

HACIA LA CONSTRUCCIÓN DE UN INSTRUMENTO PARA EVALUAR LA FAMILIARIDAD DE PACIENTES CRÓNICOS CON UNIDADES LÉXICAS RELEVANTES PARA EL AUTOMANEJO DE SU CONDICIÓN DE SALUD

TOWARDS THE CONSTRUCTION OF AN INSTRUMENT TO ASSESS THE FAMILIARITY OF CHRONIC PATIENTS WITH RELEVANT LEXICAL UNITS FOR THE SELF-MANAGEMENT OF THEIR HEALTH CONDITION.

María Natalia Castillo Fadic
Pontificia Universidad Católica de Chile
mcastilf@uc.cl

Josué David Pino Castillo
Universidad de Chile
josue.pino@uchile.cl

Resumen

Este artículo pretende contribuir al desarrollo de un instrumento para evaluar la familiaridad de personas con unidades léxicas. Para ello, se centra específicamente en determinar cuáles son los vocablos relevantes para que personas con condiciones crónicas de salud comprendan las explicaciones e indicaciones relacionadas con su autocuidado por parte del personal de salud. Para determinar cuáles son los vocablos relevantes, propone la aplicación de tests de disponibilidad léxica a expertos en torno a centros de interés específicos. Del procesamiento y análisis estadístico de los datos, se obtuvo un listado de 20 vocablos relevantes para una condición específica, la diabetes, que fueron descritos estadística, lexicológica, fonológica y fonéticamente. De acuerdo a los resultados obtenidos, se elaboró un test con escala de Likert cuya aplicación se recomienda al personal de atención primaria.

Palabras clave: familiaridad léxica, léxico disponible, literacidad en salud, diabetes.

Abstract

The purpose of this article is to contribute to the development of an instrument to assess people's familiarity with lexical tokens. To this effect, this article is centered specifically in determining which ones are the relevant terms/words for people with chronic conditions to understand the explanations and indications related to their self-care given by the healthcare professional. In order to determine which ones are the relevant terms/words, the implementation of lexical availability to experts in the context of specific interest centers is proposed. From data statistic processing and analysis, a list of 20 relevant terms/words to

a specific condition, diabetes, was gathered. These lexical units were described statistically, lexicologically, phonologically and phonetically. According to the results obtained, a test with a Likert scale was made and its application is recommended for primary healthcare professionals.

Key words: *Lexical familiarity, lexical availability, health literacy, diabetes.*

Recibido: 27 de abril de 2020

Aceptado: 30 de mayo de 2020

1. Introducción

Dada la alta prevalencia de las enfermedades crónicas cardiovasculares que según el Ministerio de Salud (*Mes del corazón* s.p), "permanecen como la principal causa de muerte en Chile y en el mundo", nos parece fundamental contribuir a una mejor comunicación entre personal de salud y pacientes. A nuestro juicio, cuanto facilite el acceso a la información de salud y ayude a incrementar la capacidad de los individuos para utilizar dicha información de una manera eficaz es un aporte a la equidad; esto es particularmente importante en un contexto como el chileno, puesto que en nuestro país un elevado porcentaje de personas con enfermedades crónicas cardiovasculares no logra realizar un automanejo eficiente que le permita compensarse metabólicamente y, por lo tanto, evitar las complicaciones de su patología (Lucero et al); para Nath, la calidad de este automanejo se relaciona directamente con la literacidad en salud (LS) de los individuos.

Literacy refers to basic skills needed to succeed in society while health literacy requires some additional skills, including those necessary for finding, evaluating and integrating health information from a variety of contexts. It also requires some knowledge of health-related vocabulary as well as the culture of the health system. Kwan, Frankish and Rootman (cit. en Kanj & Mitic 10)

Y si consideramos que la LS suele presentar una correlación positiva con el nivel sociocultural (Bustamante et al), nos encontramos con que justamente la población menos favorecida desde el punto de vista sociocultural tendrá baja literacidad en salud y, por ende, menos posibilidades de ejercer un autocuidado eficiente, lo que podrá conducirla a complicaciones de su condición y, eventualmente, a la muerte. Así, una investigación en

esta línea, resulta tener no sólo implicancias aplicadas sino también una fuerte motivación ética, que busca la justicia social y el acceso universal de la población a una salud digna.

Este artículo tiene como principal antecedente a Castillo Fadić y Oplustil Gallegos, donde se plantea la importancia de crear un instrumento que permita reconocer de manera automática si una unidad léxica es familiar o no para una persona. Para ello, se propone el diseño de un instrumento fonético-acústico basado en dos aspectos imprescindibles para su construcción: a) la existencia de un listado de unidades léxicas relevantes para un ámbito específico; y b) la existencia de una correlación significativa entre familiaridad léxica y distintas variables suprasegmentales. En el presente trabajo pretendemos avanzar hacia la construcción de este instrumento, mediante la obtención de un listado de unidades léxicas relevantes para el ámbito específico de la salud, su descripción lexicoestadística y fonológica y la descripción fonético-acústica de sus realizaciones orales.

El estudio, exploratorio, sintópico y sincrónico, considera especialmente la realidad del español de Chile. Se basa en los niveles léxico y fonológico de la lengua y considera el empleo de herramientas de lingüística de corpus, fonética acústica y lingüística computacional.

Nuestro objetivo general es generar aportes desde la lexicoestadística, la fonología y la fonética, orientados a avanzar en el diseño de un instrumento automático para realizar evaluaciones de *screening* de la familiaridad con unidades léxicas, orientado directamente al ámbito de la salud y, más específicamente, al área de las enfermedades crónicas cardiovasculares; dentro de estas últimas, nos enfocamos especialmente en la diabetes mellitus 2, dada su alta prevalencia en Chile (cf. Ministerio de Salud, *Resultados*). Así, nuestros objetivos específicos son: a) obtener un listado de unidades léxicas relevantes para pacientes con diabetes mellitus 2; b) determinar la frecuencia, dispersión y uso de estas unidades léxicas en el español general de Chile; c) describir las unidades léxicas relevantes en su estructura silábica y acentual; d) describir el comportamiento acústico de frecuencia fundamental (f_0) y duración de cada unidad léxica.

A continuación, presentamos nuestros fundamentos teóricos y metodológicos y damos cuenta de nuestros resultados y conclusiones.

2. Fundamentos teóricos

2.1 Familiaridad léxica y estadística léxica

Las unidades léxicas familiares pueden entenderse como altamente frecuentes, altamente usadas, altamente disponibles o, incluso, altamente asociadas al ámbito del hogar o de la familia, en cuyo caso *familiar* pasa a ser sinónimo de *coloquial* o *informal* (López Morales, *Sociolingüística*), como sucede en diccionarios de lengua española.

Para efectos de esta investigación, la noción de *familiaridad* pretende dar cuenta del grado de conocimiento o cercanía que un hablante experimenta respecto de una unidad léxica dada. Así, las unidades léxicas menos familiares serían aquellas que un individuo o un grupo escucha o lee por primera vez, mientras las más familiares serían aquellas que el hablante o comunidad ha oído o leído en reiteradas oportunidades. Esta noción puede sustentarse en criterios estadísticos y se entiende que, en la medida en que se da en un continuum, es eminentemente relativa. Así, si hablamos de unidades léxicas extraídas de un corpus textual o de referencia, consideramos que las unidades léxicas serán más familiares cuanto mayor sea su índice de uso (U), entendido este como el producto de la frecuencia (F) por la dispersión (D). Una unidad que simplemente sea frecuente, pero que no presente una distribución estable entre distintos mundos o géneros discursivos –es decir, una dispersión elevada–, no será familiar para el público general, sino únicamente para el público objetivo de cada mundo o género. Por ejemplo, *fonema* puede resultar frecuente en textos de fonética y fonología, del área técnico-científica, pero probablemente tendrá una frecuencia tendiente a cero en prensa o en obras dramáticas. Así, será una unidad familiar para los estudiosos de la fonética y la fonología, pero poco familiar o rara para los demás. El análisis lexicoestadístico de distintos corpus de referencia ha mostrado que los vocablos raros de tienden a presentar mayor complejidad silábica y fonológica que los vocablos de alto uso, lo que puede observarse en distintas lenguas. Los vocablos de alto uso, por el contrario, suelen ser predominantemente monosílabos, bisílabos y en menor grado trisílabos (Ávila 256), lo que facilita tanto su retención como su reconocimiento, pronunciación y escritura.

Las unidades de bajo uso podrán, sin embargo, a nuestro juicio, ser también altamente familiares cuando presenten una disponibilidad léxica alta, es decir, cuando pese a formar parte predominantemente del léxico pasivo, se activen fácil y rápidamente cuando

se proporciona al hablante el contexto para ello. Las unidades de mayor disponibilidad de una comunidad son aquellas que tienden a aparecer en los primeros rangos cuando se solicita a los hablantes enumerar unidades léxicas relacionadas con un tema específico, es decir, con un centro de interés particular. De este modo, unidades léxicas de alta disponibilidad dentro de la comunidad, serán también familiares. Unidades léxicas de baja disponibilidad y bajo uso, en tanto, serán estimadas como raras. Por último, las unidades léxicas de alto uso suelen pertenecer a familias léxicas más numerosas que las unidades raras, lo que propicia que los hablantes las reconozcan o puedan deducir su significado. Esto no sucede con las unidades raras, que tienden a ser representantes únicos de familias léxicas, o a emparentarse con otras unidades también raras, lo que no contribuye a su reconocimiento, retención o comprensión.

La necesidad de utilizar los índices de uso y disponibilidad en combinación radica en que mientras el léxico más usado es atemático, el más disponible es temático, de modo que ambos lexicones son necesarios para la comunicación y ambos conforman el llamado léxico fundamental (Lara).

Ahora bien, un estudio lexicológico supone trascender la *palabra*, unidad mínima textual y centrarse en el *vocablo*, unidad mínima del léxico. Mientras la primera se actualiza en el discurso y puede ser observada fonéticamente, el segundo implica una abstracción desde un conjunto de variantes hacia una variable que representa todo un paradigma de realizaciones y puede actuar como lema o unidad de cita; este último, el vocablo, es susceptible de análisis fonológico. Las distintas realizaciones de superficie de cada vocablo, las palabras, permiten un análisis fonético.

El vocablo, para efectos de este estudio, podrá ser univocal o plurivocal, según esté formado por una o más palabras que, en conjunto, funcionen como una sola unidad de denominación. En sentido extenso, incluimos también las colocaciones, entendidas como combinaciones frecuentes no arbitrarias de dos o más palabras, que dan cuenta de restricciones de predicación respecto de un argumento o un conjunto de argumentos (Bosque 17).

2.2 Sílaba, acentuación, duración y frecuencia fundamental

Los vocablos pueden ser descritos de acuerdo a su estructura silábica y a su acentuación. Las sílabas se definen como las unidades fonológicas inmediatamente inferior al enunciado o cadena fónica. Estas unidades se estructuran de acuerdo al sistema lingüístico de cada lengua, específicamente el fonológico. En el español, las sílabas pueden estar conformadas por una estructura simple, es decir, un solo fonema (una vocal) o por estructuras complejas, es decir, más de un fonema (vocal más consonante, por ejemplo). Cualquiera sea la estructura de la sílaba, esta siempre estará conformada por un núcleo que en el español siempre corresponderá a una vocal. Este núcleo puede ser acompañado (o no) por una o más consonantes en posición anterior o posterior al mismo núcleo. También, el núcleo puede ser acompañado por una o más vocales para conformar juntos un núcleo complejo (Martínez Celdrán, *Fonología*).

La acentuación está estrechamente vinculada a las sílabas, en cuanto estas últimas "son las unidades mínimas que soportan los tres parámetros que intervienen en la prosodia" (Martínez Celdrán y Fernández Planas, *Manual* 197). Estos parámetros se refieren a la *duración*, la *intensidad* y la *frecuencia fundamental*, parámetros acústicos que sustentan el acento al interior de un enunciado o de una unidad léxica. Este último caso, se conoce como "acento léxico", es decir, aquel relieve que se produce en una sílaba específica al interior de una unidad léxica (Gil 535). La posición que tome este acento determinará si el vocablo se cataloga como voz aguda, grave, esdrújula o sobreesdrújula.

Desde un punto de vista fonético, la voz humana produce sonidos cuya fuente puede ser sonora o áfona, debido al (nivel de) movimiento de los pliegues vocálicos. Los sonidos cuya fuente sea sonora tienen como consecuencia la *frecuencia fundamental*, también llamada f_0 . Esta suele definirse como el armónico de menor frecuencia en una onda periódica compleja.

Una cadena fónica, que esté conformada por sonidos sonoros en su totalidad o en parte de ella, tendrá una serie de frecuencias fundamentales que se moverán entre ciertos rangos frecuenciales. La serie de frecuencias conformarán lo que se conoce como curva de frecuencia fundamental (f_0). Hay tres maneras de trabajar los valores acústicos de la curva de f_0 : la primera es a partir de valores absolutos de Hz; la segunda es realizar una normalización de los valores para disminuir la variación intra e inter locutor; la tercera

consiste en analizar la curva, considerando aspectos lingüísticos (Cantero; Martínez Celdrán).

La duración es un parámetro que forma parte de lo segmental y de lo suprasegmental (Cantero; Martínez Celdrán, *El sonido*; Fernández Planas). Así lo deja entrever Gili Gaya (43) cuando dice "Este *mínimum* es distinto para cada sonido, según su naturaleza articulatoria y la de los sonidos que se hallen en contacto con él; varía además en relación a la intensidad y el tono". También se ha definido como la prolongación de un sonido o conjunto de ellos en una determinada unidad de tiempo (milisegundos, por lo general). Esta duración puede estar determinada, al igual que la frecuencia fundamental, por características propias de los sonidos. Por ejemplo, todos los sonidos consonánticos áfonos u oclusivos tienen menos posibilidades de ser prolongados o "alargados" que los sonidos consonánticos sonoros o vocálicos, debido a las características propias de los primeros versus los segundos.

En español, las realizaciones suprasegmentales tienen implicancias dialectales; así lo demuestran los diversos estudios de AMPER que se han llevado a cabo en hispanoamérica (Martínez Celdrán y Fernández Planas, *Atlas*). De ahí la necesidad de adecuar un futuro instrumento según la zona de aplicación. En este caso, en particular, el análisis suprasegmental se centrará en realizaciones propias del español de Chile.

3. Metodología

3.1 Determinación de léxico relevante y creación de un listado de vocablos

Vista la necesidad de crear un listado de vocablos relevantes específico, nos centramos en los relacionados con la diabetes mellitus 2. Así, previa firma de un consentimiento informado, se aplicó un test de disponibilidad léxica focalizado en el centro de interés *La diabetes* a 83 profesionales de la salud especializados en atención de pacientes crónicos cardiovasculares, que trabajan en atención primaria en la Región Metropolitana y en la Región de Valparaíso. Tras descartar a extranjeros, personas con perfiles sociológicos incompletos e informantes no profesionales (técnicos en enfermería de nivel superior o TENS), que pueden ser objeto de un estudio posterior, llegamos a un n=70, compuesto por médicos, enfermeros, kinesiólogos y nutricionistas chilenos, que han residido la mayor parte de su vida en Chile y han realizado sus estudios escolares y universitarios en el país.

El test se realizó siguiendo los parámetros del Proyecto Panhispánico de Léxico Disponible (DispoLex, Grupo de Investigación), basado en López Morales (*Léxico*), de modo que se entregó a cada informante una hoja en blanco y se le pidió escribir en ella todas las palabras que vinieran a su mente en relación con el tema *La diabetes*, de arriba hacia abajo como en una lista de supermercado.

Los datos fueron transcritos e ingresados en una hoja de cálculo. Cada unidad léxica se asoció a un código alfanumérico de tres dígitos que funciona como etiqueta de cada informante y a un rango que da cuenta de la posición que dicha respuesta ocupa en el listado producido por cada uno. Luego, se estandarizaron los datos en cuatro pasos: 1) corrección ortográfica y lematización en singular de formas nominales; 2) lematización en masculino de formas nominales; 3) lematización en infinitivo de verbos conjugados y adaptación de préstamos léxicos según el *DLE*; 4) estandarización de unidades fraseológicas, incluyendo no sólo locuciones y unidades fraseológicas, sino también colocaciones. Al igual que en Santos Díaz, se incluyeron todas las aportaciones de los informantes, con independencia de que estuviesen relacionadas con la temática o fuesen nombres propios. Sólo se eliminaron las repeticiones: cuando un mismo informante daba exactamente la misma respuesta en dos rangos distintos, se mantuvo únicamente la primera.

A continuación, los datos fueron exportados a *.TXT, codificación ANSI, y se reordenaron de acuerdo con los requerimientos de Dispogen para Windows Versión 1.3.1, programa escogido para el cálculo del índice de disponibilidad léxica (IDL) de cada vocablo. Interesa destacar que Dispogen para Windows Versión 1.3.1 es un software libre creado en Chile, que puede instalarse en cualquier ordenador con ambiente Windows. Se trata de

Una aplicación creada en MatLab por Max Echeverría, Paula Urzúa e Israel Figueroa. Matlab es un programa que realiza cálculos sobre la base de matrices y está especializado en el análisis estadístico de datos multivariantes, como es el caso de la disponibilidad léxica. La ventaja de este *software* es que para calcular el IDL utiliza la última fórmula matemática elaborada en México y cuyos resultados muestran mayor grado de precisión que los entregados por anteriores fórmulas. (Ferreira y Echeverría 141)

Este programa exige que los datos estén organizados como se ilustra en el ejemplo siguiente:

22234 010 03 respuesta 1, respuesta 2, respuesta 3, respuesta n

La primera posición la ocupan cinco números (en este caso, 22234). Cada uno representa una variable extralingüística que da cuenta del grupo al que pertenece el informante. Dispogen admite hasta cinco variables. Si todos los informantes pertenecen al mismo grupo, se puede ingresar un 1 en cada caso, pero no se pueden dejar campos vacíos. Luego de un espacio, en segunda posición se ubica el código que identifica a cada informante (en el ejemplo, 010), que necesariamente debe tener tres dígitos. Después de otro espacio, se ingresa el número que identifica al centro de interés. En este caso, es el centro de interés 03, que puede corresponder a cualquier centro previamente definido por los investigadores en su plan de codificación. En nuestra investigación, se trata del centro de interés *La diabetes*. Por último, y a continuación de un nuevo espacio, se ordenan las unidades léxicas mencionadas por el informante, en el orden en que fueron escritas y separadas por coma y espacio.

Este archivo *.TXT, adecuadamente ordenado, fue procesado en Dispogen para Windows Versión 1.3.1, lo que permitió ordenar las unidades léxicas de mayor a menor disponibilidad, según la fórmula de López Chávez y Strassburger Frías.

Para efectos de esta investigación, seleccionamos las primeras veinte unidades para ser postuladas como las más relevantes en relación con diabetes. El punto de corte se estableció a partir de criterios estadísticos (cf. 4. Resultados). Estos vocablos pasaron a ser nuestros vocablos meta y el núcleo del listado a partir del cual podrá evaluarse la familiaridad. Para evitar que alguno de ellos quedara en posición inicial o final, con las consiguientes consecuencias suprasegmentales en la lectura en voz alta (Castillo Fadić y Oplustil Gallegos 189-198), agregamos un distractor al principio de la lista (*coolesterol*) y otro al final (*nefropatía*), ambos extraídos de las respuestas de los informantes en relación con este centro de interés, para que no fuera evidente que no forman parte de las unidades léxicas sujetas a evaluación.

colesterol
glicemia
hemoglobina glicosilada
diabetes
insulina
compensación
hidrato de carbono
medicamento
autocuidado
actividad física
azúcar
dieta
metformina
páncreas
alimentación
ejercicio
enfermedad
alimentación saludable
control
complicación
hipoglicemia
nefropatía

Figura 1. Listado de unidades léxicas para lectura en voz alta. Centro de interés: La diabetes

3.2 Descripción estadística de las unidades léxicas relevantes

Cabe recalcar que la descripción se realizó considerando las 20 unidades léxicas relevantes y no los dos distractores.

Además de obtener el índice de disponibilidad (IDL), frecuencia (F_i) y porcentaje de aparición (%) de las unidades léxicas en el grupo de informantes, se revisó frecuencia, dispersión (D) y uso (U) de estas unidades léxicas en el español general de Chile. Para la determinación de los índices de F_i , D y U nos basamos en el *Corpus Básico del Español de Chile* © (Castillo Fadić). Aunque no contamos con estudios previos de este centro de interés en el español de Chile, nos pareció de interés revisar los resultados de Valencia y Echeverría, ya que algunos de nuestros vocablos relevantes sí aparecen en centros de interés tradicionales. Además, se revisó si las unidades léxicas estaban o no registradas en el *DLE* (Real Academia Española y ASALE); si estaban consignadas, se observó su definición y su marcación diatécnica.

3.3 Descripción lexicológica y fonológica de los vocablos relevantes

Las unidades se clasificaron en univerbales y pluriverbales; en este último grupo se incluyeron también las colocaciones.

Se dispuso de una tabla (planilla electrónica) donde se describió el acento fonológico (unidad aguda, grave o esdrújula), se contó el número de sílabas por cada unidad y se describió la estructura simple (vocal - consonante) de cada sílaba.

3.4 Obtención y procesamiento de un corpus oral de realizaciones expertas

Grabamos por medio de un teléfono móvil a 22 profesionales de la salud, principalmente médicos y enfermeros, leyendo la lista de los vocablos meta (con los dos distractores incluidos), diez veces cada uno; para ello, se imprimió una lista que ocupa una página tamaño carta, en un tipo de letra sin serif (Arial), tamaño mediano (13) e interlineado doble, donde cada vocablo ocupa una línea, de modo de facilitar la lectura (cf. Figura 1). Esto generó un corpus de 4.840 palabras, donde 4.400 corresponden a realizaciones de vocablos meta y donde cada vocablo meta tiene una $F_i=220$. Tuvimos la precaución, además, de incluir tanto hombres como mujeres, para facilitar el reconocimiento de voz (Oropeza Rodríguez) sin importar si ésta es grave o aguda, lo que se realizará en etapas posteriores de la investigación.

La razón de emplear teléfonos celulares y no grabadoras no es antojadiza: radica en lo accesible de estos aparatos, en la dificultad de contar con grabadoras en centros de atención primaria y en la dificultad de contar con espacios especialmente acondicionados para realizar grabaciones libres de ruidos externos en este tipo de centros. A esto se suma la expectativa de generar a futuro una aplicación para teléfonos móviles que permita aplicar el test en cualquier lugar, sin requerir de una infraestructura especial. La realización de estas grabaciones en ambientes no controlados permitirá evaluar la viabilidad de generar una herramienta automática utilizable en espacios comunes de centros de salud.

Dado que estas grabaciones podrían ser la base para entrenar a posteriori modelos de reconocimiento de voz (sistema experto) que permitan segmentar de manera automática las palabras en futuras grabaciones de pacientes (señales de voz desconocidas) realizamos el ejercicio piloto de segmentar cada archivo de audio en formato *.WAV, manualmente,

usando el programa Praat; para ello se etiquetaron los valores temporales que delimitan cada palabra y el silencio o ruido ambiental, tanto entre cada una de ellas como al final y comienzo de cada grabación. Ya que estas etiquetas de segmentación son imprescindibles para aplicar metodologías de reconocimiento de voz (Oropeza Rodríguez 284), este ejercicio, que por un lado constituye un avance para etapas posteriores de desarrollo del instrumento, permite también ponderar hasta qué punto la pérdida de calidad de la grabación y control de la señal redundante en pérdida de información imprescindible para el análisis fonético acústico automático, para alcanzar un razonable equilibrio entre los ambientes no controlados de un centro de salud y la obtención de archivos de audio útiles para el instrumento en desarrollo.

3.5 Descripción fonético-acústica de realizaciones expertas

La obtención de los datos acústicos de duración y frecuencia fundamental (f_0) se realizó mediante el Praat (Boersma y Weenink) a través de la utilización de un *script* (cf. 7. Anexo programa). Esta rutina obtuvo los valores de duración a partir del tiempo de inicio y tiempo final de cada vocablo en milisegundos (ms) desde las marcas establecidas previamente en el mismo programa. También extrajo los valores de la frecuencia fundamental promedio de cada unidad léxica junto con la frecuencia mínima y máxima de las mismas; todo lo anterior en hertz (Hz).

4. Resultados

4.1 Listado de léxico relevante

Los vocablos más disponibles del centro de interés *La diabetes*, que conforman el primer componente de nuestro instrumento, son los siguientes, ordenados por rango de mayor a menor disponibilidad:

- | | |
|----------------------------|---------------------|
| 1. glicemia | 7. medicamento |
| 2. hemoglobina glicosilada | 8. autocuidado |
| 3. diabetes | 9. actividad física |
| 4. insulina | 10. azúcar |
| 5. compensación | 11. dieta |
| 6. hidrato de carbono | 12. metformina |

- | | |
|------------------|----------------------------|
| 13. páncreas | 17. alimentación saludable |
| 14. alimentación | 18. control |
| 15. ejercicio | 19. complicación |
| 16. enfermedad | 20. hipoglicemia |

Como se puede observar, un número importante de ellos pertenece al ámbito técnico (cf. Lerat); en esta categoría encontramos de hecho los vocablos de los cuatro primeros rangos: *glicemia*, nivel de glucosa en la sangre; *hemoglobina glicosilada*, examen para evaluar la compensación metabólica de un paciente diabético, que mide el nivel promedio de glucosa en sangre de los últimos tres meses; *diabetes*, la condición en la que se centra este listado; *insulina*, la hormona que regula la cantidad de glucosa en la sangre. Sólo *actividad física*, *azúcar*, *dieta*, *alimentación*, *ejercicio*, *enfermedad* y *control* parecen pertenecer a la lengua común. *Compensación*, si bien puede resultar una palabra conocida, aparece aquí con un significado diferente del corriente, restringido a la compensación metabólica. *Medicamento* es la variante técnica de *remedio*. *Autocuidado*, aun cuando contiene una base léxica de alto uso, resulta también propia del ámbito de salud. *Metformina*, un hipoglicemiente oral muy utilizado por pacientes diabéticos y resistentes a la insulina, presenta un uso mínimo desde el punto de vista lingüístico, por su baja dispersión; incluso sus usuarios suelen conocerlo por alguno de sus nombres de fantasía. *Páncreas*, el órgano encargado de la producción de insulina, no se cuenta entre los más conocidos por los no especialistas, que raramente lo relacionan con la diabetes y tienen dificultades para indicar su función o su ubicación dentro del cuerpo. *Alimentación saludable*, anotada aquí separada de *alimentación* a secas, resulta técnica en la medida en que exige saber qué es *saludable* dicho de *alimentación*. *Complicación* resulta también vaga para los no especialistas; se entiende como algo negativo, pero no se distingue de *agravamiento*. *Hipoglicemia*, por último, es claramente técnico.

4.2 Descripción estadística de las unidades léxicas

En la Figura 2 resumimos los principales índices estadísticos obtenidos a partir del análisis de los tests aplicados a profesionales de la salud chilenos. El índice de disponibilidad léxica (IDL) da cuenta de los vocablos más relevantes para que un paciente

María Natalia Castillo Fadic y Josué David Pino Castillo. "Hacia la construcción de un instrumento para evaluar la familiaridad de pacientes crónicos con unidades léxicas relevantes para el automanejo de su condición de salud".

diabético comprenda las indicaciones relacionadas con su condición, en la medida en que son los que se activan más fácilmente en el grupo de profesionales estudiado cuando se habla de diabetes. La frecuencia aquí es absoluta y corresponde al número de repeticiones de la respuesta en el total de la muestra. El porcentaje de aparición, en tanto, relativiza esta frecuencia respecto del tamaño de la muestra.

Vocablo	IDL	Frecuencia	% Aparición
glicemia	0,49213725	46	65,7142878
Hemoglobina glicosilada	0,43329629	45	64,2857134
diabetes	0,34351087	26	37,1428579
insulina	0,24960192	31	44,2857146
compensación	0,21176843	22	31,4285725
hidrato de carbono	0,19647385	20	28,5714298
medicamento	0,17670389	20	28,5714298
autocuidado	0,17337595	21	30,0000012
actividad física	0,16920885	24	34,2857152
azúcar	0,13625287	13	18,571429
dieta	0,13031842	14	20,0000003
metformina	0,11501777	14	20,0000003
páncreas	0,11466284	13	18,571429
alimentación	0,11231304	13	18,571429
ejercicio	0,11087679	14	20,0000003
enfermedad	0,1100394	10	14,2857149

alimentación saludable	0,10369705	11	15,7142863
control	0,10257731	14	20,0000003
complicación	0,0980647	14	20,0000003
hipoglicemia	0,09573425	14	20,0000003

Figura 2. Índices de disponibilidad, frecuencia y porcentaje de aparición de las unidades léxicas relevantes en la muestra de profesionales de la salud.

Las unidades del listado son, como es natural en un estudio de disponibilidad léxica, fuertemente temáticas. Pese a su alta disponibilidad en el grupo encuestado, presentan bajos índices en el *Corpus Básico del Español de Chile* © (Castillo Fadić), que incluye alrededor de 500.000 palabras obtenidas de 26 años de publicaciones chilenas (Drama, Narrativa, Ensayo, Técnico-Científico y Prensa). Entre los vocablos que podríamos considerar más raros, se encuentran *hemoglobina glicosilada*, *hidrato de carbono*, *autocuidado*, *metformina* y *alimentación saludable*, que ni siquiera aparecen en el corpus de referencia consultado (F=0); les siguen *glicemia* (F=2), *insulina* (F=2), *actividad física* (F=1), *páncreas* (F=1) e *hipoglicemia* (F=2), con dispersión D=0, ya que sólo tienen ocurrencias en un tipo de texto -en el caso de los cuatro primeros, textos técnicos y científicos; en el caso del último, la sección de deportes de un diario-; a continuación, *diabetes* (F=5, D=0,2, U= 1,15) con cuatro ocurrencias en textos técnicos y científicos y una en narrativa; *compensación* (F=4, D=0,5, U= 2,12), con una ocurrencia en ensayo, dos en textos técnicos y científicos y una en prensa, todas con acepciones distintas a la relativa a salud; *azúcar* (F=4, D=0,3, U= 1,08), con tres ocurrencias en textos narrativos y una en un texto técnico-científico de un ámbito distinto a la salud, en todos los casos con un significado restringido a la sustancia dulce obtenida de la remolacha o de la caña de azúcar, que no abarca la acepción química relativa a hidrato de carbono, pertinente para el ámbito de la diabetes; *dieta* (F=4, D=0,5, U= 2,12), con dos ocurrencias en textos narrativos, una en un texto técnico y una en un diario; y *complicación* (F=8, D=0,39, U= 3,12), con cuatro ocurrencias en textos técnico-científicos del ámbito de la salud y cuatro en diarios (dos en noticias de salud). Así, mientras sólo *medicamento* (F=11, D=0,78, U= 8,58), *alimentación* (F=6, D=0,69, U= 4,14), *ejercicio* (F=51, D=0,69, U= 35,19), *enfermedad* (F=45, D=0,84, U=

37,80) y *control* (F=89, D=0,6, U= 53,4) forman parte de los cinco mil vocablos de mayor uso en el español general de Chile (Cf. Castillo Fadić, *Léxico*), los quince vocablos restantes, estimados relevantes para la comprensión de la diabetes y del autocuidado relativo a esta condición, son infrecuentes, poco dispersos o derechamente raros. En este sentido, la mayor parte de ellos será poco familiar para el hablante no especialista; otros le resultarán familiares o medianamente familiares, por afinidad fonosemántica con vocablos conocidos o por su pertenencia a familias de palabras conocidas; algunos, polisémicos, como *azúcar*, les parecerán familiares, aunque el hecho de que el significado de las unidades léxicas en el español general difiera del relativo al área de la salud podrá generar equívocos importantes.

Respecto del índice de disponibilidad en el español general de Chile (cf. Figura 3), si bien no se encuentran estudios previos del mismo centro de interés, en la revisión de los resultados obtenidos por Valencia y Echeverría en torno a los centros de interés tradicionales, se observa que la mayor parte de los vocablos relevantes de nuestro listado no aparece en el mencionado trabajo (*glicemia, hemoglobina glicosilada, diabetes, insulina, compensación, autocuidado, actividad física, metformina, alimentación saludable, complicación e hipoglicemia*); los que se encuentran presentes lo están con baja disponibilidad (*hidrato de carbono, medicamento, páncreas, alimentación*) y, en la mayoría de los casos, con diferencias semánticas de comprensión o extensión (*azúcar, dieta, control*).

Vocablo	IDL Valencia y Echeverría 1999
glicemia	No se registra. En el CI <i>alimentos</i> , se registra el término afín <i>glucosa</i> ID= 0,001. Tb aparece en el CI <i>Ciencia y tecnología</i> , con ID=0,0002.
hemoglobina glicosilada	No se registra.
diabetes	No se registra.
insulina	No se registra.
compensación	No se registra.
hidrato de carbono	En el CI <i>Alimentos</i> , presenta ID=0,001. Se registra también la variante <i>carbohidrato</i> , ID=0,002.

medicamento	En el CI <i>Ciencia y tecnología</i> , presenta ID=0,019. El término afín <i>medicina</i> presenta ID= 0,174 en el mismo CI. El equivalente semántico <i>remedio</i> presenta ID= 0,027 en el CI <i>Ciencia y tecnología</i> ,
autocuidado	No se registra.
actividad física	No se registra. En el CI <i>Juegos y entretenimientos</i> aparece el vocablo afín <i>actividad</i> .
azúcar	En el CI <i>Alimentos</i> , presenta ID= 0,100. En el CI <i>La cocina</i> , presenta ID= 0,026
dieta	En el CI <i>La cocina</i> , presenta ID= 0,00006. El término afín <i>dietista</i> aparece en el CI <i>Ciencia y tecnología</i> con ID=0,0002.
metformina	No se registra.
páncreas	En el CI <i>Las partes del cuerpo</i> , presenta ID=0,049.
alimentación	En el CI <i>Alimentos</i> , presenta ID=0,005. En el CI <i>La cocina</i> , ID= 0,0004 En el CI <i>Problemas del ambiente</i> , ID= 0,0004.
ejercicio	CI <i>Juegos y entretenimientos</i> , ID= 0,0008
enfermedad	En el CI <i>Ciencia y tecnología</i> , presenta ID=0,013.
alimentación saludable	No se registra. Cf. <i>alimentación</i> .
control	En el CI <i>Ciencia y tecnología</i> , presenta ID=0,0005. En el CI <i>Procesos mentales</i> , ID=0,002.
complicación	No se registra. En el CI <i>Procesos mentales</i> , se encuentra <i>complicado y complicarse</i> , con diferencia semántica de comprensión.
hipoglicemia	No se registra. Cf. <i>glicemia</i> .

Figura 3. Vocablos relevantes en Valencia y Echeverría 1999.

La baja disponibilidad de los vocablos relevantes del centro de interés *La diabetes* en el español general de Chile, sumada a su bajo uso, dan cuenta de su escasa familiaridad para hablantes no especializados.

4.3 Descripción lexicológica y fonológica de las unidades léxicas relevantes

Si observamos el listado de 20 vocablos relevantes para la diabetes, no hay monosílabos; solo dos unidades son bisilábicas *-dieta* y *control-*, pero ambas presentan sílabas complejas. Mientras la primera incluye un diptongo, la segunda contiene una sílaba compuesta o trabada y es, además, aguda, en una lengua como la española que tiende a las palabras graves.

Dentro de los primeros rangos, encontramos unidades pluriverbales que funcionan como unidades de denominación. Así, *hemoglobina glicosilada* e *hidrato de carbono* no sólo están constituidas por más de una palabra, sino que además presentan una serie de complejidades estructurales, desde su extensión (10 sílabas y siete sílabas) hasta la inclusión de sílabas trabadas y la presencia de al menos dos acentos de intensidad.

En los casos de *actividad física* y *alimentación saludable*, se trata de colocaciones, a saber, grupos de palabras que aparecen frecuentemente en combinación, motivados por restricciones de predicación. Si bien tanto el sustantivo *actividad* como el adjetivo *físico* pueden aparecer en distintas combinaciones, *actividad física* funciona como unidad en la medida en que no toda obra, ejecución o tarea (actividad) relativa a la naturaleza corpórea (física) constituye una *actividad física* en términos de salud. Por otra parte, aunque *alimentación* ("Acción y efecto de alimentarse", Real Academia Española y ASALE, s.v. *alimentación*, 1ª acep) aparece en nuestro listado de manera independiente, estimamos que *alimentación saludable* funciona en su conjunto como una unidad distinta, en la medida en que *saludable* ("Que sirve para conservar o restablecer la salud corporal", Real Academia Española y ASALE, s.v. *saludable*, 1ª acep)., dicho de *alimentación*, supone que existen ciertos alimentos que pueden contribuir a la salud de los individuos, de modo que la comprensión cabal de esta colocación implica conocer cuáles son esos alimentos y en qué proporción conviene consumirlos.

Del total de 20 unidades léxicas, se observa un grupo de 16 unidades univerbales y otro de 4 unidades pluriverbales. En relación al acento, el primer grupo está compuesto por 5 unidades univerbales agudas, 10 unidades graves y solo 1 unidad esdrújula. Del grupo de las unidades pluriverbales, 2 corresponden a grave-grave, 1 a aguda-esdrújula y 1 a aguda-grave.

En relación con la cantidad de sílabas al interior de cada unidad, las univerbales se encuentran en el rango de 2 a 5 sílabas por unidad; en cambio, las pluriverbales están en el rango de 7 a 10 sílabas. En cuanto a la estructura silábica, la mayoría de las unidades (17 exactamente) están compuestas por al menos una estructura consonante-vocal (CV), le sigue la estructura consonante-vocal-consonante (CVC) y luego la configuración consonante-consonante-vocal (CCV). Esta distribución es acorde a lo que ocurre en el sistema lingüístico hispánico general (Moreno Sandoval, Torre Toledano y Curto).

4.5 Descripción acústica de las unidades léxicas realizadas por personal de salud

La lectura en voz alta a repetición de las unidades léxicas relevantes por parte de profesionales de la salud permite abstraer parámetros hacia la construcción de un modelo ideal o figura referencial totalizadora. Si bien todos los informantes presentan realizaciones distintas de los vocablos del listado, y si bien existen variaciones entre las realizaciones de un mismo sujeto -“Es imposible que un locutor pronuncie dos veces exactamente igual una misma sílaba, palabra o frase” (Esparza Arellano & Ávalos Briseño, 2003)-, esta descripción pretende precisar rasgos suprasegmentales comunes que puedan dar cuenta de la familiaridad de los sujetos con las unidades léxicas. Estos rasgos podrán, en una fase posterior, ser contrastados con los de usuarios no expertos.

Respecto de la duración, se observa que las unidades léxicas con mayor duración son las unidades léxicas pluriverbales de la lista, tanto en hombres como en mujeres. También se observa que la duración más baja, tanto en hombres como en mujeres, corresponde a la unidad léxica “dieta” (cf. Figura 4 y 5). En ambos gráficos, se aprecian valores espurios en casi todas las palabras; vale decir, hay valores de duración que se alejan de la media, de sus mínimos y máximos. Aunque estos valores son minoritarios dentro de la totalidad del corpus, se evidencia que la mayoría están contenidos en las voces femeninas.

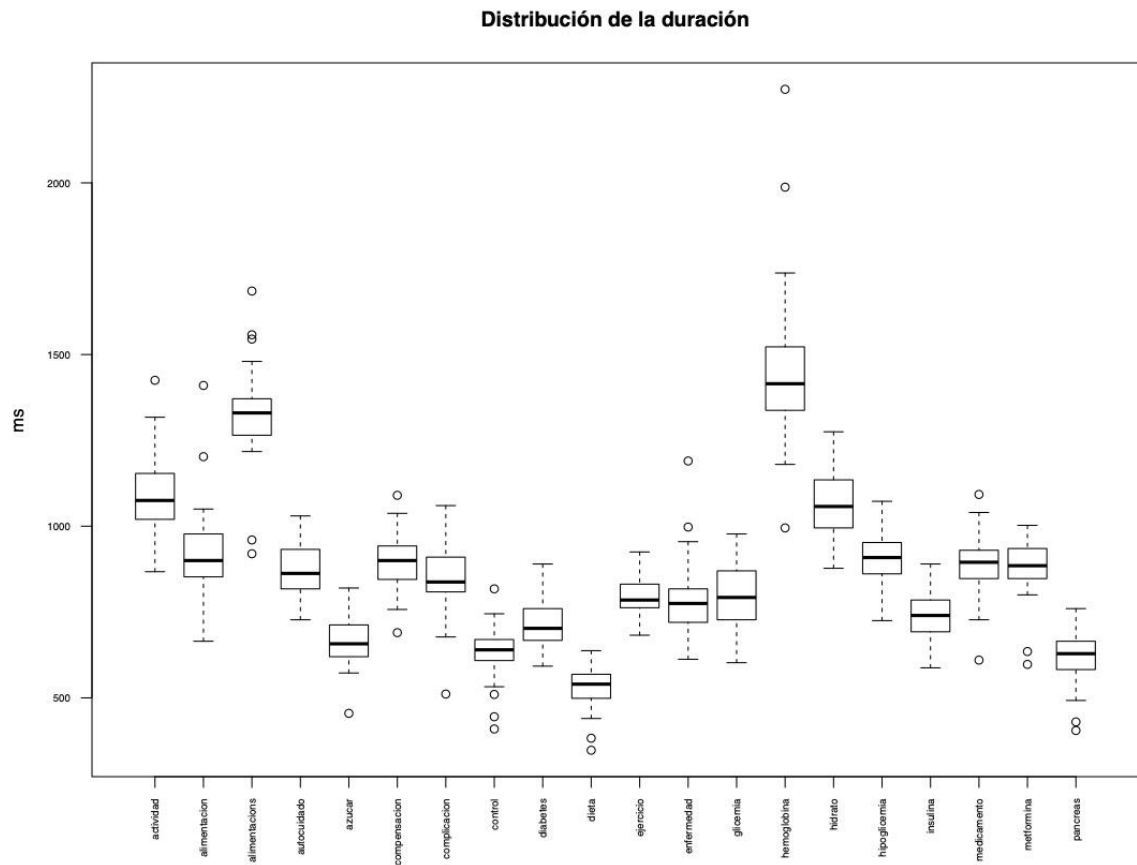


Figura 4. Distribución de la duración de hablantes hombres.

En relación a la frecuencia fundamental (cf. Figura 6 y 7), se observa que los valores de todas las palabras se mantienen dentro de un mismo rango frecuencial: en el caso de las voces masculinas, la f_0 se mantuvo sobre los 100 Hz y bajo los 180 Hz. En cambio en los valores de las voces femeninas, las frecuencias estuvieron sobre los 150 Hz y bajo los 250 Hz. En este último grupo, aparecen valores espurios, principalmente bajo los 150 Hz.

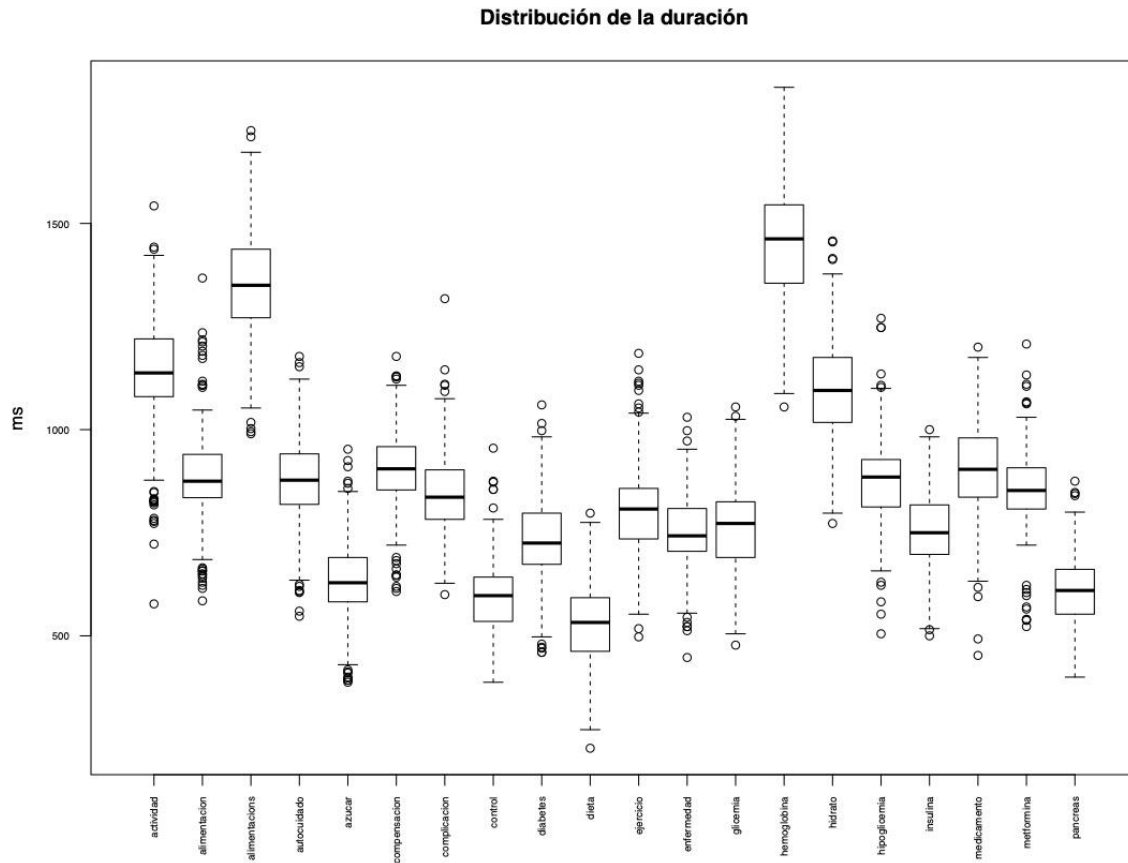


Figura 5. Distribución de la duración de hablantes mujeres.

5. Conclusiones

5.1 Alcances metodológicos.

Estimamos que la aplicación de un test de disponibilidad léxica ha sido una herramienta adecuada para obtener el léxico relevante en relación con con la diabetes. Recomendamos, por tanto, el empleo de este tipo de test para obtener léxicos relevantes de otras áreas temáticas, estudiando previamente cuál es el estímulo o seña más adecuado para el ámbito en cuestión. Al respecto, cabe destacar tanto la importancia de acotar adecuadamente el centro de interés, como la necesidad de contar con un $n \geq 50$ e idealmente cercano a 100, formado específicamente por expertos en el área de que se trate. En este sentido, consideramos que fue acertado tanto acotar el centro de interés a una condición de salud en específico como aumentar el n de expertos en relación con lo propuesto por Castillo Fadić y Oplustil Gallegos.

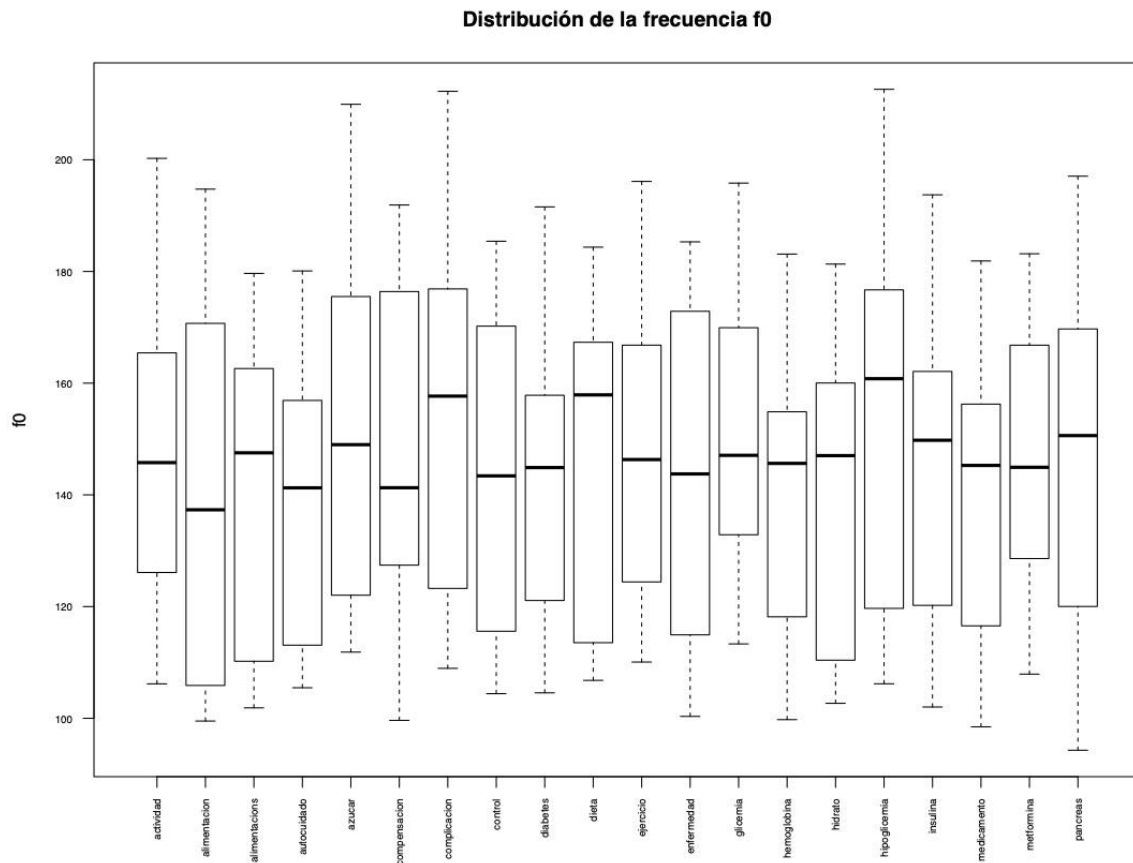


Figura 6. Distribución de la frecuencia fundamental (f0) de hablantes hombres.

Las dificultades metodológicas que implica el diseño de un instrumento de evaluación automática de la familiaridad, motivadas en gran medida por la necesidad de mayores estudios centrados en la correlación entre variables suprasegmentales y familiaridad léxica en el español de Chile, no nos han impedido avanzar tanto en el eje léxico como en el fónico de nuestro prototipo; esperamos que las mejoras realizadas al modelo (a saber, la consideración de rasgos comunes a lectores expertos) faciliten el desarrollo del instrumento en comunidades que no cuenten con investigaciones previas sobre la correlación entre variables suprasegmentales y familiaridad léxica.

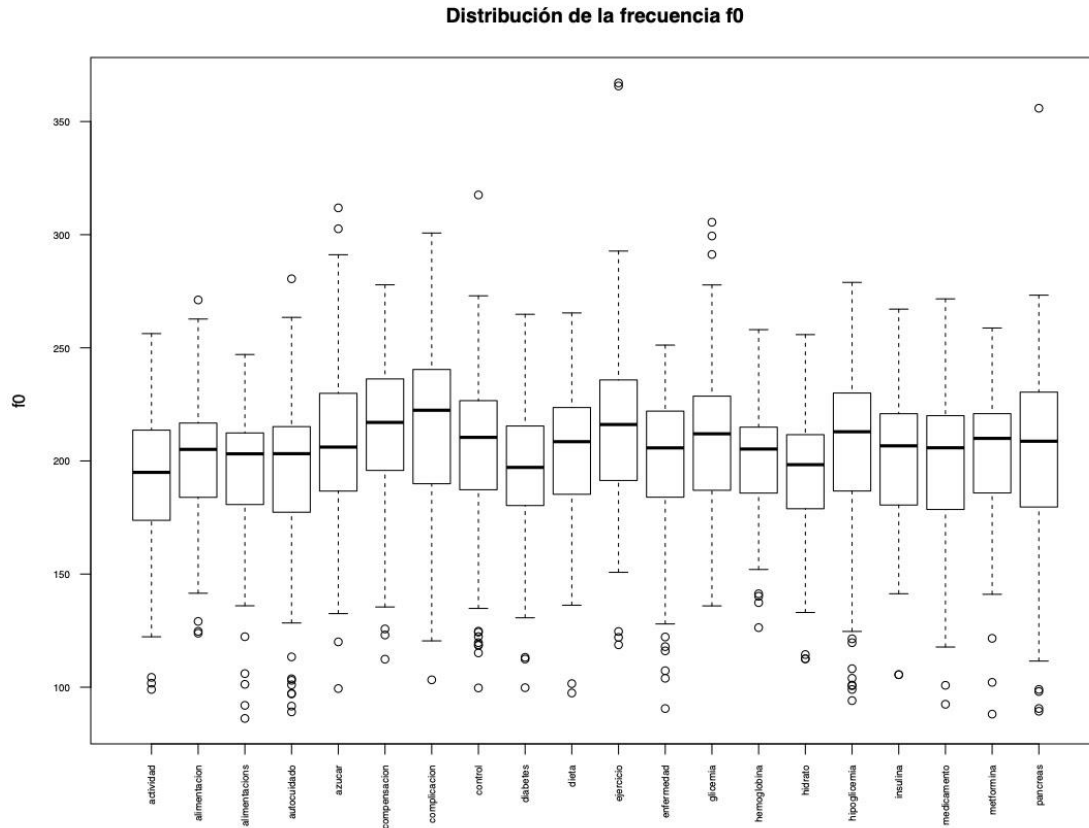


Figura 7. Distribución de la frecuencia fundamental (f0) de hablantes mujeres.

Respecto de la decisión de grabar con teléfonos celulares en ambientes no controlados, nos ha previsto una experiencia útil en lo relativo a las limitaciones de este método. Si bien sostenemos que un futuro instrumento automático debe funcionar en espacios no insonorizados, los avances actuales dificultan el funcionamiento en espacios abiertos con ruido ambiental; sería deseable buscar un término medio, como una sala de consulta cerrada.

5. 2 Principales resultados

Junto con confirmar la conveniencia de usar tests de disponibilidad léxica para obtener léxico relevante de áreas temáticas específicas, hemos obtenido un listado con las 20 unidades léxicas más relevantes en el área de la diabetes; recomendamos que sea tomado en cuenta por los profesionales de salud a la hora de explicar conceptos fundamentales para el autocuidado a sus pacientes diabéticos, puesto que, por su carácter técnico y su bajo

María Natalia Castillo Fadic y Josué David Pino Castillo. "Hacia la construcción de un instrumento para evaluar la familiaridad de pacientes crónicos con unidades léxicas relevantes para el automanejo de su condición de salud".

índice de uso y de disponibilidad en el español general de Chile, resultarán muy probablemente poco familiares a los no especialistas. En tanto no se disponga de una herramienta automática validada, proponemos emplear un breve test con escala de Likert, donde los pacientes pueden marcar su grado de familiaridad con cada uno de los vocablos relevantes (cf. Figura 8). Los vocablos con puntajes de 0 a 2 deberían ser explicados cuidadosamente a los pacientes. Esto, porque la *inseguridad lingüística*, mayor en hablantes de menor nivel sociocultural (López Morales, *Sociolingüística* 277-279), podría llevarlos a brindar respuestas que sobrevaloren su familiaridad con las unidades léxicas, para evitar ser juzgados desfavorablemente por el personal de salud o por sus acompañantes.

Listado	0	1	2	3
	No me es familiar. Es primera vez que leo o escucho esta palabra.	Me es poco familiar. He oído esta palabra pocas veces.	Me es más o menos familiar He oído o leído esta palabra varias veces.	Me es muy familiar. He oído o leído muchas veces esta palabra.
glicemia				
hemoglobina glicosilada				
diabetes				
insulina				
compensación				
hidrato de carbono				
medicamento				
autocuidado				
actividad física				
azúcar				
dieta				
metformina				
páncreas				
alimentación				
ejercicio				
enfermedad				
alimentación saludable				
control				
complicación				
hipoglicemia				

Figura 8. Cuestionario para medir familiaridad léxica

5.3 Temas para investigaciones futuras.

Proyectamos generar nuevos listados para otras condiciones crónicas de salud, usando tests de disponibilidad léxica diferenciados, de modo de ampliar las funcionalidades del instrumento para evaluar la familiaridad de pacientes con distintas condiciones crónicas con el léxico necesario para comprender adecuadamente las explicaciones del personal de salud y seguir sus indicaciones, para alcanzar una mejor calidad de vida y evitar complicaciones.

A la luz de los datos recogidos, nos parece relevante realizar un análisis estadístico pormenorizado de los datos obtenidos de los test de disponibilidad léxica aplicados a los profesionales de la salud, que incorpore variables sociolingüísticas y dialectales, así como un análisis cualitativo detallado que dé cuenta de las relaciones asociativas que establecen los profesionales de salud entre las unidades léxicas mencionadas.

Además, estimamos de alta relevancia cotejar la disponibilidad léxica de profesionales de la salud y pacientes en este y otros centros de interés. Si bien es esperable que el léxico disponible de un individuo dé cuenta de su especialización, no podemos olvidar que una persona con una condición crónica de salud, adecuadamente informada, debería contar también con mayores antecedentes respecto de su condición que una persona no especialista que no presenta esa condición ni ha estado en contacto cercano con personas que la presenten. En esta misma línea, nos parece de gran relevancia estudiar posibles correlaciones entre disponibilidad léxica en el centro de interés relativo a una condición crónica de salud y distintas variables; entre ellas, el tiempo transcurrido desde el diagnóstico, la adherencia a tratamiento y el índice de compensación metabólica. En estos estudios comparativos, nos resulta también de especial importancia cotejar los procesos de salida y reentrada del núcleo entre distintos grupos y las relaciones que establecen los hablantes respecto de cada centro de interés, ya que pueden dar pistas no sólo de las asociaciones que realizan, sino también de su dominio del vocabulario del área en estudio.

Desde una perspectiva fonético acústica, consideramos de interés tomar muestras de audio de informantes no expertos para compararlas y contrastarlas con las del personal de salud y precisar de este modo cuáles son las características de las emisiones de quienes conocen las unidades léxicas respecto de los que no las conocen.

*Proyecto de Investigación INTERDISCIPLINA N° 1/2014 VRI UC. Agradecemos a la Vicerrectoría de Investigación (VRI) de la Pontificia Universidad Católica de Chile (UC) por su patrocinio. Agradecemos también: a las profesoras Claudia Bustamante y Claudia Alcayaga (Enfermería, UC), por permitirnos aplicar tests de disponibilidad léxica a expertos en el marco de una de sus clases; a la misma profesora Alcayaga y al Dr. Jaime Sapag (Medicina, UC), por aportar grabaciones de personal de salud leyendo el listado de vocablos relevantes, lo que facilitó la obtención de nuestro corpus oral; nuevamente al Dr. Jaime Sapag por sus aportes en la revisión del cuestionario con escala de Likert para evaluar familiaridad léxica (Figura 8).

6. REFERENCIAS

- Ávila, Antonio Manuel. *Elaboración, anotación y análisis del corpus oral del Proyecto V.U.M.* Málaga: Universidad de Málaga, Departamento de Filología Griega, Estudios Árabes y Traducción e Interpretación, Área de Lingüística General, 1998.
- Boersma, P. y D. Weenink. *Praat: A System for Doing P phonetics by Computer [Software]*. 2019. <http://www.fon.hum.uva.nl/praat/>
- Bosque, Ignacio. "Bases para un diccionario de restricciones léxicas". *Moenia* 7 (2001): 11-52.
- Bustamante, Claudia et al. "Literacidad en salud para personas con condiciones de salud crónicas". *Revista Horizontes de Enfermería* 2.19 (2008): 97-103.
- Camargo Serrano, Jaime. "Sistema de reconocimiento de voz humana por hardware" Diss. Universidad Pontificia Bolivariana, 2010. Bucaramanga: UPB, 2010. Print.
- Cantero, Francisco José. *Teoría y análisis de la entonación*. Barcelona: Edicions Universitat De Barcelona, 2002.
- Carrillo Aguilar, Roberto. "Diseño y manipulación de Modelos Ocultos de Markov, utilizando herramientas HTK. Una tutoría". *Ingeniare. Revista chilena de ingeniería* 15.1 (2007): 18-26.
- Castillo Fadic, María Natalia y Pilar Oplustil Gallegos. "Diseño inicial de un instrumento fonético-acústico para evaluar la familiaridad de los sujetos con unidades léxicas". *Lengua y habla* 23 (2019): 183-201.
- Castillo Fadic, María Natalia. "Corpus Básico del Español de Chile ©.» 2012.
- Castillo Fadic, María Natalia. *Léxico Básico del Español de Chile*. Liberalia Ediciones. Fondo del Libro y la Lectura, 2020, en prensa.
- DispoLex, Grupo de Investigación. *¿Qué es el Proyecto Panhispánico? Dispolex*. 2003-2013. 2013 de 05 de 17. <http://www.disporex.com/info/el-proyecto-panhispanico>
- Echeverría, Max, Paula Urzúa y Israel Figueroa. *DispoLex para Windows Versión 1.3.1 (programa computacional)*. s.f.
- Esparza Arellano, María Elena, y J. Benito Avalos Briseño. "Reconocimiento de voz." *Conciencia Tecnológica* 22 (2003) <https://www.redalyc.org/articulo.oa?id=94402206>
- Fernández Planas, Ana María. *Así se habla. Nociones fundamentales de fonética general y española*. Barcelona: Horsori, 2011

- Ferreira, Roberto y Max S. Echeverría. “Redes semánticas en el léxico disponible de inglés L1 e inglés LE”. *Onomázein* 21 (2010): 133-153.
- Gil, Juana. *Fonética para profesores de español: de la teoría a la práctica*. Madrid: Arco/Libros, 2007.
- Gili Gaya, Samuel. *Elementos de fonética general*. Madrid: Gredos, 1975.
- Kanj, Mayagah y Mitic Wayne. “Health literacy and health promotion: Definitions, concepts and examples in the Eastern Mediterranean region”. 2009. *7th Global Conference on Health Promotion, "Promoting Health and Development: Closing the Implementation Gap", Nairobi, Kenya, 26-30 October 2009*. 2019 de Enero. <http://www.who.int/healthpromotion/conferences/7gchp/Track1_Inner.pdf>.
- Lara, Luis Fernando. *Curso de lexicología*. México: El Colegio de México, 2006.
- Lerat, Pierre. *Las lenguas especializadas*. Barcelona: Ariel, 1997.
- López Chávez, Juan y Juan Carlos Strassburger Frías. “Un modelo para el cálculo del índice de disponibilidad léxica individual: enseñanza del español como lengua materna.” López Morales, Humberto (ed.). *Actas del II Seminario Internacional sobre Aportes de la Lingüística a la Enseñanza del Español como Lengua Materna*. Río Piedras: Editorial de la Universidad de Puerto Rico, 1991. 99-112.
- López Morales, Humberto. *Léxico disponible de Puerto Rico*. Madrid: Arco Libros, 1999.
- López Morales, Humberto. *Sociolingüística*. 3a edición aumentada. Madrid: Gredos, 2004.
- Lucero, Yalda et al. “Diabetes mellitus e hipertensión arterial en un consultorio urbano de Los Andes”. *Revista Chilena de Salud Pública* 8.2 (2004): 93-100.
- Martínez Celadrán, Eugenio y Ana María Fernández Planas (coords). *Atlas Multimèdia de la Prosòdia de l'Espai Romànic*. 2003-2018.
<http://stel.uib.edu/labfon/amper/cast/index_ampercat.html>
- Martínez Celadrán, Eugenio y Ana María Fernández Planas. *Manual De Fonética Española*. Barcelona: Ariel, 2007.
- Martínez Celadrán, Eugenio. *El sonido en la comunicación humana*. Barcelona: Octaedro, 2003.
- Martínez Celadrán, Eugenio. *Fonología general y española*. Barcelona: Teide, 1989.
- Ministerio de Salud. «Resultados, Encuesta Nacional de Salud 2016-2017. Primeros resultados.» noviembre de 2017. *Ministerio de Salud. Gobierno de Chile*. <https://www.minsal.cl/wp-content/uploads/2017/11/ENS-2016-17_PRIMEROS-RESULTADOS.pdf>.
- Ministerio de Salud. *Mes del Corazón 2017*. 4 de agosto de 2017. <<https://www.minsal.cl/mes-del-corazon-2017/>>.
- Moreno Sandoval, Antonio et al. “Inventario de frecuencias fonémicas y silábicas del castellano espontáneo y escrito”. 2006. *Actas de las IV Jornadas de Tecnologías del Habla*. <<http://www.llf.uam.es/ESP/Publicaciones/LLI-UAM-4JTH.pdf>>.
- Nath, Charlotte. “Literacy and diabetes self-management”. *American Journal of Nursing* 6.107 (2007): 43-49.
- Oropeza Rodríguez, José Luis. “Algoritmos y Métodos para el Reconocimiento de Voz en Español Mediante Sílabas”. *Computación y Sistemas* 9.3 (2006): 270-286.
- Real Academia Española y ASALE. *Diccionario de la Lengua Española (DLE 23.3)*. 2020. <<https://dle.rae.es>>.
- Santos-Díaz, Inmaculada Clotilde. *El léxico bilingüe del futuro profesorado. Análisis y pautas para estudios de disponibilidad*. Frankfurt: Peter Lang, 2021.

Valencia, Alba y Max Echeverría. *Disponibilidad léxica en estudiantes chilenos*. Santiago de Chile: Ediciones Universidad de Chile - Universidad de Concepción, 1999.

7. Anexo

```
# This script goes through sound and TextGrid files in a directory,
# opens each pair of Sound and TextGrid, calculates the duration and f0 values.
# Then it saves results to a text file.
#
# This script is distributed under the GNU General Public License.
# Copyright 4.7.2003 Mietta Lennes
# Modify for Josué Pino
# 27/09/2019

writeInfoLine: "Script started."

form Analyze from labeled segments in files
  comment Directory of sound files
  text sound_directory
  sentence Sound_file_extension .wav
  comment Directory of TextGrid files
  text textGrid_directory
  sentence TextGrid_file_extension .TextGrid
  comment Full path of the resulting text file:
  text resultfile
  comment Which tier do you want to analyze?
  integer Tier 1
  comment Pitch (ac) analysis parameters
  positive Time_step 0.01
  positive Minimum_pitch_(Hz) 75 (=default)
  positive Maximum_pitch_(Hz) 600 (=default)
endform

# Here, you make a listing of all the sound files in a directory.
# The example gets file names ending with ".wav" from D:\tmp\

Create Strings as file list... list 'sound_directory$'*'sound_file_extension$'
numberOfFiles = Get number of strings

# Check if the result file exists:
if fileReadable (resultfile$)
  pause The result file 'resultfile$' already exists! Do you want to overwrite it?
  filedelete 'resultfile$'
endif

# Write a row with column titles to the result file:
# (remember to edit this if you add or change the analyses!)
titleline$ = "Filename Segment_label start end duration
... f0mean f0sd f0max f0min"
```

```

... 'newline$"
fileappend "'resultfile$" 'titleline$'

appendInfoLine: "Running analyses: this might take a while..."

# Go through all the sound files, one by one:
for ifile to numberOfFiles
    filename$ = Get string... ifile

    # A sound file is opened from the listing:
    Read from file... 'sound_directory$'filename$'

    # Starting from here, you can add everything that should be
    # repeated for every sound file that was opened:
    soundname$ = selected$ ("Sound", 1)

    # to create a Pitch object
    select Sound 'soundname$'
    To Pitch (ac): time_step, minimum_pitch, 15, "no", 0.03, 0.45, 0.01, 0.35, 0.14,
    maximum_pitch

    # Open a TextGrid by the same name:
    gridfile$ = "'textGrid_directory$'soundname$'textGrid_file_extension$"
    if fileReadable (gridfile$)
        Read from file... 'gridfile$'

        # Find the tier number that has the label given in the form:
        numberOfIntervals = Get number of intervals... tier

        # Pass through all intervals in the selected tier:
        for interval to numberOfIntervals
            label$ = Get label of interval... tier interval
            # to include all interval labels
            if label$ == "control" or label$ == "dieta" or label$ == "azucar"
or label$ == "pancreas"
            ... or label$ == "enfermedad" or label$ == "glicemia" or label$ ==
"diabetes" or label$ == "nefropatia"
            ... or label$ == "colesterol" or label$ == "insulina" or label$ ==
"complicacion" or label$ == "ejercicio"
            ... or label$ == "hipoglicemia" or label$ == "metformina" or label$
== "medicamento" or label$ == "compensacion"
            ... or label$ == "autocuidado" or label$ == "actividad" or label$
== "hidrato" or label$ == "alimentacion"
            ... or label$ == "hemoglobina" or label$ == "alimentacions"

            # if the interval has an unempty label, get its start and
end:
            start = Get starting point... tier interval
            end = Get end point... tier interval

```

María Natalia Castillo Fadic y Josué David Pino Castillo. "Hacia la construcción de un instrumento para evaluar la familiaridad de pacientes crónicos con unidades léxicas relevantes para el automanejo de su condición de salud".

```

# get fixed temporal values
start$ = fixed$(start, 2)
end$ = fixed$(end, 2)

# get duration of each intervals
duration = (end-start)*1000
duration$ = fixed$(duration, 2)

# pitch
select Pitch 'soundname$'
f0mean = Get mean... start end Hertz
f0sd = Get standard deviation... start end Hertz
f0max = Get maximum... start end Hertz Parabolic
f0min = Get minimum... start end Hertz Parabolic

f0mean$ = fixed$(f0mean, 2)
f0sd$ = fixed$(f0sd, 2)
f0max$ = fixed$(f0max, 2)
f0min$ = fixed$(f0min, 2)

# Save result to text file:
resultline$ = "'soundname$' 'label$' 'start$' 'end$'
'duration$'
... 'f0mean$' 'f0sd$' 'f0max$' 'f0min$'
...'newline$'"
fileappend "'resultfile$'" 'resultline$'
select TextGrid 'soundname$'
endif
endfor
# Remove the TextGrid object from the object list
select TextGrid 'soundname$'
Remove
endif
# Remove the temporary objects from the object list
select Sound 'soundname$'
plus Pitch 'soundname$'
Remove
select Strings list
# and go on with the next sound file!
endifor

appendInfoLine: "Finished."

Remove

```