

Representación pictográfica del lenguaje Toki-Pona para su uso en sistemas aumentativos y alternativos de comunicación

Rigoberto Cerino¹, David Pinto¹, Sergio Vergara²

¹ Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación,
México

² Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Electrónica,
México

{cerino_rigoberto, svergara2}@hotmail.com, david.pinto@correo.buap.mx

Resumen. En este artículo se presenta un análisis de la asociación de pictogramas con un lenguaje de comunicación de mínima expresión. Este análisis tiene la finalidad de determinar si es posible construir un conjunto de datos que permita realizar expresiones en lenguaje natural basados en conceptos para su uso en sistemas aumentativos y alternativos de comunicación. Para llevar a cabo esta actividad, es necesario verificar el grado de vinculación entre el conjunto de pictogramas que expresan conceptos y el lenguaje de expresión mínima elegido, en este caso, Toki Pona. Se determina el grado de vinculación utilizando cuestionarios sobre una población objetivo, lo que permite validar si efectivamente existe una asociación directa entre el vocabulario y la gramática de Toki Pona, con respecto a los conceptos de los pictogramas. Este análisis abre la oportunidad de usar el recurso como un mecanismo de expresión cuando se emplean ciertos tipos de tecnologías, tales como los dispositivos de lectura electroencefalográfica (EEG), que podrían utilizarse para deducir si una persona está pensando en un determinado concepto.

Palabras clave. Toki Pona, pictograma, concepto.

Pictographic Representation of the Toki Pona Language for Use in Augmentative and Alternative Communication Systems

Abstract. This article presents an analysis of the association of pictograms with a minimal expression

communication language. This analysis has the purpose of determining if it is possible to build a data set that allows expressions in natural language based on concepts for use in augmentative and alternative communication systems. To carry out this activity, it is necessary to verify the degree of link between the set of pictograms that express concepts and the chosen minimal expression language, in this case, Toki Pona. The degree of linkage is determined using questionnaires on a target population, which makes it possible to validate whether there is indeed a direct association between the vocabulary and the grammar of Toki Pona, with respect to the concepts of the pictograms. This analysis opens up the opportunity to use the resource as an expression mechanism when using certain types of technologies, such as electroencephalographic (EEG) reading devices, which could be used to infer whether a person is thinking about a certain concept.

Keywords. Toki Pona, pictogram, concept.

1. Introducción

El proceso de comunicación involucra llevar a cabo el intercambio de información y significado entre dos o más personas, usando signos para transmitir y recibir un mensaje [1]. Los signos son, básicamente, palabras o elementos portadores de significado que se organizan en lenguajes, los cuales pueden ser verbales o no verbales. En el primer caso, lenguaje verbal, se utilizan las

Tabla 1. Uso del vocabulario español

| Diccionario RAE | Pasivo | Activo | 5%-7% de palabras |
|-----------------|-----------------|------------------------|----------------------------------------------|
| 93,000 lemas | 25,000 palabras | 10,000-15,000 palabras | Pasivo 1250 palabras. Activo 750 palabras |

palabras, que es la forma de expresión más completa para comunicarnos que existe. En el segundo caso, lenguaje no verbal, no se utiliza la palabra para comunicar, sino otros medios tales como los dibujos, iconos, números, fotos, sonidos. En general, los lenguajes se han creado para hacer llegar el mensaje a un receptor universal de la forma más rápida y clara posible.

El ser humano a través de su evolución ha implementado distintos mecanismos para su comunicación; dentro de ellos podemos mencionar al habla, el lenguaje de señas y mensajes a través de expresiones corporales. Esos mecanismos de comunicación permiten establecer un mensaje hacia el receptor, a fin de que éste obtenga de manera satisfactoria ese mensaje por parte del emisor.

Hoy en día, existe una cantidad importante de personas que tienen problemas al comunicarse. Los trastornos del habla y del lenguaje en numerosos casos han significado que muchas personas sufran de exclusión social, ya que se encuentran sin acceso a formas de comunicación clave, como escribir, leer y hablar [2], lo que es una gran desventaja si consideramos que las nuevas tecnologías de la información requieren de un mecanismo de comunicación efectivo.

1.1. Sistemas aumentativos y alternativos de comunicación

Los Sistemas Aumentativos y Alternativos de Comunicación (SAAC) son formas de expresión distintas al lenguaje hablado, y su objetivo principal es aumentar (aumentativos) y/o compensar (alternativos) las dificultades de comunicación y lenguaje de las personas con discapacidad. Todas las personas que por cualquier causa no han adquirido o han perdido un

nivel de habla suficiente para comunicarse de forma satisfactoria, deberían usar un SAAC.

Entre las principales causas de la falta del nivel suficiente del habla se encuentran las enfermedades neurológicas tales como la esclerosis lateral amiotrófica (ELA), la esclerosis múltiple (EM) o el párkinson, la parálisis cerebral (PC), los trastornos del espectro autista (TEA), la discapacidad intelectual, las distrofias musculares, los traumatismos cráneo-encefálicos, las afasias o las pluridiscapacidades de tipologías diversas, entre otras.

Los expertos recomiendan introducir SAAC a edades tempranas, tan pronto como se observan dificultades en el desarrollo del lenguaje oral, o poco tiempo después de que cualquier accidente o enfermedad provoque un deterioro en las habilidades de comunicación.

La Comunicación Aumentativa y Alternativa (CAA) incluye diversos sistemas de símbolos, tanto gráficos (fotografías, dibujos, pictogramas, palabras o letras) como gestuales (mímica, gestos o signos manuales) y, en el caso de los primeros, requiere también el uso de otros sistemas de apoyo. Los diversos sistemas de símbolos se adaptan a las necesidades de personas con edades y habilidades motrices, cognitivas y lingüísticas en diferentes niveles [3].

La comunicación aumentativa actúa como ayuda o complemento y surge para potenciar las habilidades comunicacionales de la persona, mientras que la comunicación alternativa es utilizada directamente pensando en opciones distintas a la del habla, es decir, sustituyen al lenguaje oral [4].

1.2. Los sistemas de símbolos

Los sistemas de símbolos para la CAA se dividen en gestuales y gráficos. En ambos casos existe una graduación que va desde sistemas muy sencillos que se adaptan a personas con deficiencias cognitivas y/o lingüísticas de diversa magnitud, hasta sistemas complejos que permiten niveles avanzados de lenguaje con signos (basado en signos manuales) o asistido (basado en signos gráficos).

En el caso de los símbolos gestuales, la gradación va desde el uso de mímica y gestos de uso común hasta el uso de signos manuales,



Fig. 1. Los glifos de Toki Pona

generalmente en el orden correspondiente al lenguaje hablado; es lo que se denomina lenguaje con signos o bimodal. Las lenguas de signos utilizadas por las personas con problemas auditivos no se consideran SAAC, ya que constituyen idiomas que se han desarrollado y se adquieren de forma natural, al igual que ocurre con el lenguaje hablado.

El uso de signos manuales requiere disponer de habilidades motrices suficientes, como puede ser el caso de personas con discapacidad intelectual o TEA.

Los símbolos gráficos abarcan desde sistemas muy sencillos basados en dibujos o fotografías, hasta sistemas progresivamente más complejos como los sistemas pictográficos o la ortografía tradicional (letras, palabras y frases). Gracias a los sistemas de apoyo para la comunicación y los diversos recursos para el acceso, los sistemas gráficos pueden ser utilizados por personas con movilidad reducida, incluso en casos de extrema gravedad. Es por esto que, además de ser usados

por personas con discapacidad intelectual o TEA, es posible que lo utilicen también personas con discapacidades motoras (PC, ELA, EM, etc.) [3].

1.3. Los sistemas pictográficos de comunicación

Los pictogramas son básicamente dibujos sencillos que pueden ser fotocopiados y que representan palabras y conceptos habituales en la comunicación cotidiana. Estos dibujos están asociados por categorías, utilizando un código de colores, y pueden ser adaptados al léxico y las necesidades [4]; son signos no vocales de tipo gráfico en los que hay alguna relación física y real con el concepto representado.

Los Símbolos Pictográficos para la Comunicación no vocal (SPC) fueron creados por Roxana Mayer Johnson en 1981. Alrededor de 5000 pictogramas en blanco y negro, y en imágenes a color, fueron recogidos en un libro con su correspondiente representación en diferentes idiomas y aspectos ortográficos.

Estos pictogramas fueron clasificados en las siguientes ocho categorías: verbos, descripciones alimentos, ocio, nombres, cosas misceláneas, personas y lo social. Actualmente, es uno de los sistemas pictográficos más utilizados en el mundo gracias a su “alta definición visual” con respecto a las imágenes y a un software denominado “Boardmaker” que permite crear pictogramas a través de mensajes usados por medio de una computadora [5].

Los sistemas pictográficos de comunicación representan todo tipo de conceptos, acciones (individuales o grupales) u objetos en forma de dibujos, que implementan el lenguaje visual para verse reflejado a través de imágenes muy simples que permiten mayor entendimiento de lo que se expone ante ellos.

Existen múltiples herramientas en la literatura que usan SPC y que permiten mejorar algunas de las problemáticas analizadas, para beneficiar a las personas a poder expresar sus deseos, sus sentimientos y a satisfacer de esta manera sus necesidades. Esto conlleva a que la persona de forma gradual recupere parte de su autonomía y tenga una mejor inclusión social.

En [6], por ejemplo, se presenta una plataforma mediante la cual una persona normal se puede

comunicar con personas que tienen problemas de desarrollo intelectual en su proceso de comunicación, a través de la traducción de texto a pictogramas.

En [7], por otro lado, se diseñó un corpus basado en pictogramas que permite generar un modelo estadístico de traducción automática que brinda la posibilidad de generar una oración textual a partir de una secuencia de pictogramas con la finalidad de apoyar a personas con problemas de desarrollo intelectual en su proceso de comunicación.

Estos sistemas de apoyo a personas con cierta discapacidad se han logrado encapsular incluso en un teléfono móvil. EC+ [1], por ejemplo, es una aplicación gratuita cuya finalidad es incrementar la competencia comunicativa funcional en personas con discapacidad intelectual.

Esta aplicación presenta los presupuestos fundamentales de la comunicación aumentativa y alternativa especialmente destinada a personas con discapacidad intelectual, y detallan la aportación de los sistemas de comunicación alternativa para estos usuarios.

Además, ofrecen indicaciones para la toma de decisiones sobre la elección de dispositivos para la comunicación.

1.4. La lengua española

El Diccionario de la Real Academia Española está constituido por cerca de 93,000 lemas. No obstante, este número no considera a todas las variantes morfológicas de las palabras, lo que podría incrementar el vocabulario en aproximadamente un 33% más.

Al intentar determinar el tamaño del vocabulario de la lengua española, es importante darse cuenta de que el idioma español contiene un gran número de palabras de baja frecuencia de uso u obsoletas, y que normalmente no se encuentran en el vocabulario de un hablante promedio. Incluso aquellas personas que hablan español con fluidez, utilizan un porcentaje relativamente pequeño del vocabulario en comparación con el total de palabras del español.

En promedio, un hablante nativo del español tiene un vocabulario activo (es el que usamos al hablar o escribir) de 10,000 o 15,000 palabras. Este vocabulario incluye solamente las palabras

Tabla 2. Número de pictogramas por categoría PoS

| Categoría PoS | No. de Pictogramas |
|---------------|--------------------|
| Adjetivo | 1,780 |
| Adverbio | 254 |
| Artículo | 12 |
| Conjunción | 36 |
| Determinante | 31 |
| Número | 226 |
| Preposición | 44 |
| Pronombre | 137 |
| Sustantivo | 8,759 |
| Verbo | 2,456 |

que utilizan al hablar, ya que el vocabulario pasivo (las palabras que reconocen y entienden, pero no usan) seguramente asciende a unas 25,000 palabras.

Algunos expertos en lingüística afirman que con manejar entre el 5% y el 7% de las palabras que conoce un hablante nativo se puede ser capaz de entender del 90% al 95% de las palabras que se utilizan en un texto estándar. Esta cifra del 95% es muy significativa, ya que es el umbral a partir del cual suele ser posible comprender el mensaje perfectamente si a nuestro dominio del idioma le sumamos el conocimiento del contexto.

Este hecho es muy importante, puesto que, si se extrapolan estos datos a la lengua española, se puede decir que con un vocabulario pasivo de 1,250 palabras y un vocabulario activo de tan solo 750 palabras sería suficiente para sobrellevar la mayoría de situaciones del día a día [8].

Un vocabulario activo de 750 palabras podría resultar atractivo al tomar en cuenta el desarrollo computacional de algún proceso alternativo de comunicación humana, sin embargo, al llevar a cabo tareas que requieren la vinculación de las palabras con sus correspondientes representaciones pictográficas, el escenario ya no es tan favorable.

Consideremos, por ejemplo, la identificación de conceptos a través de señales electroencefalográficas (EEG); en este caso, se requiere estimular visualmente a las personas usando los pictogramas, con el objetivo de deducir el concepto que se encuentra en la mente de la persona. Presentar 750 o más pictogramas al usuario y realizar la captura y análisis de señales

Tabla 3. Ambigüedad en los pictogramas

| Palabra | Acepciones | PoS |
|----------|------------|-------------------------|
| Adjetivo | 2 | Verbo |
| Banco | 3 | Sustantivo |
| Carta | 1 | Sustantivo |
| Copa | 2 | Sustantivo |
| Gato | 2 | Sustantivo |
| Lima | 2 | Sustantivo |
| Lista | 2 | Adjetivo, Sustantivo |
| Pila | 2 | Sustantivo |
| Planta | 2 | Sustantivo |
| Radio | 4 | Sustantivo |
| Ruta | 2 | Sustantivo |

| | | |
|-----------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| ilo: del Esperanto <i>ilo</i> | <ul style="list-style-type: none"> sust; Herramienta | |
| insa: del Tok Pisin <i>insait</i> ; del inglés <i>inside</i> | <ul style="list-style-type: none"> sust; centro, interior, vientre adj; contenido, interno | |
| jaki: del inglés <i>yucky</i> | <ul style="list-style-type: none"> adj; sucio, nauseabundo sust; mugre, contaminación, polvo verbo tr; ensuciar, manchar interj; expresa desagrado, guacala! | |
| jan: del chino cantónes 人 <i>/jɔn21/</i> | <ul style="list-style-type: none"> sust; hombre, ser, persona, alguien adj; humano, de alguien, personal verbo tr; personificar, encarnar | |

Fig. 2. Conjunto de pictogramas del portal de ARASAAC vinculados al lenguaje Toki Pona

EEG es un trabajo agotador e incómodo para la población seleccionada, lo que requiere de mucho tiempo, concentración, estados de relajación y entrenamiento para los usuarios.

En este sentido, es necesario pensar en un vocabulario sumamente reducido, un lenguaje que se caracterice por ser de mínima expresión. De esta manera, se tendrá por consecuencia un conjunto reducido de conceptos básicos y a su vez, un número pequeño de pictogramas, lo que

facilitará las tareas de identificación de características conceptuales.

Al analizar la literatura disponible sobre el lenguaje humano, se encontró un lenguaje artificial llamado Toki Pona que permite llevar a cabo el proceso de comunicación usando tan solo 123 palabras.

A continuación, se describe a detalle la el origen y la naturaleza de este lenguaje de mínima expresión.

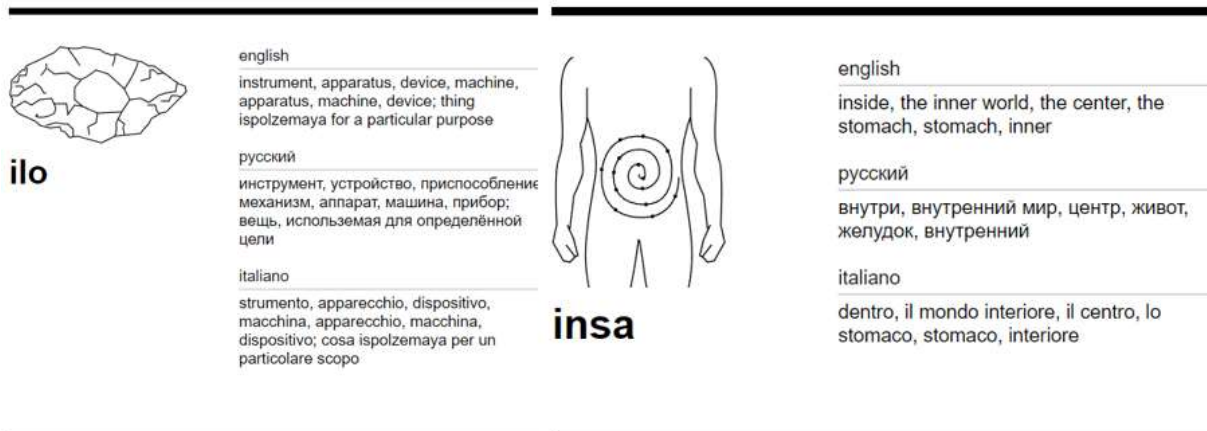


Fig. 3. Diccionario Toki Pona con imágenes



Fig. 4. Ejemplo de evaluación de conceptos basado en pictogramas (agua y número 2)

2. Lenguaje Toki Pona

Toki Pona es un lenguaje creado por la lingüista y traductora canadiense Sonja Lang como un lenguaje filosófico, con el propósito de simplificar los pensamientos y la comunicación.

Se publicó por primera vez en línea en 2001 como borrador, y posteriormente en forma completa en su obra denominada: "Toki Pona: El lenguaje del bien" en 2014. Una pequeña comunidad de hablantes de Toki Pona se desarrolló a principios de la década del 2000. La comunidad creció después del lanzamiento del libro oficial y ha seguido incrementándose desde su publicación.

Si bien la actividad se realiza principalmente en línea en salas de chat, en las redes sociales y en otros grupos, ha habido algunas reuniones organizadas en persona durante la década de

2010. Lo que caracteriza al lenguaje de Toki Pona es el minimalismo.

Se centra en conceptos universales simples, haciendo uso de muy poco para expresar al máximo. El idioma tiene 123 palabras raíz y 14 fonemas que son fáciles de pronunciar en diferentes idiomas.

Aunque inicialmente no se pensó como un idioma auxiliar internacional, puede funcionar como uno.

Este lenguaje está diseñado para ayudar a los usuarios a concentrarse en cosas básicas y promover el pensamiento positivo, de acuerdo con la hipótesis de Sapir-Whorf. A pesar de contener un vocabulario reducido, los hablantes pueden entenderse y comunicarse entre sí, basándose principalmente en el contexto y las combinaciones de varias palabras para expresar significados más específicos [12].

Tabla 4. Niveles máximos de vinculación de los conceptos del diccionario Toki Pona y los pictogramas

| Concepto | % de vinculación |
|--------------|------------------|
| Herramientas | 98.9 % |
| Casa | 97.4 % |
| Recortar | 93.3 % |
| Hojas | 92.8 % |

Tabla 5. Niveles mínimos de vinculación de los conceptos del diccionario Toki Pona y los pictogramas

| Concepto | % de vinculación |
|----------|------------------|
| Reunión | 8.3 % |
| Cantidad | 3.1 % |
| Final | 2.6 % |
| Esto | 2 % |

Tabla 6. Número de pictogramas con mejores niveles de vinculación y confiabilidad

| % de vinculación | No. De pictogramas |
|------------------|--------------------|
| >= 50 % | 96 |
| >= 60 % | 80 |
| >= 70 % | 56 |
| >= 80 % | 34 |
| >= 90 % | 11 |

Como se ha mencionado, el diccionario de Toki Pona dispone de solo 123 palabras para comunicarse. El vocabulario está compuesto de verbos, sustantivos o adjetivos, y en función de la posición que tome la palabra en la frase, se dará su significado.

Al ser un lenguaje tan compacto, algunas de las palabras incluidas contienen términos ambiguos, con múltiples significados.

En la literatura, existe un sistema vinculado con Toki Pona, denominado Sitelen Pona, el cual es un sistema ideogramático universal de escritura que genera la representación de ideas, frases o palabras por medio de ideogramas, es decir, símbolos o signos, y que toma como punto de partida el diccionario Toki Pona.

Sitelen Pona es un sistema logográfico, se usan grafemas como unidades mínimas de un sistema de escritura que por sí solas representan una palabra, lexema o morfema, es decir, una unidad con significado. Esto contrasta con sistemas de representación no estrictamente

lingüísticos como los pictogramas o ideogramas, y con sistemas fonográficos como los alfabetos.

Un gráfico de Sitelen Pona no remite a sonido alguno, no es posible en esta lengua utilizar el valor fonético de los ideogramas para crear la oralización de una palabra que no representa a un ideograma, como es el caso en las escrituras jeroglíficas (egipcias, mayas). El autor [13] señala que el vocabulario y la gramática de Sitelen Pona se pueden aprender fácilmente en ocho días. Con el objetivo de ser lo más rápido y lo más fácilmente comprensible, Sitelen Pona toma prestado de todo lo que para nosotros tiene hoy sentido gráfico, lo más universalmente posible.

En la Figura 1 se presentan ejemplos de los elementos de este sistema logográfico. Como puede observarse, se utilizan colectivamente pictogramas, paneles de señalización vial, símbolos matemáticos, emoticonos y estilizaciones fáciles de reproducir.

En [14] se presenta un diccionario Toki Pona en imágenes, el cual cuenta con las traducciones de las 123 palabras de Toki Pona al inglés, italiano y ruso. El objetivo del autor fue conocer la complejidad o simplicidad del aprendizaje de otro idioma usando como puente las imágenes y el famoso lenguaje sintáctico.

Como una restricción adicional, optó por no utilizar ningún símbolo o signo de la cultura moderna en la medida de lo posible, sino con sus propias imágenes, ya que considera que éstas ayudan a recordar el vocabulario.

Es importante mencionar, sin embargo, que sus imágenes no están validadas o estandarizadas en la literatura.

Por esta razón, es necesario analizar el conjunto de pictogramas adecuados para vincularlos con las palabras de éste lenguaje de mínima expresión.

3. Pictogramas

Ante la gran variedad de herramientas alternativas desarrolladas para el apoyo de la comunicación humana que implementan distintas bases de datos de pictogramas, es natural que surjan los siguientes cuestionamientos: ¿Qué conjunto de pictogramas, o signos, de todos los disponibles, es el mejor para usar en los procesos

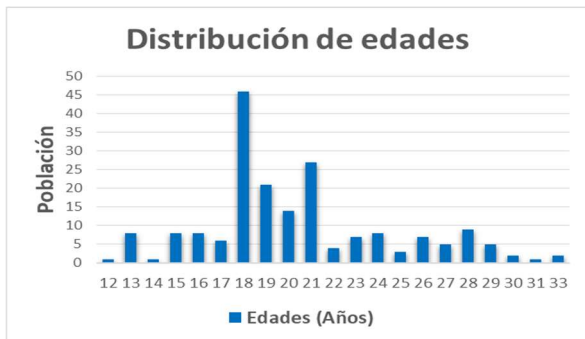


Fig. 5. Distribución de edades de la población



Fig. 6. Pictograma de "casa"



Fig. 7. Pictograma de "comer"



Fig. 8. Porcentaje de vinculación obtenido al clasificar las respuestas de los usuarios para el pictograma de "iniciar", en el que se presenta demasiada ambigüedad en el significado

alternativos de comunicación?, así como también, ¿Cuáles son los criterios para elegir entre estas opciones disponibles?

La característica de los pictogramas que ha recibido más atención en este sentido es la iconicidad, la cual puede ser entendida como "la asociación que el individuo hace entre un símbolo y su referente". La iconicidad se ha venido a conceptualizar como un continuo en el que se distingue entre símbolos cuyo significado es sugerido fácilmente (también denominados transparentes), y símbolos abstractos que tienen un significado difícil de entender (también llamados opacos) [9].

En [9] se demuestra que el conjunto de pictogramas de ARASAAC es más transparente en pruebas a distintas poblaciones en

comparación a los conjuntos de pictogramas SPC y Bliss, y por lo tanto su significado es fácil de adivinar.

La base de datos orientada al proceso de comunicación humana del portal de ARASAAC ofrece recursos gráficos y materiales para facilitar la comunicación de aquellas personas con algún tipo de dificultad en esta área. El catálogo de pictogramas contiene 16,801 pictogramas a color y 15,026 en blanco y negro, con tamaño de 500x500 píxeles cada uno y con formato png [3].

En [6], el autor realiza una clasificación de los pictogramas considerando la etiqueta de partes de la oración (Part Of Speech o PoS por sus siglas en inglés) del pictograma (ver Tabla 2). Teniendo a disposición un volumen tan extenso y representativo del lenguaje mediante los

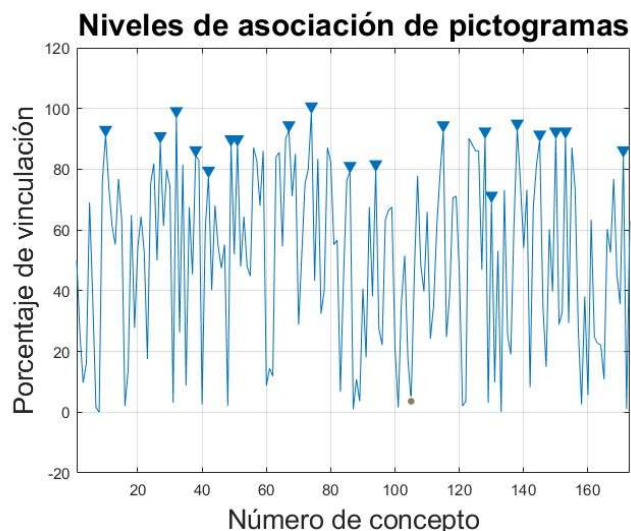


Fig. 9. Distribución del porcentaje de vinculación que generan las respuestas de todos los usuarios ante los pictogramas

(174/174) ¿Qué tan complicado le resultó responder el cuestionario anterior basado en pictogramas?
160 respuestas

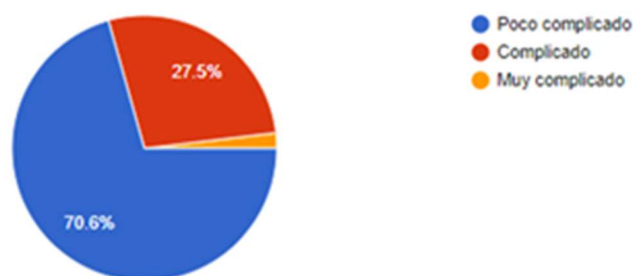


Fig. 10. Nivel de dificultad que representó el cuestionario para la población

pictogramas de ARASAAC, se procedió a vincularlos con el vocabulario de Toki Pona.

4. Análisis de los pictogramas ARASAAC y su vinculación con Toki-Pona

El ser humano siempre ha usado sus sentidos para tener una mejor comunicación con su entorno, y suele recibir información de todos lados de manera inconsciente y automática.

La percepción de la información que nos rodea proviene en un 20% de sentidos como el olfato, tacto, oído, mientras que el 80% proviene directamente de la vista.

En consecuencia, podemos afirmar que la comunicación visual es vital en nuestra vida cotidiana [10, 11], se menciona que el descubrimiento de que los conceptos podían materializarse gráficamente ha motivado al intelecto induciéndolo al logro de sistemas visuales conformados por símbolos y signos.

Los símbolos adquieren formas propias con un significado particular. Con el objetivo de vincular

las palabras que conforman el vocabulario de Toki Pona con un subconjunto de pictogramas de ARASAAC, en este artículo de investigación, inicialmente se realizó una búsqueda de pictogramas ambiguos. Se realizó un análisis profundo y exhaustivo de todos aquellos pictogramas que pudieran ser ambiguos.

En este sentido, se creó una tabla que contiene un índice para cada palabra, una palabra polisémica y sus correspondientes etiquetas PoS (ver Tabla 3).

A continuación, se separaron los pictogramas que refieren a conceptos de aquellos que hacen referencia a palabras individuales. Un pictograma puede representar una sola palabra (por ejemplo: casa, carro, libro, etc), pero otros pictogramas pueden representar una serie de palabras (por ejemplo: jugar al golf, cama de gato).

Se realizó un filtrado de palabras y conceptos dividiéndolo en dos tablas y se eliminaron los pictogramas repetidos, localizando un total de 2,817 conceptos.

Tomando en cuenta la investigación presentada y la reducción de pictogramas del portal de ARASAAC en conceptos, es necesario realizar la recopilación y reducción de pictogramas para evaluar cuáles son los que se pueden utilizar y que sean correspondientes al contenido del lenguaje Toki Pona; esto con la finalidad de conocer el grado de complejidad para una determinada población.

Se seleccionaron 173 pictogramas para las 123 palabras disponibles en el lenguaje Toki Pona. A ciertas palabras del vocabulario se le asociaron más de un pictograma, con la finalidad de abarcar la mayor cantidad de significados posibles.

En la Figura 2 se presenta un ejemplo de la recopilación de la base de datos de pictogramas asociada al lenguaje Toki Pona.

Se encontraron algunas palabras del diccionario Toki Pona cuyo significado es muy ambiguo, o la descripción de este no se puede expresar a través de un solo pictograma. Por ejemplo, expresiones como “hey” para conseguir atención, o sustantivos como “superficie”, que representan una excepción en su representación visual, ya que no se encontraron disponibles en el portal de ARASAAC.

5. Evaluación

Se diseñó un formulario con el objetivo de conocer qué tan genéricos son los pictogramas, sometiéndolo a evaluación en una población objetivo, con la finalidad de conocer el grado de vinculación que existe entre el conjunto reducido de conceptos básicos presentados mediante los pictogramas y el lenguaje Toki Pona.

Este cuestionario se aplica para conocer el grado de consistencia de una expresión pictográfica y su significado esperado. En este cuestionario se solicitan los datos del usuario como su nombre (opcional para el evaluador) y su edad, para tener una relación de las respuestas y asociarlas a estos datos.

Posteriormente se comienza con la presentación de los pictogramas, solicitando a los evaluadores escribir el primer concepto que le llegue a la mente cuando observa el pictograma¹.

El formulario fue enviado a una población total de 194 personas con un rango de edades de 12 a 33 años (ver Fig. 5) Se les solicitó escribir lo primero que pensarán al momento de observar cada uno de los 173 pictogramas.

De esta forma se obtuvieron los datos para determinar un grado de vinculación entre el conjunto de pictogramas que representan los conceptos, y el lenguaje Toki Pona.

6. Resultados y discusión

Las respuestas recibidas fueron satisfactorias, en el sentido de que la mayoría de la población pensó en el mismo concepto al observar el pictograma. Al realizar una clasificación de las respuestas se obtiene un alto porcentaje de similitud con respecto al significado original del pictograma, alcanzando valores del 75% de acuerdo (ver Figuras 6 y 7).

Existen pictogramas muy ambiguos que generan confusión, en los cuales las personas escribieron respuestas con una distribución amplia de significados distintos, alcanzando únicamente el 20% de acuerdo (ver Fig. 8).

El análisis completo de las respuestas proporcionadas por los 194 usuarios o

¹ <https://forms.gle/KokFMw6vPC7rodsx8>

evaluadores para cada uno de los 173 pictogramas que se les presentaron, generó resultados variables entre los conceptos, ya que hay pictogramas que generan confusión o no se adaptan adecuadamente al significado real.

En la Fig. 9 se muestra una distribución del grado de vinculación de todos los pictogramas del conjunto de datos propuesto, donde existen pictogramas muy generales ante una población, es decir, transmiten mayormente el significado de este y, asimismo, pictogramas que no son interpretables de manera óptima para la sociedad.

En la tabla 4 se presentan los pictogramas que arrojaron mayor porcentaje de vinculación con su significado en el diccionario Toki Pona, en donde se realizó el ajuste de las respuestas uniendo los sinónimos, por lo que se alcanzaron porcentajes de vinculación mayores.

En la tabla 5 se presentan los pictogramas que generaron el menor grado de vinculación entre el lenguaje y los pictogramas.

En la tabla 6 se despliega la cantidad de pictogramas que superan cierto grado de confiabilidad en la vinculación con el diccionario Toki Pona. Y para porcentajes menores al 50% de vinculación, este conjunto de datos tiene 75 pictogramas.

El resultado del estudio indica por una parte que existen ciertos elementos que presentan un alto grado de ambigüedad, lo cual representa un reto importante en el análisis automático del lenguaje natural.

Es un resultado bastante importante en el cual es necesario realizar un proceso de investigación posterior donde se evalúen distintas combinaciones de pictogramas, por ejemplo, agregando características visuales a estos, con el objetivo de analizar si generan una mayor comprensión a la percepción humana, enfatizando en la amplitud del intercambio comunicativo y resaltando la importancia de llegar a alcanzar un significado compartido o comunitario.

Precisamente este fue el objetivo de la encuesta, lograr identificar y conocer la generalidad de un conjunto reducido de pictogramas y su grado de entendimiento para la población, con la finalidad de observar el grado de vinculación con un lenguaje de mínima expresión.

Debido a la reducción drástica del vocabulario humano a través del lenguaje Toki Pona y

posterior a la encuesta para conocer el grado de dificultad de la comprensión de los pictogramas, se obtuvo un conjunto reducido en cuanto al número de conceptos genéricos, resultando en aproximadamente 75 conceptos no ambiguos para la sociedad.

Al final de la encuesta se realizó una última pregunta con el propósito de conocer si este formulario había presentado dificultad para responderlo. En la Figura 9 se aprecia el grado de dificultad reportado por la población objetivo. En este caso, el 70% de los encuestados declaró que fue un cuestionario de baja dificultad para responder.

7. Conclusiones

Los datos actuales muestran que, si es posible generar una vinculación entre un conjunto de pictogramas ARASAAC asociados a conceptos y el lenguaje de mínima expresión denominado Toki Pona, y que al mismo tiempo estos conjuntos pueden implementarse en procesos de comunicación humana.

Este tipo de estudios que tienen por objeto aumentar la comunicación entre los seres humanos pueden tener aplicaciones en la vida real muy concretas, por ejemplo, como apoyo a personas que perdieron el habla, personas que no escuchan, personas que han tenido un derrame cerebral y se necesita verificar si la identificación de los conceptos aún se encuentran presentes en su mente, ya que es posible convertirlas en un medio de comunicación que facilite o permita a los usuarios una adaptación real con el medio exterior, así como su aplicabilidad a sistemas robóticos, automatizados, que se convierten en sistemas de asistencia o medios de comunicación para las personas.

Referencias

1. **Calleja, M., Rodríguez-Santos, J. (2018).** La comunicación aumentativa y alternativa para hacer frente a las necesidades complejas de comunicación en usuarios de bajo perfil cognitivo. *Disability and Communication Scientific Analysis, Total Communication, Ict*

- Tools and Case Studies McGraw-Hill Education Capitulo 12, pp. 163–176.
2. **Bautista, S., Hervás, R., Hernández, A., Martínez, C., Pascua, S., Gervás, P. (2017).** AraTraductor: Text to pictogram translation using natural language processing techniques. *Interacción '17: Proceedings of the XVIII International Conference on Human Computer Interaction*, No. 28, pp. 1–8. DOI: 10.1145/3123818.3123825.
 3. **ARASAAC. (2014).** ¿Qué son los sistemas aumentativos y alternativos de comunicación (SAAC)? Portal Aragonés de la Comunicación Aumentativa y Alternativa. Available: <http://www.a-rasaac.org/aac.php#1>.
 4. **Marín, A. (2019).** Sistemas alternativos y aumentativos de comunicación. *Publicaciones Didacticas*, No. 104, pp. 235–240.
 5. **Regis-Sansalonis, P. J., Callejón-Chinchilla, M. D. (2016).** Del pictograma a la imagen: herramientas de comunicación y lenguaje en personas con síndrome de Asperger a través de recursos visuales para la inclusión social. *Arteterapia - Papeles de arteterapia y educación artística para la inclusión social*, Vol. 10, pp. 329–341. DOI: 10.5209/rev_ARTE.2015.v10.51700.
 6. **Olivares, G. (2019).** Traducción automática de texto a pictogramas. Tesis maestría, Facultad de Ciencias de la Computacion, BUAP, Puebla, México.
 7. **Vázquez, F. K. (2019).** Traducción automática de pictogramas a texto. Tesis maestría, Facultad de Ciencias de la Computacion, BUAP, Puebla, México.
 8. **Lingoda. (2019).** ¿Cuántas palabras de vocabulario español hay que aprender?. Available: <https://www.lingoda.com/es/content/vocabulario-espanol/>.
 9. **Lang, S. (2014).** Toki Pona: The language of good. Available: https://en.wikipedia.org/wiki/Toki_Pona.
 10. **Garnerin, X. (2016).** Textbook of Sitelen Pona.
 11. **Shapiro, A., Shkitov, M. (2021).** Toki Pona Visual Dictionary. Available: <https://github.com/x-ra-iz-or/vi-sual-tokipona>.
 12. **Bertola, L. E. (2017).** Análisis empírico de las características formales de los símbolos pictográficos ARASAAC. Tesis doctoral, Departamento de Psicología Evolutiva y de la Educación de la Universidad de Murcia, España.
 13. **Cámara-Montoya, K., Saldaña-Hernández, J. C. (2018).** Diseño de signos pictográficos para un proyecto de índole comercial, social o educativo. *Jóvenes en la ciencia*, Vol. 4, No. 1, pp. 2114–2119.
 14. **Cedeño, L., Coral, M., Andrade, V., Molina, P. (2018).** Lenguaje ideográfico y pictográfico (LIP) como sistema de comunicación para niños con trastorno del espectro autista. *Revista de la Universidad Cubana de Diseño*, No. 9, pp. 93–108.

*Article received on 14/09/2022; accepted on 06/12/2022.
Corresponding author is David Pinto.*