

Silvia Tamayo Moreno
Diana Pérez Marín

**Evaluación heurística de Agente
Conversacional Pedagógico
Dr. Roland para el nivel de
Educación Infantil**

Número 2017-16

Serie de Informes Técnicos DLSI1-URJC
ISSN 1988-8074
Grupo Docente de Lenguajes y Sistemas Informáticos I
Universidad Rey Juan Carlos

Índice

1	Introducción.....	5
1.1	Evaluación heurística, método de evaluación de la Interacción Persona Ordenador	5
2	Evaluación heurística del agente Dr. Roland: Educación Infantil.	8
3	Conclusiones.....	17
	Referencias	18

Evaluación heurística de Agente Conversacional Pedagógico Dr. Roland para el nivel de Educación Infantil

Silvia Tamayo Moreno y Diana Pérez Marín

Dto. de CC Comp., Arq. Comp., LSI y EIO, Universidad Rey Juan Carlos,
C/ Tulipán s/n, 28933, Móstoles, Madrid
silviatamayomoreno@gmail.com, diana.perez@urjc.es

Abstract. In this paper the heuristic evaluation of the pedagogical conversational agent Dr. Roland is carried out for the level of Pre-Primary Education following the Nielsen heuristics. The heuristic evaluation is one of the inspection methods of the techniques of evaluation of the Human Computer Interaction.

Keywords: Evaluación heurística, Agente Dr. Roland, Educación Infantil, Interacción persona Ordenador.

1 Introducción

En primer lugar se va a hacer alusión a la interacción persona ordenador (IPO). Seguidamente a técnicas de evaluación de la IPO [1], centrando la atención en los métodos de inspección. Concretamente en la evaluación heurística.

A continuación se realizará la evaluación heurística de Agente Conversacional Pedagógico Dr. Roland [23, 24] para el nivel de Educación Infantil.

1.1 Evaluación heurística, método de evaluación de la Interacción Persona Ordenador

Los sistemas interactivos son dispositivos hardware y software que actúan de interconexión entre las personas, permitiéndoles interactuar, y favoreciendo la realización de tareas y el alcance de los objetivos propuestos.

Con el desarrollo de los sistemas interactivos, y para tener una mejor comprensión de las relaciones persona-ordenador y mejorarla, surge la interacción persona ordenador (IPO). La IPO, conocida en la comunidad internacional como Human-Computer Interaction (HCI) o Computer-Human Interaction (CHI), según el grupo SIGCHI (Special Interest Group on Computer-Human Interaction, [2]) de ACM: “*Es la disciplina relacionada con el diseño, evaluación e implementación de sistemas informáticos interactivos para el uso de seres humanos, y con el estudio de los fenómenos más importantes con los que está relacionado*” [1].

Constantemente se presenta la situación en la que se diseñan productos que no se prueban con quienes van a ser los usuarios, sin prestar atención por tanto a la usabilidad de los productos y el software, como a las pruebas con usuarios. Los motivos suelen ser los costes y cambios que pueda suponer dicha evaluación. Se centran los esfuerzos y atención en la funcionalidad, y se tiene el pensamiento de que si quienes diseñan y programan, pueden usar el sistema y les gusta, es suficiente [3].

Evaluar algo implica hacer pruebas de funcionamiento y comprobar si cumple las expectativas, necesidades y requisitos. En el ámbito de los sistemas interactivos es clave y es importante que se realice durante todo el proceso de desarrollo. Los resultados de ello, obtenidos mediante la aplicación de una serie de técnicas, se pueden emplear para mejorar los sistemas. Lorés et al. [4] define la evaluación como “*La actividad que comprende un conjunto de metodologías y técnicas que analizan la usabilidad y/o la accesibilidad de un sistema interactivo en diferentes etapas del ciclo de vida del software*”.

Los prototipos se hacen para comprobar si el sistema reúne características que lo hagan usable para los usuarios finales. La evaluación del diseño conlleva comprobar si se adapta a lo esperado, permitiendo realizar las tareas, y comprobar el impacto en el usuario en aspectos como la facilidad de aprendizaje, manejabilidad, identificación de áreas que necesiten que sea recordada gran cantidad de información, etc. Con ello se identifican posibles problemas específicos del diseño, como todo aquello que cause resultados no esperados o confusos al usuario estando en su contexto de uso [5]. Según Dix [6] los objetivos de la evaluación son la comprobación de la funcionalidad del sistema, del efecto de la interfaz y la identificación de problemas específicos del sistema.

Los sistemas interactivos pueden ser evaluados en cualquier contexto que lo permita, donde se reúnan usuarios y evaluadores, desde espacios equipados como laboratorios, salas de reuniones o hasta en el entorno en el que los usuarios desarrollan sus tareas.

El abanico de métodos de para evaluar los diferentes aspectos es muy amplio, existiendo solapamiento entre algunos de ellos en cuanto a las actividades que desarrollan, no son independientes [7]. La elección de unos u otros depende de los costes y lo que se obtendrá con su uso.

En cuanto al tipo de técnica de comprobación utilizada, se distinguen:

- **Métodos de inspección:** Agrupa una serie de métodos que cuentan con expertos, llamados evaluadores, que explican el grado de usabilidad de un sistema. Se basan en la inspección de aspectos de la interfaz del sistema relacionados con la usabilidad ofrecida a los usuarios. Se considera opiniones, juicios o informes de los inspectores sobre elementos específicos de la interfaz como factor fundamental de la evaluación [8]. Permiten identificar, clasificar y contabilizar un gran número de errores potenciales de usabilidad a precio relativamente bajo, contribuyendo a ello el hecho de no usar usuarios [8, 9].

- **Métodos de indagación.** Consiste en alcanzar el conocimiento de una cosa mediante conjeturas y señales. Se llevan a cabo, hablando con los usuarios, observándolos, usando el sistema en tiempo real y observando las respuestas a preguntas, ya sea orales o escritas. La información sobre gustos del usuario, necesidades o identificación de requisitos es importante, especialmente en fases tempranas del proceso de desarrollo.
- **Métodos de test:** En ellos usuarios, normalmente representativos, usan el sistema o prototipo para tareas concretas. Los evaluadores usan los resultados para saber la forma en la que la interfaz de usuario da soporte a los usuarios con sus tareas.

Como parte de los métodos de inspección se encuentra la evaluación heurística.

Evaluación heurística

Los métodos de evaluación por inspección como la evaluación heurística, que inicialmente fueron propuestas por Molich y Nielsen [10], son técnicas que no requieren la participación de usuarios, por lo que son más económicas [11].

En el diseño son los expertos quienes lo inspeccionan y analizan buscando problemas potenciales, comprobando los requisitos que tiene que cumplir el diseño para facilitar la comprensión y el uso por el usuario final, en definitiva, los principios de diseño usable. En relación a los expertos, es aconsejable que sean 3-5, para poder realizar un informe conjunto con las opiniones ya que la opinión de menos evaluadores sería sesgada, y contar con más evaluadores supondría un coste de la evaluación sin justificación [8]. Además, cuanto más experiencia se tenga será más fácil que detecten más problemas [12, 4]. Cada uno de ellos, analizará independientemente el diseño, informando de los problemas que se detecten. Finalmente, se pondrán en común y hará un informe conjunto [13].

Diversos autores han propuesto reglas de diseño que pueden emplearse como heurísticas (las 10 reglas heurísticas de usabilidad de Nielsen [14], las ocho reglas de oro de Schneiderman [15], o Tognazzini [16]). Otros ofrecen guías formadas por criterios heurísticos más específicos, más fáciles de aplicar por evaluadores que no son expertos. Además, existen diversos casos prácticos de aplicación de evaluación heurística que pueden ser consultados [17, 18, 19].

La evaluación heurística puede realizarse en cualquier momento del ciclo de desarrollo del proyecto (antes de las pruebas de usuarios es un buen momento), y dependiendo de cuando sea, los principios o criterios a comprobar podrían variar. En relación a las limitaciones y problemas [12]:

- Detecta más problemas menores de usabilidad, detecta menos problemas mayores que otras metodologías, por lo que no sustituye a la realización del test de usuarios.
- Puede reportar alarmas falsas, identificando como problema algo que no lo es.

- Para que ofrezca resultados importantes es necesario que participen varios evaluadores.

2 Evaluación heurística del agente Dr. Roland: Educación Infantil.

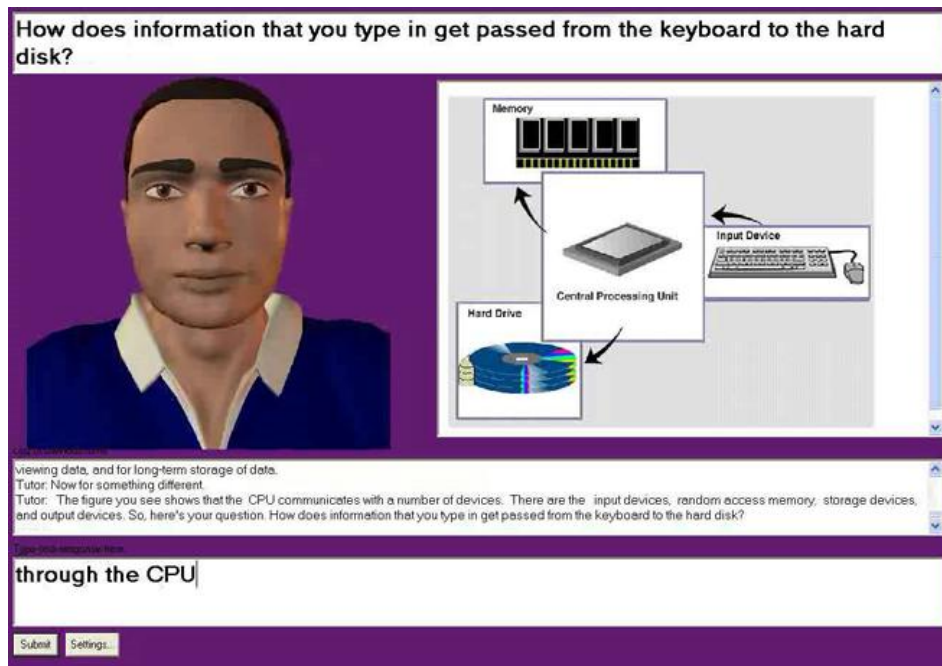


Fig. 1. Autotutor [20]

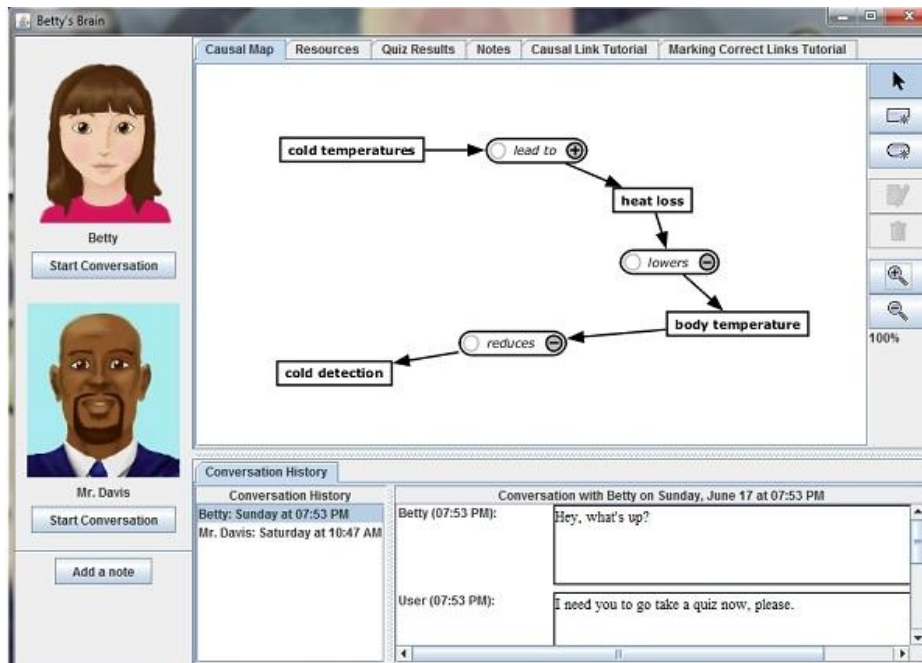


Fig. 2. Interfaz del agente Betty [21]



Fig. 3. Interfaz del agente Laura [22]

Se ha llevado a cabo la Técnica Inspección Heurística, mediante la evaluación de las 10 heurísticas de usabilidad de Nielsen [14].

H1. Visibilidad del estado del sistema: se aprecia en la Figura 4 como el sistema mantiene informado al usuario, mostrando su estado y ofreciendo respuesta a la ac-

ción de usuario, e indicándole el siguiente paso a dar. En las figuras Figura 5, Figura 6 y Figura 7 puede comprobarse en los recuadros marcados en rojo como el sistema informa del resultado de la acción previa del usuario (resaltado en el cuadro verde), y que está en estado de espera pendiente de que realice una nueva acción el usuario, indicando también cuál debe ser y ofreciendo los medios para que el usuario responda (resaltado en los recuadros azules). De esta manera, el usuario conoce que el siguiente en realizar alguna opción es él y cuál debe ser.



Fig. 4. Prototipo Infantil - Interfaz inicial

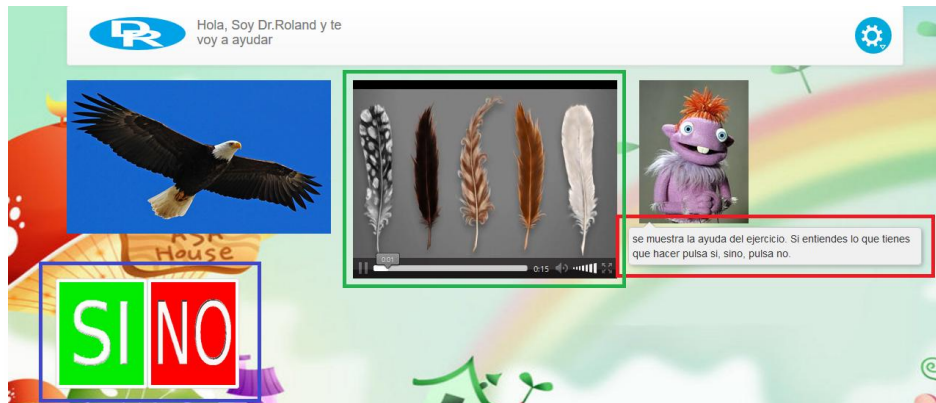


Fig. 5. Prototipo Infantil - Ejemplo 1



Fig. 6. Prototipo Infantil - Ejemplo 2

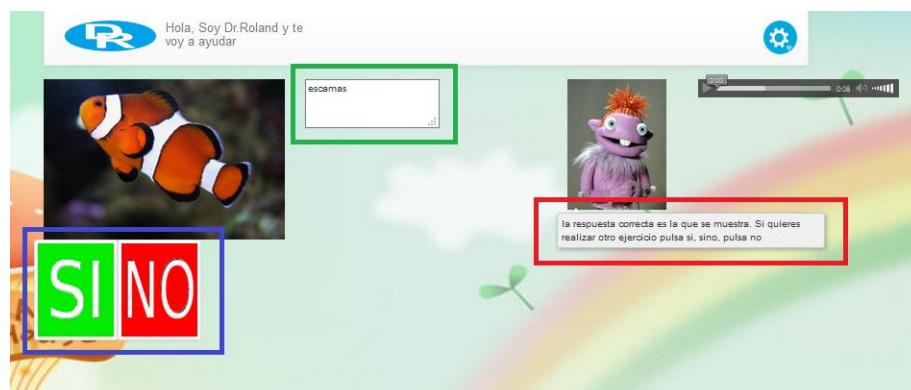


Fig. 7. Prototipo Infantil - Ejemplo 3

H2. Relación entre el sistema y el mundo real: Se puede comprobar en las figuras Figura 5, Figura 6 y Figura 7 que el sistema usa un lenguaje entendible por los usuarios. Sin embargo, en este contexto el problema es que algunos de ellos no saben leer, pero en esos casos el profesor o el propio agente realiza esa labor, leyéndolo en voz alta, por lo que es entendible para los niños. Además, el sistema se adecúa a la situación correspondiente, adaptando su apariencia, interacción y opciones que ofrece la situación en la que se encuentra y de la que informa en modo texto en todo momento. Además ofrece flexibilidad para adaptar la respuesta e interacción a las características de los niños, por ejemplo con el hecho de que el propio agente pueda comunicarse hablando (diciendo de voz aquello que además comunica de forma escrita), porque

esto también hace que sea una interacción mucho más cercana a la realidad, como indicaron las profesoras de la primera y tercera experiencia con el uso del agente.

H3. Control y libertad del usuario: el usuario cuenta con una cierta salida ante estados no deseados al que accedieron. Como por ejemplo, al tratarse de una aplicación web, siempre puede salir cuando quiera, simplemente cerrándola, o volver atrás. Además, dentro de las posibilidades cubiertas por el sistema, que es quien llevará el control principal de la interacción, permite al usuario libertad en determinadas acciones, como por ejemplo decidir si quiere realizar otro ejercicio, el modo de interacción que quiere usar para dar la respuesta a los ejercicios o si quiere que el propio agente se comunique mediante voz, etc. Sin embargo, en este punto cabe destacar que la figura del docente es fundamental, y que muchas características serán el docente quien las use, ya que al ser niños tan pequeños van a estar guiados por su docente, por lo que serán estos quienes les ayuden con su ingenio a solventar estas situaciones.

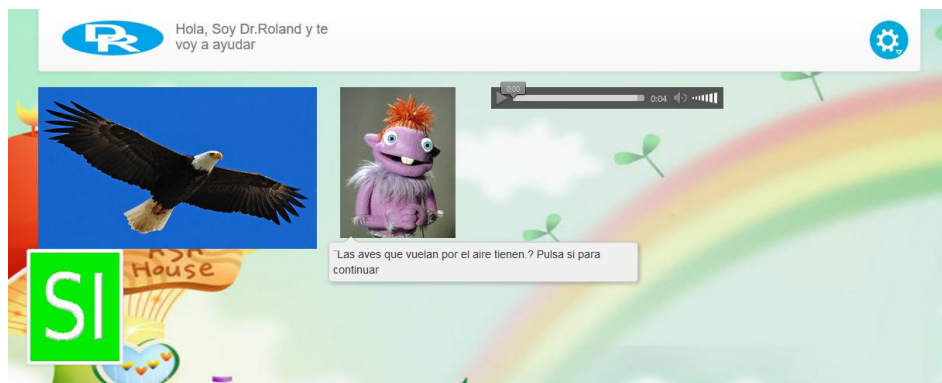


Fig. 8. Prototipo Infantil - Enunciado del ejercicio

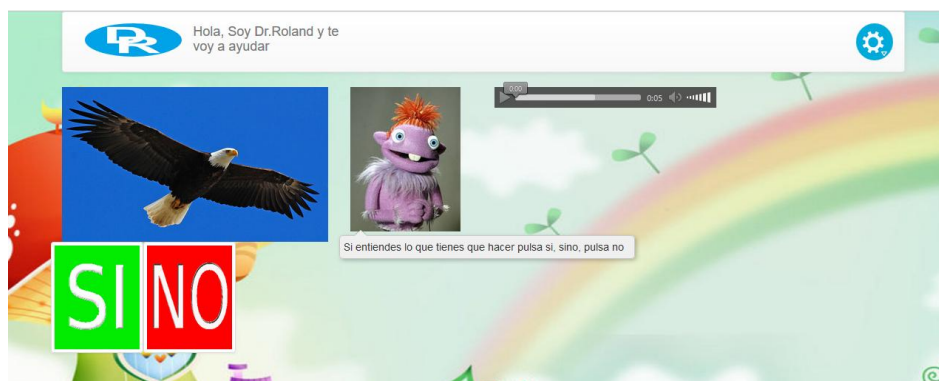


Fig. 9. Prototipo Infantil - Se pregunta si entiende lo que se tiene que hacer

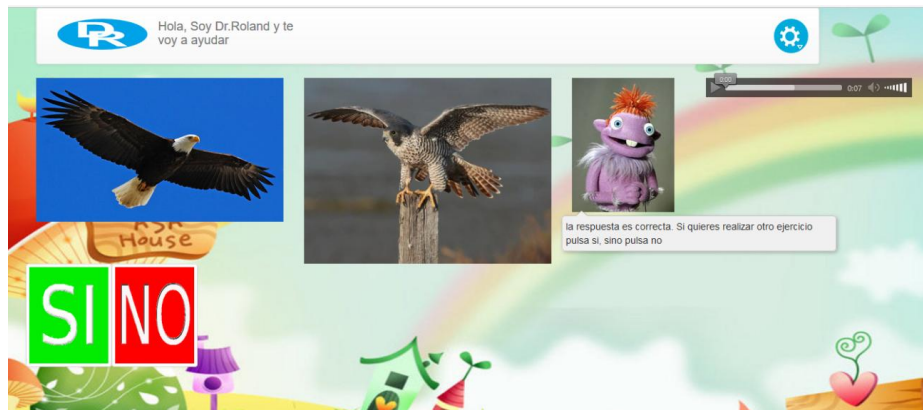


Fig. 10. Prototipo Infantil - Se informa que la respuesta es correcta



Fig. 11. Prototipo Infantil – Interacción con teclado

H4. Consistencia. En relación a la consistencia externa, que se refiere a la consistencia con otros agentes, puede comprobarse, que la mantiene respecto alguno agente como el Autotutor (Figura1), Betty (Figura 2) o Laura (Figura 3). Mantienen una distribución de los elementos similar, ofreciendo la imagen del agente en la parte izquierda, realizan la interacción con el usuario por medio de cuadros de texto en los que el usuario introduce información, y el agente le va respondiendo, lo que también es compartido en el agente Dr. Roland.

Respecto a la consistencia interna, que se refiere a la existente entre pantallas del propio agente, puede apreciarse, que la interfaz con sus elementos correspondientes según el caso de que se trate, mantiene los mismos patrones, la misma tonalidad y estilos parecidos, siendo en este caso muy coloridos y llamativos. Esto puede apreciarse en las figuras Figura 8, Figura 9, Figura 10 y Figura 11.

H4. Estándares: En relación a los estándares, como en la metodología MEDIE y en el desarrollo del agente se ha seguido el DCU, se han abarcado lo relativo a la aplicación del DCU.

H5. Prevención de errores: el sistema trata de prevenir ciertos errores, especialmente a la hora de que el usuario inserte la información. Un ejemplo de ello es que gran parte de las interacciones se realizan pulsando teclas de gran tamaño, puede observarse un ejemplo de ello en Figura 12. Además, que para avanzar, únicamente lo pueden hacer pulsando las teclas ofrecidas según el caso que corresponda, que además tanto la situación como el siguiente paso a dar por parte del usuario son explicadas por el agente, como puede apreciarse en la Figura 14.

Igualmente con la opción de resolver por voz y que sea el profesor quien l corrija, indicando al agente si ha sido correcta o no, se evita errores por parte del niño al introducir la respuesta, un ejemplo de este modo puede verse en la Figura 13.



Fig. 12. Prototipo de Infantil - Ejemplo 1 de prevención de errores

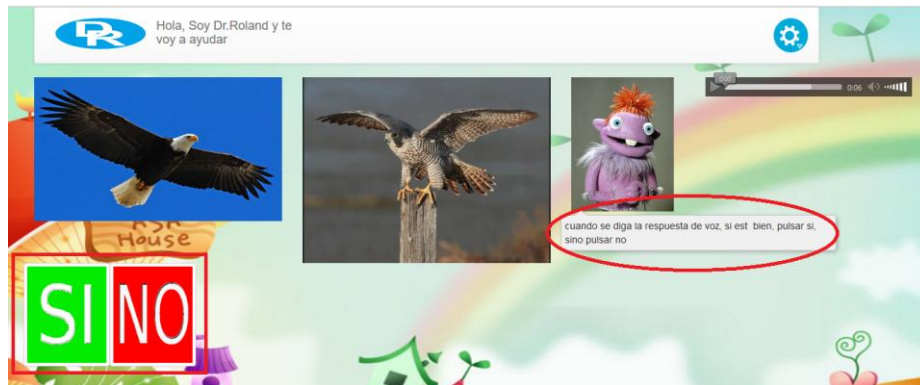


Fig. 13. Prototipo de Infantil - Ejemplo 2 de prevención de errores



Fig. 14. Prototipo de Infantil - Ejemplo 3 de prevención de errores

H6. Reconocimiento antes que recuerdo: implica todo lo que haga al usuario reconocer, y que mantenga patrones entre ello. En el agente se intenta dar visibilidad a objetos, acciones y opciones disponibles, de tal manera que el usuario no tenga que recordar la información. Por ejemplo, dispone de un botones de selección de opciones, que por su logotipo pueden ser reconocidos, estando a la vista del agente Figura 18. Otro ejemplo, es cuando en función del caso en el que esté, se le ofrecen todas las posibles acciones que puede realizar en forma de botón, para que pulse la que crea. Además, las teclas ofrecidas son representativas en sí mismas de la acción a realizar como puede apreciarse en la Figura 15, donde la de la interrogación significa ayuda, el teclado es la respuesta por teclado y la del altavoz la respuesta de voz.

Otro ejemplo es el bocadillo con texto que se muestra en la parte inferior de la

imagen del agente, 16, que el hecho de tener esa forma y ubicación permite reconocer al usuario que es la interacción por parte del agente.

O la imagen del ejercicio, y la ayuda si la tiene, que permiten al usuario reconocer en todo momento con qué está relacionado o cuál es el ejercicio que está intentando resolver Figura 17. Estos son ejemplos que evitan al usuario tener que recordar.

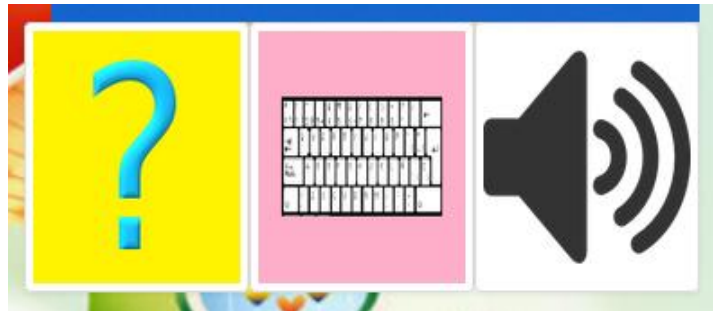


Fig. 15. Prototipo de Infantil – Ejemplo 1 reconocimiento

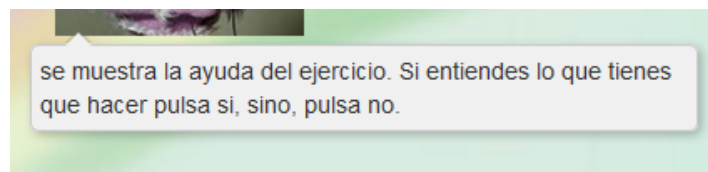


Fig. 16. Prototipo de Infantil – Ejemplo 2 reconocimiento

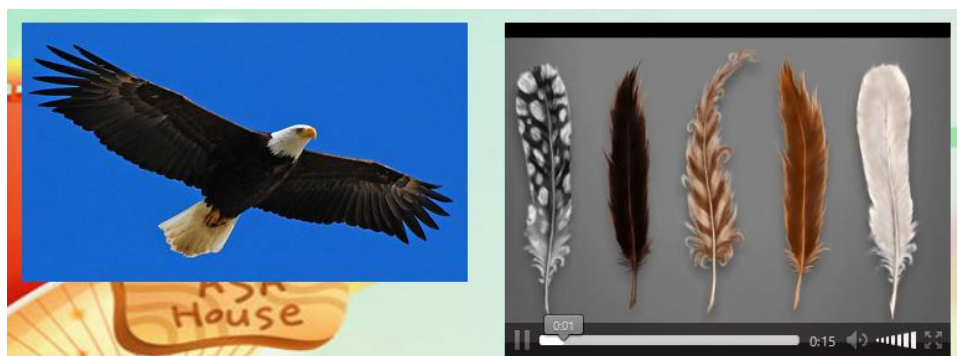


Fig. 17. Prototipo de Infantil - Ejemplo 3 reconocimiento



Fig. 18. Prototipo de Infantil - Ejemplo 4 reconocimiento

H7. Flexibilidad y eficiencia de uso: el sistema tiene cierta flexibilidad ya que permite que los usuarios adapten el sistema para usos frecuentes. Esto es algo que tiene sentido dada la particularidad de agente Dr. Roland que va dirigido a niños. La entrada se podría hacer tanto por teclado, como por ratón a la hora de responder a las preguntas y también la salida es flexible al poderse mostrar tanto en pantallas de ordenador, tabletas, móviles y pizarras digitales.

H8. Estética y diseño minimalista: los diálogos usados en la interacción contienen información concreta y relevante. Se ha tratado de no incluir información que no es útil y que puede llevar confusión.

H9. Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de errores. Lo que se ha tratado de hacer es prevenir el error. El agente responde a cada acción del usuario, habiéndose evitado la inclusión de errores, intentando redirigir la situación a alguna salida. Pero en este sentido, no se ha hecho ni incluido un plan de errores que detalle al usuario qué pasó o cómo recuperarse de ellos.

H10. Ayuda y documentación: el sistema ofrece ayuda relacionada con la resolución del ejercicio (Figuras 5 y 6), y explicando en todo momento la siguiente acción a realizar por parte del usuario (Figura 6). Pero no ofrece ayuda en cuanto a funcionamiento del propio sistema en este nivel, puesto que al ser usados por niños, probablemente no al ser tan pequeños no sabrían interpretarlo. Además, durante la interacción con el agente, en todo momento van a contar con una figura adulta, que le prestará asistencia cuando corresponda.

3 Conclusiones

Las heurísticas fueron revisadas por tres expertos, dos de ellos doctores en informática, para el agente Dr. Roland para el nivel de Educación Infantil. Tras la evaluación de las mismas, puede concluirse que el sistema es usable porque, en mayor o menor grado cumple las heurísticas evaluadas.

Referencias

1. Hewett, T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G. & Verplank, W. (1992). ACM SIGCHI Curricula for Human-Computer Interaction. Technical Report. ACM, New York, NY, USA.
2. SIGCHI Special Interest Group in Computer Human Interaction. Última fecha de consulta: 20-04-2017.
<http://www.sigchi.org/>
3. Granollers i Saltiveri, T., Lorés Vidal, J., & Cañas Delgado, J. J. (2012). Diseño de sistemas interactivos centrados en el usuario.
4. Lorés, J., Granollers, T. y Lana, S. (2002). Introducción a la interacción persona-ordenador. En J. Lorés (ed.) La interacción persona-ordenador. Lérida, 2001-2002, 20-40.
5. Saltiveri, G. (2007). MPIu+ a. Una metodología que integra la ingeniería del software, la interacción persona-ordenador y la accesibilidad en el contexto de equipos de desarrollo multidisciplinares. Universitat de Lleida.
6. Dix, A. (1993). Human computer interaction. Prentice Hall, Englewood Cliffs, NJ.
7. Rodeiro, J. (2001). Representación y Análisis de la componente visual de la interfaz de usuario. Tesis doctoral. Universidad de Vigo.
8. Nielsen (1994), "Heuristic evaluation", Usability Inspection Methods, John Wiley & Sons, New York, NY.
9. Bias, R. & Mayhew D. (1991). Cost-justifying usability. IEEE Software.
10. Molich y Nielsen, "Improving a human-computer dialogue", Communications of the ACM, 3 (33), pp. 338-348, 1990
11. Hassan Montero, Y. & Ortega Santamaría, S. (2009). "Informe APEI sobre usabilidad".
12. M.P. González, A. Pascual, J. Lorés, "Evaluación Heurística", Lorés, J. (Ed.), 2006.
13. Manchón, E. Evaluación heurística (o por expertos) de la usabilidad. Última fecha de consulta: 20-04-2017
14. Nielsen, J. (1995). 10 usability heuristics for user interface design. Nielsen Norman Group, 1(1).
15. Schneiderman, B. (1986). Eight golden rules of interface design
16. Tognazzini, B. (2003). First principles of interaction design. AskTOG.
17. Marcos, M. C. (2006). Evaluación de la usabilidad en sistemas de información terminológicos online. Hipertext. net: Anuario Académico sobre Documentación Digital y Comunicación Interactiva, (4).
18. García Gómez, J. C. (2008). Análisis de usabilidad de los portales en español para personas mayores. No Solo Usabilidad, (7).
19. Candamil-Llano, M. and Guevara-Hurtado, A. F. "Análisis de sitios web universitarios colombianos mediante evaluaciones heurísticas para el proyecto del nuevo portal web Unicauca", 2009.
20. Person, N. K., & Graesser, A. C. (2000). Designing AutoTutor to be an effective conversational partner.
21. Leelawong, K. and Biswas, G. (2008) Designing Learning by Teaching Systems: The Betty's Brain System. In International Journal of Artificial Intelligence in Education.
22. Theodoridou, K. & Yerasimou, T. (2008). "Learning spanish with"laura": The role of an intelligent agent in a spanish language course," in World Conference on Educational Multimedia, Hypermedia and Telecommunications, vol. 2008, no. 1, 2008, pp. 4907– 4912.
23. Tamayo-Moreno, S & Pérez-Marín, D. (2016). Analizando la interacción de estudiantes de educación Infantil y Primaria con un agente conversacional. Proceedings of IkaSnabar-

GUIDE 2016, 9th International Conference. Gorka J. Palazio (ed.), servicio editorial de la Universidad del País Vasco.

24. Tamayo-Moreno, S & Pérez-Marín, D. (2016). Adapting the design and the use methodology of a Pedagogical Conversational Agent of Secondary Education to Childhood Education