Still Together?: The Role of Acoustic Features in Predicting Marital Outcome

Md Nasir[1], Wei Xia[1], Bo Xiao[1], Brian Baucom[2], Shrikanth S. Narayanan[1], Panayiotis Georgiou[1]

[1]Signal Analysis & Interpretation Laboratory, Univ. of Southern California, Los Angeles, CA, USA
[2]Dept. of Psychology, University of Utah, Salt Lake City, UT, USA.

Abstract

The assessment and prediction of marital outcome in couple therapy has intrigued many clinical psychologists. In this work, we analyze the significance of various acoustic features extracted from couples' spoken interaction in predicting the success or failure of their marriage. We also investigate whether speech acoustic features can provide complementary information to behavioral descriptions or codes provided by human experts (*e.g.,* relationship satisfaction, blame patterns, global negativity). We formulate marital outcome prediction as both binary (improvement *vs.* no improvement) and multiclass (different levels of improvement) classification problem. Our experiments show that acoustic features can predict marital outcome more accurately than those based on behavioral descriptors provided by human experts. We also find that dialog turn-level acoustic features generally perform better than frame-level signal descriptors. This observation supports the notion that the impact of the behavior of one interlocutor on the other is more important than the behavior itself looked in isolation. Finally, acoustic features together with human-derived behavioral codes show the best performance in outcome prediction, suggesting some complementarity in the information captured by these behavioral representations.

Index Terms: Behavioral Signal Processing (BSP), couple therapy, outcome prediction, acoustic features, classification

1. Introduction

Behavioral Signal Processing (BSP) [1, 2] is the study of behavior from signals, based on analysis of the multimodal cues of human interaction. It connects real-world physical signals to abstract complex human behavior [3], often in social context. In BSP literature, couple therapy (notably marriage counseling) has been one of the most significant domains of application for signal-driven approaches of behavior. During last few years, there have been many attempts and progress on extraction and classification of behavior of couples in distressed relationship using information gathered from various modalities such as speech [4–7], visual gestures [8], and spoken language use [9, 10].

In observational psychology, monitoring, understanding, and even predicting the outcomes of different human conditions related to the distress, and illness—including response to interventions for these—through behavioral data is crucial, be it marital conflict, addictive behavior or neuro-cognitive developmental issues such as in Autism Spectrum Disorders. As an example, the marital outcome of a distressed couple undergoing counseling has been a subject of long standing interest in order to assess the quality and the effectiveness
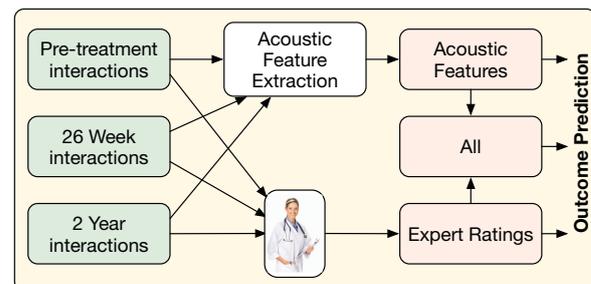

Figure 1: Overview of the work described in this paper

of marriage counseling sessions [11–13]. An objective demonstration of the quality of couple therapy and the progress of the couple towards a better relationship can be helpful to the domain experts in evaluating and adjusting their therapy methods. These aforementioned studies investigated how changes in different behavioral traits and communication patterns of the couples are linked to the final outcome of their marital therapy, *i.e.,* whether they recovered from their distress or their relationship had little or no improvement.

In many cases, the studies involved follow-up of the couples over a prolonged period of time, up to 5 years [14]. Most of the analyses are however limited to manual observational assessment or self report of behavior. The goal of this work is to investigate vocal (speech) signal cues directly in a computational setting with regard to their association to marital outcome. Having such a computer-assisted technology to indicate outcome can be beneficial to the advancement of behavioral outcome studies in psychology towards assessing quality of care and treatment methods.

In this paper, we explore the problem of predicting the outcome of couple therapy from acoustic features during couple interactions as depicted in the upper part of Figure 1. We employ a range of spectral and prosodic measures for verbal and non-verbal cues of the interlocutors for the same. For comparison, we also formulate the prediction task based on session-ratings of domain experts, as shown in the bottom path of Figure 1. This behavioral coding provided by human experts is deemed relevant for understanding the nature of the relationship and its dynamics. Finally we employ both human and machine provided information towards outcome prediction as in the middle of Figure 1.

The rest of the paper is organized as follows. The methodology of our analysis is outlined in Section 2. A brief description of the corpus used for our experimental study is given in Section 3, followed by a description of the experimental setup in Section 4. The results are presented in Section 5, followed by the conclusion and future work.

Feature Type	Feature Names
Prosody	Loudness and its derivative, Intensity, Pitch
Spectral	15 MFCCs and their derivatives, 8 MFBs
	and their derivatives, 8 LSFs and their derivatives
Voice quality	Jitter, Shimmer and their derivatives
Functionals	Minimum, Maximum
	Range (*i.e.,* Maximum − Minimum)
	Mean, Median, Standard Deviation

Table 1: *Acoustic features and six statistical functionals.*

2. Methodology

In this section, we describe the general framework of marital outcome prediction, employing and comparing two different representations of human behavior—one that is based directly on vocal cues of verbal and nonverbal communication using recordings of couple interactions while the other uses human-expert specified and derived constructs (behavioral codes).

First, we present the methodology for extracting various spectral and prosodic acoustic features which can be essentially viewed as machine-perceived cues from the speech signal.

2.1. Acoustic Feature Extraction

Acoustic feature extraction from the speech signal is preceded by two stages of preprocessing, namely Voice Activity Detection (VAD) and Speaker Diarization. Next, in order to obtain their session-level representations, we compute various functionals over individual speaker segments extracted from a conversation session. To enrich the acoustic characterization that better reflects the interaction dynamics, we use the turn-level delta features which are also described later.

2.1.1. Preprocessing of the audio signal

Voice Activity Detection: Acoustic features are extracted only for speech regions. For this purpose, we used the Voice Activity Detection (VAD) system described in [15] to separate the audio into speech *vs.* non-speech regions.

Speaker Diarization: We next performed diarization that first segments the speaker changes using Generalized Likelihood Ratio based criteria, and then clusters the speaker-homogeneous segments using agglomerative clustering [16]. This way we partition the entire interaction session into regions spoken by each of the speakers.

2.1.2. Frame-level audio features

Extraction of frame-level features: We extract a variety of acoustic features characterizing speech prosody (pitch and energy), voice quality (jitter, shimmer) and spectral envelope characteristics (MFCCs, MFBs, LSFs). In literature, pitch and other prosodic features have shown to reflect vocal characteristics relating to various behavioral aspects [5–7]. On the other hand, spectral features like MFCCs are known to be informative in various tasks related to emotion recognition [17–19] and behavioral signal processing [5, 7].

These Low-Level Descriptors (LLD) are extracted every 10 *ms* with a 25 *ms* Hamming window. We used *PRAAT* [20] to extract pitch (f_0) and intensity while the rest of the LLDs were extracted using *openSMILE* [21]. In total we have used 74 acoustic features which are summarized in Table 1.

Functionals: Next we compute functionals of all these acoustic
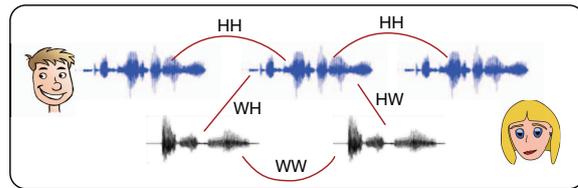

Figure 2: Turn-level delta features employ the differences between the means of features of adjacent turns in the interaction, both within an interlocutor (*e.g.,* Wife to wife turn changes) but also across interlocutors (*e.g.,* Wife to husband turn changes)

features as listed in Table 1. They are computed separately for each of the speakers (husband and wife) across the session, generating a set of session-level features for every LLD. This way we ensure capturing the acoustic cues from each of the speakers individually. For performing marital outcome prediction of couple therapy, we repeat this procedure for all the interaction sessions the couple were engaged (at multiple time points over the period of the clinical trial), so as to gather information about the interaction between spouses before and after a considerable period of counseling.

2.1.3. Turn-level delta features

Aspects of communication of one interlocutor affect the other. The acoustic features described in the previous section are based on features of one speaker in isolation, and hence cannot fully capture interaction phenomena like dyadic coordination and entrainment. To address this, turn level analysis is often adopted such as for example in the context of emotion recognition [22, 23]. Similar features are useful in the context of dyadic interactions such as in the context of behavioral entrainment [7]. This motivates the use of features which can capture such coupled changes.

Our turn-level delta feature extraction process is two-fold: first, we compute the mean of each acoustic feature over each turn of a speaker. Next, we compute the differences between the means of the features over adjacent turns of the same speaker as well as the differences of the means of features over one turn of a speaker from those over the previous turn of the other speaker, as depicted in Figure 2. Our hypothesis is that these turn-level difference or *delta* features can capture useful information about the mutual and self-influence of behavioral patterns of the speakers over time.

2.2. Behavioral Coding: Human-Provided Features

Observational psychologists often use standardized rating systems for assessment of behavioral constructs, also known as *behavioral codes*, that are relevant to characterizing the domain and problem under consideration (*e.g.,* in couple therapy, coding "blame" patterns of spouses is of interest, while in behavioral therapy for addiction the "empathy" expressed by the provider is of relevance). These codes are quantified representations of abstract behaviors of interest, typically along specific scales (*e.g.,* "positivity" of the husband on a scale of 1-9). These codes are usually dependent on the specific scope of their observation. Usually trained psychologists construct and follow specific behavior rating manuals [24, 25]. In couple therapy, in addition to individual-specific behavior constructs (*e.g.,* sadness), many of these behavioral codes also consider the

behavior of an individual towards his/her partner (*e.g.*, blame).

2.3. Classification

2.3.1. Feature Selection

The described feature extraction that aims to capture various aspects of behavior leads to a high dimensionality of the feature set, particularly compared to the training data available. We usually perform feature selection to choose a subset of the original features, that provides the maximum information in the context of a particular classification problem. In our paper, we have used the Mutual Information Maximisation [26] feature selection method available as a part of the *FEAST* toolbox [27].

2.3.2. Support Vector Machine

We use the well-known *Support Vector Machine* (SVM) for the marital outcome prediction formulated as both binary and multiclass classification problems. Even though SVM is a binary classifer by virtue, it can be extended to multiclass problems [28], and we do so in this work. We have used the one-against-all method of multiclass SVM, which, as the name suggests, decomposes the multiclass problem into a number of binary classification problems.

3. Corpus

We use the Couple Therapy corpus [29], which consists of audio-visual recordings of couples during real problem-solving interactions. This data, a part of larger longitudinal clinical study, was collected by a collaboration between the University of California, Los Angeles and the University of Washington. 134 chronically distressed couples participated in this study, who were married on average for 10.0 years (*SD*= 7.7). The recorded and coded sessions took place before therapy began, after 26 weeks, and after 2 years of therapy. Every session has two sub-sessions: one on the topic chosen by the husband and the other by the wife, and the duration of each session is ten minutes. Due to participant attrition and drop out in the study, there are some missing sessions.

The interactions were manually coded following CIRS [24] and SSIRS [25] coding manuals for couple therapy. Every behavioral code was rated by multiple human annotators on a scale of 1 to 9. These expert-derived ratings were used in our study to predict marital outcome both separately and along with acoustic features. We used the mean of different ratings given by different annotators for a specific code. In our classification experiments, these ratings are viewed as human-provided features for classification and compared to the direct signal-derived acoustic features.

The marital outcome of the couples in above mentioned corpus were observed and recorded in subsequent studies [13, 14]. These outcomes represented the recovery of the couples at the time of either 26-week or 2 years-session after the time they started the therapy (*pre* to *post*-therapy). The outcome was defined on categorical ordinal scale: 1 = deteriorated (*i.e.,* they got measurably worse over treatment), 2 = no change (*i.e.,* no meaningful improvement), 3 = improved (*i.e.,* they got measurably better over treatment, but still clinically insignificant), 4 = recovered (*i.e.,* they got measurably better over treatment and their score is above the upper cut-off for clinically significant distress). These outcome ratings are considered the reference labels for our classification tasks.

In the Couple Therapy corpus, the outcome ratings could not be recorded for some couples due reasons like dropout of couples from the study, or insufficient information. After discarding such cases from our dataset, we had 139 instances of outcome, which had i) outcome after 26 weeks relative to pre-treatment, and ii) outcome after 2 years relative to pre-treatment. Therefore we have 139 samples in our dataset, every sample belonging to one of the 4 outcome-classes (with ratings 1 though 4) shown in Table 2.

Outcome	Decline	No Change	Partial Recovery	Recovery
Rating	1	2	3	4
Count	12	26	34	67

Table 2: *Number of instances with different outcome ratings*

4. Experimental Setup

We conducted three different types of classification experiments, each using different combinations of our feature sets. The experiments are as follows:

- **Experiment 1**: Classification of all instances into 2 classes, *i.e.,* complete recovery (rating 4) *vs.* incomplete or no recovery (ratings 1,2,3 combined)

- **Experiment 2**: Classification of instances of no (or incomplete) recovery into finer levels, *i.e.,* rating 1 *vs.* rating 2 *vs.* rating 3

- **Experiment 3**: Classification of all possible outcomes *i.e.,* ratings 1 through 4.

One of the reasons behind the choice of multiple experiments as mentioned above is the sample imbalance in the dataset, as seen in Table 2. For each of these experiments, we use various combinations of different feature sets:

  (i) frame-level acoustic features,
  (ii) turn-level delta features,
  (iii) all acoustic features (frame-level and turn-level delta combined),
  (iv) behavioral codes,
  (v) all features (all acoustic features and behavioral codes combined)

Each of the aforementioned feature sets comprise corresponding features extracted from *pre* and *post*-therapy sessions. Prior to each classification task, we perform z-score normalization on every feature and use feature selection (Mutual Information Maximisation method) to select the optimal subset of features. To account for variability in the dataset, 10-fold cross-validation is performed.

As a part of our analysis, we also investigate the relative significance of different categories of acoustic features in contributing to the outcome prediction. In order to do this, we consider all acoustic features and perform leave-one-feature-type-out with Experiment 1. This is done by partitioning the acoustic feature set into the following four categories: i) MFCC, ii) MFB, iii) LSF, and iv) prosodic and voice quality features (pitch, intensity, jitter, shimmer) and removing one of these categories at a time from the original acoustic feature set.

5. Results

Table 3 shows the classification accuracy of different feature sets by performing classification using SVM. The mean

Featureset	Dim.	Expt. 1		Expt. 2		Expt. 3	
		mean	*SD*	*mean*	*SD*	*mean*	*SD*
Frame-level	3552	76.4	10.0	70.9	13.8	63.2	11.4
Turn-level Delta	5328	78.4	8.3	70.7	15.5	60.4	10.2
All Acoustic	8880	77.7	12.7	72.3	14.6	**63.4**	9.5
Behavioral Codes	264	75.6	13.5	65.4	14.7	61.8	11.2
All Features	9144	**78.7**	7.1	**74.3**	12.4	63.1	13.6

Table 3: *Classification accuracy (in terms of their mean and standard deviation over all folds of cross-validation) of different experiments (across the columns) with different feature sets (across the rows)*

Feature Set	Expt. 1	Expt. 2	Expt. 3
Frame-level	0.56	0.60	0.52
Turn-level Delta	0.62	0.58	0.50
All Acoustic	0.69	**0.63**	**0.55**
Behavioral Codes	0.68	0.49	0.48
All Features	**0.76**	**0.63**	0.48

Table 4: *F-measures (micro-averaged) of different experiments (across the columns) with different feature sets (across the rows)*

accuracy and standard deviation over all folds are reported. Moreover, the full original dimensionality of each feature set is also reported. However, it should be noted that the number of features actually used by the classifier is reduced due to feature selection prior to actual classification.

Since we have an unbalanced dataset (especially for multiclass) with respective to different outcome classes, we also reported mean F-measure (with micro-averaging) [30] of the predicted labels for each classification setup in Table 4. The F-measure values are in the interval $(0, 1)$ and higher value of F-measure signifies better quality in classification.

Based on the obtained classification accuracy and F-measures, we can make the following observations:

- We achieved higher accuracy and F-measure in all 3 experiments for the feature set of all acoustic features (frame-level and turn-level combined) than using frame-level or turn-level features separately. It suggests that turn-level delta features provide complementary information that frame-level features.

- Furthermore, turn-level features alone outperform frame-level features in Experiment 1, but their accuracy is slightly lower than frame-level features in Experiments 2 and 3. However it does not undermine their importance due to their complementary to frame-level features, as mentioned above.

- We can also infer from the results that direct signal-derived acoustic features can predict the marital outcome better than the manual ratings of behavioral codes.

- The classification results after fusing manual ratings with acoustic features are mixed. While fusion appears to help overall discrimination between improvement and no improvement cases (Experiment 1), it does not seem to help the finer classification of Experiments 2 and 3.
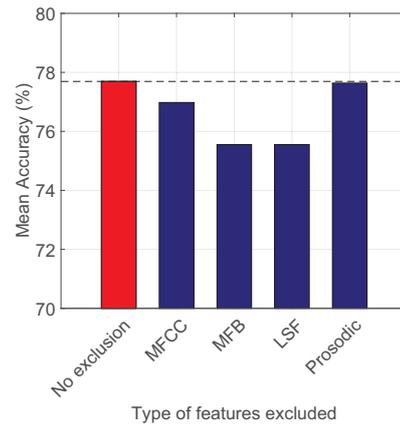

Figure 3: *Classification accuracy (%) of different acoustic feature set for leave-one-feature type-out experiment (larger difference with the leftmost bar signifies more importance of the corresponding feature type)*

Figure 3 shows the results of the leave-one-feature-type-out analysis for Experiment 1 considering all acoustic features to be the complete feature set. The leftmost bar is for reference, which is the accuracy with complete feature set and has the maximum accuracy (77.69%). As we can observe, the removal of spectral features lowers the accuracy to a much higher extent, compared to prosodic and voice quality features. MFB and LSF features are similar in significance (both cases with 75.54% accuracy), and found to be most informative in outcome prediction, based on this experiment. The figure also shows that, even if we choose to drop the prosodic and voice quality features, there is little loss of accuracy (77.63% after excluding the prosodic and voice quality features).

6. Conclusions and Future Work

This paper presents an analytical study on predicting outcomes of couples undergoing marital therapy by directly using information from speech. In the context of outcome prediction, the study also compares this automatic approach of capturing important behavioral information from signal in form of acoustic features to the traditional approach taken by psychologists, *i.e.,* manual coding of behavior.

Based on the results of our experiments, we observe that acoustic features capture more relevant information than the human-encoded behavioral constructs for predicting marital outcome. We have also found that turn-level delta features are better in terms of performance than simple frame-level acoustic features for outcome prediction.

This study opens up many opportunities for future investigation such as the analysis of the importance of various communication channels including lexical (*i.e.,* what is being spoken), visual (*e.g.,* head-movement and other face and body gestures) as well as the use of more complex and advanced representations of speech signal that can capture temporal information across sessions and dynamics of local behavioral cues of the speakers.

Although the focus of this study was marital outcome, similar approaches can be investigated for predicting outcome in other spoken interaction based interventions as well.

7. References

[1] P. G. Georgiou, M. P. Black, and S. S. Narayanan, "Behavioral signal processing for understanding (distressed) dyadic interactions: some recent developments," in *Proceedings of the 2011 joint ACM workshop on Human gesture and behavior understanding*, 2011, pp. 7–12.

[2] S. Narayanan and P. G. Georgiou, "Behavioral signal processing: Deriving human behavioral informatics from speech and language," *Proceedings of the IEEE. Institute of Electrical and Electronics Engineers*, vol. 101, no. 5, p. 1203, 2013.

[3] A. Pentland, "Socially aware, computation and communication," *Computer*, vol. 38, no. 3, pp. 33–40, 2005.

[4] M. Black, A. Katsamanis, C.-C. Lee, A. C. Lammert, B. R. Baucom, A. Christensen, P. G. Georgiou, and S. S. Narayanan, "Automatic classification of married couples' behavior using audio features." in *INTERSPEECH*, 2010, pp. 2030–2033.

[5] M. P. Black, A. Katsamanis, B. R. Baucom, C.-C. Lee, A. C. Lammert, A. Christensen, P. G. Georgiou, and S. S. Narayanan, "Toward automating a human behavioral coding system for married couples interactions using speech acoustic features," *Speech Communication*, vol. 55, no. 1, pp. 1–21, 2013.

[6] J. Gibson, A. Katsamanis, M. P. Black, and S. S. Narayanan, "Automatic identification of salient acoustic instances in couples' behavioral interactions using diverse density support vector machines." in *INTERSPEECH*, 2011, pp. 1561–164.

[7] C.-C. Lee, A. Katsamanis, M. P. Black, B. R. Baucom, A. Christensen, P. G. Georgiou, and S. S. Narayanan, "Computing vocal entrainment: A signal-derived PCA-based quantification scheme with application to affect analysis in married couple interactions," *Computer Speech & Language*, vol. 28, no. 2, pp. 518–539, 2014.

[8] B. Xiao, P. G. Georgiou, B. Baucom, and S. S. Narayanan, "Data driven modeling of head motion towards analysis of behaviors in couple interactions," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, pp. 3766–3770.

[9] P. G. Georgiou, M. P. Black, A. C. Lammert, B. R. Baucom, and S. S. Narayanan, "That's aggravating, very aggravating: Is it possible to classify behaviors in couple interactions using automatically derived lexical features?" in *Affective Computing and Intelligent Interaction*, 2011, pp. 87–96.

[10] S. N. Chakravarthula, R. Gupta, B. Baucom, and P. Georgiou, "A language-based generative model framework for behavioral analysis of couples' therapy," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Apr. 2015.

[11] M. Sevier, K. Eldridge, J. Jones, B. D. Doss, and A. Christensen, "Observed communication and associations with satisfaction during traditional and integrative behavioral couple therapy," *Behavior therapy*, vol. 39, no. 2, pp. 137–150, 2008.

[12] W. K. Halford, M. R. Sanders, and B. C. Behrens, "A comparison of the generalization of behavioral marital therapy and enhanced behavioral marital therapy." *Journal of Consulting and Clinical Psychology*, vol. 61, no. 1, p. 51, 1993.

[13] K. J. Baucom, B. R. Baucom, and A. Christensen, "Changes in dyadic communication during and after integrative and traditional behavioral couple therapy," *Behaviour research and therapy*, vol. 65, pp. 18–28, 2015.

[14] K. J. Baucom, M. Sevier, K. A. Eldridge, B. D. Doss, and A. Christensen, "Observed communication in couples two years after integrative and traditional behavioral couple therapy: outcome and link with five-year follow-up." *Journal of consulting and clinical psychology*, vol. 79, no. 5, p. 565, 2011.

[15] M. Van Segbroeck, A. Tsiartas, and S. S. Narayanan, "A robust frontend for VAD: Exploiting contextual, discriminative and spectral cues of human voice," in *INTERSPEECH*, Aug. 2013.

[16] K. J. Han, S. Kim, and S. S. Narayanan, "Strategies to improve the robustness of agglomerative hierarchical clustering under data source variation for speaker diarization," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 16, no. 8, pp. 1590–1601, 2008.

[17] C.-C. Lee, E. Mower, C. Busso, S. Lee, and S. Narayanan, "Emotion recognition using a hierarchical binary decision tree approach," *Speech Communication*, vol. 53, no. 9, pp. 1162–1171, 2011.

[18] M. Grimm, K. Kroschel, E. Mower, and S. Narayanan, "Primitives-based evaluation and estimation of emotions in speech," *Speech Communication*, vol. 49, no. 10, pp. 787–800, 2007.

[19] O.-W. Kwon, K. Chan, J. Hao, and T.-W. Lee, "Emotion recognition by speech signals," in *Proceedings of International Conference EUROSPEECH*, 2003.

[20] P. Boersma and D. Weenink, "PRAAT, a system for doing phonetics by computer," *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2001.

[21] F. Eyben, M. Wöllmer, and B. Schuller, "Opensmile: the munich versatile and fast open-source audio feature extractor," in *Proceedings of the international conference on Multimedia*, 2010, pp. 1459–1462.

[22] B. Schuller, B. Vlasenko, R. Minguez, G. Rigoll, and A. Wendemuth, "Comparing one and two-stage acoustic modeling in the recognition of emotion in speech," in *Automatic Speech Recognition & Understanding, 2007. ASRU. IEEE Workshop on*, 2007, pp. 596–600.

[23] C. M. Lee and S. S. Narayanan, "Toward detecting emotions in spoken dialogs," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 2, pp. 293–303, 2005.

[24] J. Jones and A. Christensen, "Couples interaction study: Social support interaction rating system," *University of California, Los Angeles*, 1998.

[25] C. Heavey, D. Gill, and A. Christensen, "Couples interaction rating system 2 (CIRS2)," *University of California, Los Angeles*, 2002.

[26] D. D. Lewis, "Feature selection and feature extraction for text categorization," in *Proceedings of the workshop on Speech and Natural Language*. Association for Computational Linguistics, 1992, pp. 212–217.

[27] G. Brown, A. Pocock, M.-J. Zhao, and M. Luján, "Conditional likelihood maximisation: a unifying framework for information theoretic feature selection," *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 27–66, 2012.

[28] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *Neural Networks, IEEE Transactions on*, vol. 13, no. 2, pp. 415–425, 2002.

[29] A. Christensen, D. C. Atkins, S. Berns, J. Wheeler, D. H. Baucom, and L. E. Simpson, "Traditional versus integrative behavioral couple therapy for significantly and chronically distressed married couples." *Journal of consulting and clinical psychology*, vol. 72, no. 2, p. 176, 2004.

[30] A. Özgür, L. Özgür, and T. Güngör, "Text categorization with class-based and corpus-based keyword selection," in *Computer and Information Sciences-ISCIS 2005*. Springer, 2005, pp. 606–615.