Towards a Conversational Expert System for Rhetorical and Vocal Quality Assessment in Call Center Talks

Mathias Walther[1], *Baldur Neuber*[2], *Oliver Jokisch*[3], *Taïeb Mellouli*[1]

[1]Department of Business Information Systems and Operations Research, Martin Luther University Halle-Wittenberg, Germany
[2]Department of Speech and Phonetics, Martin Luther University Halle-Wittenberg, Germany
[3]Institute of Communications Engineering, Leipzig University of Telecommunication, Germany

mathias.walther@wiwi.uni-halle.de

Abstract

This article presents the concept and development steps towards a conversational expert system for rhetorical and vocal quality assessment in call center talks. At first the state of the art in quality assessment is discussed. The influencing rhetorical and vocal factors are introduced. In our novel approach, the recognition of vocal factors is modeled by competing classification systems and combined into a multi-classifier system which is based on decision trees. Finally we propose an expert system which incorporates the generated decision rules. The system accuracy can be improved by user-adapted rule sets. Furthermore solutions to the problem of inconsistent rules are discussed.

Index Terms: conversation quality, call center, fusion, expert system, multi-classifier system

1. Introduction

Despite the rapid development of other communication channels, the telephone is still the "no. 1" communication mean in customer service. In Germany, there are 6,800 call centers and omni-channel centers with 520,000 employees, who conduct 25 million calls per day. In addition to private companies, public institutions, such as health care providers, hand over a significant proportion of their total media communication—and thus also phone calls—to the call center branch. Therefore we are confronted with an "industrialized conversation production" of gigantic magnitude [1]. In a telephone call, only one communication channel is used, i. e. all interaction is solely based on the agent's and customer's voices. Because of the huge amount of calls, a manual monitoring can only consider a very small proportion of talks. However there is no available support system for automatically monitoring conversational quality in call center talks, although automatic speech processing is on the research agenda since the 1950s. The automatic analysis and processing of prosodic and paralinguistic aspects came in focus in the mid 1990s and is still under research. A well-studied field in the paralinguistic speech processing is the emotion detection with pattern recognition methods, including prototypical applications for call centers. The customer can be directed from an interactive voice response (IVR) system either to a human agent or to a computer depending on his emotional state (see e. g. [2, 3, 4]). A significant drawback of these techniques is the lack of explanatory capabilities which are needed to measure and/or to explain the complex phenomenons of conversational quality.

This contribution presents a modeling approach for detecting vocal markers of conversation quality starting with the working hypothesis that conversation quality can be described by perceptual features which can be recognized by automatic classification systems. In parallel these features can be understood and interpreted by human experts.

1.1. System vision and methodology

Figure 1 shows the vision and the general use case of our system under development. During a call the agent's voice is recorded and analyzed. The analysis is done in real-time and has a high explanatory power to support both agent and trainer in their effort to improve conversational quality. The monitor on the right side of Figure 1 exemplarily depicts some criteria for speech quality. To strengthen usability and simplicity, the feedback in the example is given by emoticons.

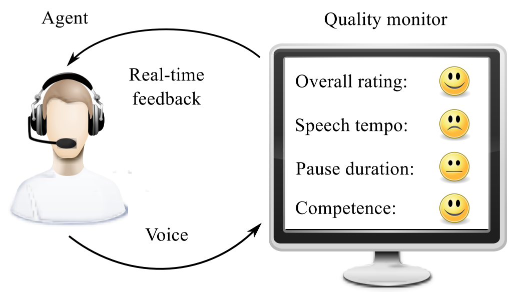

Figure 1: *System vision of an intelligent conversational expert system as a quality monitor for call center agents.*

Our classification approach consists of four steps which are described as follows:

1. Definition of markers of conversational quality and identification of speech features which are relevant for the perception of quality in call center talks.

2. Development of classification models for the automatic recognition of these features.

3. Classifier selection and combination of pre-trained models for speech features to a fusion system based on decision trees.

4. Creation of an expert system with a knowledge base that contains the fusion model's classification rules.

2. Conversation quality in a call center

2.1. State of the art

In a call center the monitoring of conversation quality is a critical part of general quality management and an essential success factor [5]. From the rhetorical and conversational point of view the main question is: How can people, who have to perform large quantities of similar calls under time constraints, always be adequately friendly, knowledgeable, and respond individually to each conversation partner? Currently, the call center branch—at least in Germany—almost exclusively relies on instruction from branch experts. The quality assessment is conducted by trainers and team leaders—the "experts" and follows best-practice methods. Hence, coaching strategies are applied that rely on personal findings and experiences. The coaching is usually based on general, introductory literature. There is no scientific foundation of conversation quality in a professional communication context. Furthermore, there is a lack of knowledge in rhetorical and negotiation abilities of the agent.

Figure 2 illustrates the consequences of this loss-making situation [1]. The "vicious cycle of telecommunication" starts

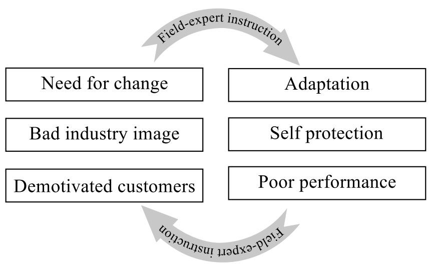

Figure 2: *The "vicious cycle of telecommunication" [1].*

with the field-expert instruction. Neglecting the fact that communication is an individual process, the agents are urged to reproduce the guidelines which make him feel uncomfortable or "artificial". This pressure on the agent often leads to a low performance. These stereotypes, especially the acted friendliness, which is perceived by the customer, lead to a bad image of the whole branch. This bad image causes a need for change, which is implemented by field-expert instruction. Thus, the cycle starts a new loop.

2.2. Need for action

With the aim of a substantiated empirically and theoretically based didactic for professionally operated phone calls on an industrial scale, we have been systematically exploring the criteria of conversation quality using authentic corpora since 2006. We rely on a combination of methods from qualitative and quantitative linguistic, phonetic and conversational approaches. In previous studies, we identified six factors for conversation quality in professional telephone calls: conversational form, emotionality, intelligibility, conversational partner orientation, personality/authenticity and situational adequateness. As seen in Figure 3, the main focus of research is speech and voice presentation since these criteria—criteria that can assessed in a qualitative way or that can be measured by its acoustic correlates—are important components of speech perception [6]. According to our present knowledge, the following factors are relevant for high conversation quality in professional telephony:

- The consideration of the conversation quality factors shown in Figure 3,
- Flexibility in the communication style,
- A deliberate speech presentation,
- Transparency,
- Polite and empathetic conversational manner,
- The ability to listen to the partner, and to recognize minimal signals.

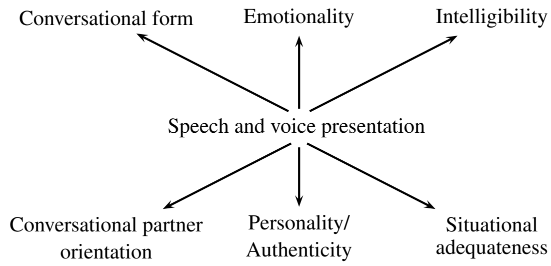

Figure 3: *Conversation quality factors [7].*

Since probably not all, but several of these factors can be parameterized, we rely on a combination of didactic implementation via interpersonal communication with computerized assistance. An ideal intelligent support system for conversations is able to process simultaneously all conversational aspects of the involved partners: form, content and meta data. This requires an—at least partially—automatic analysis of both form and function of conversation quality which will be briefly presented in the following section (detailed discussion in [8]).

3. Classification

3.1. Corpus

Starting from the factors for conversational quality and vocal characteristics as described in the previous section (see Figure 3), an annotation catalogue was developed [9, 7]. The corpus is a collection of 800 real sales talks from three outbound campaigns that have been provided by three call centers [10]. Due to resource limitations, 218 talks were selected for the final corpus, which were regarded as typical representatives in the corresponding categories [7]. For the annotation process, the factors where subdivided into assessment criteria. Table 1 summarizes the most important criteria. The labeling was done by four experts in speech science, who labelled parts of the corpus separately, i. e. each segment was rated by one expert. The classes and corresponding number of instances are shown in column two of Table 1. Except from the accent form, which has three classes, all other categories split in two classes. In this article we focus on speech and voice presentation and on competence as the only factor of conversational quality, which is shown in the last row of Table 1. Other quality factors can be modeled in a similar way.

3.2. Base models and classifier selection

For the experiments, a feature set with 2,106 features was used. The features were extracted with openEAR [11] based on the configuration file from the "Interspeech Paralinguistic Challenge 2010" [12]. This configuration has been complemented with five formants and their statistical functionals. In addition the gender of the agent has been manually labeled. During the first experiments the base models were built with Weka toolkit [13] and its classification algorithms: Naive Bayes (NB), Bayesian networks (BN), logistic model tree (LMT), RIPPER (JRip), support vector machine (SMO), AdaBoost (Ada), C4.5 tree (J48) and multilayer perceptron (MLP). The algorithmic details are described in [14] and [15]. The classification performance is measured by the recognition rate (RR) which is calculated via ten-fold cross-validation. Table 1 shows in column four the classification rates for speech and voice presentation. In the third column, we denoted the classification algorithm which produces the best models. The ranking of the algorithms was conducted via analysis of variance (ANOVA) followed by Duncan's test. The statistical methods are discussed in detail in [16].

The results in Table 1 show that the accuracy of the base models for all dichotomous classes is above 65 %. The best accuracy is reached for loudness and speech pitch. These results are expected since these voice features have acoustic correlates in pitch and intensity, which can be measured [17]. Pause type, melody jump and stress frequency show good recognition rates, too. However, having less than 10 instances per class, the results are not significant. The criterion accent type has an accuracy of 48 % which outreaches 33 %—the expected value of random guessing within three classes.

The last row of Table 1 depicts the accuracy of the base model for competence trained in same way as the models above. The best result, 78 % accuracy (rank seven among all recognition rates), is reached by the SMO. Although comparison to further studies is difficult, due to fundamental differences in corpora, feature sets and classification algorithms, it can bee seen that the competence model has a competitive accuracy to those presented in other surveys, e. g. the "Interspeech Speaker Trait Challenge 2012" [18], which set a benchmark at 70 % accuracy for personality traits (OCEAN-Model).

3.3. Classifier fusion

Besides a good prediction accuracy, which is reached in some criteria, the explanatory power of the models is a crucial factor in designing an intelligent software system as monitoring tool for call center conversations.

The classification models we presented in the previous section show a good accuracy, but they are poor in terms of their ability to explain decisions (introspection). As our results show, speech and vocal features are better recognized than competence with these base models which operate on signal features. To overcome that drawback, a special multi-classifier system was developed. The classification system has a two-layered structure, which is shown for competence in Figure 4. The low-level classifiers are trained on perceptual features a human expert can understand. The high-level model extends the concept of fusion systems by a learning algorithm, which creates a traceable decision tree. In general, fusion systems are combinations of individually trained classification models [19].

The modeling process of the fusion model is as follows: at first every instance that belongs to competence is classified by the chosen model for speech and voice presentation (see Table 1), leading to 13 independent decisions for every instance. The original class attribute, either "competent" or "incompetent", is kept in the new instance. Figure 5 shows an instance of the class "competent" in arff-format used by Weka.

The second development step is to learn a new classification model with the instances created in the first step, which is called fusion model. Due to the structure of the input data, the fusion model can be trained with any classification algorithm that can handle nominal attributes. Due to the fact that traceability of decisions is important, the well-known C4.5 algorithm

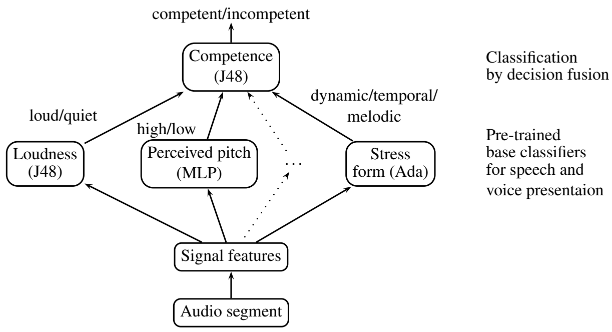

Figure 4: *Schematic diagram of the fusion model for competence.*

```
[...]
@data
loud, low, end pauses, strong, high,
interrogative, inflected, short, high, fast,
high, unpleasant, dynamic, competent
[...]
```

Figure 5: *Training instance for the fusion classifier.*

[20] is used for tree learning. Figure 6 shows a trained decision tree for competence as an example. The C4.5 algorithm reaches an accuracy of 73.2 % in ten-fold cross validation [8], which is comparable to the base model (Table 1).

A benefit of decision trees is the ability to transform the branches and leafs into decision rules [21]. During the transformation the leaves become the target classes and the branches become decisions upon the attributes' values. Accordingly the tree in Figure 6 can be transformed into 11 decision rules. Figure 7 shows the derived rules for leafs 1,2 and 8 in a formal notation. The shortest rules are 1 and 8 with two attributes each. The longest rules are 6 and 7, which cover seven attributes.

For evaluation purpose, confidence measures for each generated rule are calculated. For a given rule n the confidence c_n is the likelihood of the predicted outcome, provided that the rule has been satisfied. The support c_n is defined as the number of instances that satisfy the rule divided by the number of instances in the training set [22]. In Figure 6, confidence c and support s are shown in each leaf. The rules 2, 4, 5, 6, and 7 have the highest confidence but a small number in support, which means that they rather represent special cases than universal rules. Rule 8 has both—a high confidence as well as a high support (see Figure 6).

Our experiments show that the rules are unstable, i. e. they are sensitive to changes of the learning set. This means that the generated trees and rules are not as universal as it would be sufficient for a practical use. To some extent, this data sensitivity can be reduced by better base classifiers. Nonetheless the following external factors have a significant impact on the rule set:

- The agent's vocal and rhetorical abilities and habits,
- the trainer's perception,
- the language that is spoken,
- technical settings (recording, etc.) and

Table 1: *Classification accuracy for speech and voice presentation and competence.*

Criterion	Classes (Number of Instances)	Algorithm	RR (%)
loudness	loud (57), quiet (57)	J48	96.89
perceived pitch	high (146), low (146)	MLP	94.84
pause type	end pauses (9), inner pauses (9)	LMT	95.00
melody jump	strong (6), weak (6)	NB	90.00
pause frequency	high (37), low (37)	BN	86.43
phrase-final melodic contour	terminal (57), interrogative (57)	SMO	78.79
perceived pitch contour	inflected (120), uninflected (120)	MLP	77.08
pause duration	long (45), short (44)	MLP	75.42
accent frequency	high (8), low (8)	MLP	75.00
speech tempo	fast (44), slow(44)	MLP	75.00
speech tension	high (98), low (98)	BN	69.97
timbre	pleasant (74), unpleasant (74)	BN	65.43
accent form	dynamic (71), temporal (71), melodic (71)	Ada	48.00
competence	*incompetent (71), competent (71)*	*SMO*	*78.76*

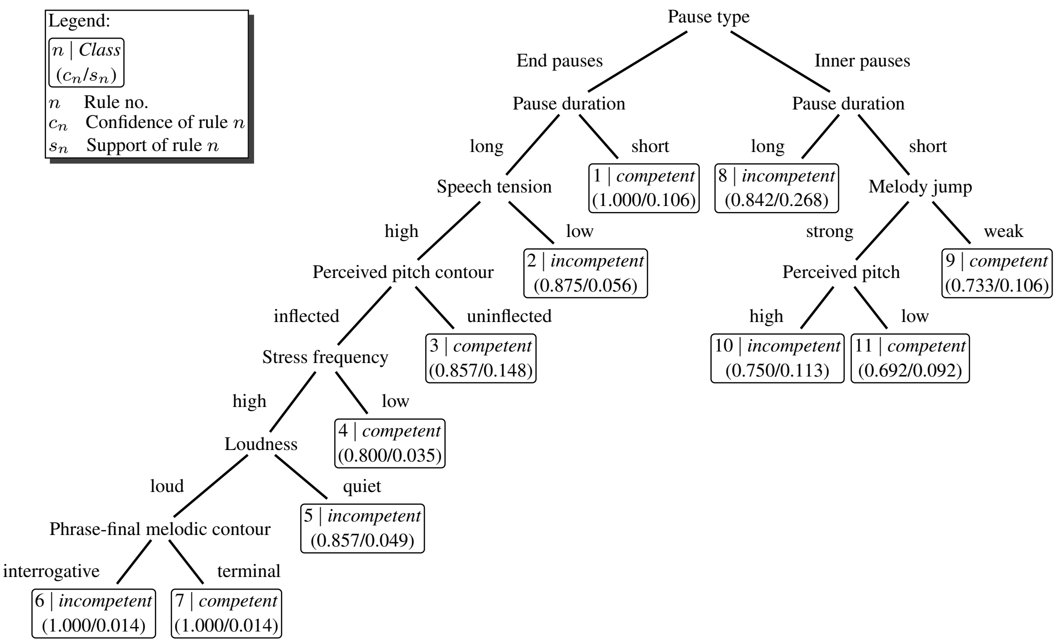

Figure 6: *C4.5 decision tree as fusion model for competence.*

```
1: competent := pause_type = end_pauses AND
   pause_duration = short
2: incompetent := pause_type = end_pauses
   AND pause_duration = long AND
   speech_tension = low
   ...
8: incompetent := pause_type = inner_pauses
   AND pause_duration = long
```

Figure 7: *Decision rules 1, 2 and 8 derived from the decision tree depicted in Figure 6.*

- the content of the conversation.

Since the user, both coach and agent, cannot train new base models to consider the external factors, the system has to be adaptable on the perception level to improve classification results in practical use.

4. Design of an expert system for conversational quality assessment

4.1. Baseline concept

As discussed in the previous section, the derived decision rules are unstable and thus have a limited accuracy in real-live scenarios. Therefore, the adaption of the system to all relevant factors is crucial to its classification performance in practical use. The presented fusion system with its ability to explain its decisions can be integrated into an expert system for quality assessment and tutoring purpose. The expert system has the following use case scenarios:

1. Feeding the knowledge base with the produced rules,
2. Adaption of the rules to the environmental setting and
3. Assistance to the agent during live calls.

At first the expert system has to be provided with decision rules. The initial rule set is generated by the fusion system, which learns decision trees. In addition to a conventional expert system, the proposed system must be provided with algorithms that process support and confidence measures, which are part of the input, too.

During the use of the system, it can continuously improve its accuracy by modifying the initial rules or generating new ones. Modification of the knowledge base is done by changing the values for confidence and support as follows:

During the adaption phase the system's decisions will be presented to the coach. While listening to a specific segment of the agent's call, he is provided with the system's high-level classification result and the decision rule that was used.

Then the coach has to judge, whether the rule is applicable in the current case. If the coach agrees with the rule's decision, the number of true positives for this particular rule will increase, and in the same vein support and confidence will increase. Otherwise theses numbers will be reduced.

Hence the values for support and confidence for each single rule change over the adaptation time. To make use of this evolving rule base, the rules need to be sorted in the knowledge base. The inference engine only uses rules with confidence or support above a certain threshold. The advantage of this use case is that the inference and the calculations are done by the system and the coach doesn't need to know internals of the decision process, he only needs a basic understanding of the involved speech features and the generated rules.

4.2. Components and architecture

Since a simple black-box approach is not suitable for explanation, a glass-box structure [23] is needed for our system goals. Figure 8 shows the adapted architecture of the proposed expert system with its components and connections [24]:

Knowledge base. The knowledge base is the main component of an expert system. It is the repository of the domain knowledge that is used for reasoning and problem solving. It contains the decision rules from the learned trees shown in section 3.3.

Inference engine. The inference engine is the rule interpreter for the decision rules in the knowledge base. The proposed system has extended capabilities to operate with numerical values for support and confidence.

Knowledge acquisition. This component's function is the transfer of the collected knowledge from the expert to the knowledge base.

Explanation facility. This component describes the system's actions that were executed to solve the given problem.

User interface. This subsystem communicates with the user and the expert. It needs to have several user interfaces to address different user roles. One user interface is the proposed real-time monitoring in Figure 1.

User. In the proposed system, there are two different user roles: the agent and the trainer instead of one user the in usual expert system.

Fusion system. The fusion system adds its knowledge, i. e. decision rules, via the acquisition component into the system. It replaces the knowledge engineer who provides the system with domain knowledge.

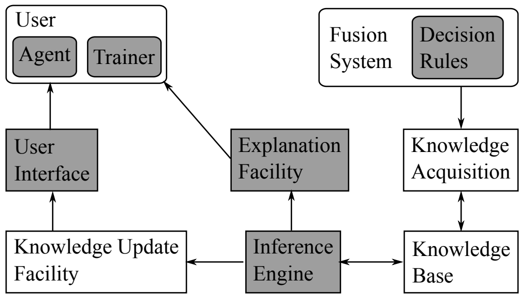

Figure 8: Architecture for a conversational expert system (adapted from [24])

4.3. Handling potential inconsistencies

Since the adaption process of the rule set involves different knowledge sources (decision trees) and different experts (agent and coach), inconsistencies can appear in the knowledge base. For the sake of accurate inference with association rules, [25] suggest to calculate the formerly introduced certainty factors (CF) [26] from confidence and support values of mined rules. They firstly show both theoretically and experimentally, that the CF framework representing measure of (increased) beliefs or disbeliefs avoids the extraction of misleading rules with high support [25]. Two examples for these misleading rules are rules no. 1 and 8 (see Figures 6 and 7), which have high support values (0.106 and 0.286). However, their base models are trained on a small amount of data with 18 instances in "pause type" and 89 in "pause duration", respectively. Because of this, the rules' classification accuracy can be low on unknown data, although confidence and support on the training data are high. The anomaly of gradual/partial inconsistencies when reasoning with mined or experts' knowledge under uncertainty is discussed more generally in [27]. In order to detect inconsistencies within argumentations, he extends the CF model by introducing complex certainty factors (CCF) able to represent and propagate belief and disbelief in a separate way. For evaluating conclusions and classifications deduced from mined and expert rules, a "skepticism factor" is introduced reflecting the amplitude of inconsistency in the actual inference state.

5. Discussion

The building process for the classifiers—for both base classifiers and fusion systems—is subject of further investigation. Since the accuracy of the rule set depends on good base classifiers, the goal is to improve their classification accuracy. Another topic for further research is the development of the rule based system. We have shown that components need to be extended with new functions. This includes the handling of contradictions within the knowledge base and the algorithmic processing of the numerical values for confidence and support in the inference machine.

The discussed CCF model as extension to the certainty factors provides ways of avoiding misleading rules and conclusions as well as recognizing exceptions. Besides inherent absolute inconsistencies, Mellouli [27] discusses the type of apparent inconsistency which can be resolved in either stronger belief or disbelief by acquiring more specific information—a very important issue for association rules equivalent to going in-depth within a decision tree.

An important factor that we did not address yet, is segmentation which is a necessary step within the prosodic speech processing [28]. In order to process paralinguistic speech information on a continuous speech signal, the segments should be long enough to contain all relevant information, but not too small [29]. In our experiments, we used manually labeled segments. Since this not practicable in a real-world application, a suitable strategy for segmentation has to be found. The experiments were carried out for German. Since the rhetoricity of conversations is to a considerable extent culturally specific, multilingual findings can not be applied.

6. Conclusion

The analysis and optimization of industrial conversational skills creates a challenging field of research and requires learning technologies and educational support. Given the huge number of calls, computerized support, e. g. analysis and assistance systems, become necessary. We showed that conversation quality in call centers, which is indiceted by vocal and rhetorical features, can be detected with classification algorithms. Despite their acceptable recognition accuracy, the models lack explanatory power to serve as a feedback system. In order to circumvent these drawbacks, we presented the concept of an expert system which is based on high-level decision rules generated by a fusion model. The proposed system improves the assessment of conversational quality by aggregation and formalization of expert knowledge. It also complements the current training and coaching with well-funded, easily understandable and continuously presented feedback. In addition to the live-monitoring features, the system can be used for long term analysis of conversations: when the rules detect negative speech and voice presentation and negative quality ratings, coaching methods—tailored to the agent and the detected deficiencies—can be applied.

7. References

[1] B. Neuber and U. Hirschfeld, "Sprechwirkungsforschung in der professionellen Telefonie," in *Klangsprache im Fremdsprache-nunterricht VII*, L. Veličkova and E. Petročenko, Eds. Woronesh State University, 2013, pp. 66–85.

[2] D. Morrison, R. Wang, and L. C. D. Silva, "Ensemble methods for spoken emotion recognition in call-centres," *Speech Communication*, vol. 49, no. 2, pp. 98–112, 2007.

[3] V. Petrushin, "Emotion in speech: Recognition and application to call centers," in *Artificial Neural Nets in Engineering. (ANNIE '99)*, 1999, pp. 7–14.

[4] F. Burkhardt, M. van Ballegooy, R. Englert, and R. Huber, "An emotion-aware voice portal," in *Proc. Electronic Speech Signal Processing ESSP*, 2005, pp. 123–131.

[5] K. Dawson, *The Call Center Handbook: The Complete Guide to Starting, Running, and Improving Your Call Center*, 5th ed., ser. Call Center Handbook. San Francisco: CMP Books, 2003.

[6] H. Fastl and E. Zwicker, *Psychoacoustics: facts and models*, 3rd ed., ser. Springer Series in Information Sciences, M. R. S. T. S. Huang, T. Kohonen, Ed. Berlin: Springer, 2007, vol. 22.

[7] S. Meißner and J. Pietschmann, "Rhetorische und phonetische Einflussfaktoren auf die Qualität von Telefonverkaufsgesprächen," in *Erforschung und Optimierung der Callcenterkommunikation.*, U. Hirschfeld and B. Neuber, Eds. Berlin: Frank & Timme, 2011, pp. 215–248.

[8] M. Walther, T. Mellouli, and O. Jokisch, "Fusion von Klassifikationsmodellen zur automatischen Erkennung von Stimmeigenschaften in der Qualitätsbewertung von Callcentergesprächen," in *Konferenzband Elektronische Sprachsignalverarbeitung (ESSV) 2015*, ser. Studientexte zur Sprachkommunikation, G. Wirsching, Ed., vol. 78. Dresden: TUDpress, 2015, pp. 188–195.

[9] U. Hirschfeld and B. Neuber, Eds., *Erforschung und Optimierung der Callcenterkommunikation.* Berlin: Frank & Timme, 2011.

[10] S. Meißner, J. Pietschmann, M. Walther, and L. Nöbel, "Innovative IT-gestützte Ansätze zur Bewertung der Gesprächsqualität in Telefonverkaufsgesprächen," in *Erforschung und Optimierung der Callcenterkommunikation.*, U. Hirschfeld and B. Neuber, Eds. Berlin: Frank & Timme, 2011, pp. 195–214.

[11] F. Eyben, M. Wöllmer, and B. Schuller, "openear - introducing the munich open-source emotion and affect recognition toolkit," in *Proc. 4th International HUMAINE Association Conference on Affective Computing and Intelligent Interaction 2009 (ACII 2009)*, vol. I. IEEE, 2009, pp. 576–581.

[12] B. Schuller, S. Steidl, A. Batliner, F. Burkhardt, L. Devillers, C. Mueller, and S. Narayanan, "The Interspeech 2010 paralinguistic challenge," in *INTERSPEECH 2010*, 2010, pp. 2795–2798.

[13] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten, "The weka data mining software: an update," *SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, 2009.

[14] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. San Francisco: Morgan Kaufmann, 2011.

[15] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, and P. S. Yu, "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2008.

[16] A. P. Bradley, "The use of the area under the roc curve in the evaluation of machine learning algorithms," *Pattern Recognition*, vol. 30, pp. 1145–1159, 1997.

[17] A. Paeschke, *Prosodische Analyse emotionaler Sprechweise*, ser. Mündliche Kommunikation. Berlin: Logos, 2003.

[18] B. Schuller, S. Steidl, A. Batliner, E. Nöth, A. Vinciarelli, F. Burkhardt, R. van Son, F. Weninger, F. Eyben, T. Bocklet, G. Mohammadi, and B. Weiss, "The interspeech 2012 speaker trait challenge," in *INTERSPEECH 2012*, 2012, pp. 254–257.

[19] H. Hertlein, *Fusion von Klassifikationssystemen für die automatische Sprechererkennung*. Berlin: Logos, 2010.

[20] J. R. Quinlan, *C4.5: programs for machine learning*. San Mateo: Morgan Kaufmann, 1993.

[21] F. Gorunescu, *Data Mining: Concepts, models and techniques*. Berlin: Springer, 2011.

[22] D. L. Olson and D. Delen, *Advanced data mining techniques*. Berlin: Springer, 2008.

[23] J. R. Anderson, "The expert module," in *Foundations of intelligent tutoring systems*, M. C. Polson and J. J. Richardson., Eds. Hillsdale: Erlbaum, 1988, pp. 21–53.

[24] K. P. Tripathi, "A review on knowledge-based expert system: Concept and architecture," *IJCA Special Issue on Artificial Intelligence Techniques - Novel Approaches & Practical Applications*, no. 4, pp. 21–25, 2011.

[25] F. Berzal, I. Blanco, D. Sánchez, and M. A. Vila, "Measuring the accuracy and interest of association rules: A new framework," *Intelligent Data Analysis*, vol. 6, no. 3, pp. 221–235, 2002.

[26] E. H. Shortliffe and B. G. Buchanan, "A model of inexact reasoning in medicine," *Mathematical biosciences*, vol. 23, no. 3, pp. 351–379, 1975.

[27] T. Mellouli, "Complex certainty factors for rule based systems–detecting inconsistent argumentations," in *WLP 2014 – 28th Workshop on (Constraint) Logic Programming*, ser. CEUR Workshop Proceedings, S. Brass and J. Waldmann, Eds., vol. 1335, 2014, pp. 81–102. [Online]. Available: http://ceur-ws.org

[28] K. Bartkova and D. Jouvet, "Automatic detection of the prosodic structures of speech utterances," in *Speech and Computer – 15th International Conference, SPECOM 2013*, ser. Lecture Notes in Computer Science, M. Železný, I. Habernal, and A. Ronzhin, Eds. Heidelberg: Springer, 2013, vol. 8113, pp. 1–8.

[29] M. Mansoorizadeh and N. Charkari, "Speech emotion recognition: Comparison of speech segmentation approaches," *Proc of IKT*, 2007.