

Received April 16, 2021, accepted May 10, 2021, date of publication May 17, 2021, date of current version May 28, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3080837

Entertainment Chatbot for the Digital Inclusion of Elderly People Without Abstraction Capabilities

SILVIA GARCÍA-MÉNDEZ¹, FRANCISCO DE ARRIBA-PÉREZ¹,
FRANCISCO J. GONZÁLEZ-CASTAÑO¹, JOSÉ A. REGUEIRO-JANEIRO¹,
AND FELIPE GIL-CASTIÑEIRA¹

Information Technologies Group, atlanTTic, Telecommunication Engineering School, University of Vigo, 36310 Vigo, Spain

Corresponding author: Silvia García-Méndez (sgarcia@gti.uvigo.es)

This work was supported in part by Ministerio de Economía, Industria y Competitividad (MINECO) under Grant TEC 2016-76465-C2-2-R, and in part by Xunta de Galicia, Spain, under Grant GRC 2018/053.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Asociación de Familiares de enfermos de Alzheimer y otras demencias de Galicia (AFAGA), and performed in line with the Professional Gerontologists' Code of Ethics.

ABSTRACT Current language processing technologies allow the creation of conversational chatbot platforms. Even though artificial intelligence is still too immature to support satisfactory user experience in many mass market domains, conversational interfaces have found their way into *ad hoc* applications such as call centres and online shopping assistants. However, they have not been applied so far to social inclusion of elderly people, who are particularly vulnerable to the digital divide. Many of them relieve their loneliness with traditional media such as TV and radio, which are known to create a feeling of companionship. In this paper we present the EBER chatbot, designed to reduce the digital gap for the elderly. EBER reads news in the background and adapts its responses to the user's mood. Its novelty lies in the concept of "intelligent radio", according to which, instead of simplifying a digital information system to make it accessible to the elderly, a traditional channel they find familiar -background news- is augmented with interactions via voice dialogues. We make it possible by combining Artificial Intelligence Modelling Language, automatic Natural Language Generation and Sentiment Analysis. The system allows accessing digital content of interest by combining words extracted from user answers to chatbot questions with keywords extracted from the news items. This approach permits defining metrics of the abstraction capabilities of the users depending on a spatial representation of the word space. To prove the suitability of the proposed solution we present results of real experiments conducted with elderly people that provided valuable insights. Our approach was considered satisfactory during the tests and improved the information search capabilities of the participants.

INDEX TERMS Text processing, aging, human-computer interaction, user experience, conversational interface.

I. INTRODUCTION

Humans are innately social and need companionship [1], [2]. Social inclusion of elderly people is essential to improve their quality of life in our ageing society. According to the Spanish National Statistics Institute, Spain had ageing indexes of 118.26% in 2017 and 120.46% in 2018. The population in their sixties grew to 19.20% in 2018 and 54.19% of these needed care.

The associate editor coordinating the review of this manuscript and approving it for publication was Eunil Park¹.

Available at <https://www.ine.es/en/welcome.shtml>, March 2021.

The rapid expansion of digital technologies has had appalling social benefits. Few will deny that these have been unbalanced across different population groups. In the case of elderly people, several works have noted their low level of engagement due to the sometimes insuperable digital gap [3]–[5]. Alleviating their digital exclusion in a way they find acceptable is the main motivation of this work.

The digital gap impedes many elderly people to access digital content and even technology-aware people eventually experience this gap when their cognitive capabilities fade. Commercial conversational systems like Siri, Google Assistant and Cortana simplify access to digital media, but they require technological skills often beyond the reach of the

senior population [6]. This population typically alleviates solitude by continuously consuming traditional broadcast media such as television and radio [7]. We have taken this behaviour into account in the design.

We are interested in user-centric approaches to close this digital gap. More specifically, this work involves the development of a user-friendly assistant to make Internet media accessible to elderly people. The idea we are pursuing is to create an assistant that will be perceived as an “intelligent radio” companion.

Chatbots appear to be a convenient solution. A chatbot, or chatting robot, is a system that enables human-to-computer interaction through text or natural language (NL) [8]. Ideally it should sustain human-like conversations, but this is clearly beyond the current state-of-the-art [9], even for chatbots based on Artificial Intelligence (AI) technologies such as Natural Language Generation (NLG).

In general, chatbots have three main components: a user interface, a normaliser and a knowledge base [10], [11]. Some famous chatbots are ALICE [12], SimSimi and Cleverbot. A common approach to chatbot development is to preprogram questions and expressions to respond to user input in the form of questions, phrases and words [13]. Additional content is automatically added by building knowledge on the conversations held.

Given the expected behaviour of the target users (who find radio and TV entertaining), we do not need the chatbot to be as intelligent and empathetic as a human. The idea is to intersperse the news items that the chatbot reads with light dialogues that act as news “connectors”, fostering engagement without demanding too much attention from the user. In brief, the chatbot serves as an intelligent radio capable of reading news and holding a limited dialogue. The chatbot applies Sentiment Analysis (SA) to user responses to adapt the dialogue flow to the user’s mood. This increases the value of chatbot-induced dialogues to gather information on user preferences. Using Google distance measurements [14], we demonstrate that even unfocused information extracted from these dialogues help conduct personalised information searches on the web in a manner that is transparent to end users.

Summing up, from a conceptual point of view, we present an “intelligent radio” by augmenting an information channel the elderly find familiar -background news- with interactions by voice dialogues. The system allows accessing digital content of interest, by combining words extracted from user answers to chatbot questions with keywords extracted from the news items. This permits defining metrics of the abstraction capabilities of the users depending on a spatial representation of the word space.

The rest of this paper is organised as follows. Section II discusses related work. Section III presents the proposed

solution. Section IV reports experimental tests with real users. Finally, Section V concludes the paper.

II. RELATED WORK

Recent AI advances have enabled chatbots to automatically generate responses to user requests [15]. As a result, they are now widely used in domains such as business [16], education [17], [18] and healthcare [19], [20]. However, they are still very far from mimicking human intelligence [9]. Given the current state-of-the-art, it is broadly agreed that chatbots should be oriented to highly focused tasks in specific application domains and that their interactions should be precise and short, as people find easier to communicate and engage for longer with short-message dialogue systems [21].

Chatbots are based on retrieval or generative models. Retrieval models apply predefined rules according to a rigid syntax. One popular technology for the rapid development of chatbot personalities is the Artificial Intelligence Markup Language (AIML) [16], [22]. Generative models, based on advanced deep learning models [23] combined with Natural Language Programming (NLP) [24]–[26], are more complex. Optimal approaches vary depending on the application domain and the demographics of the target users. From a syntactic point of view, AIML is almost error-free, but NLP techniques allow for more human-like NL dialogue flows. We decided to combine these two approaches to exploit their respective advantages.

The application in this paper is related to healthcare [23] and entertainment [27], two domains in which chatbot applications are still at an early stage. Online platforms for general medical care such as HealthTap, Oscar and Molly support interactions between patients and doctors. In [28] a solution for mental health diagnosis and clinician training on virtual patients was described, and in [29] an assistant chatbot for people with dementia was presented. The focus of chatbot applications for elderly people is typically therapeutic. One of the few exceptions was the personal assistant in the RobAlz project [30] that also reads stories and plays games with people with early-stage dementia. However, its dialogues may be considered canned-text since the chatbot employs predefined clauses and only invites users to continue talking. Conversely, we aim at providing *ad hoc* dialogues within the context of the news not only to entertain but also to increase the feeling of companionship of the “digital radio”.

Other entertainment chatbots [31]–[33] like EduRobot [34] have singing and storytelling capabilities but they have not been designed for an audience of elderly people. On the contrary, we have actively considered these target users into the design process thanks to the collaboration with Asociación de Familiares de enfermos de Alzheimer y otras demencias de Galicia (AFAGA) which provided us with gerontology expertise and valuable recommendations.

Available at <https://www.healthtap.com>, March 2021.

Available at <https://www.hioscar.com>, March 2021.

Available at <http://www.sensely.com>, March 2021.

Available at <https://afaga.com>, March 2021.

Available at <https://rebot.me/luisse-kate-barcelona>, March 2021.

Available at <https://www.cleverbot.com>, March 2021.

Even though newscaster chatbots are available [35], they have not been provided with empathy so far. Note that we are only considering chatbots. In the case of robots there has been plenty of research on empathetic capabilities [25], [36] and even commercial products are available. Sony's AIBO robotic dog [37], for example, could express happiness, anger, fear, sadness, surprise and dislike by means of tail and body movements and the colour and shape of its eyes. That is, robots typically express empathy by simulating pets, instead of by adapting generated language to user's expressions as in this work.

Considering the state-of-the-art in different research areas (health, entertainment, etc.), the main conceptual contribution of this work is a chatbot that has been realistically conceived from the onset for inclusive elderly entertainment through information delivery. It alternates newscasts with light dialogues about the news items that are adapted to the user's mood. Descriptions of news items are automatically extracted from these dialogues for content recommendation purposes. As far as we know, the extraction of search information from dialogues with elderly people to improve their abstraction capabilities –as shown in Section IV-B– has not been previously studied. We thus leverage state-of-the-art NLG and SA technologies to engage the user and reduce the digital divide. To prevent frustration the chatbot must provide flexible responses instead of repeatedly resorting to clauses such as 'Sorry, I didn't understand you' [38]. Some degree of chatbot personality is crucial for motivating user engagement [39]. Here is where SA and NLG are essential.

Regarding these underlying key technologies, NLG allows creating complete, coherent, human-like texts [40] from linguistic information (words, sentences, texts) and non-linguistic information (objective and visual data). NLG applications include automatic question generation [41], summaries [42] and text simplification [43], and their possibilities have multiplied thanks to Machine Learning. In particular, automatic text generation (whose only required inputs are meaningful words such as nouns, verbs and adjectives) has attracted great attention [44].

SA [45] is a hot field in social media research [46]–[48]. It allows inferring the polarity of a text (positive, negative or neutral). Among its many potential business applications [49], [50], it has been used to analyse consumer opinions in online media [51], politics [52] and sociology [53]. It can work at document [51], sentence [53] and even entity/aspect level [54].

III. SOLUTION DESIGN

The solution in this work is designed to entertain the elderly by leveraging the latest advances in different technologies. SA is the key building block for achieving the limited but key empathetic capabilities we need for the dialogue stages between newscasts. NLG allows varied, flexible and naturally sounding interactions with the user. To properly match keywords extracted from user speech and to keep conversation flowing and engage the user in an entertaining exchange,

the design considered personality, conversation flow, empathy/emotion, chatbot dialogue structure, graphics and accessibility. Regarding the latter, it is based on facial recognition instead of keywords, so that the system starts the dialogue when it detects that the user is present.

A. NEWS SERVICE

We chose a news service as the basis of the application because listening to the news, especially local ones, is a main pastime of elderly people. As previously mentioned, traditional entertainment media such as television and radio have a companionship effect [7].

We selected a varied range of topics that might be of interest to the target population: accessibility, environment, health, leisure, public services, retirement, social services, sport, technology and transport.

At the end, the news service allows the EBER bot engaging elderly people thanks to the varied and updated news presented to them, preventing monotony with continuous changes of dialogue topics. We remark that the dialogues are short and that their purpose is to "connect" news while, at the same time, obtaining user opinions and preferences. The system automatically gathers the news items from the Application Programming Interface (API) of the Spanish National Radio and Television (RTVE) channel with GET queries, using the *tematicas* and *noticias* services of the API, so that it is possible to extract news about certain topics.

B. NLG MODULE

Figure 1 shows the three-stage NLG module. It is based on our automatic text-to-text Spanish generator using keywords as input words [55], [56].

The first stage, the *Text Planner*, comprises content selection and text structuring tasks. Input data are meaningful words for creating an NL sentence in Spanish. The stage infers the best text structure from the information in the grammar. The grammar considers nominal and coordinated nominal, adjectival, adverbial and prepositional phrases.

The second *Sentence Planner* stage deals with lexicalisation, which consists of adding extra words automatically (once the syntactic structure has been inferred, it may be necessary to include determiners, prepositions and conjunctions). This is the reason for the feedback between the second and first stages in Figure 1.

The third stage, the *Realiser*, conducts morphological inflections to create a coherent and grammatically correct sentence in Spanish.

Thus, the main NLG actions in Figure 1 are:

- Syntax structure inference (*Text Planner*).
- Detection of affirmative, negative and interrogative sentences (*Sentence Planner*).
- Inclusion of subject (*Sentence Planner*).

Available at <https://www.rtve.es/api>, April 2021.

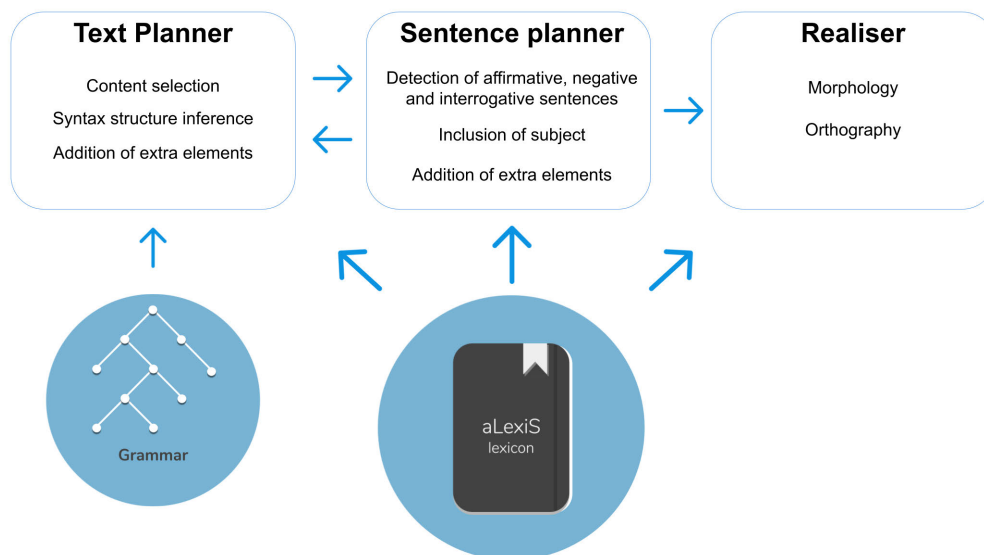


FIGURE 1. Tree-stage NLG module.

- Addition of extra elements (*Text Planner* and *Sentence Planner*).
- Morphological inflections (*Realiser*).
- Orthographic rules (*Realiser*).

These actions take linguistic knowledge from our *aLexiS* lexicon [55], [56] to adjust the sentence features (number, gender, person, tense). Specifically, the NLG module creates affirmative and negative clauses using keywords (verbs, for example), and adjective and noun phrases extracted from news content and user utterances. For that purpose, we use the spaCy parser. At the end, it creates complete, coherent and correct natural language sentences by applying the aforementioned NLG actions.

To illustrate the operation of the module, let us consider a complete example of text expansion. Assuming the user says *Me parece una noticia interesante* ‘I think this news item seems interesting’, the chatbot’s response is *Yo considero que no parece interesante* ‘I consider it does not seem interesting’. To create this reply, which is closely related to the previous user utterance, firstly the NLG module extracts the subject from the user clause *yo* ‘I’ and the main verb *parece* ‘seem’ (this linguistic information is taken from *aLexiS*). Thus, for short dialogues that are highly related to the news, the module can infer from the verbs that the users are expressing opinions. Secondly, the module extracts the adjective from the response of the user, *interesante* ‘interesting’.

C. SA MODULE

We have designed a three-stage SA module (Figure 2) to provide EBER with empathetic capabilities. It classifies short texts into three polarity levels (negative, neutral and positive). In order to evaluate the effectiveness of the module,

we trained some selected Machine Learning algorithms from the Scikit-Learn Python library, Gradient Descent, Decision Tree and Random Forest, on a dataset of 676 negative, 1,334 neutral and 918 positive clauses. Table 1 shows the features we considered: character and word *n*-grams [57] and three word dictionaries of deniers (adverbs with negative meaning), positive and negative words.

Then, we applied an attribute selector to extract the most relevant features. Specifically, we employed the SelectPercentile method from the Scikit-Learn Python library, as it outperformed other alternatives (SelectFromModel, SelectKBest, and RFECV). This method selects features according to a highest score percentile. We set a χ^2 score function and an 80th percentile threshold. Table 2 presents the performance of the Machine Learning algorithms. The Decision Tree classifier was the best alternative for the SA module. It attained 79.99% accuracy and 79.71% F_{macro} with 10-fold cross validation on class-balanced testing subsets. Figure 3 shows the scheme of this classifier. As previously said, the conversational system employs the output of the SA module to adjust the facial expression of its avatar to the user mood. Following the example in Section III-B in which the user expresses an opinion, the SA module detects that the clause *Me parece una noticia interesante* has positive polarity. Consequently, the avatar shows a happy face.

The NLG module employs SA knowledge to avoid monotony by adjusting the polarity of the dialogue depending on the polarity of user responses. With a configurable probability, they take opposite values, as in the example in

Available at <https://spacy.io>, March 2021.

Available at https://scikit-learn.org/stable/supervised_learning.html#supervised-learning, April 2021.

Available at https://scikit-learn.org/stable/modules/feature_selection.html, April 2021.

TABLE 1. Input features in the machine learning model.

Type	Feature name	Description
Textual	CHAR_GRAMS	Sequence of character n -grams
	WORD_GRAMS	Sequence of word n -grams
Numerical	NEG_ADVERBS	Amount of adverbs with negative meaning
	POS_WORDS	Amount of words with positive meaning
	NEG_WORDS	Amount of words with negative meaning

TABLE 2. Evaluation of machine learning algorithms for the SA-module.

Classifier	P_{macro}	F_{macro}	Accuracy
Gradient Descent	76.20%	74.11%	74.52%
Decision Tree	80.65%	79.71%	79.99%
Random Forest	74.82%	72.97%	73.49%

Section III-B. We remark that the dialogue is very light and that it is only intended to engage the users and obtain their opinion on the news. When the NLG module is unable to extract keywords or phrases to generate coherent empathetic text, it selects positive or negative sentences from a set of predefined templates based on the knowledge of the SA module.

D. CHATBOT DESIGN

We seek to assist elderly people with limited information abstraction capabilities, to help them find information of their interest (although this research would also be valid for other target groups with similar limitations). Traditional broadcast entertainment media has a proved “companionship” effect for the elderly [7]. This is why we focused on news complemented with short dialogues as content delivery format.

In the following sections we elaborate on the different modules of the EBER chatbot, which has been designed to reduce the digital gap for those target users. Accordingly, we have actively considered them into the design process.

Apart from well-known accessibility standards, first and foremost we have followed the advice of AFAGA. They have provided us with valuable design recommendations from their Cognitive Stimulation Program through New Technologies, an innovative experience to promote new technologies in devices for the elderly.

1) AIML CHATBOT PERSONALITY

Since the purpose of the chatbot is well defined (to obtain opinions about news from elderly people through light conversations), we decided to create our own AIML-based knowledge base rather than reuse an existing chatbot. We leveraged AIML features such as *star*, *srai* and *random* as well as pattern extraction with keywords. This allows the chatbot starting the conversation with a controlled dialogue flow, using short questions about daily routines and mood. These questions

Available at <https://afaga.com/es/servicios/taller-de-nuevas-tecnologias>, April 2021.

are extracted from predefined templates as described in sections III-D2 and III-D5.

2) CONVERSATION DECISIONS

Table 3 shows some examples of default answers to keep conversation flowing and prevent frustration. When the EBER chatbot detects that the user cannot understand something, it asks him/her to explain why. By doing so the conversation feels more natural and the user feels more involved.

Precise, short chatbot interactions are adequate for our application domain. They are also beneficial for news recommendation purposes, since they reduce meaningless concepts in the dialogues and allow the conversation to move from general to more specific topics.

3) ACCESSIBILITY AND GRAPHICS

In this regard, the system is simply activated through facial recognition (using the OpenCV library trained with an eye-detection dataset) or voice commands. All user interactions with the system are by voice, which is considered the most natural access interface for the elderly [11] and thus fosters a more engaging conversation experience [58]. For this purpose we transform Spanish speech into text and vice versa, as input and output of the conversational system, with the Google Voice Android Software Development Kit (SDK) library.

Furthermore, we keep the graphic appearance of the chatbot as simple as possible by following accessibility conventions for people with vision and hearing impairments [59], [60] (full screen mode, colour contrast, capitalisation, font and icon size, volume configuration, etc.). Additionally, we use visual references to guide the user through the dialogue stage. For example, since previous research has proved the benefits of using pets as emotional avatars for the elderly [61], the EBER chatbot interface is an animated dog, and the system presents to the user the lead paragraph of a news item as a summary.

4) EMPATHY AND EMOTION

The SA module allows guessing the user mood from his/her answers and provides feedback to the conversational system.

Available at <https://opencv.org>, April 2021.

Available at https://github.com/opencv/opencv/blob/master/data/lbpcascades/lbpcascade_frontalface.xml, April 2021.

Available at <https://developer.android.com/reference/android/speech/SpeechRecognizer>, April 2021.

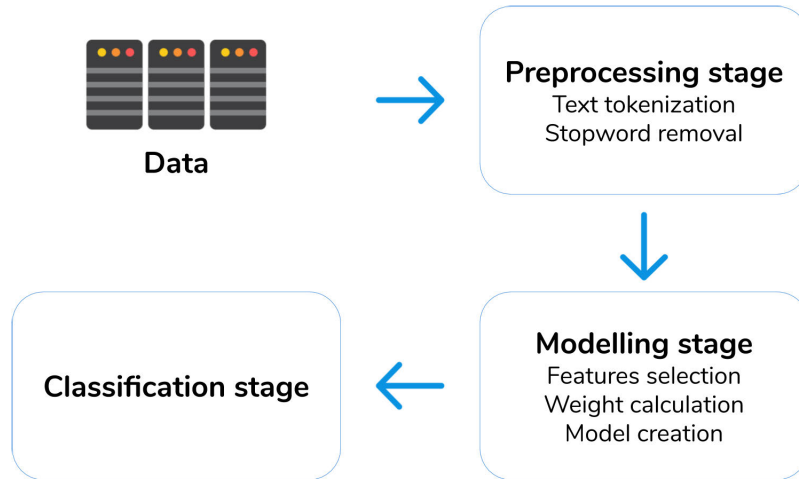


FIGURE 2. Flow diagram of the SA module.

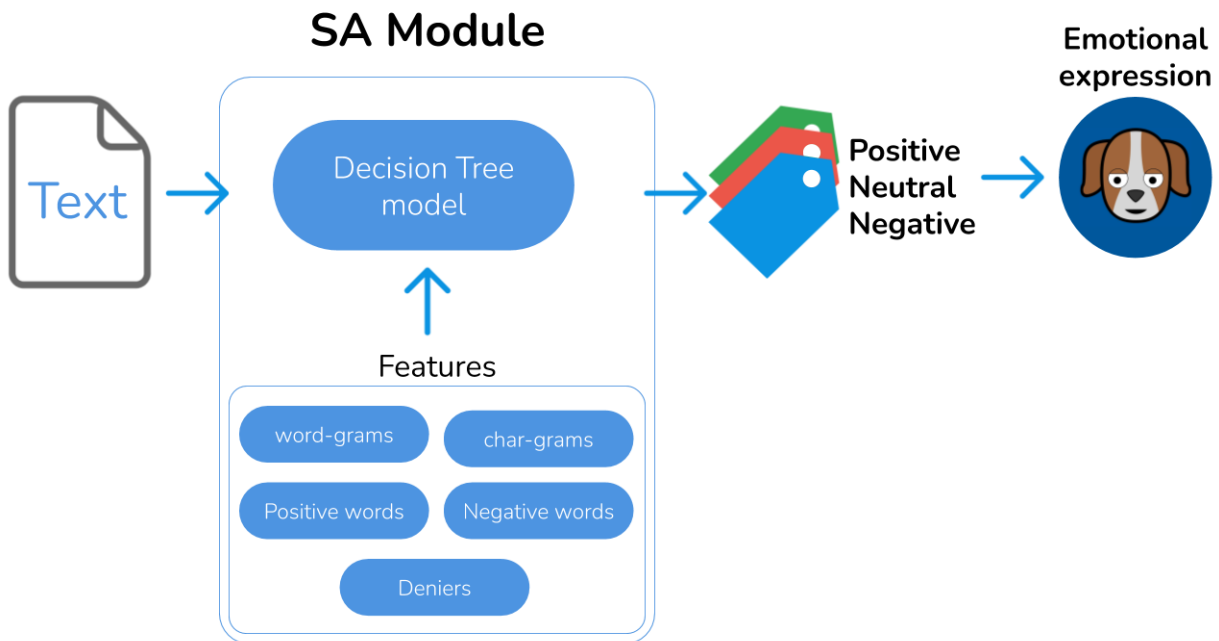


FIGURE 3. Decision Tree classifier scheme.

Based on SA knowledge, the NLG module presents questions to the user that are aligned with his/her opinion or refute it with configurable probability. Furthermore, the system adjusts the facial expression of its avatar according to the user mood. For example, the dog’s face will become sad if EBER detects a negative emotion.

5) DIALOGUE STRUCTURE

During the dialogue stage, we keep the interactions of the system precise and short due to the fact that people, and especially the elderly, find short dialogue systems simpler and keep engaged for longer [21].

Ontologies and similar models only yield good dialogue structuring results when the expected variation is minimal. In the target scenario, however, users are free to express any ideas about the news they are presented with. To handle this disparity we combined AIML, NLG and SA. Dialogues are structured into two parts: AIML-based interactions and NLG-based replies. When the EBER chatbot starts a dialogue, it randomly selects a clause for the first interaction (see sample of options in Table 4). User responses in turn are expected to contain words such as *bien* ‘well’ (positive answers) or *mal* ‘bad’ (negative answers) (see interaction 2 in the example in Table 4). The chatbot answer depends on the polarity of the user’s answer in the previous interaction (see interaction 3 in

TABLE 3. Default answers and respective chatbot connectors.

Option	Default answers
1	<i>Eres una persona muy interesante</i> ‘You’re a really interesting person’
2	<i>Me gustaría conocerte más</i> ‘I’d like to get to know you better’
3	<i>Me gustaría saber más de tí</i> ‘I’d like to know more about you’
4	<i>Cuéntame más</i> ‘Tell me more’
Option	Connector
1	<i>¿Usas mucho el teléfono móvil?</i> ‘Do you use your cell phone a lot?’
	<i>¿Tuviste alguna mascota?</i> ‘Have you ever had a pet?’
	<i>¿Cuál fue tu primer empleo?</i> ‘Which was your first job?’ [...]
2	<i>¿Cuál es tu color favorito?</i> ‘What’s your favourite colour?’
	<i>¿Cuál es tu deporte favorito?</i> ‘What’s your favourite sport?’
	<i>¿Te gusta el fútbol?</i> ‘Do you like football?’ [...]
3	<i>¿Te gusta hacer deporte?</i> ‘Do you like practising sport?’
	<i>¿Te gusta ir de compras?</i> ‘Do you like shopping?’
	<i>¿Te gusta jugar a las cartas?</i> ‘Do you like playing cards?’ [...]
4	<i>¿Te gusta leer?</i> ‘Do you like reading?’
	<i>¿A qué jugabas de pequeño?</i> ‘Which games did you play when you were a child?’ [...]

Table 4). The chatbot completes its answer with a dialogue connector (see examples in Table 5).

Before reading a news item, the chatbot announces its intention by stating *Ahora te voy a mostrar una noticia, pero antes de empezar...* ‘Now I’m going to read some news to you, but before we start...’ followed by a sentence like:

- *Tengo una duda* ‘I have a question’
- *Tengo curiosidad por saber tu opinión* ‘I’m curious about your opinion’

Once the chatbot receives the user’s answer it says *Vale, vamos entonces con noticias relacionadas con esta temática* ‘OK, then let’s listen to some news on this topic’ and reads the news item.

Finally, the NLG and SA modules support the exchange between the chatbot and the user, providing EBER with valuable information for future recommendations on news items.

Figure 5 shows a real example of a whole dialogue between EBER and a senior taken from the experiments in Section IV.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In the tests an Ubuntu server with 12 cores and 64GB RAM provided remote SA and NLG capabilities to a user handheld device. The user interface was presented on a Huawei MediaPad T5 tablet with 8 cores, 3GB RAM, 32GB hard disk and Android 8 operating system. Figure 6 shows all the components in the EBER testbed.

The experiments involved 31 users (20 women and 11 men), 75.5 ± 6.95 years old (average \pm standard deviation). In the sequel we will call a session a particular newscast and its associated dialogue. The average number of sessions per user was 4.33. Of the 31 participants, 10 had some basic technology skills (such as experience with Google or WhatsApp), 8 had hearing problems, 23 were highly focused during the experiments and 5 felt awkward or stressed during the interactions.

After the dialogue, users were asked if they felt confused, stressed or focused (yes/no in all cases). After each independent session, the users were asked which keywords they would have used to search for similar news in a news repository (logically, users with technology skills immediately identified this exercise as a Google search). When the experiment concluded, users rated their experience, in terms of satisfaction, amazement and chatbot-human likeliness on a 5-level Likert scale.

A. USER SATISFACTION AND BEHAVIOUR

User satisfaction with the service was close to 4 and the perception of chatbot-human likeliness was close to 3 on average. This indicates that even though AI technology is not yet mature enough to mimic a human conversation, EBER was entertaining to the users during the tests (in this regard, 4 users stated that they were amazed by the capability of the chatbot to adapt itself to the user mood. Table 6 presents the responses of these users about empathy). This encouraged us to consider the “intelligent radio” approach realistic and potentially useful. However, to further validate these results we studied the behaviours of the users during the tests to check if they were coherent with the demographics of the sample and rule out user bias from the experiment setup.

In this study we checked pairwise correlations between different behavioural variables using the Pearson correlation coefficient [62] (see Expression (1), where x and y are two generic variables). This metric varies between -1 and 1 and is closer to -1 if the relationship is inverse and closer to 1 if it is direct. We considered that two variables were possibly related for absolute correlations around 0.4 or above, since the goal was to obtain preliminary but valuable insights.

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2} \sqrt{\sum(y_i - \bar{y})^2}} \quad (1)$$

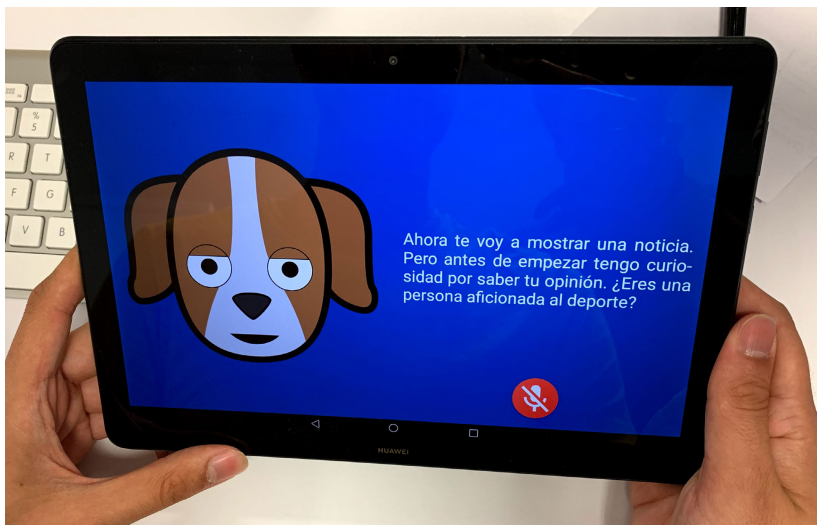


FIGURE 4. EBER chatbot interface. The muted microphone indicates that it is EBER’s turn to speak.

TABLE 4. Dialogue flow interactions, sample.

Int. #	Speaker	Options
1	Chatbot	<i>¿Has dormido bien?</i> ‘Did you sleep well?’
		<i>¿Cómo te encuentras hoy?</i> ‘How are you today?’
		<i>¿Cómo te sientes hoy?</i> ‘How are you feeling today?’
		<i>¿Qué tal te encuentras?</i> ‘How are you doing?’
		<i>¿Qué tal estás?</i> ‘How are you?’
2	User	<i>¿Qué tal está tu familia?</i> ‘How’s your family?’
		<i>^ bien</i> ‘well’ ^
		<i>^ genial</i> ‘great’ ^
		<i>^ fantástico</i> ‘awesome’ ^
		<i>^ fantástica</i> ‘awesome’ ^
		<i>^ mal</i> ‘bad’ ^
		<i>^ no muy bien</i> ‘not very well’ ^
<i>^ fatal</i> ‘really bad’ ^		
<i>^ horrible</i> ‘terrible’ ^		
3	Chatbot	<i>Me alegro por ti</i> ‘I’m happy for you’
		<i>¡Eso es genial!</i> ‘That’s great!’
		<i>¡Eso está muy bien!</i> ‘That’s brilliant!’
		<i>Bueno, poco a poco</i> ‘OK, step by step’
		<i>Espero hacerte sentir mejor</i> ‘I hope I’ll make you feel better’
		<i>¡Ánimo!</i> ‘Cheer up!’

TABLE 5. Chatbot connectors, example.

Connectors
<i>Háblame de ti</i> ‘Tell me about yourself’
<i>Cuéntame cosas sobre ti</i> ‘Tell me something about you’
<i>¿Qué tienes pensado hacer hoy?</i> ‘What plans have you got today?’
<i>¿Darás un paseo hoy?</i> ‘Are you going to go for a walk today?’
<i>¿Has quedado con algún amigo?</i> ‘Are you meeting a friend?’
<i>¿Has quedado con alguien?</i> ‘Are you meeting anybody?’
<i>¿Qué harás el fin de semana?</i> ‘What are you going to do at the weekend?’
<i>¿A dónde tienes pensado ir hoy?</i> ‘Where are you going today?’
<i>¿De qué trató el último sueño que tuviste?</i> ‘What was your last dream about?’
<i>¿En qué estás pensando ahora mismo?</i> ‘What are you thinking right now?’

Figure 7 shows the correlations between the behavioural variables in the study (user focus, stress and confusion; chatbot-human likeliness; level of amazement with the service; and user satisfaction) and the demographics of the sam-

ple. Note that gender was uncorrelated with behaviour. Age only seemed to be correlated with opinion on chatbot-human likeliness and user satisfaction: the younger people in the sample were more open to technological novelty while the

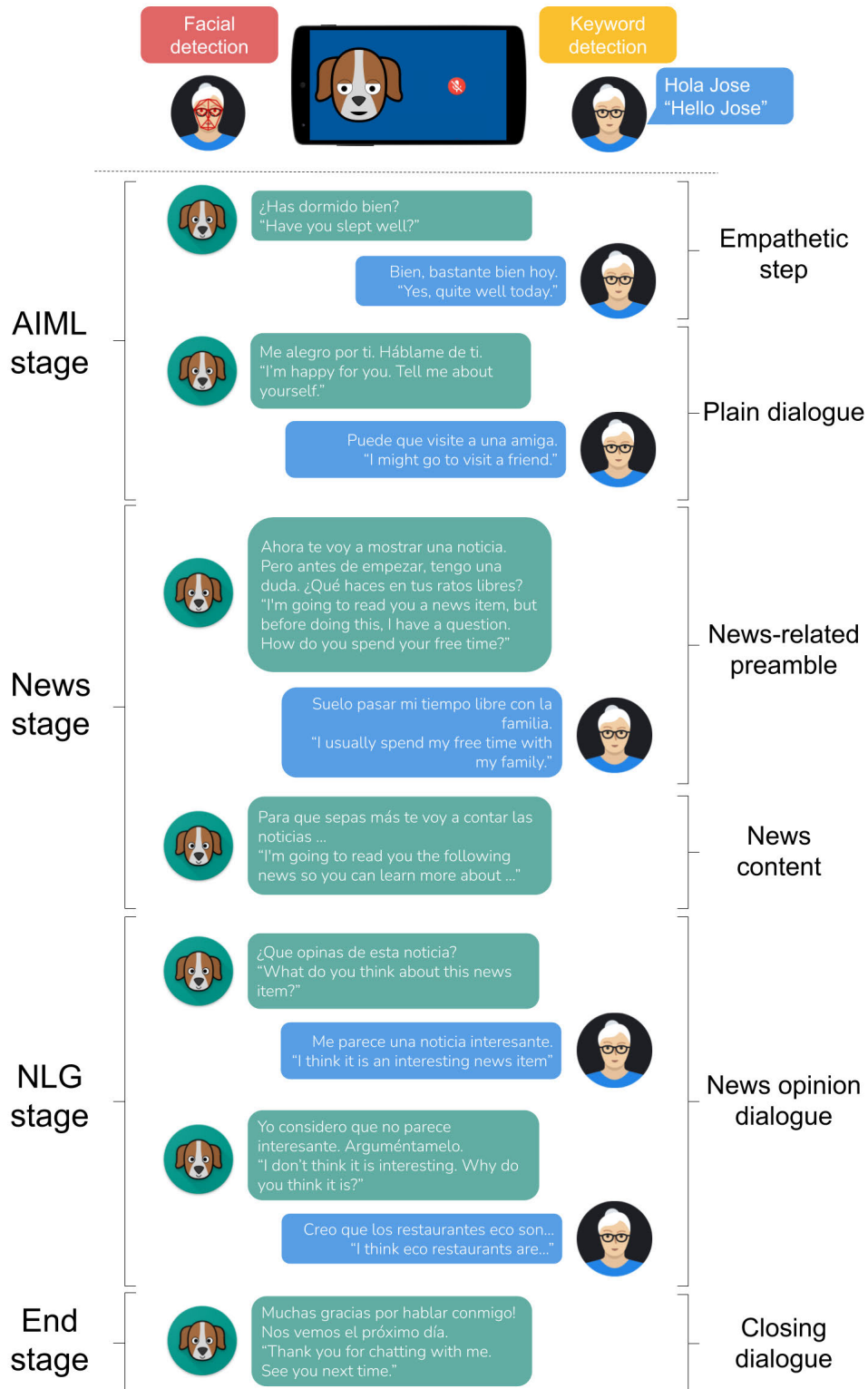


FIGURE 5. EBER chatbot flow diagram. Example of a real dialogue.

older people found the system more entertaining (which is especially interesting to us).

As expected, people with hearing impairments were more confused, less surprised and less satisfied. The level of technology skills had a productive relationship with most of the

behavioural variables (the higher that level, the more the focus, amazement and satisfaction and the less stress and confusion). However, it had no clear relationship with opinions on chatbot-human likeness, which suggests that most users were objective regarding this aspect.

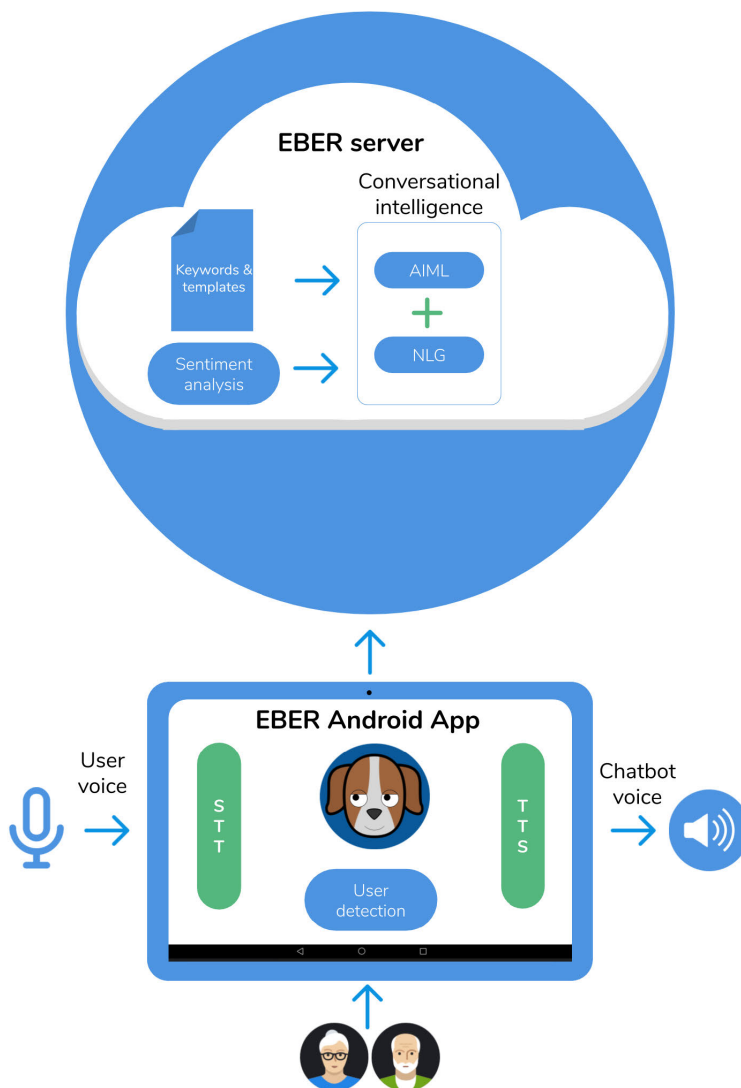


FIGURE 6. EBER testbed.

TABLE 6. User answers about empathy of the EBER chatbot.

User ID: 3	Is there anything that you found especially surprising or useful while using the bot? <i>The understanding of the bot.</i>
User ID: 5	Is there anything that you found especially surprising or useful while using the bot? <i>The system moves its eyes and mouth to show a frame of mind.</i>
User ID: 16	Is there anything that you found especially surprising or useful while using the bot? <i>The systems detects the mood of my opinions.</i>
User ID: 17	Is there anything that you found especially surprising or useful while using the bot? <i>The chatbot understands me and adapts itself to my answers.</i>

Summing up, the most important insight from these measurements was that user satisfaction scores were coherent with expected behaviour, thus supporting their validity.

B. EXPLICIT VERSUS IMPLICIT CONTENT DESCRIPTION

We measured the ability of the users to describe the content they were presented with by calculating the Normalised

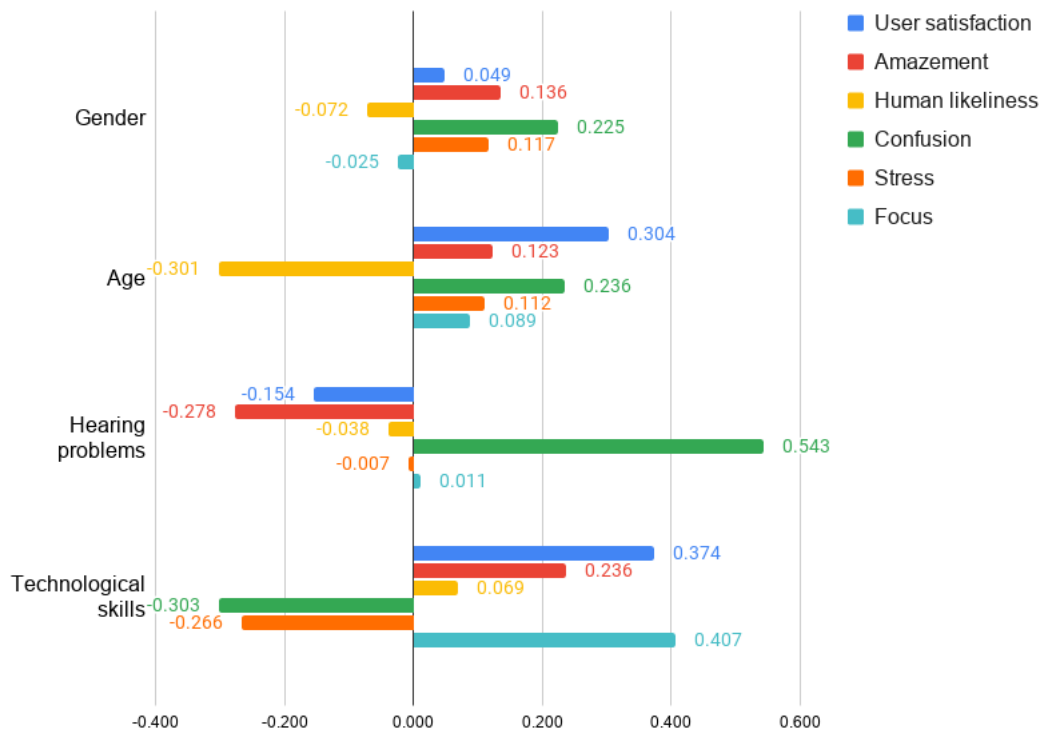


FIGURE 7. Correlations between behavioural variables and demographics of the sample.

Google distance (NGD) [14] between news metadata and the words the users explicitly chose to search for news when asked to do so. In this regard, some people tended to describe news with lengthy clauses, probably because it was difficult for them to summarise concepts with specific words [63]. In such cases we segmented clauses into their word constituents. We discarded stopwords and applied lemmatisation procedures, and the resulting words were taken as separate search terms. For example, the four keywords ‘good’, ‘beautiful’, ‘cheap’ and ‘restaurants’ would be extracted from ‘good, beautiful and cheap restaurants’.

The differences among users were notable. Focused users with technology skills selected effective words such as ‘social services, Spain, Galicia’ for news about nursing homes in Galicia, whereas stressed or confused users provided vague terms such as ‘walk carefully, elderly person, illnesses, too old’. Therefore, we divided the sample into three groups. The first group contained the seven users that were focused and relaxed, not confused, and had good hearing, and the second and third groups, with eleven and thirteen users respectively, everybody else. The tests with the second group, which had similar demographics as the other two (age statistics of 76 ± 6.75 years in the second group versus 75.23 ± 7.07 years in the other two) were performed several months later, the main difference being that the users in the third group had worse technological skills.

Table 7 shows the NGD values for user-provided explicit keywords by group. The gold standard was given by young researchers in our group with high web search skills (average

TABLE 7. NGD values, user-provided explicit keywords.

	Group 1	Group 2	Group 3
	0.61	0.64	0.54
	0.50	0.59	0.31
	0.55	0.60	0.74
	0.56	0.60	2.73
	0.58	0.69	0.56
	0.58	0.62	2.23
	0.52	0.54	0.85
		0.62	0.57
		0.63	0.45
		0.59	0.48
		0.59	0.49
			0.48
			0.49
AVG	0.56	0.61	0.84

NGD of 0.47 for the same news). The average NGD was 0.56 (only 0.09 above the gold standard) in the first user group. The technological skills of the users in the second group mitigated discomfort, and their average NGD was only a bit worse (0.61 versus 0.56). Of the 13 users in the third group, 4 had very high NGD values. Even after discarding the terms that were not directly related to the news’ topic, the average NGD was still high, of 0.84. As in the previous section, we checked the coherence of these results by analysing correlations. In particular, we found positive correlations between NGD values and stress and focus levels. More specifically, stressed and focused users had respective correlations of 0.58 and -0.44 .

TABLE 8. NGD values, by adding keywords from semantic knowledge.

	Group 1	Group 2	Group 3
	0.55	0.72	0.51
	0.46	0.69	0.47
	0.53	0.60	0.57
	0.47	0.63	2.00
	0.49	0.52	0.52
	0.48	0.55	0.47
	0.58	0.49	0.75
		0.56	0.52
		0.46	0.47
		0.54	0.58
		0.52	0.53
			0.55
			0.44
AVG	0.51	0.57	0.64

TABLE 9. Summary of average NGD values: Gold Standard (GS), group 1, group 2 and group 3.

	GS	Group 1	Group 2	Group 3
Explicit keywords	0.47	0.56	0.61	0.84
Explicit keywords + semantic knowledge	-	0.51	0.57	0.64

Finally, we evaluated the knowledge about the news in the dialogue between the chatbot and the users (see example in Figure 5). For this purpose we identified the semantic category of the words in the news (excluding stopwords) using the Multilingual Central Repository (MCR) tool [64]. Then we extracted from the dialogues the words with the highest semantic similarity with the content of the news. We considered there was a match between the words provided by the users and the words in the news when their MCR semantic categories were the same. Next, we built a new set with all the semantically related words (if there was no semantic match, we just employed the user’s response). The NGD was recomputed with this set (Table 8). Table 9 summarises the results of both experiments. Compared to the gold standard, the average NGD decreased from 179% to 136% (−0.2), 130% to 121% (−0.04) and 119% to 109% (−0.05) in groups 3, 2 and 1, respectively.

We remark that the results were consistent over time: even though the tests with user group 2 (with similar demographics) were performed much later, they also reflected a systematic improvement in news abstraction capabilities by automatically adding knowledge. In any case, the benefits of the system were more evident in users without technological skills who experienced some discomfort.

We have therefore not only proved that a specialised chatbot can engage elderly people as an “intelligent newscaster radio”, but also that it can extract keywords that help to personalise subsequent news searches if the user expresses positive sentiments about previous news.

C. USER BENEFITS

From a holistic perspective, the system may open a world of digital information to the users through a familiar channel (background voice or “radio”) with a friendly interface (voice chatbot), and it may also provide value as a tool to measure the evolution of their abstraction capabilities (and thus, their cognitive abilities). It can also enable non-intrusive monitoring from families and caregivers.

During the experiments, the users were also asked about the aspects they liked or disliked. Remarkably, in addition to human-machine communication features (visual indications of user’s turn to speak, chatbot empathy and front-end avatar), the newscaster functionality of the system was praised. Regarding the aspects the users disliked, confused users were particularly baffled by chatbot interruptions when they paused for too long. This revealed that we need to develop better ways to detect when a user has finished his/her turn.

Overall, given the high user satisfaction scores, we concluded that the “intelligent radio” approach has the potential to help elderly users by improving their abstraction capabilities, which narrows the digital gap.

V. CONCLUSION

In this work we have presented the EBER intelligent chatbot. To the best of our knowledge this is the first system to combine AIML, NLG and SA to generate short coherent contextualised dialogues as connectors between newscasts. Thanks to this combination, EBER behaves realistically as an “intelligent radio” for entertaining elderly people.

In the tests, 80% of the users gave the system a satisfaction score of 4 out of 5. The analysis of correlations between behavioural variables and sample demographics was coherent, further supporting the validity of the user satisfaction scores. By automatically extracting knowledge from connecting dialogues with positive sentiment, the system improves content characterisation, even for unfocused, stressed or confused people.

In forthcoming work, we will explore how to improve turn detection to avoid speech interruptions and add advanced personalisation features with better directed dialogue stages. Moreover, we will consider an experimental assessment of loneliness and study how the quality of users’ responses (regarding user words that complement news keywords for subsequent searches) evolves in time, by paying special attention to differences in speech related to age. We theorise that quality will not worsen thanks to the positive reinforcement users experience when new newscasts match their interests.

It may also be interesting to study the effect of other forms of language semantics such as sarcasm [65].

ACKNOWLEDGMENT

The authors are indebted to Asociación de Familiares de enfermos de Alzheimer y otras demencias de Galicia (AFAGA) for providing gerontology expertise and valuable design recommendations.

REFERENCES

- [1] P. Caligiuri and M. Lazarova, "A model for the influence of social interaction and social support on female expatriates' cross-cultural adjustment," *Int. J. Hum. Resource Manage.*, vol. 13, no. 5, pp. 761–772, Jan. 2002.
- [2] A. L. Blaakilde, "The becoming of radio bodies," *Eur. J. Cultural Stud.*, vol. 21, no. 3, pp. 290–304, Jun. 2018.
- [3] S. J. Czaja and C. C. Lee, "The impact of aging on access to technology," *Universal Access Inf. Soc.*, vol. 5, no. 4, pp. 341–349, Mar. 2007.
- [4] J. Favela and L. A. Castro, "Technology and aging," in *Aging Research-Methodological Issues*. Los Angeles, CA, USA: Springer, 2015, pp. 121–135.
- [5] A. Fleming, C. Mason, and G. Paxton, "Discourses of technology, ageing and participation," *Palgrave Commun.*, vol. 4, no. 1, p. 54, Dec. 2018.
- [6] T. Kleinberger, M. Becker, E. Ras, A. Holzinger, and P. Müller, *Ambient Intelligence in Assisted Living: Enable Elderly People to Handle Future Interfaces* (Lecture Notes in Computer Science), vol. 4555, no. 2. Berlin, Germany: Springer, 2007, pp. 103–112.
- [7] B. Östlund, "Watching television in later life: A deeper understanding of TV viewing in the homes of old people and in geriatric care contexts," *Scandin. J. Caring Sci.*, vol. 24, no. 2, pp. 233–243, Dec. 2009.
- [8] R. P. Mahapatra, N. Sharma, A. Trivedi, and C. Aman, "Adding interactive interface to E-Government systems using AIML based chatterbots," in *Proc. CSI 6th Int. Conf. Softw. Eng. (CONSEG)*, Sep. 2012, pp. 1–6.
- [9] M. Skjuve, I. M. Haugstveit, A. Følstad, and P. B. Brandtzaeg, "Help! Is my chatbot falling into the uncanny valley? An empirical study of user experience in human-chatbot interaction," *Hum. Technol.*, vol. 15, no. 1, pp. 30–54, Feb. 2019.
- [10] S. A. Abdul-Kader and D. John, "Survey on chatbot design techniques in speech conversation systems," *Int. J. Adv. Comput. Sci. Appl.*, vol. 6, no. 7, pp. 72–80, 2015.
- [11] W.-D. Liu, K.-Y. Chuang, and K.-Y. Chen, "The design and implementation of a Chatbot's character for elderly care," in *Proc. Int. Conf. Syst. Sci. Eng. (ICSSE)*, Jun. 2018, pp. 1–5.
- [12] B. AbuShawar and E. Atwell, "ALICE chatbot: Trials and outputs," *Computación y Sistemas*, vol. 19, no. 4, pp. 625–632, Dec. 2015.
- [13] S. Gupta, D. Borkar, C. De Mello, and S. Patil, "An e-commerce Website based chatbot," *Int. J. Comput. Sci. Inf. Technol.*, vol. 6, no. 2, pp. 1483–1485, 2015.
- [14] R. L. Cilibrasi and P. M. B. Vitanyi, "The Google similarity distance," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 3, pp. 370–383, Mar. 2007.
- [15] A. Xu, Z. Liu, Y. Guo, V. Sinha, and R. Akkiraju, "A new chatbot for customer service on social media," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, May 2017, pp. 3506–3510.
- [16] N. T. Thomas, "An e-business chatbot using AIML and LSA," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2016, pp. 2740–2742.
- [17] B. Robins, K. Dautenhahn, R. T. Boekhorst, and A. Billard, "Robotic assistants in therapy and education of children with autism: Can a small humanoid robot help encourage social interaction skills?" *Universal Access Inf. Soc.*, vol. 4, no. 2, pp. 105–120, Dec. 2005.
- [18] D. Carlander-Reuterfelt, A. Carrera, C. A. Iglesias, O. Araque, J. F. S. S. Rada, and S. Munoz, "JAICOB: A data science chatbot," *IEEE Access*, vol. 8, pp. 180672–180680, 2020.
- [19] V. M. Kumar and A. Keerthana, "Sanative chatbot for health seekers," *Int. J. Eng. Comput. Sci.*, vol. 5, no. 3, pp. 16022–16025, Mar. 2016.
- [20] X. Ren, G. Spina, S. De Vries, A. Bijkerk, B. Faber, and A. Geraedts, "Understanding physician's experience with conversational interfaces during occupational health consultation," *IEEE Access*, vol. 8, pp. 119158–119169, 2020.
- [21] J. Hill, W. Randolph Ford, and I. G. Ferreras, "Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations," *Comput. Hum. Behav.*, vol. 49, pp. 245–250, Aug. 2015.
- [22] A. M. Galvao, F. A. Barros, A. M. Neves, and G. L. Ramalho, "Persona-AIML: An architecture for developing chatterbots with personality," in *Proc. 3rd Int. Joint Conf. Auto. Agents Multiagent Syst. (AAMAS)*, 2004, pp. 1266–1267.
- [23] M.-H. Su, C.-H. Wu, K.-Y. Huang, Q.-B. Hong, and H.-M. Wang, "A chatbot using LSTM-based multi-layer embedding for elderly care," in *Proc. Int. Conf. Orange Technol. (ICOT)*, Dec. 2017, pp. 70–74.
- [24] C. J. Baby, F. A. Khan, and J. N. Swathi, "Home automation using IoT and a chatbot using natural language processing," in *Proc. Innov. Power Adv. Comput. Technol. (i-PACT)*, Apr. 2017, pp. 1–6.
- [25] K.-J. Oh, D. Lee, B. Ko, and H.-J. Choi, "A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation," in *Proc. 18th IEEE Int. Conf. Mobile Data Manage. (MDM)*, May 2017, pp. 371–376.
- [26] T.-H. Wen, N. Mrkšić, P.-H. Su, D. Vandyke, and S. Young, "Semantically conditioned LSTM-based natural language generation for spoken dialogue systems," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2015, pp. 1711–1721.
- [27] S. Noh, J. Han, J. Jo, and A. Choi, "Virtual companion based mobile user interface: An intelligent and simplified mobile user interface for the elderly users," in *Proc. Int. Symp. Ubiquitous Virtual Reality (ISUVR)*, Jun. 2017, pp. 8–9.
- [28] P. Kenny, T. Parsons, J. Gratch, and A. Rizzo, "Virtual humans for assisted health care," in *Proc. 1st ACM Int. Conf. Pervas. Technol. Rel. Assistive Environ. (PETRA)*, 2008, p. 6.
- [29] K. Yasuda, M. Fuketa, and J. Aoe, "An anime agent system for reminiscence therapy," *Gerontechnology*, vol. 13, no. 2, pp. 118–119, Jun. 2014.
- [30] M. A. Salichs, I. P. Encinar, E. Salichs, Á. Castro-González, and M. Malfaz, "Study of scenarios and technical requirements of a social assistive robot for Alzheimer's disease patients and their caregivers," *Int. J. Social Robot.*, vol. 8, no. 1, pp. 85–102, Jan. 2016.
- [31] D. O. Johnson, R. H. Cuijpers, K. Pollmann, and A. A. J. van de Ven, "Exploring the entertainment value of playing games with a humanoid robot," *Int. J. Social Robot.*, vol. 8, no. 2, pp. 247–269, Apr. 2016.
- [32] F. Correia, T. Ribeiro, P. Alves-Oliveira, N. Maia, F. S. Melo, and A. Paiva, "Building a social robot as a game companion in a card game," in *Proc. 11th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2016, p. 563.
- [33] I. Aaltonen, A. Arvola, P. Heikkilä, and H. Lammi, "Hello Pepper, may i tickle you?: Children's and adults' responses to an entertainment robot at a shopping mall," in *Proc. Companion ACM/IEEE Int. Conf. Hum.-Robot Interact.* Washington, DC, USA: IEEE Computer Society, Mar. 2017, pp. 53–54.
- [34] W. Budiharto, A. D. Cahyani, P. C. B. Rumondor, and D. Suhartono, "EduRobot: Intelligent humanoid robot with natural interaction for education and entertainment," *Procedia Comput. Sci.*, vol. 116, pp. 564–570, Jan. 2017.
- [35] R. Matsumoto, H. Nakayama, T. Harada, and Y. Kuniyoshi, "Journalist robot: Robot system making news articles from real world," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2007, pp. 1234–1241.
- [36] A. Augello, G. Saccone, S. Gaglio, and G. Pilato, "Humorist bot: Bringing computational humour in a chat-bot system," in *Proc. Int. Conf. Complex, Intell. Softw. Intensive Syst.*, 2008, pp. 703–708.
- [37] J. Pransky, "AIBO—The No. 1 selling service robot," *Ind. Robot, Int. J.*, vol. 28, no. 1, pp. 24–26, Feb. 2001.
- [38] M.-C. Jenkins, R. Churchill, S. Cox, and D. Smith, "Analysis of user interaction with service oriented chatbot systems," in *Proc. Int. Conf. Human-Comput. Interact.*, 2007, pp. 76–83.
- [39] P. Y. Andrews, "System personality and persuasion in human-computer dialogue," *ACM Trans. Interact. Intell. Syst.*, vol. 2, no. 2, p. 12, 2012.
- [40] E. Reiter and R. Dale, "Building natural language generation systems," *Language*, vol. 77, no. 3, pp. 611–612, 2000.
- [41] H. Wang, C. J. Waple, and T. Kawahara, "Computer assisted language learning system based on dynamic question generation and error prediction for automatic speech recognition," *Speech Commun.*, vol. 51, no. 10, pp. 995–1005, Oct. 2009.
- [42] P.-E. Genest, "Framework for abstractive summarization using text-to-text generation," in *Proc. Workshop Monolingual Text Text Gener.*, 2011, pp. 64–73.
- [43] E. Rennes and A. Jönsson, "A tool for automatic simplification of swedish texts," in *Proc. 20th Nordic Conf. Comput. Linguistics (NODALIDA)*, 2015, pp. 317–320.
- [44] S. García-Méndez, M. Fernández-Gavilanes, E. Costa-Montenegro, J. Juncal-Martinez, F. J. González-Castaño, and E. Reiter, "A system for automatic english text expansion," *IEEE Access*, vol. 7, pp. 123320–123333, 2019.
- [45] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 102–107, Mar. 2016.
- [46] T. H. Nguyen, K. Shirai, and J. Velcin, "Sentiment analysis on social media for stock movement prediction," *Expert Syst. Appl.*, vol. 42, no. 24, pp. 9603–9611, Dec. 2015.

- [47] F. Poecze, C. Ebster, and C. Strauss, "Social media metrics and sentiment analysis to evaluate the effectiveness of social media posts," in *Procedia Computer Science*, vol. 130. Amsterdam, The Netherlands: Elsevier, 2018, pp. 660–666.
- [48] M. K. Sohrabi and F. Hemmatian, "An efficient preprocessing method for supervised sentiment analysis by converting sentences to numerical vectors: A Twitter case study," *Multimedia Tools Appl.*, vol. 78, pp. 1–20, Sep. 2019.
- [49] A. P. Kirilenko, S. O. Stepchenkova, H. Kim, and X. R. Li, "Automated sentiment analysis in tourism: Comparison of approaches," *J. Travel Res.*, vol. 57, no. 8, pp. 1012–1025, Nov. 2018.
- [50] V. Vyas and V. Uma, "Approaches to sentiment analysis on product reviews," in *Sentiment Analysis and Knowledge Discovery in Contemporary Business*. Philadelphia, PA, USA: IGI Global, 2019, pp. 15–30.
- [51] R. K. Chaurasiya and U. Sahu, "Improving performance of product recommendations using user reviews," in *Proc. 3rd Int. Conf. Workshops Recent Adv. Innov. Eng. (ICRAIE)*, Nov. 2018, pp. 1–4.
- [52] S. Martín-Gutiérrez, J. C. Losada, and R. M. Benito, "Semi-automatic training set construction for supervised sentiment analysis in political contexts," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2018, pp. 715–720.
- [53] A. B. Eliacik and N. Erdogan, "Influential user weighted sentiment analysis on topic based microblogging community," *Expert Syst. Appl.*, vol. 92, pp. 403–418, Feb. 2018.
- [54] Y. Ma, H. Peng, and E. Cambria, "Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 5876–5883.
- [55] S. García-Méndez, M. Fernández-Gavilanes, E. Costa-Montenegro, J. Juncal-Martínez, and F. J. González-Castaño, "Automatic natural language generation applied to alternative and augmentative communication for online video content services using SimpleNLG for spanish," in *Proc. 15th Int. Web All Conf.*, Apr. 2018, pp. 1–4.
- [56] S. García-Méndez, M. Fernández-Gavilanes, E. Costa-Montenegro, J. Juncal-Martínez, and F. J. González-Castaño, "A library for automatic natural language generation of spanish texts," *Expert Syst. Appl.*, vol. 120, pp. 372–386, Apr. 2019.
- [57] T. Brants, A. C. Papat, P. Xu, F. J. Och, and J. Dean, "Large language models in machine translation," in *Proc. Joint Conf. Empirical Methods Natural Lang. Process. Comput. Natural Lang. Learn. (EMNLP-CoNLL)*, 2007, pp. 858–867.
- [58] J. Pereira and Ó. Díaz, "Using health chatbots for behavior change: A mapping study," *J. Med. Syst.*, vol. 43, no. 5, p. 135, May 2019.
- [59] J.-M. Díaz-Bossini and L. Moreno, "Accessibility to mobile interfaces for older people," *Procedia Comput. Sci.*, vol. 27, pp. 57–66, Jan. 2014.
- [60] M. F. M. Yusof, N. Romli, and M. F. M. Yusof, "Design for elderly friendly: Mobile phone application and design that suitable for elderly," *Int. J. Comput. Appl.*, vol. 95, no. 3, pp. 28–31, Jun. 2014.
- [61] A. Sharkey and N. Sharkey, "Granny and the robots: Ethical issues in robot care for the elderly," *Ethics Inf. Technol.*, vol. 14, no. 1, pp. 27–40, Mar. 2012.
- [62] H. Zhou, Z. Deng, Y. Xia, and M. Fu, "A new sampling method in particle filter based on Pearson correlation coefficient," *Neurocomputing*, vol. 216, pp. 208–215, Dec. 2016.
- [63] S. Kemper and K. Kemtes, "Aging and message production and comprehension," in *Cognition, Aging and Self-Reports*. Abingdon, U.K.: Taylor & Francis, 2010, pp. 229–244.
- [64] A. Gonzalez-Agirre, E. Laparra, and G. Rigau, "Multilingual central repository version 3.0," in *Proc. 8th Int. Conf. Lang. Resour. Eval. (LREC)*, 2012, pp. 2525–2529.
- [65] J. Juncal-Martínez, T. Álvarez-López, M. Fernández-Gavilanes, E. Costa-Montenegro, and F. J. González-Castaño, "GTI at SemEval-2016 task 4: Training a naive Bayes classifier using features of an unsupervised system," in *Proc. 10th Int. Workshop Semantic Eval. (SemEval)*, 2016, pp. 115–119.

•••