mismas para todos los informantes, debemos atribuir estas diferencias a la propia variación idiolectal entre estos hablantes, particularmente en lo que respecta a sus aparentemente distintos conceptos de situación formal.

### 6.3. Selección de corpus

Como ya constatamos, existe una dispersión notoria entre los hablantes: en algunos casos, las realizaciones de /s/ codal se corresponden con un estilo espontáneo y en otros se evidencia un registro muy cuidado. En la siguiente tabla se resumen los datos de todos los informantes de la tercera versión del instrumento.

Tabla 9. Realizaciones de /s/ codal, muestra III, todos los informantes

	[s]	[h]	Ø
Grupo A, informante 1	39,8%	56,3%	3,9%
Grupo A, informante 2	54,3%	45,7%	0%
Grupo A, informante 3	10,8%	77,1%	1,2%
Grupo A, informante 4	34,3%	62,9%	2,9%
Grupo B, informante 1	9,6%	77,7%	12,8%
Grupo B, informante 2	61,1%	36,1%	2,8%
Grupo B, informante 3	61,6%	36,6%	1,8%
Grupo B, informante 4	40,5%	59,5%	0%
Grupo B, informante 5	42,2%	57,8%	0%

En esta muestra el porcentaje de realizaciones aspiradas de /s/ en coda silábica va desde un 36,1% (equivalente a los datos de lectura formal) hasta un 77,7% (equivalente a los datos de entrevista sociolingüística). Una investigación que pretenda describir el comportamiento de un segmento o grupo fónico en un estilo espontáneo no podría utilizar todos los datos obtenidos con este instrumento, puesto que no todos representan ese registro. Para homogeneizar la muestra se hace necesario seleccionar el corpus más representativo del estilo buscado. Hemos visto que la variación de los datos no se explica por las repeticiones de los enunciados ni por el tipo de tarea: las mayores diferencias son las que se verifican de un informante a otro. De este modo, la selección del corpus pasa por la discriminación de los informantes, para utilizar finalmente sólo aquellas muestras más aptas.

Esta selección no significa que el investigador manipule deliberadamente la conformación del corpus, escogiendo las realizaciones de

la variable estudiada que prefiera; al contrario, la separación debiera efectuarse sin observar la variable en estudio, en base a una comparación de parámetros externos a ésta, entre los datos recogidos con el instrumento y datos de otras investigaciones. En este caso, por ejemplo, el objeto de estudio es la variación de /dʒ/, pero el parámetro de comparación con otras investigaciones, para seleccionar los datos más representativos del estilo buscado, es la distribución de las realizaciones de /s/ codal.

El procedimiento de seleccionar el corpus de entre un conjunto de datos obtenidos mediante estrategias que propician el discurso espontáneo, dejando fuera lo que no se corresponda con el estilo buscado, es recomendado también por Labov (1983 [1972]: 136-137). Este autor considera que las estrategias para estimular el discurso espontáneo son falibles, y su efectividad se comprueba sólo analizando los datos obtenidos.

Si discriminamos los datos obtenidos con el instrumento, podemos observar algunos sujetos cuyos enunciados distribuyen las variantes de manera equivalente a las mediciones de la situación de entrevista informal (Tassara, 1991 y Valdivieso y Magaña, 1991): informantes A-3 y B-1; se puede afirmar que en estos casos se ha accedido a un estilo espontáneo. En el otro extremo, existen casos en que la distribución de las variantes es semejante a la lectura formal (Valdivieso y Magaña, 1991; Pérez, 2007a): informantes A-2, B-2 y B-3. El resto de los casos son equiparables al estilo descrito como de lectura en situación familiar por Valdivieso y Magaña (1991); es decir, representan un punto medio entre estos dos extremos.

Es evidente que los datos de discurso espontáneo de los dos informantes son muy valiosos para una investigación, y el obtenerlos constituye un logro del instrumento. Asimismo, también parece evidente que los datos equivalentes al estilo formal no pueden ser considerados parte del corpus final y deben ser eliminados. Respecto de los datos equivalentes a la situación de lectura en tono familiar, puede considerarse apropiado mantenerlos en el corpus, ya que no afectan de manera importante los promedios y tienen en común con los datos equivalentes a la situación de entrevista la predominancia de la realización [h]. En la siguiente tabla se resume la distribución de las variantes de /s/ codal al eliminar los datos de los informantes que entregaron cifras similares a la situación de lectura formal, y se compara con los promedios de la muestra III completa.



Tabla 10. Realizaciones de /s/ codal en muestra 3 y corpus seleccionado

	[s]	[h]	Ø
Muestra III (completa)	40,4%	55,8%	3,8%
Corpus seleccionado muestra III	30,4%	64,7%	4,9%

Si bien este proceso de selección optimiza el corpus y está avalado metodológicamente, es necesario aún definir criterios invariables para seleccionar los datos. En este caso, parece razonable plantear una regla como: ante el predominio de [s] sobre [h], se deben eliminar los datos proporcionados por el informante. El criterio de la distribución de las realizaciones de /s/ codal es útil en este caso y probablemente en buena parte de la investigación que tenga como objeto el español de Chile; además, este es un segmento y una posición de gran frecuencia en el español, por lo que no existe dificultad para calcular su aparición en cualquier conjunto de enunciados. Para definir un procedimiento estándar de selección de corpus, que aplique a cualquier lengua, parece necesario estudiar otras aplicaciones del instrumento y definir si existe algún grado de sistematicidad en el porcentaje de los datos considerados no útiles, lo cual escapa a los alcances de este trabajo.

## 7. Conclusiones

Esta investigación se propuso desarrollar un instrumento de recolección de datos fónicos segmentales que superara algunas falencias de las técnicas de recogida de datos usadas comúnmente en la investigación fonética. Para alcanzar estos objetivos, en este trabajo aplicamos algunas estrategias fundamentadas en la teoría de la variación de registro. En primer lugar, para evitar la influencia de la escritura al momento de la elicitación, diseñamos algunas tareas con cierto nivel de complejidad consistentes en recordar, ordenar y comprender enunciados. Asimismo, para evitar la influencia del habla del investigador en el informante, utilizamos enunciados de estímulo grabados en la voz de hablantes de lengua extranjera; además, los enunciados, como parte de las tareas, estaban desordenados o modificados (ralentizados) para así lograr una menor atención en la forma fónica y mayor en el contenido semántico; para reforzar esta estrategia, presentamos imágenes alusivas a los enunciados junto con las grabaciones de estímulo.

Para buscar la disminución del estado de alerta del informante ante una situación de observación de su habla, y así acceder a un discurso espontáneo, se probaron varias estrategias. En primer lugar, con la estructuración de las tareas de elicitación como desafíos de memoria

y reconocimiento de enunciados dificultosos por distintos factores, se buscaba configurar una situación real de uso del lenguaje: el juego, la competencia. Además, siguiendo las conclusiones de Bell (1984), buscamos evitar el autocontrol del informante sobre su habla, eliminando una de las fuentes de retroalimentación que permiten ejercer mayor control sobre ésta: la presencia ante la vista de un interlocutor. Como se vio en los resultados de la primera versión del instrumento, estas estrategias por sí solas no logran su objetivo; sólo al establecer una mayor presión sobre el tiempo que los hablantes tienen para recordar los enunciados se consigue disminuir la autocorrección.

Por último, el total de los datos útiles a la investigación recogidos mediante el instrumento, se obtiene al eliminar las muestras de los informantes que, notoriamente, produjeron un habla cuidada siempre (un 33,3% en el tercer prototipo del instrumento), considerando además que del 66,7% restante, en promedio, se perdió un 14% de los datos esperados. Esta última pérdida de datos, sin embargo, puede aminorarse en la aplicación definitiva, reelaborando los enunciados más problemáticos para los informantes. El porcentaje de pérdida de datos se compensa con la economía de tiempo que ofrece el instrumento, que permite obtener una muestra de los informantes (con apariciones de gran parte de los contextos buscados) en poco tiempo de grabación (20 minutos a media hora).

A base de estos resultados, el procedimiento más efectivo del instrumento puede describirse como sigue:

- 1) Construcción: (a) seleccionar contextos a utilizar; (b) generar enunciados que contengan los contextos buscados, además de distractores (procurar agrupar los enunciados en series relacionadas semánticamente); (c) elaborar imágenes para estos enunciados; (d) grabar los enunciados de entrada en la voz de hablantes con acento claramente no hispanico (recomendamos grabar en el orden esperado); (e) alterar los enunciados para transformarlos en los estímulos que finalmente oirán los informantes; (f) organizar las imágenes y enunciados para exponerlos a los informantes en un orden determinado y asignando un tiempo limitado para que el informante elicitte cada uno (esta exposición de estímulos puede programarse en el programa *PowerPoint*, de *Microsoft*).
- 2) Aplicación: (a) instruir al informante sobre la naturaleza de las tareas en general (observar imágenes, memorizar y repetir con variaciones); (b) explicar la primera tarea, consistente en recordar un enunciado

asociado a una imagen, solicitando repetirlo dos veces; (c) realizar una serie junto con el informante, como ejemplo; (d) aplicar la tarea, monitoreando de vez en cuando el desarrollo adecuado de la misma (sin que el informante se percate); explicar la segunda tarea, consistente en organizar sintácticamente enunciados, solicitando repetirlos dos veces; (e) aplicar la tarea, monitoreando de vez en cuando; (f) explicar la tercera tarea, consistente en reconocer enunciados ralentizados, solicitando repetirlos dos veces; (g) aplicar la tarea, monitoreando de vez en cuando.

- 3) Selección de corpus: (a) comparar las muestras obtenidas con datos recogidos por instrumentos para acceder al habla espontánea (entrevista, por ejemplo), a base de parámetros externos a la variable estudiada; (b) eliminar del corpus aquellos informantes que se diferencien de forma importante de los datos de habla espontánea.

Tal como aparece descrito en este trabajo, el instrumento presenta variadas posibilidades para aplicarse a investigaciones sobre segmentos fónicos:

- a) Permite investigar unidades y combinaciones poco frecuentes en las lenguas (a las cuales es difícil acceder en la situación de entrevista informal), sin necesidad de utilizar la situación de lectura. Más aún, posibilita que un porcentaje importante de los datos resultantes puedan considerarse muestras de habla espontánea.
- b) Prescindir de la necesidad de la situación de lectura, para tomar datos precisos permite estudiar segmentos cuyo estatus fonológico sea dudoso, y en los que la escritura pueda predisponer al informante a una interpretación.
- c) El instrumento ofrece una alternativa para acceder a datos específicos en informantes con dificultades para leer (dificultades visuales, poca práctica de la situación, personas no alfabetizadas, etc.).

## Bibliografía

- Bell, A. (1984). "Language Style as Audience Design", en *Language in Society*, Vol. 13, No. 2, 145-204.
- Berko, J. y N. Bernstein (1999). *Psicolingüística*. Madrid: McGraw-Hill.
- Crespo, A. (2002). *Cognición humana. Mente, ordenadores y neuronas*. Madrid: Editorial Centro de Estudios Ramón Areces.

- Halliday, M. (1982 [1978]). *El lenguaje como semiótica social. La interpretación social del lenguaje y del significado*. Traducción de Jorge Ferreiro Santana. México, D.F.: Fondo de Cultura Económica.
- Hernández, Fernández y Baptista. (1991). *Metodología de la investigación*. México: McGraw-Hill.
- Hudson, R. (1981 [1980]). *La sociolingüística*. Traducción de Xabier Falcón. Barcelona: Anagrama.
- Labov, W. (1983 [1972]). *Modelos sociolingüísticos*. Traducción de José Miguel Marinas Herreras. Madrid: Cátedra.
- Llisterri, J. (1991). *Introducción a la fonética: el método experimental*. Barcelona: Anthropos.
- López Morales, H. (1994). *Métodos de investigación lingüística*. Salamanca: Ediciones Colegio de España.
- Martín Butragueño, P. (1993). "El control individual de las reglas sociolingüísticas", en *Revista Española de Lingüística*, Vol. 23, No 2, 333-357.
- Medina, F. (2005). "Problemas metodológicos de la sociolingüística histórica", en *Forma y función*, 18, 115-137.
- Moreno Fernández, F. (1990). *Metodología sociolingüística*. Madrid: Gredos.
- Pérez, H. (2007a). "Estudio de la variación estilística del fonema /s/ en posición implosiva en el habla de los noticieros de la televisión chilena", en *RLA. Revista de Lingüística Teórica y Aplicada*, 455, 101-115.
- \_\_\_\_\_ (2007b). "Estudio de la variación estilística de la serie /b-d-g/ en posición intervocálica en el habla de los noticieros de la televisión chilena", en *Estudios de Fonética Experimental*, XVI, 227-259.
- Rincón, L. (2007). "Metodología para un estudio dialectal urbano: el caso del español de Bucaramanga", en *Forma y función*, 20, 193-224.
- Samarin, W. (1967). *Field linguistics. A guide to linguistic field work*. New York: Holt, Rinehart and Winston.
- Silva-Corvalán, C. (2001). *Sociolingüística y pragmática del español*. Washington D.C.: Georgetown University Press.

- Suriani, B. (2008). "El tratamiento de la variación lingüística en intercambios sociales", en *Fundamentos en humanidades*, 17, 27-41.
- Tassara, G. (1991). "Variacion lingüística: /s/ implosiva en el habla semiformal culta de Valparaíso", en *Signos*, Vol. XXIV, No. 29, p. 131-141.
- Valdivieso, H. & J. Magaña (1991). "Variación fonética de /s/ en el habla espontánea", en *Revista de Lingüística Teórica y Aplicada*, 29, p. 97-115.
- Vega, M. & F. Cuetos (coords.). (1999). *Psicolingüística del Español*. Madrid: Trotta.