Empowering Requirements Elicitation Interviews with Vocal and Biofeedback Analysis

Paola Spoletini, Casey Brock, Rahat Shahwar
Kennesaw State University
Marietta (GA), USA
pspoleti@kennesaw.edu, {cbrock15, rshahwar}@students.kennesaw.edu

Alessio Ferrari
ISTI-CNR
Pisa, Italy
alessio.ferrari@isti.cnr.it

***Abstract*—Interviews with stakeholders are the most commonly used elicitation technique, as they are considered one of the most effective ways to transfer knowledge between requirements analysts and customers. During these interviews, ambiguity is a major obstacle for knowledge transfer, as it can lead to incorrectly understood needs and domain aspects and may ultimately result in poorly defined requirements. To address this issue, previous work focused on how ambiguity is perceived on the analyst side, i.e., when the analyst perceives an expression of the customer as ambiguous. However, this work did not consider how ambiguity can affect customers, i.e., when questions from the analyst are perceived as ambiguous. Since customers are not in general trained to cope with ambiguity, it is important to provide analysts with techniques that can help them to identify these situations. To support the analysts in this task, we propose to explore the relation between a perceived ambiguity on the customer side, and changes in the voice and bio parameters of that customer. To realize our idea, we plan to (1) study how changes in the voice and bio parameters can be correlated to the levels of stress, confusion, and uncertainty of an interviewee and, ultimately, to ambiguity and (2) investigate the application of modern voice analyzers and wristbands in the context of customer-analyst interviews. To show the feasibility of the idea, in this paper we present the result of our first step in this direction: an overview of different voice analyzers and wristbands that can collect bio parameters and their application in similar contexts. Moreover, we propose a plan to carry our research out.**

I. Introduction

Among the variety of elicitation techniques, interviews with stakeholders are the most commonly used one [34], [8], [13], [28], as they are considered one of the most effective ways to transfer knowledge between requirements analysts and customers [51], [21], [20], [35]. Normally, requirements elicitation interviews involve two roles: a customer and a requirements analyst. Ambiguity is considered a major obstacle for knowledge transfer in interviews [22], as it can lead to incorrectly understood needs and domain aspects and may ultimately result in poorly defined requirements.

Past works on ambiguity in requirements engineering mainly focused on natural language ambiguities in requirements documents (e.g., [45], [11], [12], [42], [32], [9], [37], [30], [16], [18], [55], [26], [25], [41]). Our previous work [27] has been the first attempt to analyze the role of ambiguity in requirements elicitation interviews. The focus of our work was to analyze how ambiguity is perceived on the *analyst* side, considering those situations in which an expression of the customer was not understood or was misunderstood by the analyst. Requirements analysts are in general trained on the problem of ambiguity, and can leverage ambiguity resolution strategies such as rephrasing, summaries or exemplification. However, the analyst's perspective is only half of the problem.

Indeed, in customer-analyst requirements elicitation interviews, the customer might not understand, or might misunderstand, some question or statement of the analyst. We refer to all these situations as *ambiguities* on the *customer* side. Notice that our definition of ambiguity is centered on the customer's perception: the same situation might be considered ambiguous by a customer, but not by another (e.g., because they have different background knowledge). Since the customers are normally not trained in dealing with the ambiguity problem, they might react in different, undesirable ways. For example, in case the ambiguity occurs in a direct *question* of the analyst, the customer might answer to the question at the best of his or her understanding, instead of asking for additional clarification. In this way, the ambiguity might be propagated in the customer's answer. In other cases, the customer might not understand a particular *statement* of the analyst, and might ignore this ambiguity, assuming that he or she does not need to understand everything. In the worst case, if ambiguities are too frequent, the customer is likely to *distrust* the analyst, and judge him or her as someone who wants to use technical language to show – but not *share* – his or her knowledge. Overall, these situations go against one of the goals of a requirements elicitation interview, which is extending the common ground [19], [31] between customer and analyst. Since the analyst might involuntarily use technical jargon or ambiguous expressions in general, it is important to provide analysts with tools that help them to spot out when the customer is not understanding, or misunderstanding, some analyst's expressions.

To support the analysts in this task, we propose to explore the relation between a perceived ambiguity on the customer side and changes in his or her voice and bio parameters. The rationale behind this idea is that the perception of an ambiguity by the customer can be generally associated with customer's emotions, which might be negative [23], and emotions can be detected by analyzing the changes in the heart-rate, temperature and other vitals [33], [54], which can be then used as cues of ambiguity on the customer side.

Indeed, humans often react to ambiguous questions to which they feel compelled to answer with an uncomfortable feeling of stress and frustration.Without entering in the debate if physical changes create emotions [36] or the opposite [14], [15], [10], everybody agree that emotion and physical reactions are *coupled*. Different emotions can be manifested in different ways (e.g., Rainville *et al.* [47] look at how cardio-respiratory activities are influenced by the feeling of different emotions). Analogously, acoustic properties of speech indicate emotional differences and can hence be used to further evaluate the current status of the customer [48].

For these reasons, we believe that a *wristband* for collecting biofeedback and a *voice analyzer* to use the voice as an additional cue are a valuable support to address this problem. The feasibility of our idea can be informally proven by analyzing the features of modern voice analyzers and wristbands and their current application contexts. The choice of these particular technologies is also motivated by the fact that, in requirements elicitation interviews, we cannot use any invasive technology, which would affect the behavior of the customer in the interview. The current paper will give a flavor of our current research in understanding the customer's side of ambiguity. At this stage, we evaluated a set of voice analyzers and wristbands and we have considered their potential to be employed in requirements elicitation interviews. In the long term, we plan to experiment with these technologies, and eventually build lightweight systems that can support analysts in identifying customer's perceived ambiguities.

The rest of the paper is organized as follows. In order to support our ideas, in Section II, we propose a comparison among state-of-art wristbands and shows how this technology has been already used for analogous tasks. Similarly, in Section III, we review the most popular voice analyzers. In Section IV, we outline our research agenda, and, in Section V, we discuss some of the risks we might encounter while implementing our research plan and, and for each risk we propose a mitigation strategy. Finally, Section VI concludes the paper.

II. BIOFEEDBACK ANALYZERS

The variety of wristbands currently available on the market able to track vital signs (e.g., heart rate and temperature) is very large and keeps increasing. Most of the available technologies are designed to track the fitness activities and performance of the users. Many of them are not customizable and are limited to evaluate the user's movements during the day and a small set of vital signs to evaluate the user's performance. For our goal, we need more advanced features since we will need to be able to measure small changes in the body reactions of a person who is not involved in any physical activity but is seated on a chair and is having a conversation. For this reason, we have considered only the subset of modern wristbands which offer advanced functionalities and possibly have been already used in other research projects. In particular, we analyzed three wearable devices, all able to collect vitals using different precision technologies: Empatica E3 (or E4) [29], Angel sensor M1 [1], and Basis Smart Watch [2]. Table I gives an overview of the three wristbands.

TABLE I: Analysis of Wristbands.

	E3, E4	**Angel Sensor**	**Basis**
PPG Sensor	+	+	∼
3-axis Accelerometer	+	+	+
Infrared Thermopile	+	+	+
EDA Sensor	+	-	-
Galvanic Skin Response	-	-	+
Internal Real-Time Clock	+	-	-
Customizable	+	++	-
Secure cloud platform	+	-	-
Built in Bluetooth Smart	+	+	∼
Splash resistant	+	+	+
data storage in time	60+ hours	7 days	7days
Battery life	2 days	7 days	4 days
Display screen	-	-	+
Water resistant to 5 ATM	-	-	+

The table is divided horizontally in two parts: in the upper part we have analyzed the sensors which are present (+), partially present (∼) or not present (-) in each wristband. The PhotoPlethysmoGraphy (PPG) sensor measures the blood volume pulse, from which heart rate, heart rate variability, and other cardiovascular features may be derived. Only Empatica E3 and E4 and Angel Sensor are equipped with this sensor. However, the Basis smartwatch is also able to measure the heart rate with different sensor. Angel sensor M1 has also a sensor to measure the blood oxygen saturation, even if this feature is currently unavailable and will be released in a new version of the firmware. All the three wristbands are equipped with an accelerometer and a sensor to measure the temperature. The ElectroDermal Activity (EDA) sensor is used to measure sympathetic nervous system arousal and to derive features related to stress, engagement, and excitement. Only Empatica E3 and E4 are equipped with such a sensor, but the Basis smartwatch can measure the Galvanic skin response, which is a measure of the conductivity of the skin. Empatica E3 and E4 are also equipped with an internal real-time clock with up to 0.2 secondstemporal resolution in streaming mode

The lower part of the table reports the main characteristics of the sensors. Both Empatica's wristbands and Angel Sensor M1 are customizable[1]. The former offer an unobtrusive and remote monitoring service either through a phone or through a desktop app. Moreover, they allow the users to develop their own app to monitor and present the collected data as needed. The latter offers a lot of flexibility through an unrestricted access to sensor data and device internals, a programmable user interface, and open source mobile SDK and apps. All the considered wristbands have more than 2 days battery life (up to 7 in the case of the Angel Sensor) and are splash resistant, but only the Basis smartwatch is water resistant. It is also the only one with a display, that, even if small, allows for immediate feedback without the need of an app.

All the considered wristbands have been used for research projects (e.g., [50], [40]). Particularly relevant are two appli-

[1]The "++" indicates that the Angel Sensor is extremely customizable.

cations of Empatica E3 for relating biofeedback and progress in developer's work. In particular, Müller [43] has used Empatica E3 for understanding the difficulty levels perceived by software developers at work by focusing on the relationship between the biometric measurements and a person's mental state. The work studied two separate issues: the possibility of perceiving developers's difficulty through the data collected with the wristband, the possibility of using the biometric sensors to detect changes in emotions while the participants were working on two different tasks. The initial results of the performed experiment were encouraging. The Empatica E3 was used in another study by Müller et al. [44] to measure the correlation between emotions in developers at work and skin, heart and temperature data. The study showed that calculating the heart rate variability using the time intervals between two consecutive heart beats allows to infer various emotional states. The study also showed a high correlation between emotions and perceived progress.

III. Voice Analyzers

Since it has been recognized that human speech contains emotional content besides the linguistic meaning and the emotions involved in the communication can be used to deduce additional information useful in many field (e.g., marketing analysis), a lot of tools have been introduced with the goal of deducing the emotions felt by a person when he/she speaks. In the following, we present the most popular tools and we highlight their main features and limitations.

One of the most popular tool available is *LVAi* (Layered Voice Analysis-based tool) [39], which is able to measure various properties of the voice. LVAi uses these properties to provide insights into the way subjects think, what troubles them, what excites them, what portions of their speech they are uncertain about, what questions require more of their attention, and what areas appear to be sensitive issues for the speaker. The tool have been used and is still currently used by insurance companies, police departments, department of work and pension, call centers, banks and airports. Notice that Lacerda [38] questioned the validity of LVA and suggested that LVA cannot possibly extract relevant information from the speech signal and the measures produced by LVA are simply an artifact of the digitization of analog speech signals. However, more recently, Elkins et al. [24] provided evidence inconsistent with this view by noting that LVA primitive variables were able to statistically differentiate between truth, deception, stressful, and cognitive dissonance induced speech.

Another popular tool based on the same kind of analysis is *Beyond Verbal* [3], a cloud-based emotions analytics technology that analyzes the tones of the human voice to determine what users are feeling. It provides information on eleven different moods: sad, lonely, anxious, self controlled, friendly, motivated, passionate, happy, domineering, critical, angry. Beyond Verbal works by analyzing voice modulation and by seeking specific patterns in the way people talk. The company claims to detect emotions with 80 percent accuracy. It might be possible to improve that by combining the technology with others, such as systems that can understand words and context. Currently, it is used in call centers.

Another significant tool is *NICE Systems* (Speech Analytics) [6], which combines current and historical speech analytics to categorize and analyze customers voice communication, aiming customer satisfaction and improved agent performance. Nice system is based on combination of technologies, such as phonetic indexing, speech to text, speaker separation, call part, talk over, emotion detection and voice biometrics. It won multiple excellence awards for customer service and is the tool most used in call centers. The tool is also uses in business consulting, education and training. In [17], Cherry confirms that NICE systems are able to detect angry emotions during calls using the changes in a voice's pitch.

OpenSMILE [7] is a speech analysis toolkit, that provides a technically solid and scientifically well evaluated core for detecting attributes from the human voice (e.g., pitch, loudness, energy). It does not detect emotions, but the features it extracts can be used for the purpose. For example, in [52], Tickle et al. use openSMILE to obtain a baseline set of 998 acoustic features from a set of emotional speech recordings to recognize emotions using a supervised neural network.

Other tools are EmoVoice [53], Good Vibrations [5], ESEDA [49], and EMOspeech [4]. EmoVoice is a machine learning-based tool that uses acoustic properties of speech to recognize emotions. Experimental evaluation has given positive results but only considering a limited set of emotions, such as happiness and anger. Good Vibrations interprets acoustic cues (e.g., static and dynamic properties of pitch, intensity, resonances, dullness, sharpness, softness, tempo, and phrasing) as emotions. The tool learns to interpret such a cues over the time by understanding what the users average voice settings are, and learns to detect the smaller hour-to-hour and day-to-day changes in the users emotional state. Good Vibrations can be used for different applications such as personal development feedback, health and gaming applications, and humanizing robots. ESEDA is a speech emotion recognition tool, based on standard supervised machine learning methods and enhanced with an additional block of classification error analysis and fixing. The tool has been only initially evaluated and is not used in practice yet. Finally, EMOSpeech is an enterprise software application that allows call centers to analyze emotion. This technology is based on a continuous psychological model that recognizes different emotion states in a wide spectrum of acoustic emotions allowing better representation and providing more flexibility to identify properties in the voice.

IV. Empowered Requirements Elicitation Interviews

To support the analysts in identifying situations in which the customer is not understanding, or is misunderstanding, some of their expressions, we propose to create a tool that exploits the relation between a perceived ambiguity on the customer side and changes in his or her bio and voice parameters, measured by wristbands and voice analyzers. The tool will provide

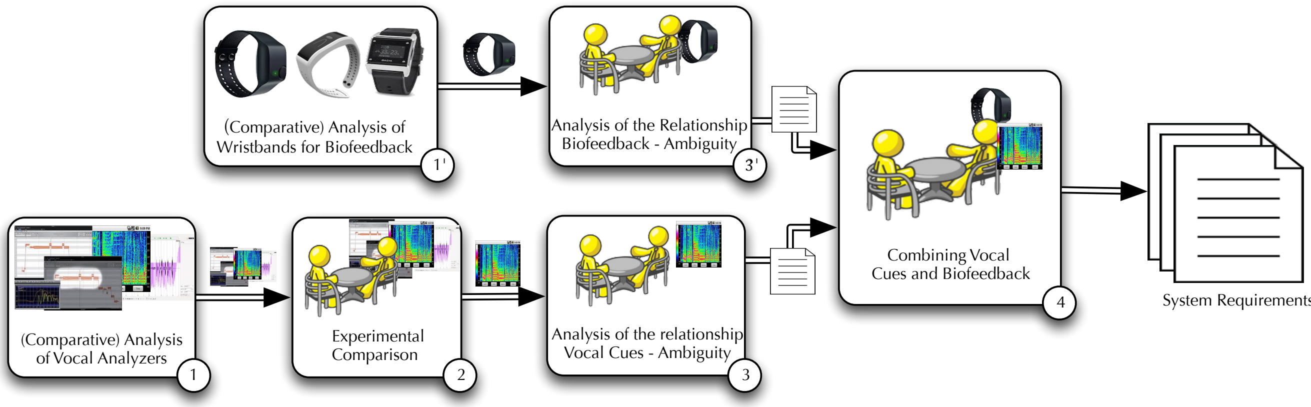

Fig. 1: The proposed research plan.

feedback at runtime to allow the analyst to clarify ambiguous situations and collet additional information, if needed.

Our research plan to reach this goal is schematically represented in Figure 1. The initial steps (1 and $1'$) consist in the analysis of modern wristbands which can precisely track changes in vital signs of the users, and of state-of-art voice analyzers, able to recognize changes in the user's emotions or to extract relevant vocal cues to recognize such changes. We have preliminarily executed these activities, as shown in Sections II and III. In our wristbands analysis, we have analyzed three among the most sophisticated and flexible wristbands on the market: Empatica E3 and E4, Angel Sensor M1, and Basis smartwatch. As a result of our analysis, by considering that our use of the wristband does not require extensive battery but asks for as richer as possible set of precise measures, and taking into account the promising results obtained with Empatica E3 in [43], [44], we consider more promising to adopt the new model of the Empatica series, called E4. As for the voice analyzers, since at the best of our knowledge these tools have not been scientifically analyzed in order to compare their features and performance, and they were very rarely used in research projects, we plan to analyze all the identified tools[2] using a set of experiments (Step 2). The experiments consist of a series of structured interviews with participants acting as customers in which some of the questions are simple and clear, some are difficult but still clear, and a set questions are on purpose ambiguous. We will record the interviews to analyze them with different tools. After the interviews, the participants will be asked to fill out a survey to identify which are the different emotions felt during the interview in correspondence to the different interview questions. We will use the results of the interviews analysis and the surveys to evaluate the most suitable tool for our problem. Since in this phase we are still not aware of which are the emotions identified by the tool that are related with ambiguities, we will look for changing in the emotions of the user (either in type or in intensity) and not for specific emotions. The choice of using structured interviews is motivated by the need to compare the participants' reactions to similar situations.

[2]We will exclude Good Vibrations from our analysis since it evaluates the emotion by learning the user characteristics over time, and this is completely out of scope with respect to our analysis.

In steps $3'$ and 3, we will analyze the *correlation between collectable data, emotions and ambiguity* for the chosen wristband and the voice analyzer, respectively. In these steps we want to relate the data that the chosen technologies can collect with the occurrences of ambiguity. To analyze this correlation we will perform two set of experiments. In the following we describe them for step $3'$. The same experiments will be performed also for step 3 without asking the participants to wear the wristband. Steps $3'$ and 3 are kept separated because we want to study the correlation between the collected feedback and emotions for the two technologies independently. Moreover, since the effect of wearing the wristband on the the reaction of the participants is still unknown, we will not reuse the interviews of step $3'$ for step 3.

- In the first set of experiments we will create three types of ad-hoc structured interviews: one type of interview will be designed to be ambiguity-free, the second will have a few questions designed to be perceived as ambiguous and the third kind will be predominantly ambiguous. The interviews will be short (10-15 minutes). We will perform 60 interviews, 20 per type. We will use only 20 participants, and each will perform all the three interviews in separate days, starting with the ambiguity-free interview and finishing with the predominantly ambiguous one. We will recruit participants who have already participated in a few interviews, so that our experiment will not be affect by the difference in familiarity with the interview process that novel participants might have from the first to the last interview and familiarity would not be a variable in our evaluation. The interviewees will be asked to wear the Empatica E4. Both the data collected by the device and the interviews will be recorded. As in the previous step, after the interviews, participants will be asked to fill out a survey to identify which are the questions that they found ambiguous or, more in general, when they felt there was a

misunderstanding situation. We will use the results of the analysis of the interviews and the surveys to categorize situations (e.g., changes in the emotions, changes in the intensity of an emotion) which may represent cues of ambiguity.
- In the second set of experiments we will repeat the same experiment performed in the first phase, but, this time, our goal is to validate the categories found in the previous experiments. If from the validation, new categories emerge, we need to evaluate also them.

In step 4, we will analyze if wearing the wristband causes any changes in the voice cues, and, hence, if we have to "re-tune" the cues found in step 3. Moreover, we want to evaluate if considering the data collected by the two technologies together can give additional cues. For the former problem, we will perform structured interviews (followed by a survey) to two set of participants: the first wearing the wristband, the second without it. The interview will be structured as in step 2. If any significant difference is found between the two sets, we need to reanalyze the cues obtained in Step 3. For the latter problem, we will perform a similar experiment to the one performed in steps 3 and $3'$ with the goal of finding how we can combine the information coming from the wristband and the voice analyzer to have additional cues. By using the correlation observed in the last three steps, we will obtain requirements of the lightweight tool that we plan to develop as a support for the analyst The tool will use the inputs from the voice analyzer and the data collected by the wristband to notify to the interviewer possible problems on the customer side. In the interviews empowered by this tool, the interviewer will use the notifications from the tool to identify possible ambiguous situation perceived by the customer, and to clarify his/her expression, or use other strategies to let the customer discuss his/her doubts. We are also considering to run focus groups with customers and real analysts, in order to discuss our findings and possibly elicit further requirements for the tool that we plan to develop.

V. Risks and Mitigations

Our research plan presents possible problems and risks. In the following, we will present some of them, together with possible mitigation strategies.

First of all, the *current tools might be not suitable for our problem*. Indeed, even if we are encouraged by the fact that biofeedback and vocal analysis has been successfully used in many different fields, such as call centers to evaluate the customer satisfactions and to analyze how emotions influence programmers productivity (see Sections II and III for more details), the available technologies might not have the needed features to detect changes of emotions in our context. In case the problem arises for the voice analyzers, we plan to use tools such as OpenSMILE to study how the raw characteristics of the voice change in correspondence of ambiguous situations, and we can use the raw features influenced in our context to develop an ad-hoc tool, thus, ignoring the correspondence between features and emotion. Moreover, we could evaluate the adoption of sentiment analysis [46], which by using natural language processing, text analysis and computational linguistics could help determine the polarity of the interviewee's sentences. This would allow us to mix verbal and non verbal cues in the decision process. However, the challenge with verbal cues, which represent the main contribution of sentiment analysis with respect to the initially proposed solution, is in collecting the linguistic information contained in any corpus of audio. This is traditionally done using text-to-speech systems, but such systems are prone to creating significant errors in the textual indexes they create and this may compromise the overall analysis. In case of problems with the chosen wristband, we will first experimentally evaluate the other two initially considered devices and, in case of negative results also with such devices, we will analyze, using more invasive and precise technologies, if the problem is the precision of the measurements or in the types of data that can be collected. In the first case, we will look for smaller changes in the data, in the second we will explore other technologies.

Another problem could be that *customers might feel that the fact that they are wearing the wristband creates an unbalanced situation* in the interview. To overcome this problem, we can establish a protocol which imposes to both the analyst and the customer to wear the wristband. They both will have access to tool which highlights possible ambiguous situations. This protocol would balance the interview and may help to create a fairness feeling in the customer.

As for the emotions, there could be the problem that *different persons may have different reactions or the same reaction with different intensity to the same situation*. This risk has to be taken into account in the list of requirements for the tool that we aim to develop. For this reason we are planning to fund the tool on a parametrized algorithm that includes different reactions for the same situation and measures the variation of intensity in the feedback of a person. The algorithm is tuned using standard questions, which help to understand the variability in the collected data for the current customer. Moreover, we could use a learning algorithm that takes advantage from the feedback from each interview to refine the algorithm.

On the emotions side, we need also to consider that *the same emotion might be caused by different factors*. For example, the level of stress of the customer might be caused not only by an ambiguous situation, but also by the length of the interview or by his or her involvement with the interview topic. During our experiments, we will use surveys to identify what the customers feel to be the reasons of each emotional situation. If we notice that the cause of some emotions cannot be identified univocally using the collected data, we will develop a set of guidelines to help interviewers to disambiguate these situations and to act accordingly to improve the quality of the interview.

Finally, *since the tool's notifications are caused by different data, they might have different probability to be correct*. To address this issue, we plan to enrich the algorithm implemented in the tool with a learning mechanism that learns which situations are more likely to be real ambiguous situations and use this information to assign a suspiciousness degree to each notification. This information can be used by the analyst to decide if he or she wants to neglect or not the notification.

VI. Conclusion

In this paper we presented a novel idea on how to identify ambiguity on the customer's side in requirements elicitation interviews. Since the customer is not trained to deal with ambiguity, our idea is to provide a tool to the analyst to recognize when the customer is not (completely) understanding, so that he or she can add additional information or rephrase to disambiguate the situation. We plan to develop such a tool by exploiting the relation between a perceived ambiguity on the customer side and changes in the voice and bio parameters of the customer, measured through wristbands and voice analyzers. We have shown the feasibility of the idea by analyzing the state-of-art voice analyzers and wristbands to estimate if the existing tools are applicable in this context. We have then presented our research plan and discussed possible risks (and ideas to mitigate them).

References

[1] Angel sensor m1. http://angelsensor.com/technology/.

[2] Basis smartwatch. http://www.mybasis.com/.

[3] Beyond verbal. http://www.beyondverbal.com/.

[4] Emospeech. http://emospeech.net/.

[5] Good vibrations. http://goo.gl/mEaZHH.

[6] Nice systems. http://goo.gl/qCgi5R.

[7] Opensmile. http://www.audeering.com/research.

[8] R. Agarwal and M. R. Tanniru. Knowledge acquisition using structured interviewing: an empirical investigation. *JMIS*, 7(1):123–140, 1990.

[9] V. Ambriola and V. Gervasi. On the systematic analysis of natural language requirements with Circe. *ASE*, 13, 2006.

[10] P. Bard. A diencephalic mechanism for the expression of rage with special reference to the central nervous system. *Am. J. Physiol*, 1928.

[11] D. M. Berry and E. Kamsties. The syntactically dangerous all and plural in specifications. *IEEE Software*, 22(1):55–57, 2005.

[12] D. M. Berry, E. Kamsties, and M. M. Krieger. From contract drafting to software specification: Linguistic sources of ambiguity, 2003.

[13] G. J. Browne and M. B. Rogich. An empirical investigation of user requirements elicitation: Comparing the effectiveness of prompting techniques. *JMIS*, 17(4):223–249, 2001.

[14] W. B. Cannon. The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *The American Journal of Psychology*, 39(1/4):106–124, 1927.

[15] W. B. Cannon. Against the JamesLange and the thalamic theories of emotions. *Psychol. Rev.*, 38:281295, 1931.

[16] F. Chantree, B. Nuseibeh, A. N. D. Roeck, and A. Willis. Identifying nocuous ambiguities in natural language requirements. In *RE'06*, pages 56–65, 2006.

[17] S. Cherry. Anger management [emotion recognition]. *IEEE Spectrum*, 42(4), 2005.

[18] A. Cimatti, M. Roveri, A. Susi, and S. Tonetta. Formalizing requirements with object models and temporal constraints. *SoSyM*, 10(2), 2011.

[19] H. H. Clark. *Using language*. Cambridge University Press, 1996.

[20] J. Coughlan and R. D. Macredie. Effective communication in requirements elicitation: a comparison of methodologies. *Requir. Eng.*, 7(2):47–60, 2002.

[21] A. Davis, O. Dieste, A. Hickey, N. Juristo, and A. M. Moreno. Effectiveness of requirements elicitation techniques: Empirical results derived from a systematic review. In *RE'06*. IEEE, 2006.

[22] A. Distanont, H. Haapasalo, M. Vaananen, and J. Lehto. The engagement between knowledge transfer and requirements engineering. *IJKL*, 1(2):131–156, 2012.

[23] T. Edwards and N. Y. U. Psychology. *Implicit and Explicit Affective and Motivational Consequences of Social Ambiguity Among Depressed Individuals with Anxiety*. New York University, 2007.

[24] A. C. Elkins, J. Burgoon, and J. Nunamaker. Vocal analysis software for security screening: Validity and deception detection potential. In *Homeland Security Affairs, Supplement*. 2012.

[25] A. Ferrari and S. Gnesi. Using collective intelligence to detect pragmatic ambiguities. In *RE'12*, pages 191–200. IEEE, 2012.

[26] A. Ferrari, G. Lipari, S. Gnesi, and G. O. Spagnolo. Pragmatic ambiguity detection in natural language requirements. In *AIRE'14*. IEEE, 2014.

[27] A. Ferrari, P. Spoletini, and S. Gnesi. Ambiguity as a resource to disclose tacit knowledge. In *RE'15l*, pages 26–35, 2015.

[28] W. R. Friedrich and J. A. Van Der Poll. Towards a methodology to elicit tacit domain knowledge from users. *IJIKM*, 2(1):179–193, 2007.

[29] M. Garbarino, M. Lai, D. Bender, R. Picard, and S. Tognetti. Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. In *EAI MOBIHEALTH'14*, 2014.

[30] B. Gleich, O. Creighton, and L. Kof. Ambiguity detection: Towards a tool explaining ambiguity sources. In *REFSQ'10*, volume 6182 of *LNCS*, pages 218–232. Springer, 2010.

[31] M. Glinz and S. A. Fricker. On shared understanding in software engineering: an essay. *CSRD*, pages 1–14, 2014.

[32] S. Gnesi, G. Lami, and G. Trentanni. An automatic tool for the analysis of natural language requirements. *IJCSSE*, 20(1), 2005.

[33] A. Haag, S. Goronzy, P. Schaich, and J. Williams. Emotion recognition using bio-sensors: First steps towards an automatic system. In *ADS'04*, pages 36–48. Springer, 2004.

[34] I. Hadar, P. Soffer, and K. Kenzi. The role of domain knowledge in requirements elicitation via interviews: an exploratory study. *Requir. Eng*, 19(2):143–159, 2014.

[35] A. M. Hickey and A. M. Davis. A unified model of requirements elicitation. *J. Manage. Inf. Syst.*, 20(4):65–84, Mar. 2004.

[36] W. James. What is an emotion? *Mind*, 9:1188205, 1884.

[37] L. Kof. From requirements documents to system models: A tool for interactive semi-automatic translation. In *RE'10*, 2010.

[38] F. Lacerda. Lva-technology the illusion of lie detection. In *Proceedings of FONETIK 2009, Dept. of Linguistics, Stockholm University*, 2009.

[39] N. Ltd. Lva. http://www.nemesysco.com/lva.html.

[40] P. Mancuso, M. Thompson, M. Tietze, S. Kelk, and G. Roux. Can patient use of daily activity monitors change nurse practitioner practice? *The Journal For Nurse Practitioners*, 10, 7 2014.

[41] A. K. Massey, R. L. Rutledge, A. I. Anton, and P. P. Swire. Identifying and classifying ambiguity for regulatory requirements. In *RE'14*, pages 83–92. IEEE, 2014.

[42] L. Mich and R. Garigliano. Ambiguity measures in requirements engineering. In *ICS'00, 16th IFIP WCC*, pages 39–48, 2000.

[43] S. C. Müller. Measuring Software Developers' Perceived Difficulty with Biometric Sensors. In *ICSE'15*, volume 2, pages 887–890. IEEE, 2015.

[44] S. C. Müller and T. Fritz. Stuck and frustrated or in flow and happy: Sensing developers' emotions and progress. In *ICSE'15*, volume 1, pages 688–699. IEEE Press, 2015.

[45] P. G. Neumann. Only his only grammarian can only say only what only he only means. *ACM SIGSOFT SE Notes*, 9(1):6, 1986.

[46] B. Pang and L. Lee. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.*, 2(1-2):1–135, 2008.

[47] P. Rainville, A. Bechara, N. Naqvi, and A. Damasio. Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International journal of psychophysiology*, 1(61):5–18, 2006.

[48] K. R. Scherer. Vocal communication of emotion: A review of research paradigms. *Speech Commun.*, 40(1-2):227–256, 2003.

[49] J. Sidorova and T. Badia. ESEDA: tool for enhanced speech emotion detection and analysis. *Procesamiento del Lenguaje Natural*, 41, 2008.

[50] R. Steele and A. Clarke. The internet of things and next-generation public health information systems. *Communications and Network*, 5(3B):4–9, 2013.

[51] A. Sutcliffe and P. Sawyer. Requirements elicitation: towards the unknown unknowns. In *RE'13*, pages 92–104. IEEE, 2013.

[52] A. Tickle, S. Raghu, and M. Elshaw. Emotional recognition from the speech signal for a virtual education agent. *Journal of Physics: Conference Series*, 450(1), 2013.

[53] T. Vogt, E. André, and N. Bee. EmoVoice - A framework for online recognition of emotions from voice. In *PIT'08*. Springer, 2008.

[54] J. Wagner, J. Kim, and E. Andre. From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification. In *ICME 2005*. IEEE, 2005.

[55] H. Yang, A. N. D. Roeck, V. Gervasi, A. Willis, and B. Nuseibeh. Analysing anaphoric ambiguity in natural language requirements. *Requir. Eng.*, 16(3):163–189, 2011.