Received: 3 June 2020 | Revised: 1 October 2020 | Accepted: 12 November 2020

DOI: 10.1002/cpp.2534

RESEARCH ARTICLE

WILEY

What do these sounds tell us about the therapeutic alliance: Acoustic markers as predictors of alliance

Aviv Nof[1] | Ofer Amir[2] | Pavel Goldstein[3] | Sigal Zilcha-Mano[1]

[1]Department of Psychology, University of Haifa, Haifa, Israel

[2]Department of Communication Disorders, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[3]School of Public Health, University of Haifa, Haifa, Israel

Correspondence
Sigal Zilcha-Mano, Department of Psychology, University of Haifa, Mount Carmel, Haifa 31905, Israel.
Email: sigalzil@gmail.com

Funding information
Society for Psychotherapy Research; Open University of Israel Research Fund; Israel Science Foundation (ISF)

Predicting the trajectories of alliance formation that the patient is likely to establish with the therapist during treatment, even before their first meeting, can help prevent the potentially harmful consequences of deterioration in alliance, such as poor outcome and premature dropout. The present study aimed to examine the ability of four pretreatment acoustic markers to predict the alliance that is likely to be formed in the course of treatment: F0 span, speech rate, pause proportion and jitter. Data from 560 observations of 38 patients were collected as part of an ongoing randomized clinical trial of short-term psychotherapy for major depressive disorder. The acoustic markers were measured using high-quality recordings at baseline, before the patient and therapist ever met or had any type of communication. A multilevel model was used to examine the ability of the four acoustic markers to predict the slopes of alliance formation in the course of treatment, all markers being introduced in the same model. The clinical utility of the acoustic markers was explored in two case studies. The model explained 22% of the variance in alliance formation. Higher levels of both jitter and pause proportion at baseline predicted less strengthening of the alliance in the course of treatment. The findings, which should be replicated in larger samples, suggest that much of the therapeutic alliance can be predicted based on the acoustic characteristics of the patient's voice in the first 3 min of their intake, before they even meet their therapist.

KEYWORDS
acoustic markers, jitter, pause proportion, psychotherapy processes, therapeutic alliance

1 | INTRODUCTION

Psychotherapy is effective for diverse mental health problems, but some patients benefit more than others from it, with approximately one third of patients not achieving significant improvement at all (Lambert, 2013). Many attempts have been made to identify the individuals for whom psychotherapy is more effective (Bohart & Wade, 2013). It has been suggested that mechanisms that are common across psychotherapies can be instrumental (Hofmann & Hayes, 2019) in this endeavour. One such mechanism is the therapeutic alliance between patient and therapist, which was found repeatedly to be a consistent predictor of outcome (Flückiger, Del Re, Wampold, & Horvath, 2018).

A large number of empirical studies conducted in the past four decades suggest that patients differ in their ability to form strong alliances (Muran & Barber, 2011), because of particular characteristics that might facilitate or interfere with the establishment and maintenance of the alliance (Constantino, Castonguay, Zack, & DeGeorge, 2010). Patients also differ in their trajectories of alliance development throughout treatment (Stiles & Goldsmith, 2010).

Traditionally, the alliance has been assessed by a single snapshot measurement (e.g. Horvath & Symonds, 1991). Recently, the use of advanced statistical tools (e.g. hierarchical linear modelling) paved the way to investigate the alliance development over multiple time points (e.g. Flückiger et al., 2020). This longitudinal approach reflects the dynamic nature of the alliance that develops and fluctuates during treatment (Kramer, de Roten, Beretta, Michel, & Despland, 2009; Zilcha-Mano & Errázuriz, 2017). Using this longitudinal approach to assess the alliance, several studies identified distinct trajectories of alliance development throughout the treatment, with some patients showing more strengthening of the alliance than others (e.g. Zilcha-Mano & Errázuriz, 2017).

The findings of the traditional snapshot assessment and the longitudinal approaches show some similarity in their findings in the ability of the alliance to predict treatment outcome. For example, Stiles and Goldsmith (2010) suggested that the alliance improvement across therapy is associated with increases in the outcome.

Predicting the alliance strength an individual will form with the therapist and its development during treatment, even before the patient and the therapist have met, could potentially help therapists in trying to reduce the negative effect of poor alliance on the outcome. Implementing effective alliance techniques (e.g. Safran & Muran, 2000) can be beneficial in this endeavour. It was suggested that patients with a weaker therapeutic alliance, measured early in treatment by a snapshot evaluation, or at several time points during treatment, are more likely to drop out of psychotherapy (for a meta-analysis, see Sharf, Primavera, & Diener, 2010). Also, successful resolution of alliance problems (ruptures) is correlated with greater retention in treatment, as measured early in treatment over the first six sessions (Muran et al., 2009).

Previous empirical studies that focused on pretreatment patient tendencies to form strong versus weak alliances have investigated pretreatment patient characteristics. An overly friendly submissive nature measured in snapshot evaluation has been positively associated with the alliance, whereas hostile-dominant problems were negatively associated with the alliance (Muran, Segal, Samstag, & Crawford, 1994).

Zuroff et al. (2000) measured the alliance over three sessions and found that increases in the alliance were large among patients with low levels of perfectionism and smaller or absent among those high in perfectionism. The attachment orientation and alliance have also been studied in a meta-analysis (Diener & Monroe, 2011) that included both snapshot and longitudinal assessments. Findings show that greater attachment security was associated with stronger therapeutic alliances (Diener & Monroe, 2011). Good interpersonal relationships (Hersoug, Høglend, Havik, von der Lippe, & Monsen, 2009) and alliance expectations (Barber et al., 2014) have also been suggested as associated with alliance growth over time.

However, other studies have yielded null or mixed results. Pretreatment symptom level was not a valid predictor of the alliance growth as measured in several time points (e.g. Gibbons et al., 2003). Patients' adaptive defences were not associated with the alliance as measured in several time points (Hersoug, Sexton, & Høglend, 2002). Constantino, Castonguay, Zack, and DeGeorge (2010) reported that studies focusing on patients' demographic characteristics, such as age, gender and education level, have not consistently predicted alliance quality. Therefore, evaluation of facilitating and blocking factors in the formation of the therapeutic alliance remains an important objective.

Given the beneficial effect of a strong alliance on the outcome (e.g. Zilcha-Mano, Dinger, McCarthy, & Barber, 2014), clinicians are encouraged to establish strong alliances in different treatment modalities (Castonguay, Constantino, McAleavey, & Goldfried, 2010). However, therapists differ in their abilities to form favourable alliances and in the associated outcome (Baldwin, Wampold, & Imel, 2007). Also, due to potential difficulties in identifying drops in the alliance, a therapist's evaluation of the alliance is not always accurate, especially concerning withdrawal ruptures (Eubanks, Muran, & Safran, 2018).

Key Practitioner Message

- This study is one of the first, if not the first, to show the utility of acoustic markers as potential indicators of the progress of the alliance.
- Much of the therapeutic alliance can be predicted based on the acoustic characteristics of the patient's voice in the first 3 min of their intake, before they even meet their therapist.
- If replicated in future studies, these findings may guide clinicians to use acoustic markers for planning and choosing their initial empirically based interventions aiming to prevent alliance deterioration.

Not all therapeutic encounters lead to favourable alliance formation (e.g. Ackerman & Hilsenroth, 2001), and some could end in premature dropout (Sharf, Primavera, & Diener, 2010). Therefore, there is a growing need for supplementary tools that would predict the alliance trajectories. As such, the current study aims to present the use of another source of information (the acoustic markers) about the alliance trajectory.

In recent years, there has been a growing interest in applying novel technologies that will assist in evaluating psychotherapy processes, such as the alliance (Imel, Caperton, Tanana, & Atkins, 2017). Several studies have shown the benefits of applying automated objective markers of psychotherapy phenomena. For example, it has been shown that patient and therapist body movement synchrony are indicative of relationship quality (Ramseyer & Tschacher, 2014), that vocally encoded arousal synchrony is indicative of therapist empathy (Imel et al., 2014) and that skin conductance is indicative of perceived therapist empathy (Marci, Ham, Moran, & Orr, 2007).

One promising group of automated and computer-based objective markers that can capture the therapeutic alliance are the vocal acoustic markers obtained from the patient's speech. Acoustic markers quantify the physical properties of the sound and the temporal characteristics of the speech (Rochman & Amir, 2013), and they were found to indicate the emotional state of the speakers and their social intent towards others (Diamond, Rochman, & Amir, 2010; Juslin & Laukka, 2003).

Based on the rich emotional and social information that acoustic markers can convey (e.g. Juslin & Laukka, 2003), it has been suggested that, ‘Just as eyes are often considered a gateway to the soul, the human voice offers a window through which we gain access to our fellow human beings' minds—their attitudes, intentions and feelings’ (Brück, Kreifelts, & Wildgruber, 2011, p. 383). There is broad consensus that most vocal parameters are powerful indicators of arousal (Diamond et al., 2010; Giddens, Barron, Byrd-Craven, Clark, & Winter, 2013; Scherer, Johnstone, & Klasmeyer, 2003). High sympathetic arousal is associated with increased muscle tension in the vocal folds and respiratory tract, whereby that increased tension drives up the corresponding acoustical vocal expression (Rochman & Amir, 2013; Scherer, Johnstone, & Klasmeyer, 2003).

Furthermore, it has been suggested that a good therapeutic alliance is associated with the emergence of emotional arousal and deep emotion processing (Greenberg & Pascual-Leone, 2006). It is therefore possible that acoustical markers, which have been suggested as indicators of arousal, can be suggested also as predictors of alliance. In our study, some acoustic markers that were previously associated with arousal (e.g. Scherer, Johnstone, & Klasmeyer, 2003) are investigated as predictors of the alliance.

The present study investigated whether pretreatment acoustic markers obtained from patients' recordings can predict who would form strong alliances during treatment. We used data from an ongoing psychotherapy trial for depression, a leading cause of disability worldwide (World Health Organization, 2018). The four acoustic markers in the current study were chosen to reflect the main aspects of prosody, the melody of speech. Prosody commonly includes aspects of dynamic changes in timing, pitch, loudness, duration and voice quality while speaking (Barth-Weingarten, Reber, & Selting, 2010).

Based on these aspects and consistent with previous studies, we focused on four well-established acoustic markers identified as promising in previous psychotherapy research (Levitt, 2001; Mundt, Snyder, Cannizzaro, Chappie, & Geralts, 2007; Rochman & Amir, 2013): (i) fundamental frequency span, (ii) speech rate, (iii) pause proportion and (iv) jitter. We excluded loudness measures (intensity and shimmer) from the study because of the potential inconvenience of wearing a headset in order to measure them.

The fundamental frequency (F0) represents the vibration rate of the vocal folds during phonation, and it is subjectively perceived as the speaker's pitch (Rochman & Amir, 2013). The F0 span captures the extent to which a prosody pattern is restricted, perceived as monotonic speech, or widely varied across the frequency spectrum, perceived as lively and playful speech (Knowles & Little, 2016). Previous studies have suggested that F0 span markers are associated with the arousal (Juslin & Laukka, 2003) and tend to decrease during depressed states (Cummins, Sethu, Epps, Schnieder, & Krajewski, 2015). Relating to our hypothesis, the F0 span has been shown to convey empathy (Couper-Kuhlen, 2012) and to be associated with either agreement or disagreement stances during personal interaction (Ogden, 2006). In psychotherapy research, the F0 span has been shown to be associated with an empathic therapeutic stance of validating and tuning to the patient's emotions (Weiste & Peräkylä, 2014).

Pauses are silent speech intervals during speech. Previous studies have shown that pauses are valid markers of the treatment alliance in psychotherapy (Daniel, Folke, Lunn, Gondan, & Poulsen, 2018; Levitt, 2001). Pauses were found to be associated with reluctance to open up to therapists, which has been suggested to hinder therapeutic engagement (Daniel, 2011).

Speech rate is defined as a prosodic feature that captures the number of syllables uttered during a given period (Amir, 2016). Previous studies have associated speech rate with emotional arousal (Scherer, Johnstone, & Klasmeyer, 2003). Concerning the alliance hypothesis, speech rate has been suggested as a valid maker of empathy, emotionally 'in-tune' attitude (Lawrence et al., 2007) and a perceived clinical stance of caring and sympathy (McHenry, Parker, Baile, & Lenzi, 2012).

Jitter is an F0 perturbation (instability) deriving from small cycle-to-cycle variations during voicing, which might be perceived as hoarseness (Rochman & Amir, 2013). Jitter is considered an important marker of emotional distress states (Cummins, Sethu, Epps, Schnieder, & Krajewski, 2015), but little is known about its relation to the therapeutic alliance.

2 | METHOD

2.1 | Participants

2.1.1 | Patients

Thirty-nine adults (producing 560 observations) participated in the training phase and the main trial phase of an ongoing randomized controlled trial (RCT). The study compared two forms of time-limited (16 sessions) psychodynamic therapy for major depression disorder (MDD): supportive-expressive (SE) and supportive therapy. Assignment to treatment arm was conducted by an outside institution, based on a minimization algorithm. The purpose of the minimization algorithm is to minimize differences between treatment conditions in patients' baseline variables (such as symptom severity) by balancing these variables between treatment conditions. Following the general requirement in psychotherapy research not to break the blindness to conditions before the main outcome paper is published, in this study, similarly to other studies in the literature, the two conditions are analysed together.

All patients provided written and oral informed consent. Patients were self-referred, in response to advertisements offering free treatment, in the central region of Israel. The candidates attended several assessment meetings with the research team to verify the MDD diagnosis and that the patient meets inclusion/exclusion criteria (Zilcha-Mano, Dolev, Leibovich, & Barber, 2018). The data obtained from one patient could not be acoustically analysed because of poor recording conditions. The remaining 38 patients (producing 560 observations) consisted of 23 women and 15 men, with a mean age of 32.68 (SD = 10.38); 95% were native Hebrew speakers, and the remaining 5% were fluent Hebrew speakers, with either Russian or French as their first language.

2.1.2 | Therapists and treatments

Six therapists, with an average of 13.8 years of clinical experience, treated six patients each, on average. The therapists underwent 20 h of training in SE techniques. After achieving a sufficient adherence level in two pilot treatments, they moved into the trial phase. They received weekly personal and group supervision, provided by two experienced licensed clinical psychologists, who themselves received supervision from an international SE expert. Both types of treatments (SE and supportive) are based on the same manualized protocol (Luborsky, 1995), but expressive techniques are forbidden in the supportive treatment.

2.2 | Measures

2.2.1 | Acoustic markers

The pretreatment acoustic markers were measured at the pretreatment evaluation phase, at the beginning of the first intake session, conducted by a clinical diagnostician who met the patients before treatment with the psychologist began. The recordings were performed to meet the high standards of audio recording recommended in psychotherapy (Rochman & Amir, 2013), with a 44.1-kHz sampling rate, 16-bit and output as wav files with a Zoom H5 digital audio recorder. The microphones were of unidirectional condenser type, set at a 90° angle towards the speaker.

The target for the analysis was the first 10 utterances of the patients. An utterance was defined as a sequence of at least three words, conveying an idea with grammatically acceptable structure (Amir, 2016). All acoustic markers were extracted and analysed in a three-step procedure. First, the 10 utterances were transcribed according to the psychotherapy transcription protocol (Mergenthaler & Stinson, 1992). Next, 10 separate audio files were trimmed and normalized using the Audacity software, open-source audio editing software, version 2.0.2 (Audacity Team, 2018). Finally, each file (i.e. utterance) was acoustically analysed using the Praat software, version 6.0.24 (Boersma & Weenink, 2009).

Vocal fundamental frequency span (F0 span)

We extracted the F0 span by measuring F0 standard deviation in a three-step analysis, using the Praat software, version 6.0.24 (Boersma & Weenink, 2009). The first step was to create a pitch object, using the system settings. The second step consisted of human monitoring and fixing the errors of Praat F0 identification (e.g. octave errors) by adjusting the lower and upper threshold values for F0 in each segment, until all outliers were resolved and deleted (Rochman & Amir, 2013). The third step consisted of measuring the F0 standard deviation and transforming it into a logarithmic scale that represents the listener's perception of intonational span (Ogden, 2006) and is a preferred procedure for preventing normality violation biases (e.g. Awan & Roy, 2005).

Pause proportion marker

The pauses were extracted using a three-step procedure. First, we applied a noise reduction filter. Second, the pauses were identified by the Praat software, using the setting of 0.25 s as the minimum silent interval and −25 dB as the silence threshold. The 0.25-s interval was set according to the standard threshold of pauses based by Goldman-Eisler (1968) and used in speech rate or silences analysis studies (e.g. Amir & Grinfeld, 2011; Stanislawski, Bilgrami, Sarac, Cecchi, & Corcoran, 2019). Third, the generated silence object was checked by a trained research assistant to correct the errors (e.g. false pause identifications). Fourth, we calculated the pause proportion (percentage) by summing the total pause durations and dividing it by the total duration of the sentence.

Speech rate marker

We calculated speech rate using a human decoding procedure: we began by transcribing the recording according to the standard psychotherapy transcription protocol (Mergenthaler & Stinson, 1992); next, we counted the syllables; and finally, we calculated speech rate by dividing the number of syllables by the total duration of the utterance.

Jitter marker

We targeted syllables containing the vowel /a/ (pronounced as /Ahhhhh/) in the fluent speech of patients. We looked for /a/ vowels whose durations were typically longer than 30 ms, trimmed them out of their phonation context, saved them as separate files and measured the jitter value. Given that continuous speech contains a combination of vowels, consonants and voiced and unvoiced segments and is also greatly affected by intonation, jitter measurements that were obtained from continuous speech are considered inaccurate. Therefore, jitter is typically extracted from voice samples of steady phonations (Rochman & Amir, 2013). The jitter was calculated by the Praat as the average absolute difference between consecutive periods (i.e. jitter local and absolute), using this derivative of the measure as previous studies did (e.g. Gregory, Chandran, Lurie, & Sataloff, 2012).

2.2.2 | The working alliance inventories

The Working Alliance Inventory (WAI; Horvath & Greenberg, 1989) is the 12-item version (Tracey & Kokotovic, 1989). Each item was rated on a 7-point Likert scale, ranging from 1 (*never*) to 7 (*always*). The alliance ratings were collected after session, after each of the 16 sessions of therapy. The total number of observations was 560.

2.3 | Data analysis

We used a Markov chain Monte Carlo sampler for linear mixed models, assuming censored Gaussian distribution for the model outcome using R package MCMCglmm (Hadfield, 2010). The data were hierarchically nested on three levels: assessments nested within patients, nested within therapists (560 observations). To account for the resulting non-independence of assessments and to prevent inflation of the effects, we added the patient and therapist as random effects. To measure the amount of variance in the WAI due to the random effects of the therapist and patient, we calculated intra-class correlations (ICCs). A model of fixed effect of log of time, random intercept and random slope in log of time was used to predict WAI development over time.

To investigate whether the acoustic markers can predict the slopes of alliance formation, we examined the interaction between all the acoustic markers and time (sessions) in a single model, as predictors of WAI individual trajectories from Sessions 1 to 16 (end of treatment). Following Brockmann, Drinnan, Storck, and Carding (2011), we controlled for potential gender effect, so that the predictors included in the model were the interaction effects, the main effects and gender.

3 | RESULTS

The rate of missing data in our sample was low (2.5%). Based on the guidelines for data screening (Tabachnick & Fidell, 2013), we found outliers, defined as at least three standard deviations above or below the mean, in one patient's jitter scores. We excluded these values from the analyses, and similarly to previous studies, we did not impute missing values (e.g. Shalom et al., 2018).

The estimated variance of the therapist's random intercept effect for the alliance formation slopes was null, whereas the variance of the patient's random intercept effect was significant (S^2 = 0.93, 95% confidence intervals [CIs] [0.41, 1.55], ICC = 0.8), and the patient's random slope was significant (S^2 = 0.09, 95% CIs [0.03, 0.15], ICC = 0.08). The model for predicting alliance formation resulted in two significant predictors (Table 1 and Figures 1 and 2): both higher levels of jitter and pause proportion predicted less strengthening of the alliance. The model explained 22% of the variance (Nakagawa & Schielzeth, 2013). The analysis suggests that gender had no effect on the model (β = −0.19, 95% CIs [−1.08, 0.72], P = 0.65).

3.1 | Sensitivity analysis

In a sensitivity analysis, we tested the unique contribution of the acoustic markers in predicting the alliance above and beyond the following patient baseline characteristics: (i) symptom severity (measured by Hamilton Rating Scale for Depression [HRSD]; Hamilton, 1967), (ii) co-morbidity with anxiety (measured by MINI International

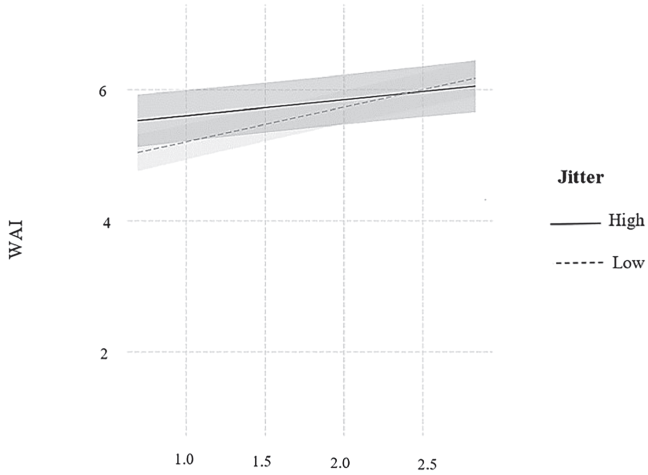

FIGURE 1 Predicting WAI slopes for high and low values of jitter. N = 38. Number of observations = 560. WAI, Working Alliance Inventory

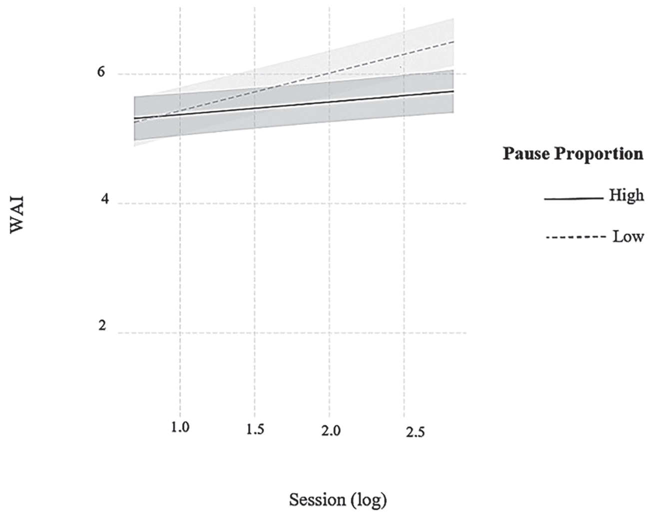

FIGURE 2 Predicting WAI slopes for high and low values of pause proportion. N = 38. Number of observations = 560. WAI, Working Alliance Inventory

TABLE 1 Acoustic markers as predictors of the slopes of alliance formation from Weeks 1 to 16

Acoustic markers (interaction with time)	Estimate	95% confidence interval	PMCMC
Jitter	−0.003	[−0.006, −0.0008]	0.014[a]
Pause	−0.02	[−0.04, −0.001]	0.022[a]
Speech rate	−0.15	[−0.36, 0.056]	0.174
F0-Std	−0.14	[−0.65, 0.28]	0.532

Note: N = 38. Number of observations = 560.
Abbreviation: MCMC, Markov chain Monte Carlo.
[a]PMCMC = 0.65 < 0.05. All acoustic measures were introduced in the same model. The model also included the main effects and gender (not shown in the table because of shortage of space). The estimates presented refer to those of the interaction of each acoustic marker with time in predicting the weekly assessments of alliance, from Weeks 1 to 16.

Neuropsychiatric Interview; Sheehan et al., 1998), (iii) co-morbidity with personality disorders (measured by the Structured Interview for DSM-IV Personality Disorders—Hebrew version [SIDP-IV]; Pfohl, Blum, & Zimmerman, 1997, 1 = PD on any type, 0 = none) and (iv) attachment orientation (measured by Experiences in Close Relationships [ECR] scale; Brennan, Clark, & Shaver, 1998).

The results indicated that the acoustic marker effects remained stable above and beyond the patient's characteristic variables. The findings appear in the Supporting Information. We also examined whether these results can be explained by systematic differences between the intake interviewers who conducted the pretreatment assessments. The results indicated that the estimated variance of the intake interviewers' random intercept effect for the alliance formation was null ($S^2 = 0$, 95% CIs [0, 0], ICC = 0). The model for predicting alliance formation remained significant as both higher levels of jitter and pause proportion predicted less strengthening of the alliance. The findings appear in the Supporting Information.

3.2 | Illustrative case studies

3.2.1 | The case of Aaron: Small improvement in the alliance

Aaron (identifying details have been disguised), an engineer in his early 40s, is married and has three children. At the pretreatment evaluation session, he met a clinical diagnostician (other than his therapist he would meet 3 weeks later) and reported feelings of sadness, anhedonia, loneliness and hopelessness. The HRSD (Hamilton, 1967) scores indicated major depression.

Aaron's pretreatment acoustic markers were obtained before he met the therapist, from his first 10 sentences answering the question: ‘Why do you seek therapy?’. Aaron's acoustic marker's profile was characterized, as seen in Figure 3, by relatively prolonged and frequent pauses ($Z = 0.63$) and high jitter values ($Z = 0.6$), relative to Mary's acoustic profile (the second case presented). The current study findings might imply that patients, such as Aaron, who present pretreatment high levels of jitter and pauses may progress slowly in forming an alliance with their therapist. Moreover, previous studies have shown an association between high pause levels and a tendency to disengage from the therapy, a dynamic that can undermine alliance formation (Daniel, Folke, Lunn, Gondan, & Poulsen, 2018; Levitt, 2001).

During the therapy, the alliance formed with the therapist progressed slowly, as seen in Figure 4, and followed a pattern that was anticipated, based on Aaron's acoustic signature measured before the treatment started. From a clinical perspective, the alliance was influenced by Aaron's core conflict relational pattern, as presented to him by the therapist when the therapy started: his wish to become emotionally involved with others and be loved could not be fulfilled because of his conflictual anticipation that others will be emotionally detached from him and disrespect him. As a result of this conflict, he tended to react in an avoidant stressful manner, as the acoustic marker reflected this tendency.

This core conflict was enacted within the therapeutic relationship and checked the strengthening of the alliance. For example, in the third session, following the therapist's genuine concern about Aaron's coughing, he reacted by saying: ‘You're just doing your job asking me how I feel. I'm not sure you actually do care about me’. Subsequently, there were several exchanges that reflected his avoidant tendency for

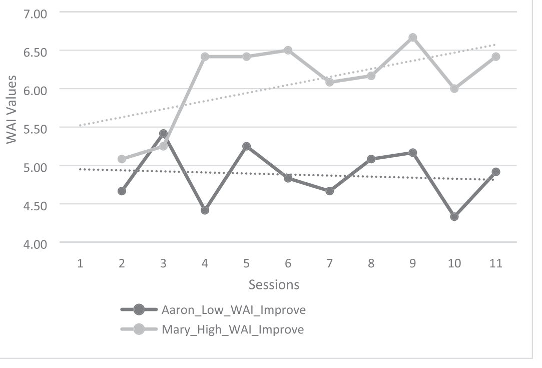

FIGURE 4 WAI scores of Aaron (low improvement) and of Mary (high improvement). WAI, Working Alliance Inventory (Horvath & Greenberg, 1989; Tracey & Kokotovic, 1989)

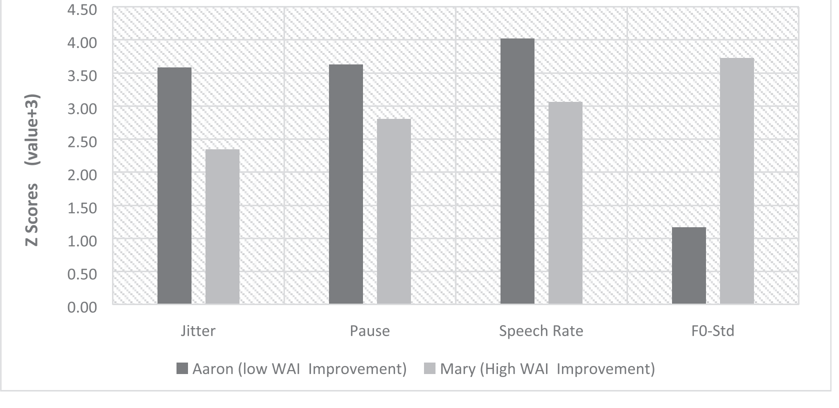

FIGURE 3 Pretreatment acoustic markers of Aaron and Mary. The raw acoustic marker values were converted into Z values (relative to the entire sample). To facilitate the presentation, the Y-axis values were linearly transformed from the original Z scores to $Z + 3$ (e.g. original $z = -2.5$ was converted to 0.5). WAI, Working Alliance Inventory

minimal responses and long silences (e.g. 'I had a bad fight with my wife last night and hardly slept, but I don't want to share it with you. I don't know if it would help me … (long pause)'). The therapist's countertransference, which is another source of information about the alliance (Newhill, Safran, & Muran, 2003), was characterized by a feeling of caring, on one hand, and of frustration and stress, on the other. This case demonstrates that, based on pretreatment jitter and pause values, we were able to anticipate the alliance that was formed in practice with the therapist in the course of the therapy that began 3 weeks later.

3.2.2 | The case of Mary: Strong improvement in the alliance

Mary (identifying details have been disguised), an attorney in her mid-50s, is married and has two children. At the pretreatment evaluation session, she met a clinical diagnostician (other than the therapist she would meet later) and reported the feelings of sadness, fatigue and lack of motivation to meet people. The HRSD (Hamilton, 1967) scores indicated major depression. Mary's pretreatment acoustic marker profile, as seen in Figure 3, was characterized by relatively short and infrequent pauses ($Z = -0.2$) and with relatively low jitter values ($Z = -0.7$).

The current study findings might imply that patients, such as Mary, who present short and infrequent pauses and low jitter values may progress rapidly in forming an alliance with their therapist. Previous studies have shown an association between infrequent pauses and verbal fluency (Martins, Vieira, Loureiro, & Santos, 2007) and a tendency to engage and bond easily (Daniel, Folke, Lunn, Gondan, & Poulsen, 2018; Levitt, 2001). Also, Mary's low pretreatment jitter values reflect a steady, non-trembling voice quality, which previous studies have associated with low distress and low anxiety levels (Juslin & Scherer, 2005).

During the therapy sessions, Mary indeed formed a solid and favourable alliance with the therapist, which progressively strengthened, as seen in Figure 4, consistent with what we anticipated based on pretreatment jitter and pause values. Mary was self-regulated, easygoing and considerate towards the therapist in a way that helped establish a good therapeutic alliance. For example, at Session 5, the therapist apologized for being late because of a traffic jam. Mary reassured the therapist: 'It's OK, I fully get it, as one who's stuck in traffic jams almost every day, so you shouldn't feel bad about it. Besides, in the meantime I managed to answer all my emails (laughing together).' The therapist, who felt comfortable with Mary, thanked her for being so empathic, and they were able to continue exploring other meanings of her agreeable attitude.

The strong alliance further helped Mary to gain insight about the meaning of her therapeutic relationship and to step out of the extreme caring stance she rigidly adopted towards others. With these insights, Mary could start to effectively assert her emotional needs and interpersonal wish to be cared for, rather than concealing them under the appeasing stance.

4 | DISCUSSION

The present findings suggest that pretreatment acoustic markers obtained before the patient and therapist have met serve as significant predictors of the alliance formed during the course of treatment. Jitter was found to be a significant predictor of alliance formation, with higher levels being associated with less strengthening of the alliance over the 16 sessions of treatment. This finding is consistent with previous reports demonstrating an association between higher levels of jitter and higher levels of distress during patient–physician interaction (Postma-Nilsenová, Holt, Heyn, Groeneveld, & Finset, 2016). From a clinical perspective, high in-session jitter levels may indicate unresolved anger (Diamond et al., 2010), which in turn hinders and burdens alliance formation.

Another acoustic marker that predicted the unique variance of the alliance was the pause proportion, with higher values of pause proportion being associated with less strengthening of the alliance. This finding is consistent with previous ones suggesting a significant association between high rates of pausing and poorer alliance, measured during treatment using qualitative (Levitt, 2001) and quantitative approaches (Daniel, Folke, Lunn, Gondan, & Poulsen, 2018). From a clinical perspective, in-session silences may indicate a disengagement dynamic, with possible negative consequences for the alliance (Levitt, 2001). We brought two case studies to show that the acoustic markers of jitter and pause can provide valuable clinical indications regarding the future trajectory of the patients' alliance with the therapist.

Speech rate and the F0 span markers did not explain the unique variance in alliance formation. The null finding about speech rate is consistent with the mixed findings regarding this measure in the literature (as mentioned in Guyer, Fabrigar, & Vaughan-Johnston, 2019) and with some previous findings associating speech rate with the linguistic rather than the emotional characteristics of the speaker (e.g. verbal fluency; Martins, Vieira, Loureiro, & Santos, 2007).

The null finding about the F0 span, however, is not consistent with previous findings about the association between F0 span (variance) and empathic communication trajectory (Weiste & Peräkylä, 2014). This inconsistency may be explained by the small sample size, which limited our ability to detect smaller effects, or to detect heterogeneity in the effect that may result from potential moderators (e.g. symptom severity; Cummins, Sethu, Epps, Schnieder, & Krajewski, 2015).

The sensitivity analysis suggests that the pretreatment acoustic markers make a unique contribution to the prediction of alliance formation above and beyond the patient's baseline characteristics of symptom severity, attachment orientation and co-morbidity. These findings suggest that acoustic markers make a unique contribution beyond the data commonly obtained in a routine intake assessment. Also, a post hoc analysis indicated that the findings cannot be attributed to the differences between intake interviewers.

In our study, some acoustic markers that were previously associated with arousal (e.g. Scherer, Johnstone, & Klasmeyer, 2003) were found as predictors of alliance. An indirect association may therefore exist between acoustic markers, alliance formation and arousal. Future studies can explore this possible association. The evidence from this study suggests that acoustic markers are signals that carry information about the preliminary tendency of the patient to form an alliance with the therapist. From an evolutionary perspective, vocalization might have provided an important communication medium for our prelinguistic ancestors (Thompson & Balkwill, 2006) and as an effective compensation for the low-visibility environments they lived in (Fedurek & Slocombe, 2011).

Behaviours that facilitated and maintained bonds between group members, such as acoustic signalling about collaboration, would therefore be crucial to survival (Bourgeois & Hess, 2008). It is reasonable to assume that selective evolutionary pressures might have promoted vocalization as one of the preferred communication strategies that signal alliance formation intentions. As a result, it might have eventually become widespread throughout the population due to its survival advantage of enhancing social grouping.

Although acoustic analysis (vocal analysis) is mentioned as one of the technologies that can move psychotherapy research into the technology-inspired revolution (Imel, Caperton, Tanana, & Atkins, 2017), it should be noted that not all acoustic procedures are yet fully automated. In the current study, some procedures were more automatic (e.g. pause detection), whereas others required intensive human investment (e.g. jitter analysis). The automatization of acoustic analysis combined with machine learning procedures is a desired goal, which has already been demonstrated in several fields of science. For example, acoustic markers were instrumental in identifying cardiac arrests (Chan, Rea, Gollakota, & Sunshine, 2019).

An important limitation of the present study is the small sample size, the result of the fact that the study is part of an ongoing RCT. The small sample size limited our ability to include more relevant acoustic markers in our prediction model. Moreover, because this is an ongoing clinical trial, we could not further explore the potential association between the acoustic markers and the outcome measures. Finally, we chose one recommended way of measuring and calculating the F0 marker, whereas other studies used different methodologies (e.g. Imel et al., 2014). A larger sample would be required to check the relative merits of the different methodologies.

The current findings are among the first, if not the first, to suggest that patients' pretreatment acoustic markers, obtained before patients and therapists met or communicated, can predict the alliance that would be formed between them. If replicated in future studies, these findings may have important clinical implications because of the consistent association between alliance and therapy outcome (Flückiger, Del Re, Wampold, & Horvath, 2018), enabling clinicians to use acoustic markers for planning and choosing their initial empirically based interventions in a way that prevents alliance deterioration (e.g. Barlow et al., 2017).

This approach can additionally be applied to online psychotherapy platforms, which became the new reality of mental health delivery during the Covid-19 pandemic of 2020, while writing this manuscript. Given the diminished visual input at online video encounters (e.g. the invisible body posture), the auditory cues come to the forefront as a prominent source of information about the alliance. Future studies will explore the utility of real-time feedback system to analyse the patient's acoustic markers and provide therapists with real-time feedback information about the patient's alliance tendency. Until such technology is achieved, clinicians can trust their own natural innate capacity to process the emotional prosody information (Wiethoff et al., 2008) and wonder 'What do these sounds tell us about the alliance'.

ACKNOWLEDGEMENTS

The writing of this paper was supported by the Israel Science Foundation (ISF), the Open University of Israel Research Fund and the Society for Psychotherapy Research.

CONFLICT OF INTEREST

The authors have no conflict of interest.

DATA AVAILABILITY STATEMENT

When this research was carried out, the informed consent form for the participants stated that we would keep the data strictly confidential. Therefore, if uploading data, we must seek consent from our participants and the consent of the ethics committee. Therefore, the data are not currently available.

ORCID

Aviv Nof https://orcid.org/0000-0003-1410-5586
Ofer Amir https://orcid.org/0000-0002-2873-954X
Sigal Zilcha-Mano https://orcid.org/0000-0002-5645-4429

REFERENCES

Ackerman, S. J., & Hilsenroth, M. J. (2001). A review of therapist characteristics and techniques negatively impacting the therapeutic alliance. *Psychotherapy: Theory, Research, Practice, Training*, *38*(2), 171–185. https://doi.org/10.1037/0033-3204.38.2.171

Amir, O. (2016). Speaking rate among adult Hebrew speakers: A preliminary observation. *Annals of Behavioural Science*, *02*, 1–9. https://doi.org/10.21767/2471-7975.100016

Amir, O., & Grinfeld, D. (2011). Articulation rate in childhood and adolescence: Hebrew speakers. *Language and Speech*, *54*(2), 225–240. https://doi.org/10.1177/0023830910397496

Audacity Team. (2018). Audacity®: free, open source, cross-platform audio software.

Awan, S. N., & Roy, N. (2005). Acoustic prediction of voice type in women with functional dysphonia. *Journal of Voice*, *19*(2), 268–282. https://doi.org/10.1016/j.jvoice.2004.03.005

Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology*, *75*(6), 842–852. https://doi.org/10.1037/0022-006X.75.6.842

Barber, J. P., Zilcha-Mano, S., Gallop, R., Barrett, M., McCarthy, K. S., & Dinger, U. (2014). The associations among improvement and alliance expectations, alliance during treatment, and treatment outcome for major depressive disorder. *Psychotherapy Research*, *24*(3), 257–268. https://doi.org/10.1080/10503307.2013.871080

Barlow, D. H., Farchione, T. J., Sauer-Zavala, S., Latin, H. M., Ellard, K. K., Bullis, J. R., … Cassiello-Robbins, C. (2017). *Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide*. New York: Oxford University Press.

Barth-Weingarten, D., Reber, E., & Selting, M. (Eds.) (2010). *Prosody in interaction* (Vol. 23). Amsterdam: John Benjamins Publishing.

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer [Computer program].

Bohart, A. C., & Wade, A. G. (2013). The client in psychotherapy. In *Bergin and Garfield's handbook of psychotherapy and behavior change* (Vol. 6) (pp. 219–257). Hoboken, New Jersey: John Wiley & Sons, Inc.

Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology*, *77*(3), 343–352. https://doi.org/10.1016/j.biopsycho.2007.11.008

Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult romantic attachment: An integrative overview. In R. W. S. Simpson (Ed.), *Attachment theory and close relationships* (pp. 46–76). New York: Guilford Press.

Brockmann, M., Drinnan, M. J., Storck, C., & Carding, P. N. (2011). Reliable jitter and shimmer measurements in voice clinics: The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *Journal of Voice*, *25*(1), 44–53. https://doi.org/10.1016/j.jvoice.2009.07.002

Brück, C., Kreifelts, B., & Wildgruber, D. (2011). Emotional voices in context: A neurobiological model of multimodal affective information processing. *Physics of Life Reviews*, *8*(4), 383–403. https://doi.org/10.1016/j.plrev.2011.10.002

Castonguay, L. G., Constantino, M. J., McAleavey, A. A., & Goldfried, M. R. (2010). The therapeutic alliance in cognitive-behavioral therapy. In *The therapeutic alliance: An evidence-based guide to practice* (pp. 150–171). New York, NY: Guilford Press.

Chan, J., Rea, T., Gollakota, S., & Sunshine, J. E. (2019). Contactless cardiac arrest detection using smart devices. *NPJ Digital Medicine*, *2*(1), 1–8.

Constantino, M. J., Castonguay, L. G., Zack, S. E., & DeGeorge, J. (2010). Engagement in psychotherapy: Factors contributing to the facilitation, demise, and restoration of the therapeutic alliance.

Couper-Kuhlen, E. (2012). Exploring affiliation in the reception of conversational complaint stories.

Cummins, N., Sethu, V., Epps, J., Schnieder, S., & Krajewski, J. (2015). Analysis of acoustic space variability in speech affected by depression. *Speech Communication*, *75*, 27–49. https://doi.org/10.1016/j.specom.2015.09.003

Daniel, S. I. (2011). Adult attachment insecurity and narrative processes in psychotherapy: An exploratory study. *Clinical Psychology*, *18*(6), 498–511.

Daniel, S. I., Folke, S., Lunn, S., Gondan, M., & Poulsen, S. (2018). Mind the gap: In-session silences are associated with client attachment insecurity, therapeutic alliance, and treatment outcome. *Psychotherapy Research*, *28*(2), 203–216. https://doi.org/10.1080/10503307.2016.1177673

Diamond, G. M., Rochman, D., & Amir, O. (2010). Arousing primary vulnerable emotions in the context of unresolved anger: "Speaking about" versus "speaking to". *Journal of Counseling Psychology*, *57*(4), 402–410. https://doi.org/10.1037/a0021115

Diener, M. J., & Monroe, J. M. (2011). The relationship between adult attachment style and therapeutic alliance in individual psychotherapy: A meta-analytic review. *Psychotherapy*, *48*(3), 237–248. https://doi.org/10.1037/a0022425

Eubanks, C. F., Muran, J. C., & Safran, J. D. (2018). Alliance rupture repair: A meta-analysis. *Psychotherapy*, *55*(4), 508–519. https://doi.org/10.1037/pst0000185

Fedurek, P., & Slocombe, K. E. (2011). Primate vocal communication: A useful tool for understanding human speech and language evolution? *Human Biology*, *83*(2), 153–173. https://doi.org/10.3378/027.083.0202

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, *1*(2), 23951.

Flückiger, C., Rubel, J., Del Re, A. C., Horvath, A. O., Wampold, B. E., Crits-Christoph, P., … Errázuriz, P. (2020). The reciprocal relationship between alliance and early treatment symptoms: A two-stage individual participant data meta-analysis. *Journal of Consulting and Clinical Psychology*, *88*(9), 829–843. https://doi.org/10.1037/ccp0000594

Gibbons, M. C., Crits-Christoph, P., de la Cruz, C., Barber, J. P., Siqueland, L., & Gladis, M. (2003). Pretreatment expectations, interpersonal functioning, and symptoms in the prediction of the therapeutic alliance across supportive-expressive psychotherapy and cognitive therapy. *Psychotherapy Research*, *13*(1), 59–76. https://doi.org/10.1093/ptr/kpg007

Giddens, C. L., Barron, K. W., Byrd-Craven, J., Clark, K. F., & Winter, A. S. (2013). Vocal indices of stress: A review. *Journal of Voice*, *27*(3), 390–e21.

Goldman-Eisler, F. (1968). Psycholinguistics: Experiments in spontaneous speech.

Greenberg, L. S., & Pascual-Leone, A. (2006). Emotion in psychotherapy: A practice-friendly research review. *Journal of Clinical Psychology*, *62*(5), 611–630. https://doi.org/10.1002/jclp.20252

Gregory, N. D., Chandran, S., Lurie, D., & Sataloff, R. T. (2012). Voice disorders in the elderly. *Journal of Voice*, *26*(2), 254–258. https://doi.org/10.1016/j.jvoice.2010.10.024

Guyer, J. J., Fabrigar, L. R., & Vaughan-Johnston, T. I. (2019). Speech rate, intonation, and pitch: Investigating the bias and cue effects of vocal confidence on persuasion. *Personality and Social Psychology Bulletin*, *45*(3), 389–405.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, *33*(2), 1–22.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social & Clinical Psychology*, *6*, 278–296. https://doi.org/10.1111/j.2044-8260.1967.tb00530.x

Hersoug, A. G., Høglend, P., Havik, O. E., von der Lippe, A., & Monsen, J. T. (2009). Pretreatment patient characteristics related to the level and development of working alliance in long-term psychotherapy. *Psychotherapy Research*, *19*(2), 172–180. https://doi.org/10.1080/10503300802657374

Hersoug, A. G., Sexton, H. C., & Høglend, P. (2002). Contribution of defensive functioning to the quality of working alliance and psychotherapy outcome. *American Journal of Psychotherapy*, *56*(4), 539–554. https://doi.org/10.1176/appi.psychotherapy.2002.56.4.539

Hofmann, S. G., & Hayes, S. C. (2019). The future of intervention science: Process-based therapy. *Clinical Psychological Science*, *7*(1), 37–50. https://doi.org/10.1177/2167702618772296

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, *36*, 223–233. https://doi.org/10.1037/0022-0167.36.2.223

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, *38*(2), 139–149. https://doi.org/10.1037/0022-0167.38.2.139

Imel, Z. E., Barco, J. S., Brown, H. J., Baucom, B. R., Baer, J. S., Kircher, J. C., & Atkins, D. C. (2014). The association of therapist empathy and synchrony in vocally encoded arousal. *Journal of Counseling Psychology*, *61*(1), 146–153. https://doi.org/10.1037/a0034943

Imel, Z. E., Caperton, D. D., Tanana, M., & Atkins, D. C. (2017). Technology-enhanced human interaction in psychotherapy. *Journal of Counseling Psychology*, *64*(4), 385–393. https://doi.org/10.1037/cou0000213

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Juslin, P. N., & Scherer, K. R. (2005). Vocal expression of affect. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 65–136). New York: Oxford University Press.

Knowles, K. K., & Little, A. C. (2016). Vocal fundamental and formant frequencies affect perceptions of speaker cooperativeness. *The Quarterly Journal of Experimental Psychology*, *69*(9), 1657–1675. https://doi.org/10.1080/17470218.2015.1091484

Kramer, U., de Roten, Y., Beretta, V., Michel, L., & Despland, J. N. (2009). Alliance patterns over the course of short-term dynamic psychotherapy: The shape of productive relationships. *Psychotherapy Research*, *19*(6), 699–706. https://doi.org/10.1080/10503300902956742

Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances.

Lawrence, E. J., Shaw, P., Baker, D., Patel, M., Sierra-Siegert, M., Medford, N., & David, A. S. (2007). Empathy and enduring depersonalization: The role of self-related processes. *Social Neuroscience*, *2*(3–4), 292–306. https://doi.org/10.1080/17470910701391794

Levitt, H. M. (2001). Sounds of silence in psychotherapy: The categorization of clients' pauses. *Psychotherapy Research*, *11*(3), 295–309. https://doi.org/10.1080/713663985

Luborsky, L. (1995). Supportive-expressive dynamic psychotherapy of depression: A time-limited version. In J. P. Barber, & P. Crits-Christoph (Eds.), *Dynamic therapies for psychiatric disorders* (pp. 41–83). New York: Basic Books.

Marci, C. D., Ham, J., Moran, E., & Orr, S. P. (2007). Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. *The Journal of Nervous and Mental Disease*, *195*(2), 103–111. https://doi.org/10.1097/01.nmd.0000253731.71025.fc

Martins, I. P., Vieira, R., Loureiro, C., & Santos, M. E. (2007). Speech rate and fluency in children and adolescents. *Child Neuropsychology*, *13*(4), 319–332. https://doi.org/10.1080/09297040600837370

McHenry, M., Parker, P. A., Baile, W. F., & Lenzi, R. (2012). Voice analysis during bad news discussion in oncology: Reduced pitch, decreased speaking rate, and nonverbal communication of empathy. *Supportive Care in Cancer*, *20*(5), 1073–1078. https://doi.org/10.1007/s00520-011-1187-8

Mergenthaler, E., & Stinson, C. (1992). Psychotherapy transcription standards. *Psychotherapy Research*, *2*(2), 125–142. https://doi.org/10.1080/10503309212331332904

Mundt, J. C., Snyder, P. J., Cannizzaro, M. S., Chappie, K., & Geralts, D. S. (2007). Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. *Journal of Neurolinguistics*, *20*(1), 50–64. https://doi.org/10.1016/j.jneuroling.2006.04.001

Muran, J. C., & Barber, J. P. (Eds.) (2011). *The therapeutic alliance: An evidence-based guide to practice*. New York, NY: Guilford Press.

Muran, J. C., Safran, J. D., Gorman, B. S., Samstag, L. W., Eubanks-Carter, C., & Winston, A. (2009). The relationship of early alliance ruptures and their resolution to process and outcome in three time-limited psychotherapies for personality disorders. *Psychotherapy: Theory, Research, Practice, Training*, *46*(2), 233–248. https://doi.org/10.1037/a0016085

Muran, J. C., Segal, Z. V., Samstag, L. W., & Crawford, C. E. (1994). Patient pretreatment interpersonal problems and therapeutic alliance in short-term cognitive therapy. *Journal of Consulting and Clinical Psychology*, *62*(1), 185–190. https://doi.org/10.1037/0022-006X.62.1.185

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Newhill, C. E., Safran, J. D., & Muran, J. C. (2003). *Negotiating the therapeutic alliance: A relational treatment guide*. New York, NY, US: Guilford Press.

Ogden, R. (2006). Phonetics and social action in agreements and disagreements. *Journal of Pragmatics*, *38*(10), 1752–1775. https://doi.org/10.1016/j.pragma.2005.04.011

Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured Interview for DSM-IV Personality: SIDP-IV*. Arlington, VA: American Psychiatric Pub.

Postma-Nilsenová, M., Holt, E., Heyn, L., Groeneveld, K., & Finset, A. (2016). A case study of vocal features associated with galvanic skin response to stressors in a clinical interaction. *Patient Education and Counseling*, *99*(8), 1349–1354. https://doi.org/10.1016/j.pec.2016.03.006

Ramseyer, F., & Tschacher, W. (2014). Nonverbal synchrony of head-and body-movement in psychotherapy: Different signals have different associations with outcome. *Frontiers in Psychology*, *5*, 979.

Rochman, D., & Amir, O. (2013). Examining in-session expressions of emotions with speech/vocal acoustic measures: An introductory guide. *Psychotherapy Research*, *23*(4), 381–393. https://doi.org/10.1080/10503307.2013.784421

Safran, J. D., & Muran, J. C. (2000). Negotiating the therapeutic alliance.

Scherer, K. R., Johnstone, T., & Klasmeyer, G. (2003). *Vocal expression of emotion*. New York, NY, US: Oxford University Press.

Shalom, J. G., Gilboa-Schechtman, E., Atzil-Slonim, D., Bar-Kalifa, E., Hasson-Ohayon, I., Van Oppen, P., … Aderka, I. M. (2018). Intra-individual variability in symptoms consistently predicts sudden gains: An examination of three independent datasets. *Journal of Consulting and Clinical Psychology*, *86*(11), 892–902. https://doi.org/10.1037/ccp0000344

Sharf, J., Primavera, L. H., & Diener, M. J. (2010). Dropout and therapeutic alliance: A meta-analysis of adult individual psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, *47*(4), 637–645. https://doi.org/10.1037/a0021175

Sheehan, D., Lecrubier, Y., Harnett Sheehan, K., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The MINI-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59*, 22–33.

Stanislawski, E., Bilgrami, Z., Sarac, C., Cecchi, G., & Corcoran, C. (2019). S19. Analyzing negative symptoms and language in youths at risk for psychosis using automated language analysis. *Schizophrenia Bulletin*, *45*(Supplement_2), S312–S313.

Stiles, W. B., & Goldsmith, J. Z. (2010). The alliance over time. In J. Muran, & J. Barber (Eds.), *The therapeutic alliance: An evidence-based guide to practice* (pp. 44–62). New York, NY: Guilford Press.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.

Thompson, W. F., & Balkwill, L. L. (2006). Decoding speech prosody in five languages. *Semiotica*, *2006*(158), 407–424.

Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *1*(3), 207–210. https://doi.org/10.1037/1040-3590.1.3.207

Weiste, E., & Peräkylä, A. (2014). Prosody and empathic communication in psychotherapy interaction. *Psychotherapy Research*, *24*(6), 687–701. https://doi.org/10.1080/10503307.2013.879619

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody—Influence of acoustic parameters and arousal. *NeuroImage*, *39*(2), 885–893. https://doi.org/10.1016/j.neuroimage.2007.09.028

World Health Organization. (2018). *Depression, fact sheet*. Geneva: World Health Organization.

Zilcha-Mano, S., Dinger, U., McCarthy, K. S., & Barber, J. P. (2014). Does alliance predict symptoms throughout treatment, or is it the other way around? *Journal of Consulting and Clinical Psychology*, *82*(6), 931–935. https://doi.org/10.1037/a0035141

Zilcha-Mano, S., Dolev, T., Leibovich, L., & Barber, J. P. (2018). Identifying the most suitable treatment for depression based on patients' attachment: Study protocol for a randomized controlled trial of supportive-expressive vs. supportive treatments. *BMC Psychiatry*, *18*(1), 362. https://doi.org/10.1186/s12888-018-1934-1

Zilcha-Mano, S., & Errázuriz, P. (2017). Early development of mechanisms of change as a predictor of subsequent change and treatment outcome: The case of working alliance. *Journal of Consulting and Clinical Psychology*, *85*(5), 508–520. https://doi.org/10.1037/ccp0000192

Zuroff, D. C., Blatt, S. J., Sotsky, S. M., Krupnick, J. L., Martin, D. J., Sanislow, C. A. III, & Simmens, S. (2000). Relation of therapeutic alliance and perfectionism to outcome in brief outpatient treatment of depression. *Journal of Consulting and Clinical Psychology*, *68*(1), 114–124. https://doi.org/10.1037/0022-006X.68.1.114

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Nof A, Amir O, Goldstein P, Zilcha-Mano S. What do these sounds tell us about the therapeutic alliance: Acoustic markers as predictors of alliance. *Clin Psychol Psychother*. 2021;28:807–817. https://doi.org/10.1002/cpp.2534