

Diseño y Evaluación de una Experiencia Conversacional con un Asistente Virtual para Usuarios Sordos

Design and Evaluation of a Conversational Experience with a Virtual Assistant for Deaf Users

Ana Julia de Oliveira

Departamento de diseño

Universidade Federal de Pelotas

Pelotas, Rio Grande do Sul, Brazil

anajdeoliveira@gmail.com

Tobias Mulling

Departamento de diseño

Universidade Federal de Pelotas

Pelotas, Rio Grande do Sul, Brazil

tobias.mulling@ufpel.edu.br

Recibido: 15.10.2023 | Aceptado: 01.12.2023

DOI: <https://doi.org/10.65234/interaccion.97>

Palabras Clave

Interfaz conversacional

Interacción gestual

Experiencia del usuario

Accesibilidad digital

Asistente virtual

Resumen

El uso de asistentes virtuales ha crecido exponencialmente, pero a pesar de la facilidad que trajo la interacción de voz, esta tecnología no es accesible para todos al no incluir a las personas sordas que utilizan la lengua de signos, especialmente la Lengua Brasileña de Señas (Libras). Este trabajo tiene como objetivo diseñar una experiencia conversacional entre un asistente virtual y usuarios sordos, donde la entrada de datos se producirá a través de gestos en el aire y retroalimentación realizada mediante el uso de una interfaz gráfica (GUI). La intención es brindar al usuario sordo que utiliza Libras una interacción más natural, teniendo en cuenta el uso de su lengua materna para establecer la comunicación, ofreciendo a las personas sordas mayor independencia, acceso a información y servicios.

Keywords

Conversational interface

Gestural interaction

User experience

Digital accessibility

Virtual assistant

Abstract

The use of virtual assistants has grown exponentially, but despite the facility that voice interaction brought, this technology is not accessible to all, not including deaf people who use sign language, especially the Brazilian Sign Language (Libras). This work aims to design a conversational experience between a virtual assistant and deaf users, where data input will occur through mid-air gestures and feedback performed by the use of a graphical interface (GUI). The intention is to provide the deaf user who uses Libras a more natural interaction, taking into account the use of their native language to establish communication, offering deaf people greater independence, access to information and services. Future work should include usability evaluation with the target audience.

1. Introducción

Los asistentes virtuales son cada vez más populares. Mediante comandos de voz y altavoces, realizan tareas y proporcionan información, ayudando a los usuarios en actividades diarias y aumentando potencialmente su productividad. Los asistentes virtuales son programas capaces de realizar acciones interpretadas mediante inteligencia artificial. Algunos ejemplos

son Siri (Apple), Alexa (Amazon), Google Assistant (Google) y Cortana (Windows), entre otros.

A pesar de la facilidad de uso que el comando y respuesta por voz ha traído, esta tecnología no atiende a todos, ya que no cubre a las personas sordas que utilizan la Lengua Brasileña de Señas (Libras) para comunicarse. Según la Encuesta Nacional de Salud (Pesquisa Nacional de Saúde, en el original) - PNS (IBGE, 2019), entre las personas de 5 a 40 años que

presentaban mucha dificultad para oír, el 22,4% utilizaba Libras. También según la encuesta, entre las personas del mismo grupo de edad que no oían en absoluto, el porcentaje era del 61,3%. Teniendo esto en cuenta, las personas con pérdida de audición severa no podrían beneficiarse de la herramienta, ya que no podrían recibir la respuesta de la interfaz a través del sentido del oído, y las personas sordas que utilicen la lengua de signos no podrían interactuar a través del habla. Al fin y al cabo, "la accesibilidad de un producto consiste en considerar la diversidad de sus posibles usuarios y las peculiaridades de su interacción con el producto" (TORRES; MAZZONI, 2004, p.152).

Una interfaz gestual puede permitir una comunicación más sencilla y natural a los usuarios sordos cuya lengua materna es la Libras, haciendo posible la interacción mediante la lengua de signos. Dado que ya existen tecnologías que permiten interactuar mediante gestos, este estudio es factible.

Las interfaces de signos pueden interpretar una amplia gama de gestos, lo que, junto con las tecnologías existentes y otros recursos de accesibilidad, puede ayudar a la comunidad sorda. De esta forma, esta investigación tiene como objetivo comprender las características de la estructura de Libras y la interacción conversacional, con el fin de proponer una experiencia conversacional con un asistente virtual a través de la interacción gestual - donde los gestos se utilizan como elementos de entrada de datos, combinados con otros elementos gráficos en el proceso de salida de datos (pantalla) - uniendo la tecnología a la accesibilidad y buscando satisfacer las necesidades de esta gama de usuarios. En este sentido, este proyecto no se caracteriza necesariamente como un dispositivo o recurso de accesibilidad, sino como una propuesta exploratoria que podría aplicarse a un asistente virtual ya existente (añadiendo hardware para el mapeo de gestos) o a un asistente virtual específico diseñado para este fin.

Aunque la comunidad sorda es diversa y está formada por personas que utilizan audífonos, personas que leen los labios y personas que hablan, por ejemplo, el enfoque de este proyecto se centra en las personas sordas que utilizan Libras, y no abarca completamente el diverso grupo de personas que existen dentro de la comunidad sorda.

2. Referencial teórico

2.1 UX para asistentes virtuales

Para definir una interacción conversacional, es necesario entender cómo interpretan el lenguaje los asistentes virtuales. El portugués tiene verbos, sustantivos y otras clases gramaticales, y este lenguaje necesita ser reinterpretado para que pueda ser entendido por el asistente. La interacción conversacional comienza cuando el usuario utiliza una wakeword, es decir, una palabra para "despertar" al asistente de modo que el dispositivo entienda que el usuario está listo para comunicarse. La figura 1 muestra la estructura de una

conversación con Alexa, donde se puede ver parte de la terminología utilizada en el desarrollo de skills para Alexa. Las skills son aplicaciones que añaden nuevas funcionalidades a los dispositivos conectados a Alexa.

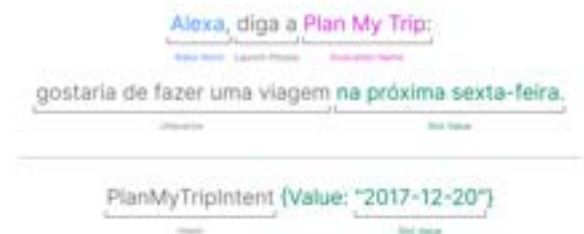


Figura 1: Estructura de la interacción conversacional. Fuente: Autores.

Es esencial comprender la estructura de una interacción conversacional, ya que el software del asistente virtual traduce el lenguaje humano en su forma oral al lenguaje de la máquina.

- Palabra de lanzamiento (launch phrase): actúa como palabra de conexión entre la wakeword y el nombre de apertura. Suelen ser verbos como: decir, preguntar, abrir;
- Nombre de apertura (invocation name): nombre que dirán los usuarios cuando quieran interactuar con una skill concreta. En el ejemplo anterior, el nombre de apertura es "Plan My Trip";
- Enunciado (utterance): cómo se formula la petición, indica la intención detrás del discurso del usuario. En el ejemplo: "Me gustaría ir de viaje";
- Slot: variable que proporciona a Alexa más información sobre la petición del usuario. En este caso, "el próximo viernes" es la variable de fecha;
- Función (intent): interpretación del enunciado para la acción que se va a realizar, atendiendo a la petición del usuario.

La comprensión de estas terminologías permite estructurar interacciones funcionales con el asistente. Esta investigación pretende comprender cómo la interacción conversacional de las personas sordas -mediada por signos - puede ser asimilada por un asistente virtual. En este sentido, si la interpretación de los datos de entrada es correcta, la función (intent) será entonces ejecutada y activada por el asistente virtual.

2.2 Transposición de la interacción de signos en Libras a un asistente virtual

Saffer sugiere que un gesto es "cualquier movimiento físico que un sistema digital pueda detectar y al que pueda responder, sin la ayuda de un dispositivo señalador tradicional como un ratón o un bolígrafo digital" (2008, p. 14). Así, moverse las manos, un toque, un movimiento de cabeza o incluso una ceja levantada pueden ser un gesto. Para interactuar con la interfaz, es necesario rastrear los signos utilizados por el usuario sordo.

Actualmente existen en el mercado tecnologías que permiten interactuar mediante gestos, como el seguimiento del mano ejemplificado en la Figura 2.

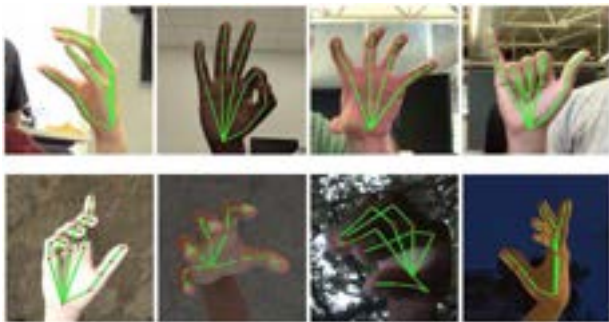


Figura 2: Seguimiento de manos en tiempo real. Fuente: Google.

Según Silva (2013, p.12), este modelo de interacción permitiría a los usuarios sordos interactuar de forma más natural con las interfaces digitales, ya que utilizan los gestos como medio. De esta forma, podrían utilizar su primera lengua para interactuar. Entre los problemas de comunicación e interacción que enfrentan las personas sordas que utilizan Libras está la dificultad de traducirlo al portugués:

"[...] Pienso en Libras y necesito escribir en portugués. Así que es como si el oyente brasileño que piensa en portugués ¡necesitara escribir en japonés! Cuando mis pensamientos se "traducen" de Libras al portugués, confundo las reglas gramaticales. Esta dificultad no ocurre solo al escribir. Incluso leer es difícil, porque la sintaxis de la lengua portuguesa es muy diferente de la sintaxis de Libras. Además, la propia lengua portuguesa presenta diferencias si consideramos sus formas oral y escrita."

(CARNEIRO, 2016, p. 26.

Traducción por autores)

La Lengua Brasileña de Señas (Libras) utiliza gestos y expresiones faciales y corporales para la comunicación, donde la información lingüística es recibida por los ojos y producida en el espacio. Según Schlindwein y Aquino (2023, p.93), el orden básico de Libras es la estructura SVO (sujeto + verbo + objeto) y otras construcciones se derivan de este ordenamiento sintáctico.

A pesar de la diferencia entre las lenguas de signos y las lenguas orales, según Pereira, ambas "siguen los mismos principios en lo que se refiere al hecho de que tienen un léxico, es decir, un conjunto de símbolos convencionales, y una gramática, es decir, un sistema de reglas que rigen el uso y la

combinación de estos símbolos en unidades mayores". (2011, p.5).

3. Procedimientos metodológicos

Se trata de una investigación cualitativa exploratoria, que comprende un estudio bibliográfico y una investigación de usuarios. Según Gil, "puede decirse que el principal objetivo de esta investigación es mejorar las ideas" (2002, p. 41). Así, tuvo como objetivo comprender las mejores prácticas de Experiencia de Usuario (UX) para el desarrollo de asistentes virtuales, cuáles podrían ser los posibles desafíos de la interacción a través de gestos en Libras y, con la comprensión de estos aspectos, desarrollar el proyecto de interacción para una interfaz conversacional, utilizando la metodología del Doble Diamante (Figura 3).

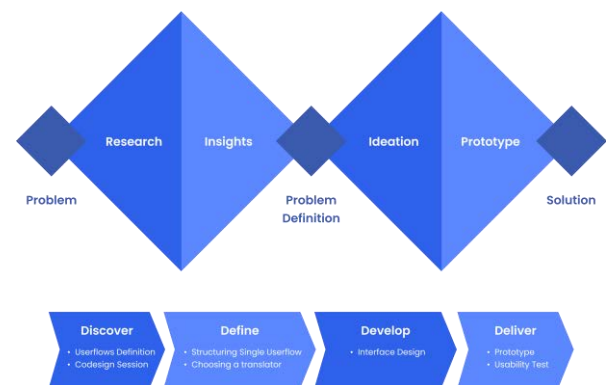


Figura 3: Metodología del doble diamante. Fuente: Autores.

La fase de Descubrimiento tuvo como objetivo comprender el problema, es decir, identificar las interacciones con asistentes virtuales que podrían conformar una experiencia conversacional para un asistente virtual, con entrada de datos gestuales y salida de datos a través de una interfaz gráfica (GUI). Para ello, se definieron como objeto de investigación tres flujos de interacción conversacional asociados a interacciones recurrentes con asistentes virtuales: la previsión meteorológica, la consulta de las noticias y la realización de una videollamada. Tras esta etapa, se celebró una sesión de codiseño junto a un profesor de Libras para comprender las especificidades del lenguaje y aplicar las mejores soluciones al proyecto. En la etapa de Definición, los flujos previamente producidos fueron reevaluados y reestructurados en un único flujo con el fin de proporcionar una prueba final más concisa, delimitando así las funcionalidades necesarias para desarrollar la interacción. En la etapa de Desarrollo, se llevó a cabo el diseño visual de la interfaz y en la última etapa, la de la Entrega, se desarrolló el prototipo de la interfaz y se comprobó con usuarios sordos mediante una prueba de usabilidad, utilizando el enfoque del Mago de Oz (MOZ).

4. Desarrollo de interacciones conversacionales para la comunidad sorda

4.1 Contextualización del proyecto

Para desarrollar interacciones conversacionales dirigidas a usuarios que utilizan Libras, se diseñaron tres flujos: consultar la previsión meteorológica, ver las noticias principales y realizar una videollamada. La propuesta fue explorar diferentes formas de estructurar el enunciado de cada una de esas proposiciones, que pudieran servir como guía para implementar futuras interacciones.

Esta investigación está dirigida a personas sordas que utilizan Libras como su principal forma de comunicación y se configura como un recurso de accesibilidad que podría integrarse en asistentes virtuales existentes, como Alexa, dado que la línea de productos del asistente ya cuenta con dispositivos con pantalla.

Los datos de interacción se introducirán mediante gestos en el aire; el asistente decodificará el lenguaje y devolverá el resultado utilizando principalmente recursos visuales y un avatar 3D (Maya, de la plataforma *Hand Talk*⁴), minimizando los elementos textuales y buscando que el contenido sea más fácil de asimilar para las personas sordas, que pueden tener dificultades con el lenguaje escrito. En las siguientes secciones se ofrecen enlaces a los entregables de diseño utilizados en el proceso de desarrollo; recomendamos acceder a ellos para comprender mejor el diseño de las interacciones conversacionales propuestas.

4.2 Descubrimiento

Los flujos de interacción sirven para representar las etapas de interacción del usuario con el producto, predecir errores, puntos de atención y establecer un recorrido para las pruebas de usuario. Los flujos desarrollados en este trabajo se basaron en algunas acciones recurrentes utilizadas en los asistentes de voz. Se utilizaron tres flujos diferentes: consultar la previsión meteorológica, ver las principales noticias y realizar una videollamada. Estos flujos pueden consultarse a través del [enlace de Figjam](#).

En esta primera fase, la intención era explotar el mayor número posible de expresiones del usuario y de errores del asistente. En el primer flujo, la comprobación de la previsión meteorológica (Figura 4), el objetivo era comprender cómo el asistente podría tratar los errores cuando se trata de la ubicación del usuario, dado que en Libras cada ciudad tiene un signo específico, pero los traductores actuales tienen pocos registros de estos signos. En el segundo flujo, la visualización de las principales noticias,

el objetivo era comprender cómo el asistente podía tratar con cantidades expresivas de texto y cuál era la mejor forma de presentar este contenido. En el tercer flujo, la realización de una videollamada, el objetivo era comprender cómo podía responder el asistente en caso de malinterpretar signos muy similares, como algunos signos utilizados para dactilología.

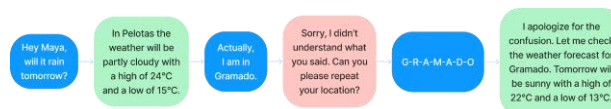


Figura 4: Detalle del flujo "Consultar la previsión meteorológica".
Fuente: Autores.

Se celebró una sesión de codiseño de una hora de duración con un profesor de Libras para comprender las especificidades del lenguaje y cómo aplicar las mejores soluciones al proyecto. Según Lupton (2017, p.96) esta técnica "hace hincapié en la experiencia del usuario como resultado final del diseño en lugar de en las características físicas de un objeto, sitio web u otro resultado".

Al participante se le mostraron los flujos y se le pidió que los tradujera al lenguaje de signos (Figura 5). Cuando se le preguntó por las dificultades de traducir del portugués a Libras, dijo que Libras es diferente del portugués en cuanto a la estructura gramatical, por lo que la traducción es compleja, como en cualquier otra lengua.

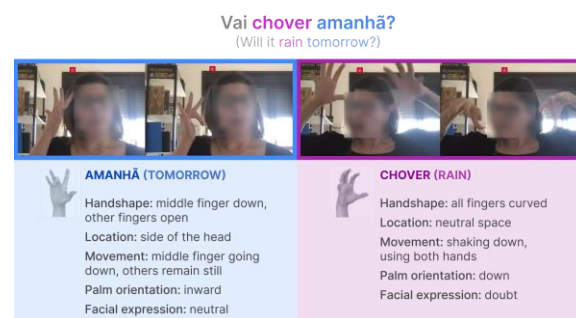


Figura 5: Estructura de la frase "¿Lloverá mañana?" en Libras.
Fuente: Autores

En el primer flujo, la tarea elegida fue "consultar la previsión meteorológica". En primer lugar, la participante señalaba los enunciados en azul en el flujo y, a continuación, se le mostraba la respuesta de la interfaz, preguntándole si le parecía adecuada. En este mapeo se pudo confirmar, por ejemplo, que los usuarios sordos prefieren los signos y sólo utilizan la dactilología en los casos en que la comunicación por signos no resulta eficaz. Como indican los colores de la Figura 5, se observa que el orden de las palabras cambia e incluso algunas se suprimen. Cuando se formulan preguntas, se utiliza la expresión facial de duda junto con el signo. Esta etapa fue muy relevante para la investigación, permitiendo una mayor comprensión de la percepción de los usuarios sordos al realizar

⁴ Disponible en la herramienta Hand Talk - <https://www.handtalk.me/br>

diálogos y el consecuente mapeo de enunciados, fundamentales para traducir el lenguaje verbal a una máquina (por ejemplo, Alexa).

4.3 Definición

En esta fase, se planteó la necesidad de establecer un nuevo flujo que combinara los tres anteriores (consultar la previsión meteorológica, ver las noticias principales y realizar una videollamada), delimitando así las interacciones en base al feedback de la sesión de codiseño realizada previamente. Para ello, se redujeron las tareas, obteniéndose como resultado el flujo que se puede visualizar en [este enlace](#). Para poder probar el flujo con los usuarios, fue necesario establecer un breve onboarding de introducción al producto, explicando cómo activar el asistente y que todas las tareas formaran parte de una misma experiencia de usuario, por lo que cuando se completa una tarea el usuario sigue "hablando" con la interfaz para solicitar la siguiente tarea.

Se eligió Hand Talk entre las plataformas de traducción de lengua de signos disponibles. Después de probar otras herramientas como VLibras y SENAI Libras, se descubrió que Hand Talk tenía mejores capacidades de traducción. Según el sitio web de la empresa⁵, Hand Talk también fue votada por la ONU como "la mejor aplicación social del mundo". La aplicación tiene dos avatares digitales, Hugo y Maya.

Se eligió a Maya porque el signo del personaje se hace con una sola mano, mientras que el signo de Hugo se hace con ambas. Pensando en el mecanismo de activación de la interfaz en diferentes contextos de uso, una persona que tiene una mano ocupada aún puede señalar "Maya". En este punto, se definió la "wakeword" del asistente, de forma similar a otros asistentes ("Hey Google", "Hey Cortana" y "Hey Siri") que se llamará gesto de activación, representado por signos que en portugués sería el equivalente a "Hey Maya".

El personaje Maya también transmite una imagen menos formal, mientras que Hugo lleva una corbata, así que su elección se dio buscando acercarse al público y porque está estéticamente más cercana a la interfaz. Además, si pensamos en los asistentes existentes, podemos ver el patrón de "personajes" femeninos (Alexa, Cortana y Siri, por ejemplo). Otro punto importante es que Hand Talk lanzó el nuevo traductor virtual con el objetivo de aumentar la diversidad de su producto, buscando representar a las mujeres y a las personas de raza negra.

4.4 Desarrollo

Tras estudiar los puntos planteados en las fases anteriores, se desarrolló la interfaz gráfica utilizada para mediar la

interacción conversacional. Se puede acceder a las pantallas diseñadas a través del [enlace Figma](#). En la página de inicio (Figura 6) el objetivo era presentar a la asistente, mostrar su gesto de activación y proporcionar una lista de sugerencias de lo que podría hacer, con el fin de guiar al usuario si no supiera qué preguntar. También aparece un icono animado que indica que se puede mover la página utilizando el gesto o tocando la pantalla. El objetivo al mantener a la intérprete siempre presente en la interfaz, incluso en las pantallas de contenido, es crear proximidad con el usuario, ya que en la comunidad sorda mirar a la persona que se comunica es esencial.



Figura 6: Página de inicio de la interfaz. Adaptado por los autores a partir de la ilustración del asistente Maya, desarrollada por Hand Talk.

Tras esta presentación, el asistente pregunta: "¿En qué puedo ayudar?" En este punto, el usuario empieza a interactuar, señala lo que necesita y la asistente sigue observando e interpretando las señales, lo que se muestra al usuario mediante el anillo azul que gira alrededor de la intérprete virtual. En esta pantalla, la asistente hace pequeños movimientos con la cabeza para mostrar que está "prestando atención".

En las pantallas de error, la asistente señala "Lo siento, no lo he entendido" y describe el error y/o sugiere cómo puede proceder el usuario. En la pantalla proyectada, sugiere "¿Podría deletrear su ubicación?" (Figura 7). En este caso, se espera que el usuario sordo prefiera utilizar el signo para la ciudad de "Gramado", pero como la asistente no tiene registrado este signo, le resulta difícil interpretarlo, por lo que el usuario recurre a la dactilología, donde deletrea la palabra utilizando los signos de Libras que corresponden al alfabeto.

⁵ Disponible en: <https://www.handtalk.me/br>



Figura 7: Pantalla de error de interfaz. Adaptado por los autores a partir de la ilustración del asistente Maya, desarrollada por Hand Talk.

Tras el flujo meteorológico, el usuario pregunta "Maya, ¿puedes mostrarme las principales noticias de Brasil?", y en la pantalla de presentación de noticias se utilizan iconos animados que representan los números en Libras. De este modo, el usuario puede seleccionar la noticia que desea leer con sólo hacer una señal (Figura 8). El objetivo es que la interacción sea fundamentalmente gestual y que el usuario sólo toque la pantalla como último recurso. La noticia seleccionada se representa mediante la selección azul alrededor de la tarjeta y, tras abrir la noticia, la asistente dice "Abriendo noticia 1. ¿Quieres que te la traduzca?".

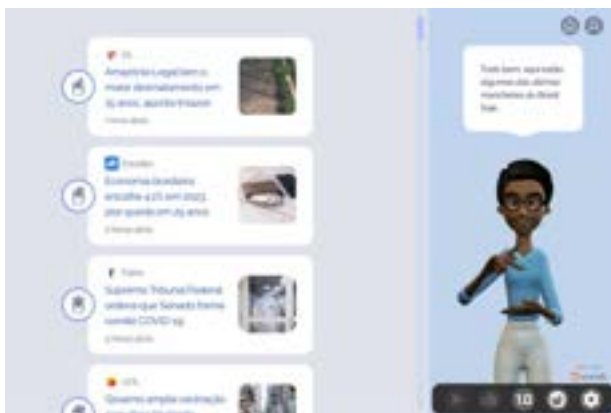


Figura 8: Pantalla de noticias. Adaptado por los autores a partir de la ilustración del asistente Maya, desarrollada por Hand Talk.

La traducción se muestra mediante una barra azul que resalta la frase que se está traduciendo (Figura 9). Es importante destacar aquí que la traducción se hace por frases y no por cada palabra del portugués, ya que, como en otros idiomas, la traducción no se hace literalmente, sino adaptada a la dinámica organizativa de cada lengua. La traducción se hace secuencialmente sin necesidad de que el usuario seleccione cada frase, sólo haciendo una breve pausa entre cada párrafo traducido, y puede interrumpirse en cualquier momento mediante un comando gestual.



Figura 9: Maya traduciendo las noticias. Adaptado por los autores a partir de la ilustración del asistente Maya, desarrollada por Hand Talk.

El último flujo consiste en que el usuario pide llamar a un contacto diciendo "Maya, haz una videollamada con Ana Lucia en Zoom". En este punto, la asistente tiene dificultades para entender la dactilología y sugiere "¿Querías decir Ana Julia?" y, como prevención de errores, aparece una pantalla de confirmación, con la foto y la dirección de correo electrónico y el número de teléfono del contacto. El usuario repite entonces el comando diciendo "No, he dicho Ana Lucia". El asistente encuentra 3 contactos con el mismo nombre y pregunta "Correcto, he encontrado tres contactos con el nombre Ana Lucia. ¿A cuál quieres llamar?" En esta pantalla, vuelve a aparecer la misma función mencionada anteriormente, con iconos que representan los números en Libras para su selección (Figura 10).

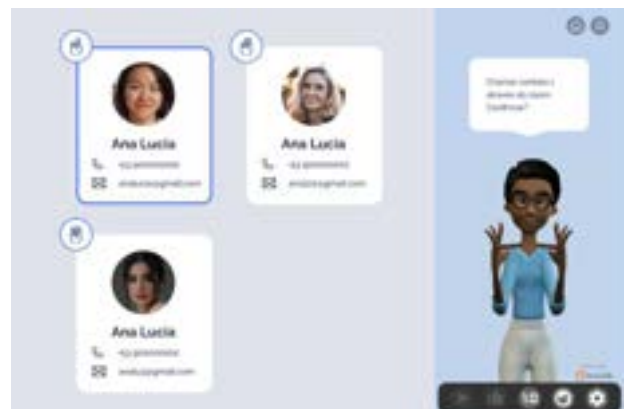


Figura 10: Pantalla de selección de contactos. Adaptado por los autores a partir de la ilustración del asistente Maya, desarrollada por Hand Talk.

Tras seleccionar el contacto, para evitar errores, el asistente pregunta "Llamar al contacto 1 mediante zoom. ¿Confirmar?". El usuario confirma y aparece la pantalla de llamada, con un botón rojo de "colgar", a continuación aparece una persona saludando en la pantalla y así finaliza el flujo.

4.5 Entrega

A partir de las soluciones trazadas en las fases anteriores, el prototipo se desarrolló utilizando el enfoque del Mago de OZ (MOZ) para crear la ilusión de un producto funcional. El enfoque MOZ permite simular el uso de la interfaz, en la que una persona maneja el prototipo, dando al usuario la ilusión de que el sistema funciona. La interacción fue simulada por los investigadores mediante comandos de teclado, y el flujo se guionizó. Al tratarse de una interfaz conversacional, este enfoque permitió probar el prototipo sin necesidad de desarrollar un proyecto 100% funcional. Para analizar los resultados, se grabó el procedimiento con el consentimiento de los participantes. A partir de ahí, fue posible evaluar la experiencia de los posibles usuarios sordos y recoger los comentarios de una prueba de usabilidad.

4.5.1 Prueba de usabilidad

Las pruebas de usabilidad son una herramienta importante para evaluar la eficacia y eficiencia de un producto o sistema a la hora de satisfacer las necesidades y expectativas de los usuarios. Permite evaluar los problemas y puntos de mejora que repercuten en la experiencia del usuario y tomar decisiones basadas en los datos recogidos. La prueba inicial, de carácter formativo, contó con la participación de una persona oyente, profesora de Libras en la UFPel. Se le pidió que leyera las frases cuando se le pidiera, que realizara la señalización y que esperara a que la interfaz respondiera. En este punto, la intención era comprender si el prototipo funcionaría realmente para lo que se proponía y validar el diseño de las pantallas. No se encontró ningún problema y la participante hizo una valoración muy positiva, pero insistió en que sería necesario realizar las pruebas con personas sordas, ya que tienen Libras como lengua materna.

Tras esta primera fase de pruebas, se organizó una prueba de usabilidad, que se llevó a cabo en un ordenador portátil de 14 pulgadas y en la que participaron dos personas, profesores de la UFPel, que tienen Libras como lengua materna, pero que también se consideran buenos conocedores del portugués. El participante 1 tiene 47 años y el participante 2 tiene 36 años. La prueba se realizó a través de la traducción de un profesor oyente. Ambos dijeron que podían entender claramente el portugués básico, como acceder a noticias y leer novelas, pero tenían más dificultades cuando se trataba de textos más formales y académicos.

Cuando se les preguntó por el uso de asistentes virtuales, ninguno de los participantes sabía qué eran los asistentes de voz ni cómo funcionaban. Cuando se les preguntó qué aplicaciones utilizaban, el participante 1 mencionó Youtube, aplicaciones bancarias, Icon (una herramienta que facilita la comunicación entre personas oyentes y sordas en tiempo real, a través de una videollamada, con la mediación de un intérprete de Libras), Google (para buscar y traducir a otros idiomas),

Instagram y Facebook (para acceder a producciones artísticas y culturales) y Hand Talk para traducir a Libras. El participante 2 también mencionó redes sociales como Instagram, Facebook, Whatsapp, Youtube (para buscar palabras en libras) y Hand Talk para la traducción.

Para realizar la prueba se utilizó un guion en el que cada discurso se colocaba en una hoja de papel A4 y se mostraba a los usuarios tal y como estaba prevista la interacción en el prototipo, pudiendo los participantes señalar las frases como lo harían en su vida cotidiana. Los usuarios se sorprendieron al darse cuenta de que la interfaz "respondía" a sus preguntas.

La participante 1 completó la prueba sin dificultad. Se pudo comprobar que realmente quería hablar con la interfaz, porque en los momentos del flujo en los que se produjo un error, ya volvió a hacer señas antes incluso de que se recogiera la hoja con las frases. Utilizaba mucho sus expresiones faciales para hacer señas. Al final de la prueba, la participante comentó la palabra "legal" que en la pantalla de noticias (Figura 9) Maya traduce como "interesante, guay" y no el aspecto legal de "legalidad" al que se refería la noticia. Comenta que le pareció interesante que, cuando el asistente no entendía algo, se le pedía que volviera a hacer la seña utilizando la dactilología - citando el momento en que el asistente no reconoció la seña de la palabra Gramado (Figura 7)- y cuando ella volvió a hacer la seña, el asistente lo entendió. La participante dice que recomendaría el producto a otros usuarios, pero que se da cuenta de los errores de vocabulario. Cree que tanto la señalización como la lengua portuguesa son esenciales y comenta que le pareció interesante el uso de una figura femenina, ya que está acostumbrada a los avatares masculinos. La participante prefirió hacer la prueba de pie, y se ve que tuvo que acercarse a la interfaz varias veces. Comentó que el personaje parecía borroso y que tal vez ampliaría la imagen y haría el fondo más claro.

Al igual que la primera participante, la participante 2 no tuvo dificultades durante la prueba. Hizo la prueba sentada, pero a veces tenía que inclinarse más hacia la interfaz para ver mejor. En la pantalla de previsión meteorológica (Figura 7), la participante también se decantó por el signo de la palabra Gramado, y se observó que utilizaba muchas expresiones no manuales, como asentir con la cabeza para confirmar, por ejemplo, y sólo entonces hacer el signo de "sí". También anticipó algunas respuestas a la interfaz, como cuando el asistente le preguntó "¿Querías decir Ana Julia?", ella espontáneamente hizo el signo "no, A-N-A L-U-C-I-A". La participante comenta la falta de expresividad de la asistente, y que éste sería un factor a tener en cuenta a la hora de recomendar la asistente a otros usuarios. También comenta la falta de algunos signos y que sería interesante adaptar los signos en función del contexto en el que la persona los vaya a utilizar, según el regionalismo.

4.5.2 Evaluación de las pruebas, posibles mejoras y aplicaciones

Tras evaluar las pruebas, nos dimos cuenta de que la mayoría de los comentarios se centraban en cuestiones de traducción, expresividad y tamaño de la interfaz. En cuanto a la traducción y la expresividad, se entiende que son limitaciones de la herramienta de traducción, pero que en el caso de que se desarrollara realmente el producto, sería un punto a priorizar para mejorar. Por el momento, la interfaz se ha diseñado para el tamaño de una tableta, teniendo en cuenta los asistentes existentes en el mercado; no obstante, podría estudiarse su aplicación en distintos soportes. De acuerdo con el alcance de este proyecto, algo que se podría hacer es mejorar la iluminación del personaje y su contraste con el fondo, como sugirió uno de los participantes.

Una posible aplicación de la interfaz sería en dispositivos similares al Echo Show 10, que tiene una pantalla de 10 pulgadas. Una de las ventajas de este dispositivo es que la pantalla se mueve automáticamente hacia el usuario. La desventaja de este modelo sería el pequeño tamaño de la pantalla, dado que en el mercado existen modelos con pantallas más grandes, como el Echo Show 15, que tiene una pantalla de 15 pulgadas.

Con la conclusión de esta etapa, los 3 flujos de interacción propuestos fueron diseñados, desarrollados y validados, demostrando que las interacciones conversacionales definidas son una alternativa importante para el desarrollo de futuros asistentes virtuales enfocados en la comunidad sorda que utiliza la lengua de signos - en el contexto brasileño el uso de Libras, pero permitiendo que el proyecto sea adaptado a otras lenguas.

5. Consideraciones finales

Aunque existen tecnologías que podrían utilizarse para reconocer gestos, estas no se han incorporado a los asistentes más populares del mercado (por ejemplo, Alexa, Google Assistant). Como resultado, los usuarios que utilizan Libras para comunicarse dependen del lenguaje escrito para interactuar, acaban enfrentándose a barreras en la comunicación e interacción con las interfaces debido al lenguaje, y un asistente virtual que entienda y utilice el lenguaje de signos puede ayudar a llenar este vacío, ya sea como una función de accesibilidad o un dispositivo que soporte de forma nativa la lectura de signos utilizada en Libras.

En la prueba de usabilidad, fue posible recoger las opiniones de dos usuarios sordos potenciales y trazar un mapa de sus experiencias, que en general fueron muy positivas. De este modo, la investigación alcanzó satisfactoriamente su objetivo general, dando como resultado un producto digital que puede ponerse en práctica y que puede ayudar a las personas sordas.

Uno de los participantes en la prueba destacó: "Creo que es muy importante y ayuda a la comunidad sorda, por ejemplo, cuando las personas sordas están estudiando, si no entienden algo, pueden utilizar la herramienta para buscar información" (Participante 1). Al final de la prueba, también comentó: "Estoy muy contenta de que os preocupéis por buscar mejoras en la accesibilidad para las personas sordas, así que yo, en representación de la comunidad sorda, os doy las gracias por tener este interés" (Participante 1). Del mismo modo, el otro participante también destacó aspectos importantes relacionados con el uso potencial del proyecto: "Creo que las personas sordas y oyentes pueden utilizarlo, por ejemplo, los CODA, que son hijos de padres sordos, también pueden hacer señales con el asistente, interactuar en familia" (Participante 2).

Sin embargo, hay que destacar que la investigación tiene algunas limitaciones, como el reducido número de participantes, la falta de pasos de iteración adicionales y la escasa participación de expertos, debido al corto plazo de tiempo para su desarrollo. Un posible reto a la hora de desarrollar el asistente virtual sería garantizar que el sistema pueda interpretar y responder con precisión a una gran variedad de signos, muchos de ellos con una configuración de manos muy similar (lo que podría implicar retos técnicos a la hora de implementar el software de reconocimiento del lenguaje de signos). Otra cuestión es garantizar que las respuestas del asistente virtual sean precisas, dada la falta de vocabulario y expresiones faciales en la propia herramienta de traducción. Durante las pruebas, se observó que el tamaño y la distancia de la interfaz con respecto al usuario son factores muy importantes para la experiencia y que sería necesario realizar un estudio más profundo para definir el soporte de la interfaz (actualmente, la mayoría de los dispositivos disponibles en el mercado con pantalla tienen el tamaño aproximado de una tableta).

Como perspectivas de trabajo futuro, se han enumerado algunas posibilidades: identificar si los usuarios prestan más atención al avatar traductor que a los demás elementos de la interfaz; definir nuevos flujos comunes de uso del asistente además de los 3 diseñados; analizar la capacidad de registrar nuevos signos en la base de datos del asistente (por ejemplo: añadir el signo al nombre de un contacto); definir el mejor soporte para la interfaz, estructurando nuevas herramientas para obtener métricas de evaluación de las pruebas de usuario. Un asistente virtual dirigido a personas que utilizan la lengua de signos puede proporcionar a las personas sordas una mayor independencia y acceso a información y servicios, ofreciendo una experiencia inclusiva para estos usuarios y contribuyendo a derribar las barreras que dificultan el uso de estos servicios. Aunque esta investigación es de carácter exploratorio y limitado, los resultados identificados con las personas sordas usuarias de Libras han sido muy positivos, inspirando la continuación de proyectos relacionados.

Referencias

- Carneiro, M. I. N. (2016). O uso social das tecnologias de comunicação pelo surdo: limites e possibilidades para o desenvolvimento da linguagem (Tese de Doutorado). Universidade Estadual de Maringá, Maringá, PR, Brasil.
- Gil, A. C. (2002). Como elaborar projetos de pesquisa (4ª ed.). Atlas.
- Google. (2022, 4 de outubro). On-Device, Real-Time Hand Tracking with MediaPipe. <https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html>
- Hand Talk. (2023, 7 de agosto). <https://www.handtalk.me/br/>
- IBGE. (2023, 27 de abril). PNS 2019: país tem 17,3 milhões de pessoas com algum tipo de deficiência. <https://censos.ibge.gov.br/2013-agencia-de-noticias/releases/31445-pns-2019-pais-tem-17-3-milhoes-de-pessoas-com-algum-tipo-de-deficiencia.html#:~:text=Cerca%20de%201%2C3%25%20da,%2C6%25%20tinham%20defici%C3%Aancia%20auditiva>
- Lupton, E. (2013). Graphic Design Thinking: Intuição, ação, criação. G. Gili.
- Pereira, M. C. C. (2011). As Línguas de Sinais: Sua importância para os Surdos. In M. C. C. Pereira (Org.), LIBRAS: conhecimento além dos sinais. PearsonPrentice Hall.
- Saffer, D. (2008). Designing Gestural Interfaces. O'Reilly Media.
- Schindwein, A. F., & Aquino, A. (2021). Aspectos Sintáticos da Libras. https://cesad.ufs.br/ORBI/public/uploadCatalogo/12242928072021Aula_08.pdf
- Silva, R. K. (2013). Interfaces naturais e o reconhecimento das línguas de sinais.
- Torres, E. F., & Mazzoni, A. A. (2004). Conteúdos digitais multimídia: o foco na usabilidade e acessibilidade. Ciência da Informação, 33(2), 152–160.