Towards a computer-assisted aesthetics of user response

Florentina Armaselu (ORCID)

Luxembourg Centre for Contemporary and Digital History (C^{2}DH), Université du Luxembourg, Esch-sur-Alzette, Luxembourg

Abstract

The article focuses on the analysis of the user as an aesthetic category and proposes a methodology for evaluating user response within a framework that combines theoretical background from different areas, the theory of aesthetic response, psycholinguistics, appraisal theory, dialogism, and affective stylistics, with the application of digital tools for corpus linguistics and sentiment analysis. Four user types were derived from the corpus linguistics analysis referred to as immersed, distant, sceptical, and enthusiastic users. Each type may encompass a certain degree of intentionality and convey an attitude, implying features such as commitment and honesty, objectivity and engagement with the audience, critical reflection and circumspection, openness to technological novelty, and enjoyment. This assumes that the users involved in usability testing are not neutral or undifferentiated informational entities placed in an experimental context but individuals that respond to the same stimuli and express themselves differently in light of psycholinguistic factors and rules of social interaction. On the other hand, the results of sentiment analysis showed that an experiential analysis, centred not only on the artefact but also on the response and the experience it generates, may enable understanding the user as involved in a hermeneutic process of interpretation during his/her interaction with the studied artefact. Given the small scale of the analysed data, the study does not intend to provide evidence for general or definitive statements but formulates and illustrates a set of interpretative hypotheses and methodological directions for further enquiry aiming at developing an 'aesthetics' of user response.

Correspondence:
Florentina Armaselu,
Universite du Luxembourg,
L-4366, Esch-sur-Alzette,
Luxembourg.
Email:
florentina.armaselu@uni.lu

1 Introduction

User experience in the digital medium is often studied within the Human–Computer Interaction (HCI) field in relation with the construction of user models or the performance of usability tests in order to support the design and evaluation of digital artefacts. User modelling research has mainly focused on a variety of characteristics that inform such models from demographic information (age, gender, native language) and relevant experience (novice, advanced, expert) to interests, goals, and plans (general interest categories, task-related objectives/sequences of actions), or contextual information (location, time, physical environment; Sosnovsky and Dicheva, 2010, pp. 33–34). The aim of this type of research is to facilitate the construction of 'usable' and 'useful' tools that provide users with 'experiences fitting their specific background knowledge and objectives' (Fischer, 2001, p. 65). Many approaches in this area merge

cognitive science and artificial intelligence (Webb et al., 2001; Biswas and Robinson, 2010; Mohamad and Kouroupetroglou, 2013), while usability testing, as a technique from user-centred design, often involves the iterative refinement of a prototype based on user feedback (Massanari, 2010) or evaluates how a tool is actually used (Brown and Hocutt, 2015), exploring constructs such as ease of use and learnability.

Other types of research from the fields of philosophy of technology or digital hermeneutics go beyond the usefulness and usability aspects of technology, trying to address questions related to the 'human, social, cultural, ethical, and political implications of those technologies' (Fallman, 2007, p. 296), to the 'self-interpretation of human beings' through the impact of digital code on 'all kinds of processes', particularly the 'societal ones' (Capurro, 2010, p. 10), or to the 'remediation' of the self through digital media (Bolter and Grusin, 2000). Further directions of investigation propose a re-orientation of HCI as 'humanistic HCI' (Bardzell and Bardzell, 2016) or as an 'aesthetic field' (Bertelsen and Pold, 2004, p. 23), which refers in part to user experience. Studies of the relationships between aesthetics and usability show that 'apparent beauty' or 'aesthetic aspect' may influence the user's evaluation of the functional aspects of an interface (Kurosu and Kashimura, 1995, p. 293). Moreover, it is considered that 'aesthetics may considerably affect system acceptability' and 'aesthetic perception and its relation to HCI relevant constructs are culturally dependent' (Tractinsky, 1997, p. 121). By combining theoretical ground from cultural studies, in particular aesthetics and critical theory, critical and aesthetic approaches to interaction have been proposed (Bardzell, 2009). Although the need for a more 'holistic approach towards understanding how people experience and judge information systems' (Tractinsky, 1997, p. 121) has already been articulated and general viewpoints on the 'representational model' of the user, bringing together 'digital representation' and 'real self' have been defined (Bardzell, 2009, p. 2363), the analysis of the user as an 'aesthetic' category in itself remains an aspect that appears to have hitherto received little attention.

The article will focus on the latter aspect. Given the small scale of the analysed data, the study formulates a set of interpretative hypotheses and methodological directions for further enquiry with the aim of developing an 'aesthetics' of user response. The contribution will consist of a methodology for evaluating user response within a framework that combines areas such as the theory of aesthetic response (Iser, 1980), psycholinguistic analysis (Pennebaker, 2011), appraisal theory (White, 2015), dialogism (Bakhtin, 1981), and affective stylistics (Fish, 1970) with the application of digital tools for corpus linguistics and sentiment analysis. Intended to provide a hermeneutic perspective on the data available for analysis, the research will mainly target the following questions. How is the user's experience and self-representation in his/her interaction with digital artefacts expressed through language, and what linguistic patterns can be discerned? Can a 'typology' of users be derived by means of these digital and theoretical tools? What may be learnt from a *computer-assisted aesthetics of the user response* that aims at bridging the fields of HCI and Digital Humanities and understanding different types of user response beyond the traditional HCI-oriented interpretation of usefulness and usability? To illustrate the proposed methodology and derived typology of users, the article elaborates on previous work (Armaselu and Jones, 2016, 2017a,b; Armaselu et al., 2016) and focuses on a case study of usability tests lead by the author within the framework of interface design for digital historical editions. Two other data sets were analysed applying the same methodology and will occasionally be mentioned as general reference observations. Detailed descriptions of these data sets and corresponding results are, however, out of the scope of this article.

The remainder of the article is structured as follows. Section 2 provides an overview of the general methodology, the data considered in the study, and the various tools used for pre-processing and analysis. Section 3 presents the specific types of analysis applied in the project, by indicating the targeted category of observations, describing the linguistic patterns observed, and interpreting the observations within the framework of an aesthetic theory. Section 4
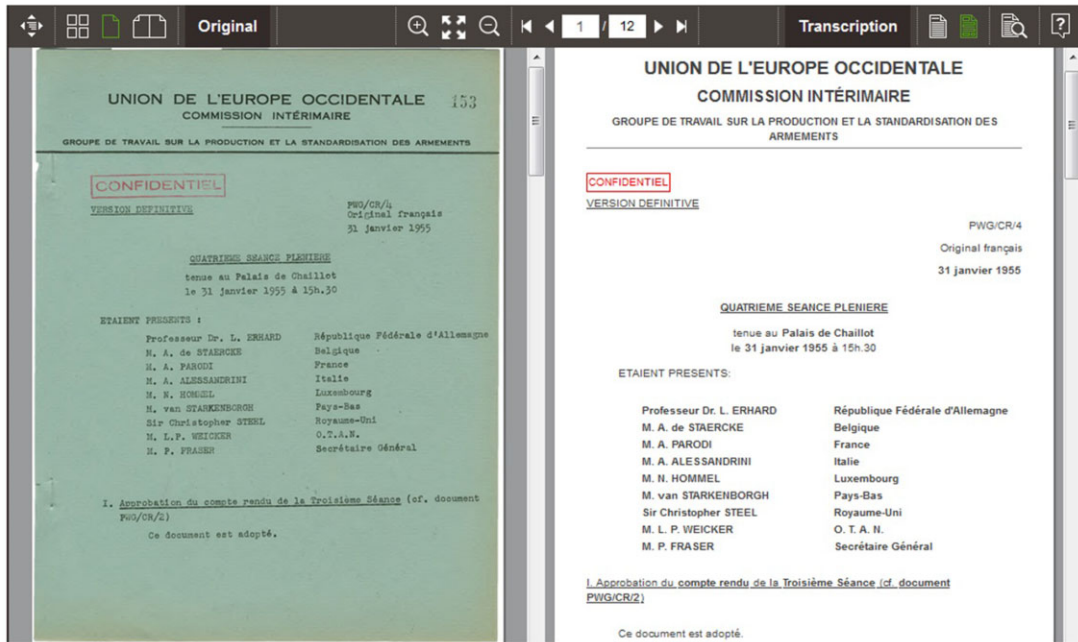
Fig. 1 Side-by-side view of facsimile versus transcription in the Transviewer

concludes the article with a discussion of the findings and future directions of study.

2 Methodology

The case study evaluated in this article considered the design and implementation of an XML-TEI-based platform known as Transviewer (Fig. 1).[1] The purpose of the interface is to facilitate the exploration of historical documents as a digital edition through features such as page-by-page navigation, side-by-side view (facsimile/transcription), synchronised scroll, free text and named entity search, and structural navigation.

The user responses gathered in this case in 2015 were based on usability tests inspired by previous studies (Nielsen, 2000; Jones and Weber, 2012) and involved a user group of eight researchers in history, political science, and linguistics, four males and four females, age range between twenty-five and sixty-four. All were working with digital history documents in a research institute focused on the history of European integration. The experiment protocol asked them to think aloud (Lewis, 1982) as they used the Transviewer interface to complete seventeen tasks incrementally building in complexity. On average the tasks took between 30 and 45 min to complete. At the beginning of the experiment, participants were asked to consent to being observed and for their audio and screen interactions to be recorded and their anonymised responses to be used for research purposes. Following the completion of the tasks, they were invited to fill in a form derived from the Usefulness, Satisfaction, and Ease of use (USE)-based questionnaire (Lund, 2001), which also contained three open-ended questions. The common language of the experiment was English, although only one of the participants was a native speaker. Owing to time and resource constraints, snippets of the think-aloud material were transcribed for each task according to the following categories: reflections on the experience, suggestions for improvement, and expressions of disorientation or frustration. Answers from the open-ended questions in the questionnaire were also included for analysis.

The case study provided qualitative data by applying think-aloud user testing and open-ended

questionnaire protocols. The analytical objectives were (1) to understand the linguistic patterns of user responses and (2) to investigate the different types of sentiment expressed in the responses. Through these forms of analysis, it was intended to explore the user experience with more depth and reflection. Transcribing the think-aloud records made it possible to gather data that could be treated as text for analysis. Pre-processing of the texts was carried out using Oxgarage[2] to convert them into an XML format compatible with the software for text analysis. Plain text format was used for the sentiment analysis phase. The texts were then processed using TXM[3]—a statistical tool for 'textometric' analysis (Heiden, 2010), and sentimentr[4] and Syuzhet[5]—sentiment analysis packages (Jockers, 2017; Rinker, 2019a,b) via RStudio.[6] The Transviewer data set was therefore based on transcribed think-aloud recordings and questionnaire answers (11,267 words) and a sampling that featured each user's response (eight users) as described in the next section. Similar procedures were applied for the two reference cases (5,974 words and sixteen users and, respectively, 33,190 words and ten users) based on data collected during usability testing via questionnaires and think-aloud protocols for the assessment of two applications (a cultural heritage game for mobile devices and an online map editor), in the prototyping and, respectively, post-release phases (Jones and Weber, 2012; Armaselu and Jones, 2017a,b; Jones et al., 2017).

3 Analysis of User Responses

The user responses from the Transviewer think-aloud audio recordings and questionnaires were transcribed, if they were not already in electronic text form. The response snapshots were pre-processed (storage as TXT, corrections, conversion to XML) according to the format required by the analysis phase (via TXM and RStudio).

3.1 Textometric analysis

TXM was chosen for its comparative, contrastive, and quantitative/qualitative features and the possibility of importing different format types (XML, TXT) and reshaping the corpus for analysis according to structural or file-naming properties. These features were used to construct sub-corpora and partitions, to compute specificity scores and concordance tables, and to display document views for context analysis.[7] For instance, in the case of Transviewer, XML-TEI encoding made it possible to divide the analysed corpus into structural elements so as to separate data from metadata in a sub-corpus and construct a user-based partition in the sub-corpus using <body> and <div> elements. The number of words per user in the partition varied from 556 (user R1-PIL02) to 2,795 words (user R1-INT06; Fig. 2).

For the purposes of uniformity and concision, the lemma property[8] was used (instead of word) when computing specificities for the partitions in the Transviewer and reference cases. Relevant lemmas (with a frequency of occurrence >10 in each data set) were chosen to be represented in the diagrams. The software allowed contrasting the specificity scores corresponding to each user, in terms of overuse/underuse of words/properties. The specificity method is based on a probabilistic model (Lafon, 1980) that enables us to compare the occurrences of a word/property in a sub-corpus (versus the parent corpus) or the parts of a partition, with each other and against a 'banality' threshold (referring to textual events that have stronger probabilities of appearing according to the model, for example, with >5% chance). The specificities focus on weak probabilities (events that are less likely to happen and thus specific to a sub-corpus or a part of a partition). The specificity scores refer to a probability's order of magnitude (e.g. a score of 3.0 corresponds to a probability of the event of about $1/10^3$), with a positive or negative sign indicating that the observed number of appearances of the event is higher or lower than an expected value according to a hypergeometric distribution model.[9] Figures 3–5 show specificity diagrams based on these measures for three groups of linguistic categories: pronouns; conditionals, negation, and uncertainty markers; and appreciative adjectives and superlatives. These linguistic elements were assumed to capture some of the users' characteristics while they interacted with the technology and expressed themselves during this interaction. Scores above/below a banality threshold (by default, $+/-2.0$; represented as red horizontal lines in the figures) indicated higher specificity, as either overuse or underuse, for responses from particular types of respondent, which enabled hypotheses to
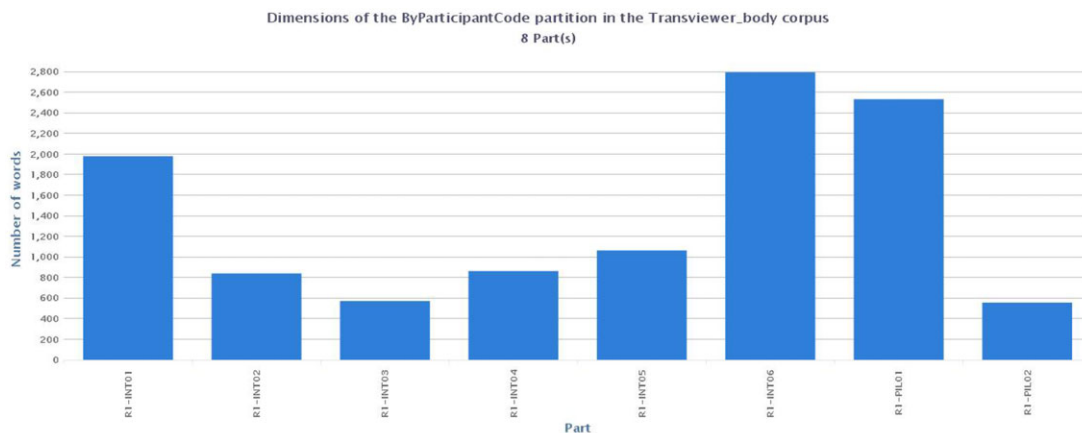
Fig. 2 TXM: Partition dimensions (body sub-corpus) by user and number of words
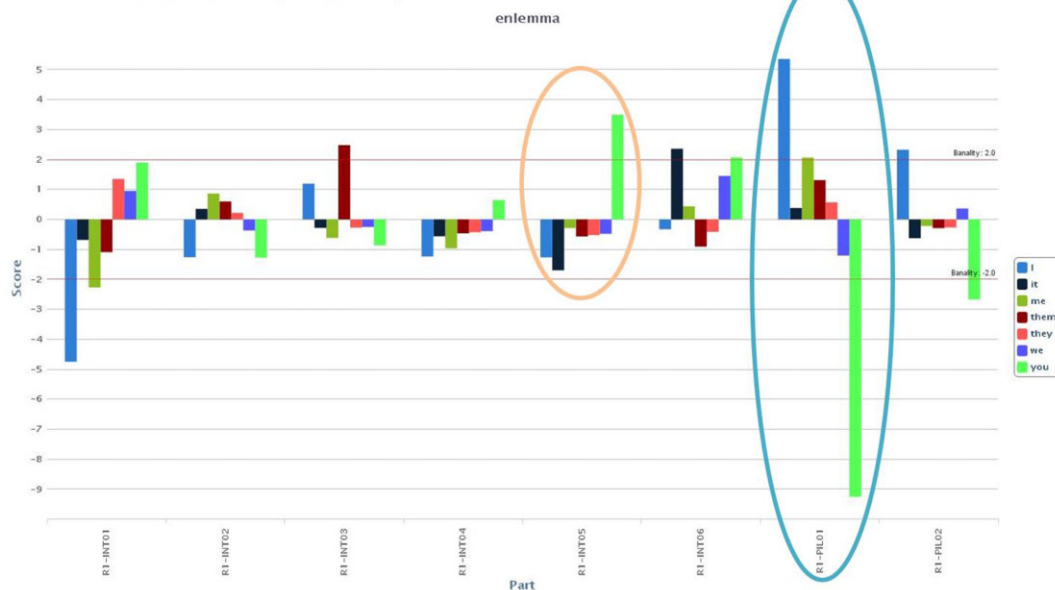

Fig. 3 TXM: User response specificities. Pronouns. *Immersed/distant* user

be made about a potential 'typology' of users that can be described within the case study.

3.1.1 *Pronouns*
The first linguistic category analysed was pronouns, considered not so much as substitutes for nouns but more as a means of defining roles in the communication process: first person as the *speaker*, second person

as the *addressee* or the 'person designated by the speaker as recipient of the communication', and *other roles* including 'other relevant entities, other than speaker or addressee' (Halliday and Hasan, 2013, p. 45). Pronouns have been studied so far from multiple perspectives, e.g. in literary texts through the angle of various disciplines such as linguistics, stylistics, narratology (Gibbons and Macrae, 2018),

Transviewer: *again, annoy, could, know, maybe, n't, no, not, perhaps, should, strange, understand, would*
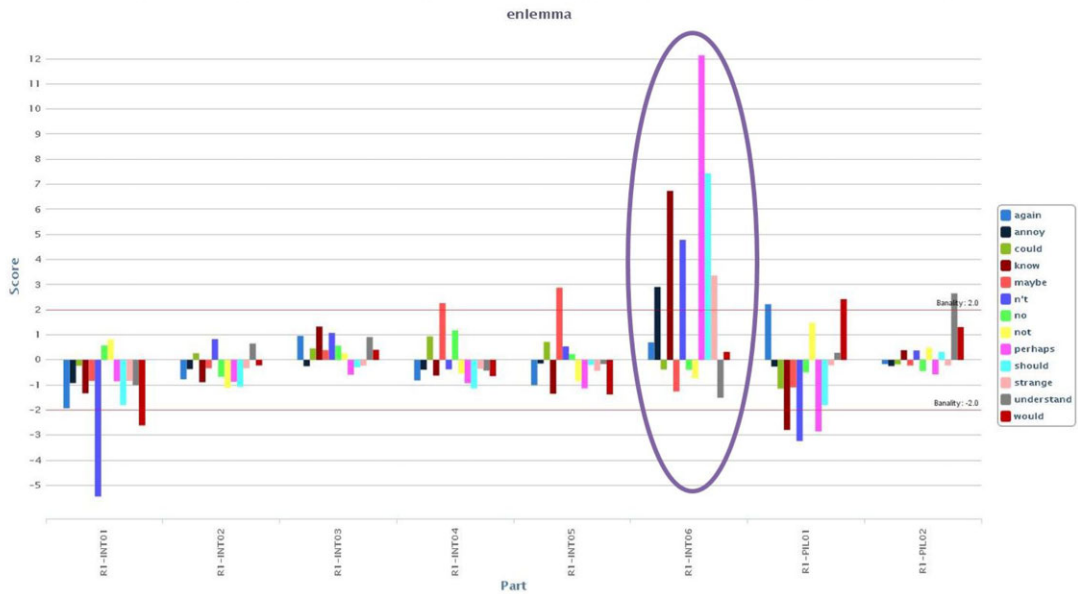

Fig. 4 TXM: User response specificities. Criticism markers. *Sceptical* user

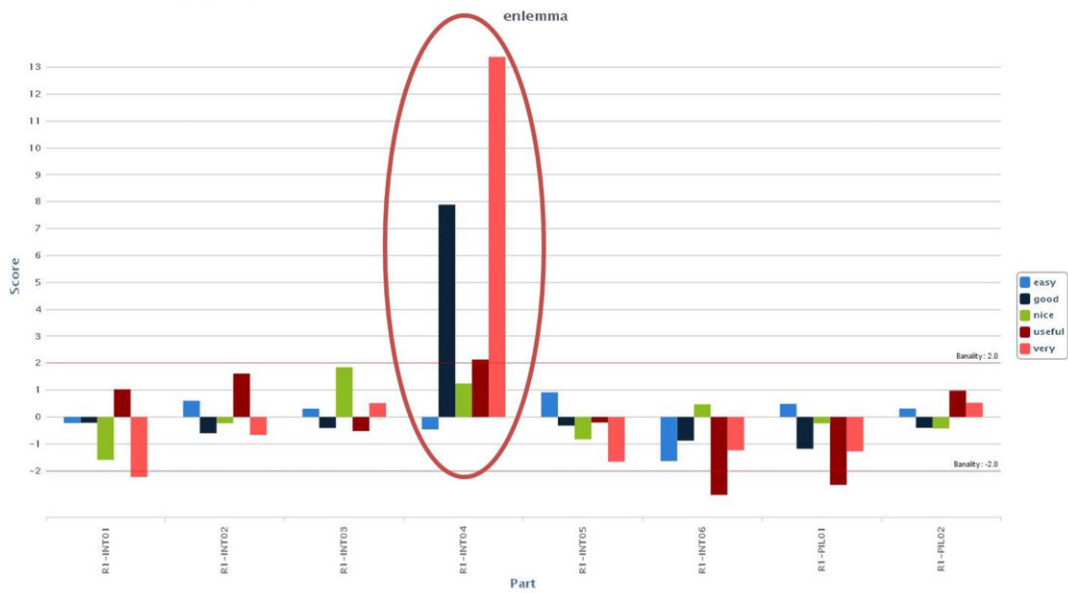**Transviewer:** *easy, good, nice, useful, very*

Fig. 5 TXM: User response specificities. Appreciative markers. *Enthusiastic* user

in political discourse to look at how they are used to construct identity (Bramley, 2001), or in psychology and social sciences to investigate topics related to personality, gender, deception, leadership, and history (Pennebaker, 2011).

The present study will focus on what pronouns may reveal about the experience of users playing with an interface when they express themselves in a think-aloud or questionnaire-based experiment. Figure 3 illustrates two types of pattern observed, referring to specific use of first- and second-person pronouns as overuse or underuse, sometimes opposed to each other. In the case of Transviewer, user R1-PIL01 (circled, right) shows a high positive score for the pronoun *I* and a high negative score for the pronoun *you*, while R1-INT05 (circled, left) exhibits a rather opposite tendency, with a lower but still significant positive score for *you* and a negative value slightly under the banality threshold for *I*.

A closer look at the TXM concordances for the *I, me, you* pronominal forms and the users highlighted above provides more insight into the mini-contexts of usage and the words that tend to co-occur with these pronouns. For instance, user R1-PIL01 mainly shows usage of *I, me* together with a variety of verb types, e.g. verbs of cognition (*expect, guess, know, realise, think, understand*): *I would expect to see some metadata; now, I realise I can scroll on the right text representation; this is nice and I think, I would be using it*; verbs of perception (*see, look, annoy*): *I'm scrolling down and I see . . .; now I'm looking at the one page view; it annoys me that there are two dimensions for scrolling*; and verbs of action (*click, move, scroll, press*): *if I click again on the thick grey bar . . .; until I move it all the way to page 8; I press enter*. On a few occasions, *you* is used: *of course, you can retrieve the information from . . .; like a 'next' button so you can jump*. Comparatively, for user R1-INT05, there are instances where *I* generally appears accompanied by cognition and perception verbs or expressions: *I observe that the size of the transcription is . . .; I think it's not the point; I don't know how you could do that*, or by verbs of direct action: *when I move the original . . .*. By contrast, *you* typically occurs in action-reflection or more explanatory contexts: *when you double click on one page . . .; once you have selected an option . . .; it reassures you that you are reading the right transcription; forcing you to reclick on . . .; if you zoomed and you are seeing less than one page . . .*.

Similar patterns were observed for the reference cases. Generally, *I* and *me* were used in cognitive, appreciative and action-related contexts and in open expressions of feelings and failure (interjections; self-dissatisfaction, frustration, annoyance). *You* appeared with verbs of perception and action to express general assertions, descriptions and explanations of actions or instruction from the experiment task sheet and the help menu, or as a sort of 'dialogue' with a generic human/technology-embodied interlocutor.

All these observations may be interpreted in a context reminiscent of Aarseth's (1997) concept of 'ergodic literature', which requires from the user/reader 'a selective movement' in a 'work of physical construction', and 'nontrivial effort [. . .] to allow the reader to traverse the text' (p. 1) during the 'cybertextual process'. The *I, me* contexts seem to reveal how this 'physical' construct is built during the experience with the interface, by using action, cognition, and perception verbs/expressions. On the other hand, *you* appears when the experiencer describes and reflects on his/her actions, repeats fragments from the experiment instructions or interface-generated messages, as a generic form for addressing the designer of the experiment, the piece of software to be evaluated or the machine, or expresses general statements. Further insight may be inferred from a closer look at the use of pronouns in other contexts, as discussed below.

Previous studies have pointed out the interpretative intricacies related to the usage of pronominal forms in different types of discourse. In conversations, *I* is simply characterised as a term of self-reference when the speaker is referring to him/herself (Sacks, 1995, p. 675). According to Pennebaker (2011), I-words have 'tremendous social and psychological significance'. Their use may be interpreted as an 'identity statement' that reflects 'self-awareness' and 'self-attention' and 'provokes honesty' (pp. 160, 161). Moreover, in oral autobiographical narrative, the narrator seems to '*enact* a characteristic type of self', the 'interactional positioning' of narrator versus audience enabling 'narrative self-construction' (Wortham, 2000, p. 2). In political interviews, *I* is central to the 'representation of self', when politicians recount actions, express opinions, knowledge, authority, etc., in order to show themselves in a 'positive light' (Bramley, 2001, p. 28). The use of *you* is more ambiguous since it may involve either singular or

plural interpretation. Sacks (1995, p. 349) asserts the 'openness' of *you* in conversation as a way of 'talking about "everybody," including "me,"' that stands 'for the set of terms: "everyone," "someone," "people," etc'. Pennebaker (2011) assumes that a higher use of *you* as compared with *I* reflects 'where people are paying attention'; that is, 'looking at or thinking about their audience' more than 'focusing on themselves' (p. 174). Additionally, 'not using I-words' makes the writer/speaker appear 'slightly less accessible and more distant' (p. 180). In political interviews, the interviewee may use generic *you* to construct him/herself as a 'typical member of a category' and 'give more validity to what is said'. *You* in the singular can also be used when addressing the interviewer (Bramley, 2001, pp. 130–131). In their study of the use of the second person in a literary work, de Hoop and Hogeweg (2014) discern two patterns which they argue to be functionally related to the development of the story: *generic* and *deictic* use, the latter with an *interactive* and *descriptive* reading.

In the context of this study, it may be considered that the user is actually 'enacting' a certain 'representation of self' in reporting his/her interaction with technology. This representation is determined by the formal framework of the experiments themselves and the particular circumstances of being observed or consciously producing a response intended for subsequent analysis. Given the patterns described above, the overuse of *I* as rendered by TXM specificity scores can be related to an *immersive* type of user or an *immersed* user, i.e. a user predominantly expressing his or her explorative, reflective, and affective experience while dealing with a technological artefact through the use of first-person pronouns (user R1-PIL01). While the use of I-words in the context of the experiments can be associated with notions of 'self-reference', 'self-awareness', 'self-attention', and 'honesty', the use of the second-person pronouns may be interpreted as a preferred means of expression of a 'more distant' user. The *distant* user would therefore report his/her experience by incorporating the expression of self into a multi-layered account involving several types of addressee, as shown in the examples cited above (user R1-INT05). Irrespective of the type, I-words may refer to a speaker who completely

assumes what is said in his/her experiential-evaluative utterance. Use of second-person pronominal forms is more complex. We will distinguish a *generic* and *interactive* versus *descriptive deictic* reading in the employment of second-person pronouns. (1) The *generic* reading refers to contexts where *you* is pointing to a more general audience, including the experiencer, the observer of the experiment, the analyst or the reader of results, etc. This plural interpretation, standing for 'me', 'someone', 'everyone', 'people', 'audience', may encompass several explanations. First, it can be related to models of what Fludernik (1993, pp. 231, 235) calls '"How to . . ." literature', which includes 'cookbooks', 'self-help manuals', 'tourist guidebooks', instructions on 'how to use or set up technical equipment', etc. This usage appears especially when the user is reading fragments from the help menu, dialogue boxes from the interface, or instructions from the task description in the experiment. Secondly, *you* is used while speaking in a more general sense, intending to cover more than the particular circumstances of the experiment, as if uttering more general truths or relying on the audience's implicit validation or approval. (2.1) In the *descriptive-deictic* sense, *you* is used instead of *I* to get the audience involved, to increase the engagement or identification of the reader/observer with the experiencer. This refers to explanatory contexts when the user describes actions and reflects on the effects of these actions. The examples discussed above seem to suggest that *I* appears more in exploratory contexts, where the action is still ongoing. In contrast, *you*, in its *descriptive-deictic* meaning, which can be read as an intermediary form between *I* and generic *you*, relates to a completed sequence of action or to the consequences of an action and the experiencer's reflection on it. (2.2) On the other hand, the *interactive-deictic* interpretation supposes an addressee who may be identified sometimes with the observer or the designer of the experiment and sometimes with the piece of software to be evaluated or the machine itself. Of course, the *immersive* and *distant* types do not completely define a user, but a more prominent facet of oneself as expressed through language while interacting with the technology. In reality, the user represents a versatile entity and can be characterized by a combination of features. Additional user types

and an appraisal theory-based reframing of these cases are discussed in the following sections.

3.1.2 Criticism markers

The second linguistic category used to compare specificity scores was that of criticism indicators, including modal verbs, conditionals, negation, uncertainty, or confusion and frustration markers, as illustrated in Fig. 4.

Within this category, user R1-INT06 shows an overuse of *annoy, know, n't, perhaps, should* and *strange* as compared with the other users from this case. A closer look at the mini-contexts (concordances) of these words indicates that *know* and *n't* are mostly used together in negative constructions such as *I don't know how/if/what/where* or as an independent sentence to express confusion, disorientation or that something doesn't work or look as expected: *I don't know how to call it; I don't know if it's a technical problem; I don't know what happened; I don't know where the document ends; I don't know. Ah, it is just slow. Should* appears in contexts containing suggestions or advice, hinting at the unsatisfactory functionality or look of the interface, and is often accompanied by *perhaps* or *I don't know* for attenuation: *I think we should adapt the window [. . .], perhaps, I don't know, to the scanned version; you should have perhaps the page numbering made . . .; it should perhaps when you click . . .. More* assertive or affective markers with negative connotations are *strange* and *annoy(ing)*, often co-occurring with *a bit*, apparently for attenuation purposes as well: *it's/that is a bit strange; it/this is a bit annoying.*

An overuse of *could* and *not* were also displayed for a user in one of the reference cases. These forms appeared together or separately to express uncertainty about the meaning of certain items, inability to achieve goals, assumptions, or simple negative statements. In the second reference case, it was a user exhibiting usage above the banality threshold for *know* and *n't*. As observed for user R1-INT06 above, the two forms co-occurred in negative contexts that expressed uncertainty or confusion. Other negative contexts referred to situations conveying inability, dissatisfaction, or frustration when dealing with the interface; this seemed to be personified to a certain degree—the user's intention has been

'misunderstood', not enough help has been provided and the user was thus prevented from completing the tasks.

In her article on the pragmatic functions of *I don't know*, Tsui (1991, p. 607) points out that as well as the central meaning of this expression that may be understood as a 'declaration of insufficient knowledge', the utterance presents a variety of other functions in conversational environments. It is argued that its pragmatic motivation resides in functions such as 'avoidance of making an assessment', of 'explicit disagreement', or of 'commitment', as well as 'minimisation of impolite beliefs' or 'marker of uncertainty'.

In novels, Mizzau (2012) distinguishes different 'modalisation markers' characterizing the 'degree of uncertainty' related to the 'source of information' as adverbs (*maybe, probably*), verbs (*think, seem, believe*), conditionals (*would, could*), or approximation terms (*towards, around*) (p. 69). These markers can indicate 'a shift from the figure of the narrator asserting something', a way of 'distanc[ing] himself from the assertion' (p. 71). On the other hand, when expressed from the perspective of the 'empirical writer', i.e. the 'real person who is writing', uncertainty appears as a result of the 'identification' with the reality described that he 'invents as if he were witnessing it' or as a 'sort of imagination bordering on perception' (p. 75). Other studies draw attention to the relationship between politeness strategies in political discourse and the deontic use of modal verbs and expressions such as *should* and *perhaps* to convey 'moral obligation (advice)' or to soften 'imposition illocutionary force' (Boicu, 2007, p. 10). Huschovà (2015), on the other hand, explores the use of modal verbs in academic discourse and highlights the primary 'epistemic sense' of *may/might* in balancing objective reporting and 'subjective evaluation', and the 'root sense' of *can* in expressing potentiality and its 'possibility, ability, or permission interpretations' (p. 45). More generally, within the larger context of the 'appraisal theory' developed by Martin and White, a set of three systems was devised to describe the 'speaker's/writer's personal, evaluative involvement in the text' (White, 2015, p. 1). The first system, ATTITUDE, refers to evaluative meanings understood in terms of 'positive or negative view vis-à-vis experiential phenomena'. This is further divided into three sub-categories: *affect*, involving

'assessment as emotional reaction', *judgement*, concerning attitudes towards 'human behaviour', and *appreciation*, dealing with evaluation of artefacts, happenings, state of affairs, etc. by reference to 'aesthetics' or other means of 'social valuation' (pp. 2–3). The second system, GRADUATION, is related to parameters such as *force* and *focus* that enable different degrees of personal investment by intensifying/mitigating or sharpening/blurring propositions and semantic boundaries (p. 4). The third, ENGAGEMENT, positions the speaker/writer vis-à-vis alternative voices and viewpoints within a 'dialogic space' (p. 5) as inspired by Bakhtin's (1981) notion of 'dialogism' and 'dialogic imagination'. Oteíza (2017) provides an overview of applications of the appraisal framework in discourse analysis, taking into account different genres, approaches, and languages.

In this study, the uncertainty markers highlighted above concerning users' interaction and perception or 'declaration of insufficient knowledge' generally fall into the ATTITUDE—*appreciation* category, expressing usability, learnability, utility, and aesthetic assessment of the evaluated artefact during the 'experiential' process. Thus, within this appreciation framework, instances of *don't know* can be interpreted as having negative connotations, meaning *I don't know* because something is missing, wrong, unclear, unexpected in a confusing way, etc. in the functionality and/or design of the interface or experiment simulation. The same can be assumed for negative statements such as *can't, don't want, won't, did not see* or hypothetical constructs as *could be* which convey a sense of inability to do or assert something with more certainty in the given circumstances of the experimental environment. The ATTITUDE—*affect* aspect is also represented by explicit utterances such as *strange, annoy(ing), don't like* or negative statement tag questions such as *can you?* that may be interpreted as words/phrases expressing a certain degree of frustration, distaste, or disagreement. At the same time, these forms seem to convey more a nuanced meaning that also pertains to both the GRADUATION and ENGAGEMENT systems. For instance, the deontic usage of *should, perhaps* and certain instances of *don't know* may evoke the *force* parameter in the former, i.e. implying attenuation, politeness, or avoidance of disagreement or commitment. Moreover, Pennebaker (2011) highlights the social and

emotional dimensions of what he calls 'discrepancy verbs' including words like *should, could, ought, must*, and *would*, used when 'people suggest some kind of subtle discrepancy between how the world is and how it could, should, or ought to be' (p. 165). On the other hand, all the forms discussed above, as observed for the *immersed* and *distant* user in the previous section, may refer to a 'social' dimension and a user 'enacting' a role within a 'dialogic space'. This underlies the presence of (and implicitly the engagement with) other 'voices', e.g. that of an implied addressee or interlocutor such as the observer of the experiment, the designer of the interface, other users, the analyst, or reader of the evaluation report, etc., or even that of the machine or piece of software itself. We therefore interpret the specific overuse of such forms by a particular user from a case study as a marker of doubt, discrepancy, and criticism regarding the evaluated artefact within a dialogic context, and as a characteristic of a so-called *sceptical* user. The following section refers to another, contrasting type.

3.1.3 Appreciative markers

While the criticism markers were considered suggestive of a 'sceptical' point of view, experiences described through appreciative adjectives and superlatives may be indicative of an opposite standpoint. Figure 5 shows the usage of such forms.

For user R1-INT04, we can observe an overuse of *good, useful*, and *very*. Looking at the contexts again reveals that *very* co-occurs, sometimes repetitively, with positive polarity adjectives or agreement markers in appreciative assertive phrases such as: *this view is very, very useful; the transcription is very, very good; this could be very helpful for/to . . .; which is very useful/good; that's very nice*, etc. On a few occasions, *very* appears to intensify a negative implication or, on the contrary, to attenuate a negative statement: *the scan is very small this time; that's a bit, not very complicated but not very user friendly. Good, better* also occur in negative or dubitative contexts, but only rarely: *that's not so good. I prefer to . . .; I don't know but maybe it would be better if this . . .*. An overuse of *nice* and, respectively, *cool* was observed for the reference cases.

The choice of words from this category can also be interpreted within the appraisal framework discussed above, including the ATTITUDE—appreciation, affect, GRADUATION and ENGAGEMENT systems,

assuming the expression of a predominantly positive stance through positive polarity adjectives and intensifiers, assertive and agreement forms in the evaluation of the artefacts assessed by the experiments. This is what corresponds in the proposed typology to an *enthusiastic* user.

In his analysis of words associated with honesty and deception in conversational environments, Pennebaker (2011) relates 'optimism and overconfidence' with 'many forms of deception'. He states that 'people who are deceptive make more references to other people and rely on more positive emotion words'. He places on a trustworthy scale (from more to less) forms such as I-words, negative emotions, *you*, discrepancies and positive emotions (pp. 162, 164). While the context of honesty and deception is different here from the situations described by Pennebaker, it can however be assumed that overuse of appreciative forms in a user's account involves a certain degree of intentionality in showing agreement and positively assessing 'how the world is'. As already discussed in the previous section, this intention is not devoid of a social and dialogic dimension, positioning the experiencer vis-à-vis one or several 'implied' interlocutors. From this perspective, we can reconsider the proposed user typology in light of what Bakhtin (1981, p. 279) calls the 'dialogic orientation of discourse' in the novel toward the '"already uttered," the "already known," the "common opinion."' Literary language is said to be saturated with 'intentions and accents', all words being expressions of their 'socially charged life' and thus being 'populated by intentions' (p. 293). If we transpose this view to the context of our analysis, the *immersed, distant, sceptical,* and *enthusiastic* types may be understood, beyond the explicit function of informant in the experiment, as enactments of certain kinds of features, such as: (1) commitment and honesty; (2) objectivity and engagement with the audience; (3) critical reflection and circumspection; (4) openness to technological novelty and enjoyment; intended to position the speaker with regard to the 'voices' and assessment of the implied interlocutors.

Combining the intentional aspects with Pennebaker's truthfulness scale, we may also assume different degrees of intentionality within the typology from less to more, with the *immersed* user at the lower end of the scale, the *enthusiastic* at the higher, and the *sceptical* and *distant* in between.

3.2 Sentiment analysis

Although TXM specificity scores and concordances combined with the theoretical appraisal framework and other linguistic theoretical devices helped with the task of identifying and characterizing possible types of users, it was considered that further insight into the four categories and the attitude tokens expressed in the text may be acquired by means of a sentiment analysis tool. For this purpose, an open-source environment, RStudio, was chosen, as well as open-source packages for sentiment analysis in R, given the availability of code and documentation. The *sentimentr* package (Rinker, 2019a,b) was selected since it allows text polarity sentiment to be calculated based on a dictionary lookup approach at the sentence level, which was considered appropriate for the study. Moreover, the chosen methods take into account valence shifters in computing sentiment scores, i.e. negators (*not, can't, didn't*), amplifiers or intensifiers (*greatly, huge, really*), de-amplifiers or downtoners (*hardly, kind of, little*), and adversative conjunctions (*but, despite, however*).

Specific methods from R packages were used to export results (Excel files and diagrams) for further analysis. For instance, a series of methods were used to: extract sentences from text (*sentimentr::get_sentences*); approximate the sentiment polarity of text per sentence (*sentimentr::sentiment*); compute metrics for average scores, e.g. general sentiment, sentence ratio by score—positive, negative, neutral, using arithmetic and vector R functions (*mean, sum, length*); apply a discrete cosine transformation (DCT) to smooth the sentiment vector for a better graphical representation (*syuzhet::get_dct_transform*) (Jockers, 2017); draw diagrams using the transformed vector and a function for plotting R objects (*plot*).

More detailed analysis of the computation of scores was needed for a better understanding of the methods. Intermediate results were exported into Excel tables to store the sentiment scores for each sentence from the four documents corresponding to the four types of users, as well as the polarity words, valence shifters, and corresponding values contributing to the calculation. The analysis of these intermediate results enabled adjustment of the polarity and valence shifters lexicons[10] dynamically, at run time, by means of specific functions

(*update_polarity_table*, *update_valence_shifter_table*) from the *sentimentr* package. A list of stopwords to be dropped from the polarity lexicon was created for each case when these words were too polysemous and introduced noise (*like, right, well*) or when they mostly appeared with a different meaning that was not sentiment- or assessment-related. For instance, *diplomatic, highlight, kind* referring to *diplomatic* transcription, *highlighting* different entities in the text, *kind* of document/information (Transviewer). A category necessitating particular treatment was that of interjections. As a general rule, interjections, such as *ah, oh, ok, okay, yeah*, were added to the polarity dictionary at run time.[11]

Various studies have drawn attention to the meaning of interjections considered as 'suggestive of sensations, attitudes and states of mind' (Jovanović, 2004, p. 18), 'markers of subjectivity and, in particular, of emotions'[12] (Fraisse and Paroubek, 2015, p. 2), or as playing an important role in 'communicative practices and social interaction' (Ameka and Wilkins, 2006, p. 2). In their analysis of data from think-aloud protocols, Caron-Pargue and Caron (1995) highlight the role played by interjections not only as 'expressive markers of the speaker's affective state' but also as 'traces of differentiated cognitive operations'[13] (p. 120). Moreover, it is assumed that interjections act as indicators of a 'change in the state of knowledge'[14] of the speaker (Caron-Pargue and Caron, 2000, p. 56) and as 'signals' intended to 'inform' the interlocutor of this state of mind[15] (p. 54).

In this study, the polarity scores assigned to interjections and dynamically added to the polarity lexicon were mainly devised by means of two online dictionaries, Wiktionary[16] and the VidarHolen Dictionary of Interjections.[17] First, the general meaning of an interjection was interpreted by similarity with other forms according to these two dictionaries and to the contexts where the interjection occurred in the experiments. A polarity score was then assigned to an interjection by looking for these similar forms in the polarity lexicon from the R package, and computing an average value of their polarities.[18] For instance, *ah* and its variations *aahh, aaah, ahh*, etc. are described in the dictionaries as an 'expression of relief, relaxation, comfort, confusion, understanding, wonder, awe, etc. [...]' (Wiktionary) or as 'realisation, understanding' (VidarHolen). Also taking into account the contexts where this interjection appeared in the users' accounts from the experiments[19] by looking at the TXM concordances, the general, average meaning of it was interpreted as *understanding + slight confusion* which according to the rule of polarity assignment described above produced an approximate value of 0.44.[20] The polarity values were approximated manually in this way for all the observed interjections and then used as polarity words and polarity value vectors via the dedicated function that we wrote in R to update the polarity lexicon at run time.

3.2.1 Sentiment scores

Table 1 shows the resulting sentiment measures for the case study, sorted by decreasing order of the general score calculated for each type of user as the average of the sentiment values corresponding to each sentence. Other computed measures were the average positive and negative scores and the ratio, as a percentage, of sentences with positive, negative, and neutral (0 polarity) scores.

Not surprisingly, the respondent with the highest general average score was represented by the *enthusiastic* type. This result may be explained by the highest values of the positive and the lowest of the negative scores (columns 3 and 4), and the highest ratio of sentences with positive polarity (column 5) observed for this category (similarly for the reference cases). This aspect appeared correlated with a lowest ratio for negative polarity sentences (also observed for one of the reference cases, while for the second, this class of users exhibited the lowest ratio of neutral sentences, as compared with the other categories).

Table 1. Transviewer: Sentiment scores at the sentence level by user type, sorted by average general score

User type	avg_gen_score	avg_pos_score	avg_neg_score	pos_sntce_ratio	neg_sntce_ratio	neu_sntce_ratio
Enthusiastic	0.203781	0.219448	−0.01567	48.61111	9.722222	41.66667
Distant	0.088556	0.111446	−0.02289	31.14754	16.39344	52.45902
Sceptical	0.077146	0.111105	−0.03396	34.29752	20.24793	45.45455
Immersed	0.028392	0.065912	−0.03752	25.16129	22.58065	52.25806

The relative order of scores obtained for the three other types was more surprising since the sceptical user would have been expected to appear in last place and the other two in between. Instead, the immersed followed the sceptical with lower values, and the distant user unexpectedly occupied either the second (Transviewer) or last position (reference cases).

Of course, given the relatively small size of the analysed data sets, it is hard to talk about generalizations. However, some observations may be formulated. While the dominance of the average positive score and positive sentence ratio can determine a higher general average score and thus a higher rank, sometimes a lower number of sentences with higher negative scores may reverse the order, despite a higher positive score or ratio. Nevertheless, the average values are probably not fine-grained enough to allow more insight based on these figures only.

3.2.2 Sentiment-derived plot arcs

For further analysis, the DCT from the Syuzhet package (Jockers, 2017) was considered to compute sentiment-derived 'plot arcs' based on the sentiment scores corresponding to the response of the four categories of users of the studied case. Figure 6 shows the arcs obtained by applying the *syuzhet::get_dct_transform* function to the vector of sentiment values per sentence calculated, as described earlier, using the *sentimentr* package. The low-pass filter parameter[21] was set to five elements (default value) to obtain the smooth, simplified shape of the emotional valence along with the experiment time from the beginning to the end (sequence of sentences in the user response). A second parameter was used to normalize the x-axis to 100 units, so that texts of different lengths can be compared. The plots displayed in Fig. 6 are approximations used to illustrate general tendencies, so the intervals represented on the y-axis may not correspond exactly to the real minima and maxima of the sentiment values.

Previous studies applied to fiction presumed that 'fluctuations in sentiment' can act as indicators of 'fluctuations in plot movement' (Jockers, 2015a) and may reveal what Vonnegut (2010) initially called the 'shapes of stories'. Different graphical representations, labels, and short descriptions have been proposed for these shapes, e.g. 'man in hole', 'man on hill', 'from bad to worse', 'Cinderella', 'rags to riches', 'riches to rags', 'Oedipus', and, respectively, 'a steady rise from bad to good fortune', 'a fall from good to bad, a
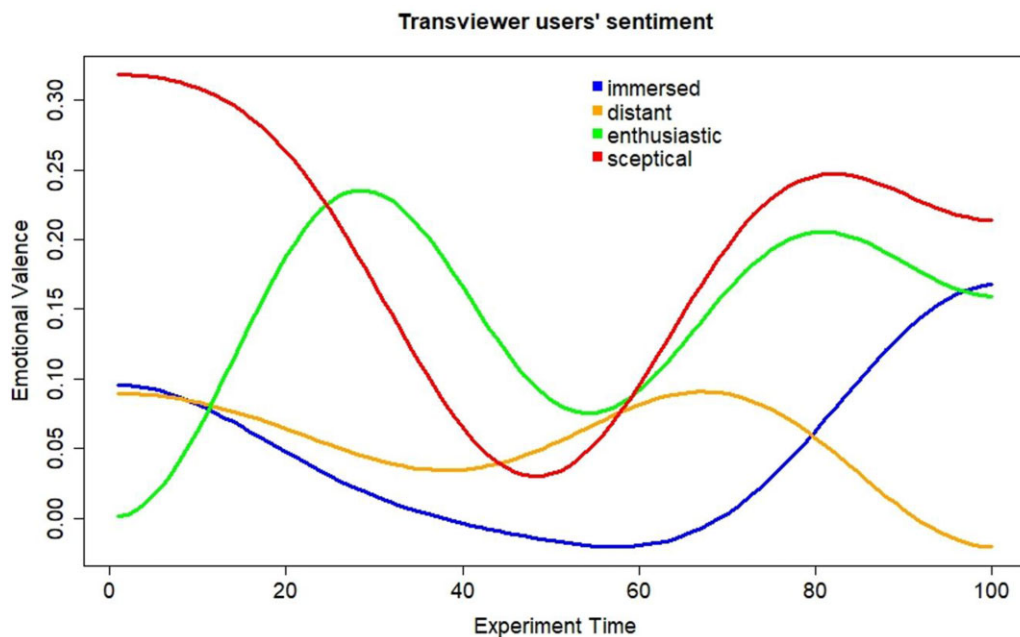

Fig. 6 Emotional valence by user type, DCT transform, low_pass_size = 5

tragedy', 'fall, rise', 'rise, fall, rise', etc. (Jockers, 2015b; Eliam, 2018; Quick, 2018).

It was assumed in the study that this type of graphical representation may also reveal sentiment variation in the users' accounts collected for the experiment. For simplicity and uniformity, the following labels will be used: 'fall' (F), 'rise' (R) and their combinations to describe the fluctuations in the arc plots observed in Fig. 6. That is, FR for *immersed* (blue, bottom); FRF for *distant* (orange, second from bottom) *and sceptical* (red, top), RFRF for *enthusiastic* (green, second from top). We also created a function to compute the local minima and maxima[22] of the set of transformed sentiment values and to approximate the position corresponding to these minima/maxima in the original set of sentiment values, with the aim of shedding light on the sentences occurring around these points.[23] For higher values of the low-pass filter, more fluctuations are made visible and the min/max points in the original set can be determined more precisely. However, even simplified curves such as the ones presented below can convey a certain sense of how different users build up a representation of their experience while interacting with the same artefact. While no general pattern was applied for each type of user, it was observed that the simplest curves (two to three fluctuations) belonged to the *immersed* and *sceptical* categories (Transviewer and reference cases—FR, FRF), whereas the more convoluted (three to four fluctuations) characterized either as the *enthusiastic* (Transviewer—RFRF and one reference case—FRFR) or as the *distant* user (the other reference case—RFRF). This aspect is worth investigating further.

In their proposal for a re-orientation of HCI as an 'aesthetic field', Bertelsen and Pold (2004) assume that as well as an involvement in a 'rational process of exploratory learning when trying to use the considered artefact', the user is 'more often engaged in a hermeneutic process of interpretation' (p. 23). Moreover, aesthetics is understood as a new perspective on the interface, as a shift from a 'use-oriented' to an 'experience-oriented' approach (p. 24). This enables critics, using for instance data from think-aloud tests and other empirical sources, to identify 'challenges to expectations in the interface' and to apply in their analysis not only notions of 'perception and cognition' but also aesthetic concepts of 'recognition and experience' (p. 30). Similarly, in his 'affective stylistics', Fish (1970, p. 127) focuses on the 'temporal flow of the reading experience' as the basis of his analysis method. Meaning therefore becomes an 'event', or 'something that is happening between the words and in the reader's mind' (p. 128). In this type of 'experiential analysis', the analytical attention shifts away from the 'work as an object to the response it draws, the experience it generates' (pp. 134, 149).

By combining these two perspectives, we consider the data from the think-aloud tests and questionnaire responses as expressions of the user's experience or as a representation of what McCarthy and Wright (2004) call users 'making sense of the situations they encounter' but also 'making sense of themselves in the process' (p. 196). Therefore, the local min/max points in the 'plot' arcs from Fig. 6 may be indicators of noteworthy interpretative events in the user's experience when interacting with the interface. These points can mark different experiential aspects such as challenges in expectations, as shown by the following sentences approximately corresponding to these locations on the experiential time axis. Some refer to observation/discovery or surprise: *And, yeah, it shows me all the dates . . .* (sceptical), *Both sides move at the same time. Which is very good* (enthusiastic); some to annoyance or disappointment: *. . . I'm just wondering whether it annoys me that there are two dimensions for scrolling . . .* (immersed). Other elements emphasized by the local min/max points may be related to successful or unsuccessful actions, confusing results or disorientation: *Yeah, it reassures you that you are reading the right transcription* (distant), *Well, it is strange because the cursor indicates me that I am on page nine but it [the page] shows sixteen. So, I don't really know on which page I am* (sceptical); to aesthetic assessments: *. . . it's irritating that the scans have a black background and the other one [transcriptions] has a white background* (immersed), *. . . it looks great but it's just this black page that is a bit strange* (sceptical); or to positive and negative closing remarks: *Look forward for more!* (immersed), *Need just some improvements like make it a bit more user-friendly [. . .] Good job!* (sceptical), *The interpretative transcription type is hiding a part of the document* (enthusiastic).

Similar aspects were observed for the local min/max points in the reference cases, for instance: positive remarks on particular features including expressions

of surprise; comments on unfulfilled expectations usually expressed by negative verbal forms; expressions of satisfaction or failure through interjections or verbs conveying disorientation; dissatisfaction with the appearance of the produced outcome that lacked certain qualities, e.g. symmetry; comments on features considered useless or not fulfilling expectations, e.g. to engage the user in a real conversation.

In his theory of aesthetic response, Iser (1980, p. 21) assumes that a literary work has 'two poles': the 'artistic pole' which is the 'author's text', and the 'aesthetic [pole]' which is the 'realisation accomplished by the reader'. For this realization to be accomplished in the process of reading, 'a textual structure anticipating the presence of a recipient' is necessary. This structure has been defined by the concept of 'implied reader' (p. 34), which also supposes the 'response' elicited by this structure, i.e. both a 'textual structure and a structured act' joined together in a dynamic process and related in the same way as 'intention and fulfilment' (p. 36). On the other hand, the presence of the reader in the text is described by the 'wandering viewpoint', a point where 'memory and expectation converge' and that permits the 'reader to travel through the text' (p. 118). Moreover, readers have the ability to 'perceive [themselves] during the process of participation' in the text, which represents an 'essential quality of the aesthetic experience' (p. 134).

Similarly, we can consider an interface as including, by default, an implied user, i.e. an entity the designer had in mind while building the artefact or simulation. However, it is during the interaction that the response to this latent structure anticipating a recipient is actually built. Experiment data such as those from think-aloud approaches or questionnaires enable the analyst to investigate how the user reflects on and represents this experience and how he/she constructs an image of him/herself within this representation. We see this act of reflection and self-reflection, assessment and affective involvement as presenting certain similarities with the reader's aesthetic experience and engagement when reading a literary text. This view has elements in common with Bardzell and Bardzell's (2016) focus on aesthetic approaches to user experience in their 'humanistic HCI', Buchenau and Suri's (2000) 'experience prototyping' in terms of 'designing an integrated experience' (p. 425), or McCarthy and Wright's (2004) 'technology as experience' and their 'aesthetic approach to technology' positioned against an 'exclusively functional approach to understanding relations between people and technology' (p. 66).

4 Conclusion and Future Work

The article describes a study in interface design and deals with language-oriented theoretical approaches and text analysis applied to user response. Given the small scale of the analysed data, the study does not intend to provide evidence for general, overarching conclusions; instead it formulates and illustrates a set of interpretative hypotheses and methodological directions for further enquiry aiming at developing an 'aesthetics' of user response.

First, an examination of different categories of linguistic markers detected in the responses suggested that, besides their position as informants in the experiments, users tend to 'enact' or build up a certain representation of self, intended to position the speaker with regard to one or several implied interlocutors. A typology was devised by means of a corpus linguistics, textometry tool that was used to compare different vocabularies in terms of overuse and underuse of lexical forms and properties, such as pronouns, negations, conditionals, appreciative adjectives, and superlatives. The four user types derived from the analysis were referred to as *immersed*, *distant*, *sceptical*, and *enthusiastic* users. Additional analysis including theoretical considerations from areas such as psycholinguistics, appraisal theory, and dialogic orientation of discourse led to the assumption that each type may encompass a certain degree of intentionality and convey an attitude implying features such as: commitment and honesty, objectivity and engagement with the audience, critical reflection, and circumspection, openness to technological novelty and enjoyment. This demonstrates that the users involved in usability testing are not neutral or undifferentiated informational entities placed in an experimental context but individuals that respond to the same stimuli and express themselves differently in light of psycholinguistic factors and rules of social interaction.

Secondly, the study of responses through sentiment analysis in R enabled us to quantify the four

types of user experience in terms of positive and negative polarity scores. The only pattern that clearly emerged, however, was the leading position of the *enthusiastic* user. With regard to the three other types, *sceptical* and *immersed* users exhibited a relatively stable order and *distant* users a more variable position. Since the amount of analysed data was not large, these results should be tested on additional cases to formulate more general statements. The graphical representation of the sentiment-derived plot arcs computed through the DCT of the sentiment values offered another perspective on the four user types studied. However, at this stage, we cannot associate a specific shape with a user type for all cases, but possible variation schemes needing further investigation were observed. The diagrams may thus inform an experiential analysis that focuses not only on the artefact but also on the response and the experience it generates, in ways that evoke methods from affective stylistics and humanistic or aesthetics-oriented HCI; that is, understanding the user as involved in a hermeneutic process of interpretation during his/her interaction with the studied artefact. This process can be approximated and visually represented and can draw attention to particular interpretative events related to the user's journey through experimental time and space.

The digital tools therefore seem capable of supporting a procedural analysis of the representation of user experience that, together with traditional analytical instruments, may lead to an aesthetics of user response. Although the discussion focused on four types of users, it was considered that, in reality, the user represents a versatile entity and can be characterized by a combination of features, including those from the showcased categories. More experiments with new cases are expected to further support, test, and possibly validate and enrich the proposed user typology and interpretation modalities, which might in future inform humanistic interface design and the approach to users and user experience models.

In addition, we assume that this type of analysis, beyond its usability-oriented value, may inspire new avenues of reflection on user self-projection and remediation in the digital space, at the intersection of digital hermeneutics, text analysis, linguistics, cognitive science, and the theory of aesthetic response. Possibly, a new way of contemplating how, according to Bolter and Grusin (2000, p. 231), 'we see ourselves today in and through our available media'.

References

Aarseth, E. J. (1997). *Cybertext: Perspectives on Ergodic Literature.* Baltimore and London: The John Hopkins University Press.

Ameka, F. K. and Wilkins, D. P. (2006). Interjections in collaboration with Eline Versluys. In Östman, J. O. and Verschueren, J. (eds), *Handbook of Pragmatics.* Amsterdam/Philadelphia: John Benjamins Publ. Co.

Armaselu, F. and Jones, C. E. (2016). Towards a Digital Hermeneutics? Interpreting the User's Response to a Visualisation Platform for Historical Documents. *DHBenelux 2016: Conference Abstracts.* Belval: Luxembourg.

Armaselu, F. and Jones, C. E. (2017a). From Usability Testing and Text Analysis to User Response Criticism. *Digital Humanities 2017: Conference* Abstracts. McGill University & Université de Montréal, Montréal, Canada, pp. 148–51.

Armaselu, F. and Jones, C. E. (2017b). From Usability Testing and Text Analysis to User Response Criticism. *Conference presentation. Digital Humanities 2017 Conference*, McGill University & Université de Montréal, Montréal, Canada, August 8–11, 2017.

Armaselu, F., Rosselli Del Turco, R., Jones, C., Wieneke, L., Alzetta, C. and Di Pietro, C. (2016). Prototypes as Thinking through Making. Decision Points and Evaluation in Prototyping a Visualisation Framework for Historical Documents. *Digital Humanities 2016: Conference Abstracts.* Kraków, Poland: Jagiellonian University & Pedagogical University, pp. 123–7.

Bakhtin, M. M., Caryl, E. and Michael, H. (1981). *The Dialogic Imagination.* Texas: University of Texas Press.

Bardzell, J. (2009). Interaction Criticism and Aesthetics. *Conference Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009*, Boston, MA, USA.

Bardzell, J. and Bardzell, S. (2016). Humanistic HCI. *Interactions*, **23**(2): 20–9.

Bertelsen, O. W. and Pold, S. (2004). Criticism as an Approach to Interface Aesthetics. *NordiCHI '04*, Tampere, Finland, 23–27 October 2004, pp. 23–32.

Biswas P., and Robinson, P. (2010). A Brief Survey on User Modelling in HCI. *Proceedings of the International Conference on Intelligent Human Computer Interaction (IHCI) 2010, Indian Institute of Information Technology, Allahabad, India, 16th to 18th January 2010.*

Boicu, R. (2007). "Modal verbs and politeness strategies in political discourse". MPRA Paper No. 45913, posted 7 April 2013 02:28 UTC. https://mpra.ub.uni-muenchen.de/45913/1/MPRA_paper_45913.pdf (accessed 14 July 2021).

Bolter, J. D. and Grusin, R. (2000). *Remediation: Understanding the New Media*. Massachusetts: The MIT Press.

Bramley, N. R. (2001*). Pronouns of Politics: The Use of Pronouns in the Construction of 'Self' and 'Other' in Political Interviews*. PhD thesis, Australian National University.

Brown, M. E. and Hocutt, D. L. (2015). Learning to use, useful for learning: a usability study of Google apps for education. *Journal of Usability Studies*, **10**(4): 160–81.

Buchenau, M. and Suri, J. F. (2000). "Experience Prototyping" *In 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*. New York: ACM, pp. 424–33.

Capurro, R. (2010). Digital hermeneutics: an outline. *AI & Society*, **35**(1): 35–42.

Caron-Pargue, J. and Caron J. (1995). La fonction cognitive des interjections. *Faits de Langues*, **6**: 111–20.

Caron-Pargue, J. and Caron, J. (2000). *Cahiers de Praxématique 34*. Montpellier: Pulm, pp. 51–76.

de Hoop, H. and Hogeweg, L. (2014). The use of second person pronouns in a literary work. *Journal of Literary Semantics*, **43**(2): 109–25.

Eliam, M. (2018). The Shapes of Stories by Kurt Vonnegut. https://tenderhuman.com/shapes-of-stories-infographic.

Fallman, D. (2007). Persuade into what? Why Human-computer interaction needs a philosophy of technology. In de Kort, Y., IJsselsteijn, W., Midden, C., Eggen, B., and Fogg, B.J. (eds), *Persuasive Technology. PERSUASIVE 2007. Lecture Notes in Computer Science, Vol. 4744*. Heidelberg: Springer, pp. 295–306.

Fish, S. (1970). Literature in the reader: affective stylistics. *New Literary History*, **2**(1): 123–62.

Fischer, G. (2001). User modeling in human–computer interaction. *User Modeling and User-Adapted Interaction*, **11**(65): 65–86.

Fludernik, M. (1993). Second person fiction: narrative you as addressee and/or protagonist. *Arbeiten aus Anglistik und Amerikanistik*, **18**(2): 217–47.

Fraisse, A. and Paroubek, P. (2015). Les interjections pour détecter les emotions, 22e Conférence sur le Traitement Automatique des Langues Naturelles (TALN 2015), June 2015, Caen, France.

Gibbons, A. and Macrae, A. (eds) (2018). *Pronouns in Literature: Positions and Perspectives in Language*. London: Palgrave Macmillan.

Halliday, M. A. K. and Hasan, R. (2013). *Cohesion in English*. London and New York: Routledge.

Heiden, S. (2010). The TXM platform: building open-source textual analysis software compatible with the TEI encoding scheme. In: Otoguro R, Ishikawa K, Umemoto H, Yoshimoto K, Harada Y (eds.), *24th Pacific Asia Conference on Language, Information and Computation - PACLIC24*, pp. 389–398. Institute for Digital Enhancement of Cognitive Development, Waseda University, Sendai, Japan. https://halshs.archives-ouvertes.fr/halshs-00549764/en.

Huschovà, P. (2015). Exploring modal verbs conveying possibility in academic discourse. *Discourse and Interaction*, **8**(2): 35–47.

Iser, W. (1980). *The Act of Reading: A Theory of Aesthetic Response*. USA: The John Hopkins University Press.

Jockers, M. (2015a). Revealing Sentiment and Plot Arcs with the Syuzhet Package. http://www.matthewjockers.net/2015/02/02/syuzhet/ (accessed 2 February 2015).

Jockers, M. (2015b). The Rest of the Story. http://www.matthewjockers.net/2015/02/25/the-rest-of-the-story/ (accessed 25 February 2015).

Jockers, M. (2017). Introduction to the Syuzhet Package. https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html (accessed 13 December 2017).

Jones, C. E. and Weber, P. (2012). Towards usability engineering for online editors of volunteered geographic information: a perspective on learnability. *Transactions in GIS*, **16**(4): 523–544.

Jones, C. E., Liapis, A., Lykourentzou, I., and Guido, D. (2017). Board game prototyping to co-design a better location-based digital game. *CHI'17 Extended Abstracts*, Denver, CO, USA, 6–11 May 2017.

Jovanović V. Ž. (2004). The Form, Position and Meaning of Interjections in English. *Facta Universitatis Series: Linguistics and Literature*, **3**(1): 17–28.

Lafon P. (1980). Sur la variabilité de la fréquence des formes dans un corpus. *Mots*, **1**: 127–65.

Lewis C. H. (1982). *Using the "Thinking Aloud" Method In Cognitive Interface Design*, No. RC-9265, IBM Thomas J Watson Research Center, Yorktown Heights, NY.

Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability Interface*, **8**(2): 3–6.

Kurosu, M. and Kashimura, K. (1995). *Apparent usability vs. inherent usability. Experimental analysis on the determinants of the apparent usability*, CHI' Companion 95, Denver, Colorado, USA, ACM 0-89791-755 -3/95/0005.

Massanari, A. L. (2010). Designing for imaginary friends: information architecture, personas and the politics of user-centered design. *New Media & Society*, **12**(3) 401–16.

McCarthy, J. and Wright, P. (2004). *Technology as Experience*. Massachusetts: The MIT Press.

Mizzau, M. (2012). **Uncertainty markers in novels**. In *Proceedings of the International Conference The Communication of Certainty and Uncertainty: Linguistic, Psychological, Philosophical Aspects*, University of Macerata (Italy) 3–5 October 2012, under the auspices of the University of Verona (Italy).

Mohamad, Y. and Kouroupetroglou, C. (2013). User modelling. *Research and Development Working Group Wiki.* https://www.w3.org/WAI/RD/wiki/User_modeling (accessed 10 May 2013).

Nielsen, J. (2000). Why You Only Need to Test with 5 Users. NN/g, Nielsen Norman Group, Evidence-Based User Experience Research, Training, and Consulting. https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users.

Oteíza, T. (2017). The appraisal framework and discourse analysis. In Bartlett, T. and O'Grady, G. (eds), *The Routledge Handbook of Systemic Functional Linguistics*, Chapter: 28. England: Routledge, pp. 457–72.

Pennebaker, J. W. (2011). *What our Words Say about Us.* New York, Berlin, London, Sydney: Bloomsbury Press.

Quick, M. (2018). *Every story in the world has one of these basic plots*. BBC Culture. http://www.bbc.com/culture/story/20180525-every-story-in-the-world-has-one-of-these-six-basic-plots (accessed 25 May 2018).

Rinker, T. (2019a). Calculate Text Polarity Sentiment. *Package 'sentimentr*. https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf (accessed 22 March 2019).

Rinker, T. (2019b). Lexicons for Text Analysis. *Package 'lexicon'.* https://cran.r-project.org/web/packages/lexicon/lexicon.pdf (accessed 21 March 2019).

Sacks, H. (1995). *Lectures on Conversation*. Malden, MA: Blackwell.

Sosnovsky, S. and Dicheva, D. (2010). Ontological technologies for user modelling. *International Journal of Metadata, Semantics and Ontologies*, **5**(1): 32–71.

Tractinsky, N. (1997). Aesthetics and apparent usability: empirically assessing cultural and methodological issues, *CHI 97*, Atlanta GA, USA.

Tsui, A. (1991). The pragmatic function of *I don't know*. *Text*, **11**(4): 607–22.

Vonnegut, K. (2010). *Kurt Vonnegut on the Shapes of Stories*. Short lecture by Kurt Vonnegut on the 'simple shapes of stories'. https://www.youtube.com/watch?v=oP3c1h8v2ZQ.

Webb, G. I., Pazzani, M. J. and Billsus, D. (2001). Machine learning for user modeling. *User Modeling and User-Adapted Interaction*, **11**: 19–29.

White, P. R. R. (2015). Appraisal Theory. In Tracy, K., Ilie, C., and Sandel, T. (eds), *The International Encyclopedia of Language and Social Interaction*, First Edition. New Jersey: John Wiley & Sons, Inc.

Wortham, S. (2000). Interactional positioning and narrative self-construction. *Narrative Inquiry*, **10**: 157–84.

Notes

1 Open source version of the Transviewer: https://github.com/C2DH/open-tei-transviewer. Sample accessible via Transviewer (Fig. 1): https://www.cvce.eu/obj/compte_rendu_de_la_quatrieme_reunion_du_groupe_de_travail_sur_la_production_et_la_standardisation_des_armements_paris_31_janvier_1955-fr-5ab2103a-9b3c-4ba5-b6dd-bf22433451ca.html. Source: Union de l'Europe occidentale. Commission intérimaire. Groupe de travail sur la production et la standardisation des armements. Quatrième séance plénière tenue au Palais Chaillot le 31 janvier 1955, à 15h30. Paris: 31.01.1955. PWG/CR/4. pp. 1–6; 14–19. Archives nationales de Luxembourg (ANLux). http://anlux.lu/. Western European Union Archives. Armament Bodies. CPA/SAC. Comité permanent des armements. File CPA-033. Volume 1/1.Copyright: (c) WEU Secretariat General - Secrétariat Général UEO. This document is available at www.cvce.eu (CVCE.EU by UNI.LU).

2 http://oxgarage.tei-c.org/.

3 http://textometrie.ens-lyon.fr/?lang=en.

4 https://github.com/trinker/sentimentr.

5 https://github.com/mjockers/syuzhet.

6   https://www.rstudio.com/.

7   For more details about these features, see the TXM User Manual, Version 0.7, ALPHA, February 2018. http://textometrie.ens-lyon.fr/files/documentation/TXM%20Manual%200.7.pdf.

8   According to the TXM data model, a property refers to information pertaining to a lexical or structural unit. Examples of properties for an English corpus are *word*, *enlemma* and *enpos* (part of speech).

9   See the TXM User Manual, pp. 94–103.

10 *lexicon::hash_sentiment_jockers_rinker*; *lexicon::hash_valence_shifters* (Rinker, 2019b).

11 A substantial usage of interjections was observed for English native speaker participants in one of the reference cases as compared with the sparse occurrences of this type of word in the non-native accounts from Transviewer and the other case.

12 Fr. 'marqueurs de subjectivité et particulièrement d'émotion' (Fraisse and Paroubek, 2015, p. 2).

13 Fr. 'non pas simplement comme des marques ″expressives″ de l'état affectif du sujet parlant, mais comme des traces d'opérations cognitives différenciées' (Caron-Pargue and Caron, 1995, p. 120).

14 Fr. "changement dans l'état des connaissances" (Caron-Pargue and Caron, 2000, p. 56).

15 Fr. "signal pour le partenaire, d'un état mental du locuteur [...]. [...] sa fonction [de l'interjection] est d'informer le partenaire de cet état mental" (Caron-Pargue and Caron, 2000, p. 54).

16 https://en.wiktionary.org/wiki/Wiktionary.

17 https://www.vidarholen.net/contents/interjections/.

18 De-amplifiers were also considered. For example, if *slight, moderate or almost* were thought appropriate to modify a form with a given polarity value in the lexicon, then that value was multiplied by 1/4, $^1/_2$, or 3/4, respectively.

19 *Ah, it is just slow* (R1-INT06; Transviewer).

20 In the *lexicon::hash_sentiment_jockers_rinker,* 'understanding' has a polarity value of 1, while 'confusion' has -0.5. The computed value for the interjection *ah* and its variants was therefore devised as the average: $[1 + (-0.5)/4]/2 = 0.4375 \sim 0.44$. The value for "confusion" was divided by 4, given the influence of the de-amplifier 'slight'.

21 Allowing only the low frequency components of the transformed signal to pass.

22 Using spatialEco::local.min.max, https://github.com/cran/spatialEco/blob/master/R/local.min.max.R.

23 Given the discrete cosine approximation and the low pass filter applied, 4–5 sentences before and after the detected local min./max. points were usually considered for analysis.