

DiapixCL: Adaptación al español de una herramienta de elicitación de habla espontánea



Mauricio Figueroa y Gastón Salamanca

maufigueroa@udec.cl | gaston.salamanca@gmail.com

Departamento de Español
Facultad de Humanidades y Artes
Universidad de Concepción

Trabajo desarrollado en el marco del proyecto Fondecyt 1161466, titulado “Realizaciones fónicas segmentales en habla subestándar: una contribución a la construcción de un inventario inclusivo y actualizado de fonemas y alófonos del español de Chile”.

Elicitación de habla espontánea

Introducción

Algunos requisitos para la pertinencia de un instrumento o tarea de elicitación.

(Llisterri, 1991; Cifuentes y Salamanca, 2012)

Tipos de instrumentos de elicitación de habla: ¿habla espontánea o corpus ad-hoc?

(Moreno Fernández, 1990; Llisterri, 1991; López Morales, 1994; Silva-Corvalán, 2001)

Diapix y DiapixUK

Introducción

Diapix: tarea interactiva y colaborativa de buscar las diferencias en parejas; exclusivamente verbal.

(van Engen, Baese-Berk, Choi y Bradlow, 2010)

Incluye una etapa de familiarización y la tarea propiamente tal, hasta lograr grabaciones de ± 20 minutos.

DiapixUK: adaptación de Diapix para su uso con hablantes de inglés británico.

(Baker y Hazan, 2011)

DiapixUK

Introducción

El objetivo primario de la adaptación fue facilitar la elicitación de /p, b, s, ʃ/ del inglés en habla espontánea.
(Baker y Hazan, 2011)

Principales modificaciones realizadas a Diapix:

- (a) aumento del número de pares de imágenes a 13;
- (b) inclusión de imágenes de cuatro “dominios”;
- (c) homologación del nivel de dificultad de las láminas;
- (d) transformación a formato de imagen con capas.

DiapixUK

Introducción



Ejemplo de una pareja de láminas de DiapixUK: láminas “1” del dominio “playa”. En el panel izquierdo la forma “A” y en el derecho la forma “B”.

DiapixCL

Adaptación al español

Objetivo primario: adaptar DiapixUK para su uso en el español.

Procedimiento: revisión de las láminas, elaboración de listado de textos por traducir y elementos culturales por adaptar, y envío a ilustradores para implementación de cambios (en total, 197).

Principales adaptaciones:

- (a) traducción libre de textos del inglés al español;
- (b) adaptación de elementos culturales.

DiapixCL

Adaptación al español



Ejemplo de la adaptación de una lámina de DiapixUK a DiapixCL (dominio “calle”, láminas 2, forma “B”).

DiapixCL

Estudio piloto de producción

Participantes: 20 hablantes nativos del castellano chileno (11 hombres y 9 mujeres; edad promedio: 20 años).

Equipamiento y seteos: registro mediante micrófonos de cabeza Sennheiser EW-152-G3, enviado a grabadoras Tascam DR40 (44.100 Hz, 24 bits); grabación en formato estéreo (un canal por miembro de la pareja).

Nota: En la etapa de toma de muestras, participó como coordinadora y entrevistadora la estudiante Daniela García, de Pedagogía en Español y tésista del primer autor.

DiapixCL

Estudio piloto: procedimientos

Se solicitó a los participantes completar la actividad mediante interacciones verbales (los colaboradores estaban de espaldas) y de manera equitativa.

Se realizó primero una etapa de entrenamiento con las láminas del dominio “parque”.

Luego se realizó la elicitación con las láminas experimentales, hasta completar 10 minutos de grabación.

Durante la ejecución de la tarea, no hubo intervención de los investigadores.

Nota: En la etapa de toma de muestras, participó como coordinadora y entrevistadora la estudiante Daniela García, de Pedagogía en Español y tesista del primer autor.

DiapixCL

Estudio piloto: procesamiento de señales

Las señales de audio de cada uno de los 20 participantes fueron transliteradas en ortografía corriente en TextGrids del programa *Praat*.

(Boersma, 2001)

Luego, se procedió a segmentar los niveles de palabra, silábicos y segmentales de manera automatizada mediante el paquete de rutinas del alineador forzado *EasyAlign*.

(Goldman, 2011)

Nota: La transliteración de las muestras fue realizada por la estudiante Daniela García, de Pedagogía en Español y tesista del primer autor.

¿Qué tan buenos son nuestros materiales?

DiapixCL

Estudio piloto: cantidad de material obtenido

Luego de 10 minutos de grabación, cada participante aportó, en promedio, 733 palabras ($\sigma = 194.04$).

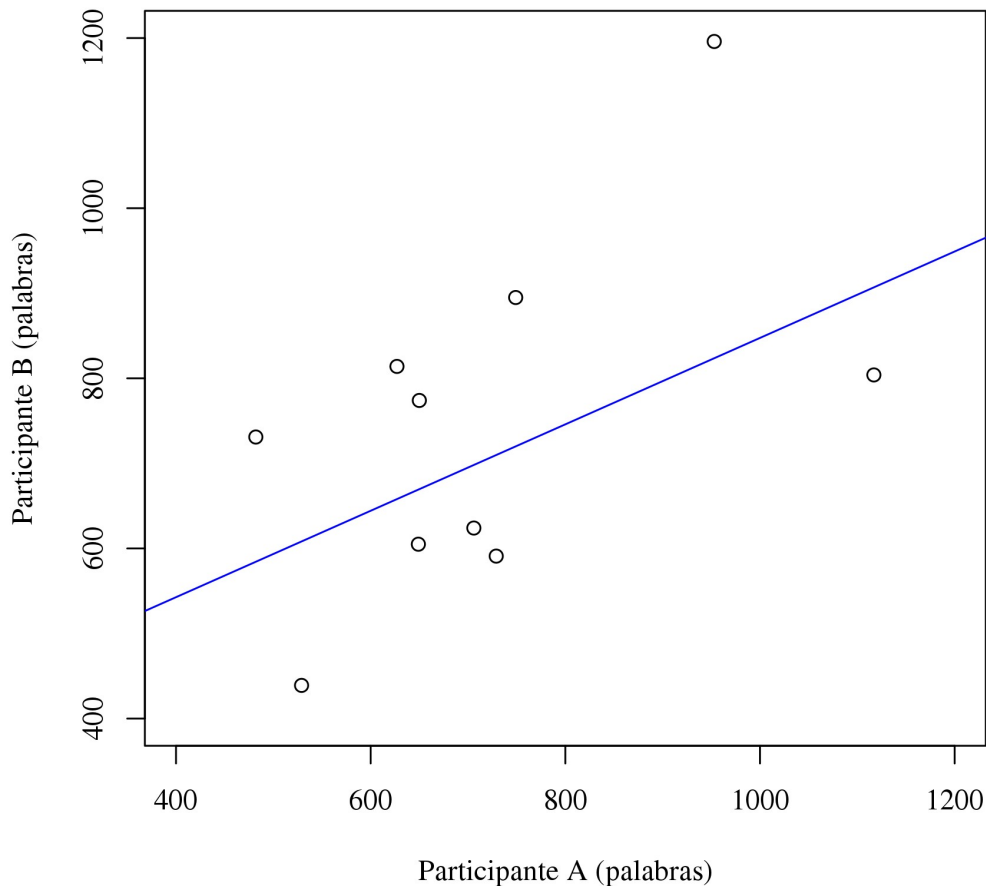
La duración promedio de habla por participante, sin contar silencios y pausas vacías, fue de 3.23 minutos ($\sigma = 0.82$).

La comparación con precedentes de investigación sugiere que DiapixCL provee suficientes materiales lingüísticos.

(Baker y Hazan, 2011)

DiapixCL

Estudio piloto: balance de las contribuciones



Se observa una tendencia a contribuciones relativamente balanceadas. Sin embargo:

Coefficiente de correlación: $r(8) = .55$, $p = .096$.

Porción de la varianza de participantes "B" explicada por la de los participantes "A": $R^2 = .22$, $F(1, 8) = 3.55$, $p = .097$.

Regresión lineal simple: $b = .51$, $t(8) = 1.884$, $p = .0963$.

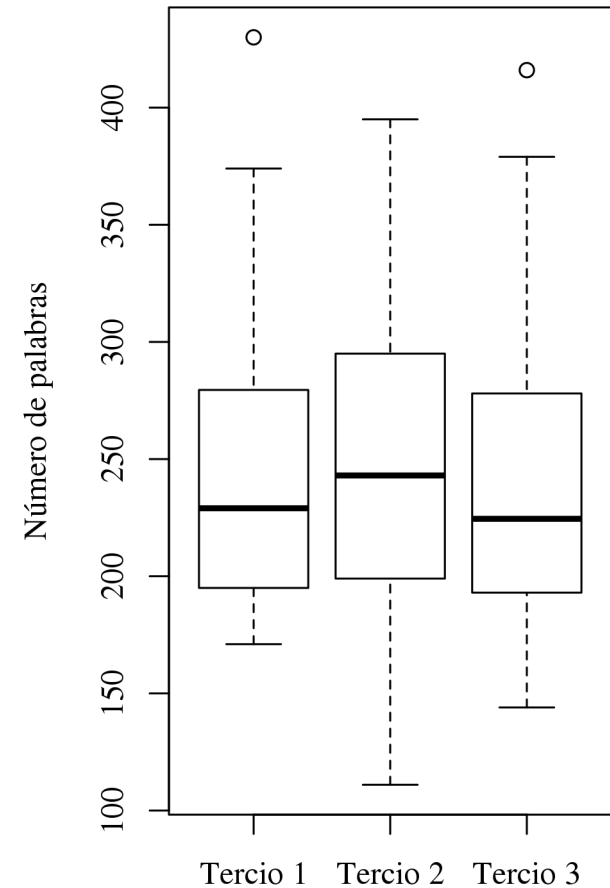
DiapixCL

Estudio piloto: efectos de cansancio y entrenamiento

La contribución de cada participante se dividió en tres porciones de igual duración, y se calculó la cantidad de palabras de cada sección.

Un efecto de aprendizaje o de entrenamiento debería conducir a diferencias entre los “tercios”.

No se encontró un efecto principal de la variable “tercios” en el número de palabras al aplicar una prueba ANOVA simple: $F(2, 57) = .063, p = .939$.



DiapixCL

Estudio piloto: diferencias por dominio

Tanto DiapixUK como DiapixCL se construyeron bajo la asunción de que las láminas presentan un grado de dificultad similar.

El dominio “playa” generó 7.044 palabras, el dominio “calle” 5.034 y el dominio “granja” solamente 2.586.

Hipótesis: puntos focales; experiencialidad y compromiso afectivo.

DiapixCL

Conclusiones

La cantidad de materiales que provee DiapixCL es adecuada.

La tarea elicitó muestras de habla espontánea.

El corpus recogido presenta una amplia gama de estructuras morfológicas, sintácticas, entonativas, etc.; en consecuencia, DiapixCL puede ser utilizado en investigaciones con variados focos de interés lingüístico.

DiapixCL

Conclusiones

Una de sus ventajas es que se plantea como un desafío lúdico a los participantes y requiere muy poca lectura.

Las láminas modificadas para construir DiapixCL pueden ser fácilmente adaptadas para adecuarse a las realidades lingüísticas de cualquier dialecto del español.

La metodología que subyace a las pruebas Diapix___ puede considerarse en la construcción de tareas de elicitación en lenguas poco familiarizadas con la escritura.

Referencias bibliográficas

- Baker, R, Hazan, V. (2011). DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods*, 43(3), pp. 761-770.
- Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5(9/10), pp. 341-345.
- Cifuentes, E. y Salamanca, G. (2012). Desarrollo y aplicación piloto de un instrumento para recoger datos del nivel fonético-fonológico. *Literatura y Lingüística*, 25, 221-248.
- Goldman, J.P. (2011) EasyAlign: an automatic phonetic alignment tool under Praat. En *Proceedings of InterSpeech*, Septiembre de 2011, Firenze, Italy
- 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.
- Moreno Fernández, F. (1990). *Metodología Sociolingüística*. Madrid: Gredos.
- Silva-Corvalán, C. (2001). *Sociolingüística y pragmática del español*. Washington D.C.: Georgetown University Press.
- Van Engen, K. J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., y Bradlow, A. R. (2010). *The Wildcat Corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles*. *Language and speech*, 53(4), 510-540.