

# Tracing change during music therapy for depression: Toward a markers-based understanding of communicative behaviors

Musicae Scientiae

1–18

© The Author(s) 2022



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/10298649221116024

[journals.sagepub.com/home/msx](https://journals.sagepub.com/home/msx)

**Sarah Knight**   
University of York, UK

**Neta Spiro**  
Royal College of Music, UK

## Abstract

This article focuses on *behavioral markers*—changes in communicative behaviors that reliably indicate the presence and severity of mental health conditions. We explore the potential of behavioral markers to provide new insights and approaches to diagnosis, assessment, and monitoring, with a particular focus on music therapy for depression. We propose a framework for understanding these markers that encompasses three broad functional categories fulfilled by communicative behaviors: *semantic*, *pragmatic*, and *phatic*. The disordered interactions observed in those with depression reflect changes in many types of communicative behavior, but much research has focused on pragmatic behaviors. However, changes in phatic behaviors also seem likely to be important, given their crucial role in facilitating interpersonal relationships. Given the strong phatic element of music-making, music represents a fertile context in which to explore these behaviors. We argue here that the uniquely multimodal and profoundly interactive environment of music therapy in particular allows for the identification of changes in pragmatic and phatic communicative behaviors that reliably indicate depression presence/severity. By identifying these behavioral markers, we open the door to new ways of assessing depression, and improving diagnosis and monitoring. Furthermore, this markers-based approach has broad implications, being applicable beyond depression and beyond music therapy.

## Keywords

music, mental health, behavioral markers, interaction

In this theoretical article, we aim to highlight the potential of changes in communicative behavior to inform us about others' mental health. We first discuss the concept of behavioral

---

## Corresponding author:

Sarah Knight, Department of Psychology, University of York, York YO10 5DD, UK.

Email: [sarah.knight3@york.ac.uk](mailto:sarah.knight3@york.ac.uk)

markers of ill health. Next, we outline the nature and function of communicative behaviors, first in general and then in the specific context of speech and music. Finally, we present an exploration of one context in which these behaviors are likely to be informative: the occurrence of communicative changes in depression, and in particular the way these changes may be harnessed during music therapy both to understand the nature of depression better and to enhance the efficacy of music-therapeutic approaches. In doing so, we hope to demonstrate that a deeper understanding of changes in communicative behaviors has the power to enrich both the theoretical and practical aspects of our current approaches to tackling mental health conditions, and to encourage further research on this topic.

## **Changes in communicative behaviors during ill health as behavioral markers**

Communicative behaviors can inform us about variations in a person's state within the normal range. For example, tone of voice can indicate excitement or tiredness (Nolan, 2006) and the ways in which we alter our tone of voice relative to that of our interlocutor can indicate agreement or disagreement (Ogden, 2006). However, our communicative behaviors are also affected by variations beyond the normal, for example, physical illness, neurological conditions, and mental health conditions. As a result of these changes, such behaviors become an informative place to seek information about people's well-being.

The association between communicative behaviors and mental state is well established, having been documented at least as early as 1921 by Emil Kraepelin, regarded as the founder of modern psychiatry (Kraepelin, 1921, cited in Cummins et al., 2015). As discussed further below, current evidence suggests that communicative behaviors are not only capable of revealing the presence of a given condition but also vary in a systematic fashion with its severity, allowing changes to be monitored over time. For example, several studies report that the mean pitch and/or pitch range of the speaking voice do not only differ between depressed speakers and healthy controls but also correlate with depression severity. This observation suggests that vocal pitch may provide a means of tracing change in depression severity over time (Cummins et al., 2015). We call these behaviors—those which both indicate the presence of a condition and vary with its severity—*behavioral markers* (after Cummins et al., 2015).

In the case of mental health, the potential importance of the information gleaned from behavioral markers is revealed when considering existing procedures for diagnosis and assessment. Many measures rely on either patient self-report or clinician judgments of symptom severity. At best, measures relying on the opinions of individual clinicians require considerable training and practice before acceptably reliable results are produced. At worst, such measures may be susceptible to systematic bias (Mundt et al., 2007) and even so-called gold standard assessments have significant psychometric weaknesses (Santor & Coyne, 2001; Zimmerman et al., 2005). Patient-reported measures, meanwhile, rely not only on patients' understanding and experience of their own symptoms, which may be highly personal in nature (Mundt et al., 2007), but also on the ability and desire of a patient to communicate their symptoms when mental health problems by their very nature may impair outlook and motivation (Cummins et al., 2015). This is in addition to broader issues related to self-report in mental illness, such as disempowerment (Bibb & McFerran, 2017). Current assessments could therefore be greatly enhanced and enriched by including information derived from the measurement and analysis of relevant behavioral markers.

## The nature of communicative behaviors

The wide range of communicative behaviors we produce are typically thought of in terms of specific activities such as speaking, making music, gesturing, and dancing. However, our communicative behaviors often share common properties both within and across different activities; for example, pitch fluctuation is a key characteristic of both spoken language and many types of music, while temporal predictability can characterize music, speech, or gesture (Lidji et al., 2011; London, 2004; Maricchiolo et al., 2005). Given these common properties, it may be more useful to think about the range of communicative behaviors we employ during interaction in terms of their function: that is, what they communicate—a consideration that is at least partially separate from how they communicate it (i.e., the external form of the communicative activity). It is possible to think about these functions in terms of a number of different frameworks. Below we present an overview of our proposed framework, which encompasses three broad functional categories fulfilled by communicative behaviors: *semantic*, *pragmatic*, and *phatic* (Jakobson, 1980; Malinowski, 1994; Wharton, 2009).

### Functions of communicative behaviors

**Semantic behaviors.** One core group of communicative behaviors helps people transmit specific concepts and ideas (i.e., *semantic* content). This group is referred to by Jakobson (1980) as *referential/denotative* and termed *ideational* by some gesture researchers (e.g., Hadar & Pinchas-Zamir, 2004). The linguistically encoded meaning found in speech is perhaps the most obvious example of such a semantic cue, but nonlinguistic phenomena such as intra-utterance prosody and representational gestures also serve important semantic functions. For example, prosody can encode important semantic information by clarifying the linguistic meaning with which it co-exists and this clarification can take various forms, including lexical clarification through stress (e.g., *permit* [n.] vs. *permit* [v.]), clarification of grammatical structure (e.g., using a lower pitch and quieter voice for a subclause), focus (highlighting the important or novel elements in an utterance), and clarification of discourse function (e.g., differentiating a question from a statement; Gussenhoven, 2002; Nolan, 2006).

**Pragmatic behaviors.** A second group of behaviors carries what might be thought of as *pragmatic* information (i.e., related to the cognitive and affective state of the interacting individuals). These behaviors reflect states internal to the interactants, and which are at best loosely reciprocal between them, but which may be completely temporally disconnected. Examples of this type of behavior include posture, vocal timbre, and speech rate, behaviors that make salient details about the interaction context apparent to the interacting individuals. Put another way, these behaviors show each interactant the other's cognitive and affective states. In the case of pragmatic behaviors that co-occur with semantic content, such as vocal pitch or timbre during speech, these behaviors guide our construal of the semantic content, informing and constraining possible interpretations (Nolan, 2006). For example, the phrase "how exciting" can be interpreted in very different ways depending on whether the speaker is upright, smiling, and speaking fast with a wide vocal pitch range, or slouching, frowning, and speaking in a monotone. These interpretive processes thus result in a rich and nuanced understanding of not only the linguistic sense of the speaker's words but also their intentions, motivations, and the more general cognitive and affective context surrounding the interaction (Wharton, 2009).

*Phatic behaviors.* Despite the importance of pragmatic behaviors, a successful interaction does not rely simply on participants' cognitive states being made mutually apparent. Rather, it requires the establishment, reinforcement, and communication of a common cognitive context and of shared goals and action plans. This outcome cannot be achieved through pragmatic behaviors alone. Communication of this kind requires behaviors that are dynamic, reflect something about a participant's relationship to their interaction partner, and are tightly temporally tied between interactants. This kind of communicative behavior, rather than simply showing, needs to be related to sharing; rather than carrying information, it needs to create social bonds and interpersonal understanding. These cues will be termed here *phatic*, after the work of Malinowski (1994). Malinowski conceived of such cues as carrying no information in and of themselves. However, a subtler understanding highlights the fact that, although these cues do not prioritize semantic or pragmatic meaning, it is nevertheless present, coexisting with higher-level meanings, and specifically with higher-level meanings related to the nature of the interaction (Senft, 2009). That is, these behaviors gain their meaning through the interaction of which they are simultaneously both a part and on which they are commenting; they are both a form of interaction and a statement about that interaction; they emerge from the context that contains them in a dynamic, real-time fashion and they therefore cannot be separated from their interaction and remain meaningful. Examples of phatic behaviors include synchrony, imitation, and turn-taking (e.g., Chartrand & Bargh, 1999; Hove & Risen, 2009; Wilson & Wilson, 2005). For example, it has been demonstrated that mimicry of posture and gesture smooths social interactions and increases liking between participants (Chartrand & Bargh, 1999). In this sense, they reflect a particular use of communicative behaviors: They are behaviors used relationally, with a focus on establishing interpersonal cohesion rather than exchanging information per se. These phatic behaviors allow us to make inferences about our relationship with others and to judge in real time how successfully our interactions with others are proceeding.

Taken together, our communicative behaviors constitute a rich network of cues, with considerable redundancy and overlap. When we interact with others, we use these behaviors to make complex inferences about what other people mean in their communications (Grice, 1957; Sperber & Wilson, 1995; Wharton, 2009), whether the interaction is proceeding smoothly and successfully, and to generate experiences of rapport, affiliation, similarity, and shared experience (Bargh & Chartrand, 1999; Lakin & Chartrand, 2003).

*Communicative behaviors in practice: Comparing speech and music.* Considering communicative behaviors in terms of the three categories outlined above allows us to identify meaningful, functional similarities and differences between different communicative activities, as opposed to superficial resemblances and disparities. To illustrate this, we will briefly consider the cases of two such activities—speech and music.

In almost all the communicative activities that we would commonly class as speech, semantic information is strongly emphasized, whereas in those types we consider music this function is typically de-emphasized. Most obviously, music per se does not involve words, so cannot convey specific referential meanings as language does. Speech does, however, contain prosody—the timing, loudness, and voice quality of spoken information. As detailed above, these qualities can encode important semantic information by clarifying the linguistic meaning with which they co-exist. Comparable surface features may be found in music, such as the use of a perceptual accent to distinguish a structural component or highlight an event that is musically important (e.g., Drake & Palmer, 1993; Sloboda, 1983). Nevertheless, these cues cannot be said to have the same function as their linguistic counterparts, since music lacks the linguistic or

conceptual meaning that in speech such behaviors work to clarify. Thus, music largely cannot be said to embrace truly semantic cues.

Both speech and music generally afford particular prominence to pragmatic cues, since both rely for their communicative success on conveying information about the mental state of a communicator, real or perceived. That is, both of these communicative types seek to fulfill a pragmatic function, and thus both recruit a similar body of communicative cues to achieve this aim. Indeed, pragmatic cues in speech, such as intonation and rhythm, form much of what is often invoked as the music of language, while it is suggested that structures in music such as melodic contour contribute to music's meaning through their similarity to pragmatic cues in speech, allowing human agency and intention to be attributed to the music (e.g., Cross & Woodruff, 2008; Watt & Ash, 1998). Speech prosody, already discussed above, fulfills important pragmatic functions during spoken interactions. The first of these is discourse regulation. For example, pitch, loudness, and relative duration and location are used to mark the end of a speaker's turn, thus allowing for a smooth alternation between the roles of speaker and listener (Gussenhoven, 2002; Local & Walker, 2012; Wilson & Wilson, 2005). The second is to carry information to the listener regarding the speaker's attitudinal and physical states. For example, given the same linguistic content, prosody serves to differentiate boredom from excitement or to communicate fatigue (Crystal, 1969; Gussenhoven, 2002; Nolan, 2006). It is this second pragmatic function that seems to be mirrored closely in music. Timbre, pitch contour, and articulation are thought to communicate an emotional/physical state or character trait of some virtual persona, thus giving music one of its many potential meanings (Cross & Woodruff, 2008; Maus, 1988; Watt & Ash, 1998). It is worth reiterating that, in the case of speech, semantic and pragmatic cues are closely bound together, with each functioning to guide and constrain possible interpretations of the other during complex inferential processes. As discussed above, music strongly de-emphasizes semantic information. In music, then, pragmatic cues are still working to convey attitudinal information, but without the application to—and indeed, one could argue, the constraints of—concurrent semantic information.

However, it is not only pragmatic cues which speech and music share: Both also emphasize phatic cues. These phatic cues allow both speech and music to function as useful tools for the development and maintenance of social bonds, which in turn reflects the importance of such cues in achieving rapport and a sense of shared goals. Spoken interactions afford mimicry of syntax, prosody, posture, and gesture, alongside the types of ritualized verbal exchange that constitute what Malinowski calls *phatic communion* (e.g., "How's it going?"). Music in and of itself does not allow for any verbal cues. However, it is able to exploit all the same nonverbal phatic behaviors as speech, such as melodic and rhythmic mimicry. Furthermore, music tends to possess a relatively strict underlying periodic structure and this temporal predictability affords prolonged and accurate synchrony between interactants (Drake et al., 2000; Kirschner & Tomasello, 2009). The effects of such synchrony appear to be similar to, but more powerful than, simple mimicry (Hove & Risen, 2009; Wiltermuth & Heath, 2009), rendering music particularly effective at promoting positive social judgments (Knight et al., 2016) and fostering positive interpersonal relationships (Kirschner & Tomasello, 2010; Miles et al., 2010, 2011). This explains music's appearance in what Cross terms *situations of social uncertainty*—contexts in which the creation and maintenance of social bonds is particularly important (Cross & Woodruff, 2008). By contrast, strict periodicity is rare in everyday conversational speech (Classé, 1939, cited in Crystal, 1969; Grabe & Low, 2002; Nolan, 2006; Ramus et al., 2000). Thus, everyday speech, although it is successful at communicating pragmatic information non-verbally, is characterized by powerful semantic—and specifically linguistic—cues, which can



occlude the phatic functionality of utterances. This, plus the absence of perceptible temporal regularity, gives everyday speech a weaker phatic function than music.

However, this apparent contrast between music and speech in the phatic domain becomes less stark if we consider situations in which these two activities overlap in their communicative functions. For example, consider infant-directed speech (IDS), the distinctive style of communication used by adults speaking to very young children who have not yet acquired language. Relative to adult-directed speech, IDS is characterized by a higher mean vocal pitch, larger and smoother pitch excursions, longer pauses, shorter utterances, a more rhythmic structure, and more prosodic repetition (Fernald & Kuhl, 1987; Fernald & Simon, 1984). In practice, these characteristics produce a sing-song, music-like quality, to the extent that IDS is often referred to as *musical speech* (Trainor et al., 2000). The particular characteristics of IDS are suggested to fulfill several functions, one of which is to help language acquisition, for example, by providing cues to word and phrase segmentation (Thiessen et al., 2005), and by maintaining attention (Kaplan et al., 1995). However, IDS is also suggested to communicate emotion, promote shared affective states, and build infant-caregiver bonds (Nakata & Trehub, 2004; Trainor et al., 2000; Trevarthen & Malloch, 2000). Indeed, when the caregiver's emotional state is affected, for example, during depression, the characteristic features of IDS become less pronounced: utterances are longer, repetition is reduced, and timing becomes less predictable, reducing synchrony between caregiver and infant (Field, 2010; Robb, 1999). Furthermore, this change in communication style has been linked to socioemotional difficulties among children of depressed caregivers, perhaps due to the lack of supportive coordination during communicative interactions (Murray et al., 2015). Notably, the nonlinguistic functions associated with IDS are also widely associated with music. Music has long been understood as a powerful tool for emotional expression and it has even been suggested that we experience music as an attempt by a virtual other to communicate their mood (Watt & Ash, 1998). Further to this, music has been suggested not simply to express emotion but also to play an important role in emotion regulation (Saarikallio, 2011). Finally, as noted above, music plays a key role in fostering interpersonal bonds and social cohesion, thanks to its powerful phatic content: Unencumbered by linguistic information, it affords not only rich expressivity but also interpersonal mimicry and tight temporal synchrony between participants.

It is clear, then, that IDS should be conceived of as approaching music in terms of its pitch, timing, and phrase structure. This convergence of features explains the sense of musicality in IDS perceived by many observers. More importantly, though, this example highlights the functional level of the perceived similarity, suggesting not only that music and IDS have common surface features, but that particular features are shared because these communicative activities also share particular functions. This point is reinforced by research that explicitly makes the link between the two: For example, Wigram and Gold (2006) describe how musical improvisation for therapeutic purposes “. . . can emulate a mother-infant interaction, where reciprocity in rhythmic, melody and dynamic style is analogous to the way the therapist [is] responding to the child” (p. 536). Furthermore, although the focus in this section has been on IDS, such functional overlap between music and speech can be found more broadly: For example, work on question-and-answer pairs in spoken English has demonstrated that answers relate to questions in more strongly music-like ways—including a shared rhythmic framework and production of across-turn musical intervals—when those answers are aligned/preferred (i.e., agreement) than when they are disaligned/dispreferred (Hawkins et al., 2013; Robledo et al., 2016).

In short, there is a tendency to divide the two aural communicative phenomena of music and speech conceptually, on the basis of linguistically encoded semantic content: Music does

not have it, while speech does. However, this is too stark a division; speech contains many non-verbal qualities we would recognize as musical, such as pitch and rhythm, while music is meaningful, albeit in a nonlinguistic and highly subjective fashion, and both are clearly in some sense communicative. Furthermore, similarities and differences between the two reflect their respective functions: As discussed above, referential communication regarding specific situations and objects is a primary function of language—and something music is rarely capable of achieving—whereas social bonding is often seen as music's primary function. However, as the functions of the two align, they grow in similarity with respect to relevant communicative features. For example, music and IDS are suggested to have a greater overlap of communicative purpose than music and everyday speech, in that both have among their primary functions the expression and regulation of affect and the promotion of interpersonal bonds. In the case of music, these functions help with the navigation of situations of social uncertainty, while in the case of IDS they help to create and sustain the caregiver–infant relationship and support the infant's emotional development. As a result, we see features in IDS that make it appear more music-like.

As this section demonstrates, the proposed framework of communicative functions encourages us to view interaction and communication not in terms of particular activities but in terms of certain behaviors that are found across multiple activities and contexts according to the function of the communication and the aims of the interactants. Such an approach opens up new communicative contexts in which to explore interaction—contexts that may be less common than everyday activities such as conversational speech, but which share relevant behaviors due to their overlapping communicative functions. Indeed, some of these contexts may go beyond more common communicative activities in terms of the emphasis afforded to certain types of behavior and may thus increase our power to observe underlying patterns in such behaviors. In particular, these novel contexts may prove fruitful places to seek communicative behaviors that are especially informative with regard to well-being: behavioral markers that robustly indicate the presence and, when examined within an individual over time, the severity of mental health disorders such as depression. In the following section, we explore one communicative context that seems highly likely to provide new insights into behavioral markers—that of music therapy—and consider its relevance to the diagnosis and monitoring of depression. As discussed further below, we have selected music therapy as our focus because it is simultaneously multimodal and profoundly interactive, thus foregrounding a range of behaviors likely to have relevance to mental well-being. However, it is worth noting that music therapy is not unique in this regard; other communicative contexts exist that display similar characteristics, albeit with a different balance of linguistic and musical (or music-like) content (e.g., other communication-based therapies, such as drama/dance therapy and even some talking therapies) or different goals (e.g., music in health activities). We would therefore suggest that at least some of the arguments presented below may be applicable more widely. However, in the interests of space, we focus here on music therapy. In the section below, we first contextualize depression-related changes in communicative behaviors and then outline their relevance to music therapy.

## **Depression and music therapy**

### *Communicative changes during depression*

The ways in which we communicate and interact with others seem to change during depression. In recent years, researchers have been trying to establish whether these changes can be

measured and used to predict depression presence and severity—that is, if these changes can function as behavioral markers of depression (see Cummins et al., 2015, for a review).

Existing studies have identified a range of behavioral markers of depression in the pragmatic domain, including a lower vocal pitch (Mundt et al., 2007), slower speech rate (Cannizzaro et al., 2004), and longer pauses (Alpert et al., 2001). As well as these absolute changes, depressed communicative behaviors also tend to display atypical variability, becoming generally less variable, for example, a monotonous voice (Cummins et al., 2015), less variable head movement (Girard et al., 2014), and reduced facial expressivity (Scherer et al., 2013). Based on these findings, a substantial and ongoing attempt by researchers is underway to create automated systems that analyze prosodic changes as a means of objectively assessing depression, allowing not just the diagnosis of its presence but also tracking changes in severity within individuals over time (see Cummins et al., 2015, for a review). Prototype systems have produced promising results (e.g., Shannon & Lan, 2016). As well as its efficacy and reduced subjectivity, such an approach is appealing for other reasons. For example, prosodic measures can be obtained noninvasively, nonintrusively, and relatively cheaply and many clinicians already make subjective assessments of prosody during diagnosis, making such measures a natural extension of existing practices (Cummins et al., 2015).

As well as changes to pragmatic behaviors, there also appear to be changes in phatic behaviors in those with a diagnosis of depression. Specifically, there appears to be reduced adaptation and interpersonal congruence. Examples include reduced eye contact (Segrin, 2000) and verbal backchannel (Fiquer et al., 2013), and poorer temporal synchronization (Perilli, 1995). However, less is known about this interactive aspect; although existing studies typically use data from clinical interviews rather than solo tasks, they overwhelmingly focus on the behavior of the interviewee, without examining the interviewer's behavior or the interactional and/or adaptive aspects of the conversation. This is despite the fact that interviewer behaviors and interactive features can predict depression severity above and beyond the interviewee's behaviors (Bouhuys & van den Hoofdakker, 1991; Yang et al., 2013). Existing studies also typically examine behaviors in a single modality, despite the multimodal nature of real-world communication. In recent years, the importance of multimodality has been increasingly recognized, but studies that explore multimodality nevertheless do so only within speech-based, interview-style interactions (Bhatia et al., 2017; Dibeklioglu et al., 2015).

We suggest that these lacunae in our understanding of depression could be addressed by examining a communicative context that is more strongly interactive and more richly multimodal than clinical interviews, and preferably one in which pragmatic and phatic communicative behaviors are foregrounded. We will argue here that one such communicative context is music therapy. Music therapy can take many forms, including listening to music and actively making music. Improvisational music therapy may make use of existing music, but in the United Kingdom, it more typically involves improvisation. It is this particular context that we focus on here, in which the uniquely rich environment of improvisational music therapy allows multiple multimodal channels of communication to be examined simultaneously. Improvisational music therapy is also profoundly interactive. Since speech is not prioritized, communication largely takes place through activities that are not only reciprocal, but coordinated, interwoven, and adaptive. As a result, an examination of the interactive aspects of the therapist–client relationship is vital to understand the communicative behaviors and processes taking place (Spiro & Himberg, 2016). Before exploring these aspects in more detail, we will first introduce improvisational music therapy and discuss its applications.



## *Music therapy*

In improvisational music therapy (hereafter music therapy), the client and therapist improvise music together for therapeutic purposes (e.g., Nordoff & Robbins, 1977; Wigram, 2004). There are many aspects to improvisational music therapy, but it can be most simply characterized as follows. No musical training is required on the part of the client and the instruments used by clients are typically relatively simple, such as drums and tuned percussion. Therapists usually use instruments that allow them to give harmonic support to the client, such as piano and guitar, but depending on the practicalities of the situation they may use other instruments. During the improvisation itself, the therapist listens carefully to all the sounds created, attunes their music to this, and offers holding or containing structures to support the sounds created by the client. Once a musical relationship has been established, the therapist may use musical techniques to expand or challenge the musical contributions (Wigram, 2004). As well as, or instead of, playing instruments, clients may sing, vocalize, and/or move along with the music. In some cases, and where possible, the therapist and client will also discuss the client's experiences of making music, including their thoughts, feelings, images, and experiences of the therapist. This combination of verbal and nonverbal music-making, spoken language, and gesture, all bound together in the context of a carefully monitored interaction, means that music therapy encompasses a huge variety of communicative behaviors that span the semantic, pragmatic, and phatic domains. The centrality of the client–therapist relationship, meanwhile, ensures that there is a particular emphasis on pragmatic and, crucially, phatic interactions.

It is generally thought that music therapy works toward positive change with respect to the relevant therapeutic goals. However, as Aalbers et al. (2017) identify in their recent Cochrane report, high-quality evidence supporting the efficacy of music therapy is limited and more specific studies are needed. This is discussed further below.

## *The use of music therapy for depression*

People with a diagnosis of depression constitute one client group that accesses music therapy. The potential mechanism(s) of action of music therapy on depression are still debated (Aalbers et al., 2017; Maratos et al., 2011). However, there is a strong focus on communication in music therapy sessions and it has been proposed that, through the co-created musical relationship, music therapy helps to engage the client physically and emotionally, creating meaning and facilitating a [re]discovery of self and one's relationship to others (Maratos et al., 2008, 2011; Odell-Miller, 1995).

There is some evidence supporting the efficacy of music therapy for depression, including randomized control trials (RCTs; for discussions of RCTs see, for example, Aalbers et al., 2017; Erkkilä et al., 2011; Gold et al., 2009; Maratos et al., 2008). However, the evidence is far from comprehensive and the number of high-quality studies is limited (Aalbers et al., 2017; Maratos et al., 2008). This relatively small evidence base is attributable to a number of factors. First, although some music therapists and music therapy researchers do carry out RCTs and other large-scale controlled studies, others prefer to focus on the uniqueness of each client and/or session and thus tend to publish individual case studies. Second, it can be difficult to find large or homogeneous enough participant groups to participate in RCTs. Moreover, the therapeutic interventions themselves can be heterogeneous and attempts to control them open up a potentially damaging gap between research and clinical practices (Rolvsjord et al., 2005).

In addition to these issues, existing music therapy assessment tools are subject to the same limitations as the mental health assessment tools discussed above, and to an even greater

degree. Despite the existence of a range of outcome measures, few have had their psychometric quality thoroughly assessed (Spiro et al., 2017). Furthermore, many rely on observational ratings by the therapist, which are prone to subjective bias; not only are individuals engaged in a musical interaction likely to have considerably different perspectives on what has taken place (Schober & Spiro, 2014), but biases are also introduced by an awareness of the aims of an activity (e.g., Kuhlen & Brennan, 2013). In short, the quality of existing tools constitutes a further barrier to establishing an evidence base; there is a need for objective, reliable tools for describing and monitoring change during music therapy.

### *A markers-based approach to music therapy for depression*

Music therapy is thought to help depressed clients by offering opportunities for the co-creation of a meaningful and engaging musical relationship (Maratos et al., 2008, 2011; Odell-Miller, 1995). However, there are other potential communication-related avenues of change. It has been suggested by researchers in other fields that addressing issues linked to the depression-related prosodic changes discussed above, such as interpersonal timing, could form part of the therapeutic approaches that emphasize social communication in recovery from depression (Yang et al., 2013). In its improvisational, active form, music therapy supports interpersonal interaction, emotional- and self-expression, and provides a framework to structure interpersonal and communicative timing (Aigen, 2014; Nordoff & Robbins, 1977; Wigram, 2004). As such, music therapy seems to constitute just such a social communication-oriented therapeutic approach, assisting depressed clients with the production and regulation of pitch- and timing-related prosodic features by inviting, supporting, and developing their use in the domains of music, gesture, and sometimes speech. More generally, the use of expressive prosodic features in music is thought to be strongly linked to the same expressive behaviors in speech (Juslin & Laukka, 2003).

With these ideas in mind, we would argue that just as pragmatic behaviors in speech (e.g., prosody) can be used to detect depression and track its severity, so the communicative behaviors forming clients' musical interactions might provide the basis for a tool to assess and track change in individuals with a diagnosis of depression over the course of their therapy sessions. Furthermore, as discussed above, music therapy seems a promising context in which to investigate not just nonverbal but also, specifically, phatic behaviors—aspects of communication that are not so easily explored in the less interactive and/or strongly speech-based context of clinical interviews and comparable data sources.

*The proposed markers-based approach.* As outlined above, both pragmatic and phatic communicative behaviors appear to undergo changes during depression. Although these changes have been identified primarily in the speech/conversational domain, in all cases comparable musical behaviors can be found that are relevant to music-therapeutic practices and strategies. As discussed above, existing evidence suggests that basic aspects of communication and their intrapersonal variability, within client or therapist, may be affected; for example, both vocal pitch (Mundt et al., 2007) and spoken pitch range (Cummins et al., 2015) of an individual have been found to be associated with the presence of depression and, at least in some studies, to correlate with depression severity. In the music therapy domain, a client's use of pitch, sung or instrumental, could be examined along similar lines, with mean musical pitch, pitch range, and level of pitch variability, all accessible, measurable, and potentially informative features of a music-therapeutic interaction. Similarly, depression has also been linked to a slower speech rate (Cannizzaro et al., 2004) and longer within-turn pauses (Alpert et al., 2001). These changes may be

mirrored by a slower musical pulse and increased within-turn pause duration during music-making. As discussed above, communicative behaviors during interaction are often examined only for the individual with a clinical diagnosis, despite compelling evidence that the behaviors of the person with whom they are interacting—such as a therapist—are also informative (Bouhuys & van den Hoofdakker, 1991; Yang et al., 2013). In our approach, relevant measures, such as pitch and temporal features, could be obtained for both client and therapist, enabling the communicative behaviors of both participants in the musical interaction to be examined as fully as possible. In addition to highlighting the potentially informative nature of therapists' behaviors, existing research suggests that variability in interpersonal communicative behaviors and behavioral adaptation between client and therapist may also be affected by depression presence and/or severity—for example, the duration and variability of switching pauses (Yang et al., 2013), occurrences of verbal backchannelling (Fiquer et al., 2013), and accuracy of temporal synchrony (Perilli, 1995) are all suggested to change during depression. Musical equivalents, including turn-taking behaviors, imitation, and synchrony (entrainment), are all available for examination and measurement.

Although these hypothesized markers are based closely on existing findings from spoken interactions, considerable empirical work is needed to determine whether or not they do in fact serve as markers of depression in the music-therapeutic context: that is, whether or not these behaviors can not only indicate the presence of depression but also correlate with depression severity. Should some subset be shown to be robust behavioral markers of depression, however, then the measurement and examination of these behaviors over time will constitute a powerful tool for assessment, allowing changes in the client's well-being to be traced over time.

*The possibility of automation.* As well as identifying behavioral markers of depression, many researchers are now attempting to automate their measurement and analysis (e.g., Girard & Cohn, 2015; see also Rana et al., 2019). The degree of automation varies, but full automation is possible—that is, software capable of analyzing behavioral information to produce numeric and/or graphical indicators of relevant markers with only minimal input from users. Such an approach may be of great value here. Automation enhances reliability, reduces subjective bias, and allows analysis protocols to be easily shared, helping to standardize diagnosis and monitoring. In addition to these general benefits, automation would enable music therapists to avoid situations where they are required to make nuanced judgments in real time about adaptive interactions in which they themselves are participants—a problematic process (Schober & Spiro, 2014)—or undertake laborious manual analyses of session recordings, for which they often lack the time and/or resources (Streeter, 2010). Attempts have been made to facilitate and investigate the computational analysis of music therapy sessions (Erkkilä, 2007; Storm, 2013; Streeter, 2010). However, it is unclear whether or not the features included for analysis in these systems and studies are meaningful with respect to any given condition; that is, it is unclear whether or not these features are actually behavioral markers. Existing tools also tend to prioritize the analysis of individual as opposed to interactive behaviors and do not usually allow for the inclusion of speech or gesture. Moreover, existing systems often involve relatively high levels of supervision, as the user is required to provide considerable manual input. It is also important to note that these approaches are now considerably outdated: Recent years have seen huge advances in signal detection and music information retrieval, and apps are now available that allow even phones to perform advanced real-time audio analysis (Marchi et al., 2016). Indeed, there has been an upsurge in the availability and use of digital psychiatry tools more generally (Torous et al., 2021). Many of these tools are designed for remote monitoring of symptoms and/or remote/virtual delivery of interventions, whereas the markers-based

approach advocated here focuses primarily on the content of real-time interactive behaviors within face-to-face music therapy sessions. Nevertheless, the growth of digital psychiatry highlights the fact that smartphones and other devices are more capable than ever of capturing and analyzing potentially relevant information—from vocal pitch to physical movements—thus making a markers-based approach a possibility even for therapists without access to expensive recording equipment. Such tools would also allow for the remote monitoring of certain musical and/or nonmusical behaviors between music therapy sessions, which may help researchers and practitioners to better understand the nature of any changes taking place. It is therefore highly desirable to examine the current state-of-the-art software to determine whether or not it is capable of producing sufficiently accurate measures of relevant behavioral markers identified in the music-therapeutic domain with minimal input from the user. Such a project would not only identify weaknesses in existing software, which any future application would need to overcome, but would also streamline an otherwise unwieldy problem by focusing only on demonstrably meaningful features.

## Implications

### *Benefits of a markers-based approach*

From the perspective of music therapists and their clients, the identification of behavioral markers has two obvious practical benefits. First, it would provide a powerful additional tool for music therapists to help them assess change and progress in their clients. It is not suggested that the information derived from behavioral markers should replace the therapist's judgment or training, but rather that it can provide a fresh perspective on the interactions that take place—a perspective not available to those involved in the interaction itself. Furthermore, if analyses of behavioral markers can be automated, they are also likely to provide considerable detail regarding subtle variation in clients' and/or therapists' behaviors, which would only otherwise be available through extensive and time-consuming analysis of video and audio data. Second, music therapists are under increasing pressure to provide evidence of effectiveness. The act of identifying behavioral markers is not in and of itself indicative of treatment efficacy: Indeed, behavioral markers may provide evidence against the benefits of a given therapy. However, identifying markers and developing the tools to analyze them are important steps toward developing a user-friendly way for therapists to collect high-quality data regarding music therapy's potential efficacy. This will therefore be highly relevant to music therapy and mental health service providers who seek to base their provision and funding on evidence-based practices. Clients of music therapists would therefore benefit not just from improved therapeutic practices but also potentially from enhanced availability of services, should behavioral markers allow a body of empirical evidence to be accumulated which encourages wider provision of music therapy. The identification of behavioral markers also has the potential to contribute to our understanding of depression and communication more broadly. There is some evidence that certain patterns of frontal cortical activity might act as biomarkers for depression and anxiety. Specifically, measures of Frontal Alpha Asymmetry (FAA) and Front Midline Theta (FMT) appear not only to differentiate depressed and/or anxious individuals from healthy individuals but have also been shown to index change over time for a group of depressed clients receiving music therapy (Fachner et al., 2013). In this study, the music therapy intervention was also associated with self-reported improvements in communication. Taken together, these results are suggestive of a complex web of biomarkers and behavioral markers related to emotional processing and expression (see also Odell-Miller et al., 2018). A better understanding of

the behavioral markers relevant to depression in the music-therapeutic context would afford a deeper insight into this network, thus broadening our understanding of communication during depression and, ultimately, allowing for improved therapeutic practices.

*Beyond depression.* Music therapy is accessed by a wide range of client groups of all ages, including those with emotional or mental health needs, learning and/or physical disabilities, developmental disorders, life-limiting conditions, neurological conditions, and physical illnesses. The therapeutic aims of music therapy vary considerably depending on the therapist, client, and context. For example, in acute psychiatric in-patient settings, music therapy tends to focus on engaging with patients, creating immediate effects such as reduction in arousal and enabling short-term management of symptoms (Carr et al., 2013). With dementia sufferers, aims may range from short-term management of mood and aggression to the accessing of autobiographical memories and enhancement of speech fluency and verbal memory (Spiro, 2010). Although we have focused here on depression, there is evidence for similar behavioral markers in other conditions, such as autism spectrum disorders (e.g., Kim et al., 2009) and schizophrenia (e.g., Pavlicevic et al., 1994). There is therefore good reason to expect that the behavioral markers approach will be helpful beyond depression—even if the specific set of behaviors that change, and the ways in which they do so, are specific to each condition.

## Conclusion

In this article, we have introduced the concept of behavioral markers: behaviors that are informative with respect to the presence and severity of mental health conditions such as depression. We have argued that an approach to tracing changes in well-being based on the use of behavioral markers has the potential to form a powerful, efficient, and evidence-based tool with applications across a variety of individuals and contexts. We have explored improvisational music therapy as one context in which, due to the multimodal and profoundly interactive nature of the activity, relevant behavioral markers are likely to be present, identifiable, and robust. In particular, we have focused on potential music therapy-derived behavioral markers of depression, but such an approach could be relevant to a range of conditions. The identification of robust behavioral markers of depression would greatly enrich, and indeed enhance, existing methods for depression diagnosis and monitoring, which typically rely on subjective judgments and as such are prone to bias.

There are several important caveats to bear in mind. First, we are not arguing for the measurement and analysis of behavioral markers to replace the judgments of therapists or clinicians. Instead, we envisage a markers-based approach as constituting a powerful additional tool, to be used in conjunction with clinicians' existing skills, expertise, and understanding of their clients. Second, much empirical work is needed before such an approach can be reliably implemented; detailed exploration of behavioral data and rigorous testing of potential markers must necessarily precede any clinical application. Finally, it is possible that a markers-based approach may not prove useful in all cases, due to the idiosyncratic nature of conditions such as depression. However, current findings in the speech domain are sufficiently robust to suggest that such an approach will nevertheless be of value in many cases.

In conclusion, we envisage a markers-based approach as having the potential to constitute a powerful and empirically supported means of tracing change, and we hope further research will bring such a tool to fruition.



## Acknowledgements

We are grateful to the music therapist and researcher Dr Catherine Carr and to Professor of Human Interaction Pat Healey (both at Queen Mary University of London), and to SUGAR (Service User and Carer Group Advising on Research). Dr Carr advised on aspects of music therapy. Both Dr Carr and Professor Healey participated in the development of proposals for empirical work in this area, the discussion of which affected some ideas in this paper. Members of SUGAR discussed the proposed project with the research team and provided feedback.

## Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

## ORCID iD

Sarah Knight  <https://orcid.org/0000-0002-5013-9364>

## References

- Aalbers, S., Fusar-Poli, L., Freeman, R. E., Spreen, M., Ket, J. C. F., Vink, A. C., Maratos, A., Crawford, M., Chen, X. J., & Gold, C. (2017). Music therapy for depression. *Cochrane Database of Systematic Reviews*, 11, Article CD004517. <https://doi.org/10.1002/14651858.CD004517.pub3>
- Aigen, K. (2014). Music-centered dimensions of Nordoff-Robbins music therapy. *Music Therapy Perspectives*, 32(1), 18–29.
- Alpert, M., Pouget, E. R., & Silva, R. R. (2001). Reflections of depression in acoustic measures of the patient's speech. *Journal of Affective Disorders*, 66(1), 59–69.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54, 462–479.
- Bhatia, S., Hayat, M., Breakspear, M., Parker, G., & Goecke, R. (2017, May 30–June 3). A video-based facial behaviour analysis approach to melancholia. In *12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)* (pp. 754–761). IEEE.
- Hammal, Z., Yang, Y., & Cohn, J. F. (2015)
- Bibb, J., & McFerran, K. S. (2017). The challenges of using self-report measures with people with severe mental illness: Four participants' experiences of the research process. *Community Mental Health Journal*, 53(6), 747–754.
- Bouhuys, A. L., & van den Hoofdakker, R. H. (1991). The interrelatedness of observed behavior of depressed patients and of a psychiatrist: An ethological study on mutual influence. *Journal of Affective Disorders*, 23(2), 63–74.
- Cannizzaro, M., Harel, B., Reilly, N., Chappell, P., & Snyder, P. J. (2004). Voice acoustical measurement of the severity of major depression. *Brain and Cognition*, 56(1), 30–35.
- Carr, C., Odell-Miller, H., & Priebe, S. (2013). A systematic review of music therapy practice and outcomes with acute adult psychiatric in-patients. *PLOS ONE*, 8(8), Article e70252.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76, 893–910.
- Cross, I., & Woodruff, G. E. (2008). Music as a communicative medium. In R. Botha & C. Knight (Eds.), *The prehistory of language* (pp. 113–144). Oxford University Press.
- Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge University Press.
- Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, 71, 10–49.
- Dibeklioğlu, H., Hammal, Z., Yang, Y., & Cohn, J. F. (2015). Multimodal detection of depression in clinical interviews. In *Proceedings of the 2015 ACM International Conference on Multimodal Interaction* (pp. 307–310). Association for Computing Machinery.
- Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, 77, 251–288.

- Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception, 10*(3), 343–378.
- Erkkilä, J. (2007). Music Therapy Toolbox (MTTB): An improvisation analysis tool for clinicians and researchers. In T. Wosch & T. Wigram (Eds.), *Microanalysis in music therapy: Methods, techniques and applications for clinicians, researchers, educators and students* (pp. 134–148). London: Jessica Kingsley.
- Erkkilä, J., Punkanen, M., Fachner, J., Ala-Ruona, E., Pöntiö, I., Tervaniemi, M., Vanhala, M., & Gold, C. (2011). Individual music therapy for depression: Randomised controlled trial. *The British Journal of Psychiatry, 199*(2), 132–139.
- Fachner, J., Gold, C., & Erkkilä, J. (2013). Music therapy modulates fronto-temporal activity in rest-EEG in depressed clients. *Brain Topography, 26*(2), 338–354.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*(3), 279–293.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*(1), 104–113.
- Field, T. (2010). Postpartum depression effects on early interactions, parenting, and safety practices: A review. *Infant Behavior and Development, 33*(1), 1–6.
- Fiquer, J. T., Boggio, P. S., & Gorenstein, C. (2013). Talking bodies: Nonverbal behavior in the assessment of depression severity. *Journal of Affective Disorders, 150*(3), 1114–1119.
- Girard, J. M., & Cohn, J. F. (2015). Automated audiovisual depression analysis. *Current Opinion in Psychology, 4*, 75–79.
- Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S. M., Hammal, Z., & Rosenwald, D. P. (2014). Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses. *Image and Vision Computing, 32*(10), 641–647.
- Gold, C., Solli, H. P., Krüger, V., & Lie, S. A. (2009). Dose–response relationship in music therapy for people with serious mental disorders: Systematic review and meta-analysis. *Clinical Psychology Review, 29*(3), 193–207.
- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology, 7*, 515–546.
- Grice, H. P. (1957). Meaning. *The Philosophical Review, 66*(3), 377–388.
- Gussenhoven, C. (2002). *The phonology of tone and intonation*. Cambridge University Press.
- Hadar, U., & Pinchas-Zamir, L. (2004). The semantic specificity of gesture: Implications for gesture classification and function. *Journal of Language and Social Psychology, 23*(2), 204–214.
- Hawkins, S., Cross, I., & Ogden, R. (2013). Communicative interaction in spontaneous music and speech. In M. Orwin, C. Howes, & R. Kempson (Eds.), *Music, language and interaction* (pp. 285–329). College Publications.
- Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition, 27*(6), 949–961.
- Jakobson, R. (1980). *The framework of language*. University of Michigan Press.
- 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.
- Kaplan, P. S., Goldstein, M. H., Hucceby, E. R., Owren, M. J., & Cooper, R. P. (1995). Dishabituation of visual attention by infant-versus adult-directed speech: Effects of frequency modulation and spectral composition. *Infant Behavior and Development, 18*(2), 209–223.
- Kim, J., Wigram, T., & Gold, C. (2009). Emotional, motivational and interpersonal responsiveness of children with autism in improvisational music therapy. *Autism, 13*(4), 389–409.
- Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology, 102*, 299–314.
- Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior, 31*(5), 354–364.
- Knight, S., Spiro, N., & Cross, I. (2016). Look, listen and learn: Exploring effects of passive entrainment on social judgements of observed others. *Psychology of Music, 45*(1), 99–115.

- Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, 20(1), 54–72.
- Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioural mimicry to create affiliation and rapport. *Psychological Science*, 14(4), 334–339.
- Lidji, P., Palmer, C., Peretz, I., & Morningstar, M. (2011). Listeners feel the beat: Entrainment to English and French speech rhythms. *Psychonomic Bulletin & Review*, 18(6), 1035–1041.
- Local, J., & Walker, G. (2012). How phonetic features project more talk. *Journal of the International Phonetic Association*, 42, 255–280.
- London, J. (2004). *Hearing in time*. Oxford University Press.
- Malinowski, B. (1994). The problem of meaning in primitive languages. In J. Maybin (Ed.), *Language and literacy in social practice* (pp. 1–10). Multilingual Matters in association with the Open University.
- Maratos, A., Crawford, M. J., & Procter, S. (2011). Music therapy for depression: It seems to work, but how? *British Journal of Psychiatry*, 199(2), 92–93.
- Maratos, A., Gold, C., Wang, X., & Crawford, M. (2008). Music therapy for depression. *Cochrane Database of Systematic Reviews*, Issue 1, Article CD004517. <https://doi.org/10.1002/14651858.CD004517.pub2>
- Marchi, E., Eyben, F., Hagerer, G., & Schuller, B. W. (2016). Real-time tracking of speakers' emotions, states, and traits on mobile platforms. *INTERSPEECH*, 2016, 1182–1183.
- Maricchiolo, F., Bonaiuto, M., & Gnisci, A. (2005). *Hand gestures in speech: Studies of their roles in social interaction* [Conference session]. 2nd International Society for Gesture Studies Conference, 15–18 June 2005, Lyon, France.
- Maus, F. E. (1988). Music as drama. *Music Theory Spectrum*, 10, 56–73.
- Miles, L. K., Griffiths, J. L., Richardson, M. J., & Macrae, C. N. (2010). Too late to coordinate: Contextual influences on behavioral synchrony. *European Journal of Social Psychology*, 40, 52–60.
- Miles, L. K., Lumsden, J., Richardson, M. J., & Macrae, C. N. (2011). Do birds of a feather move together? Group membership and behavioral synchrony. *Experimental Brain Research*, 211(3–4), 495–403.
- Mundt, J. C., Snyder, P. J., Cannizzaro, M. S., Chappie, K., & Geralt, 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.
- Murray, L., Fearon, P., & Cooper, P. (2015). Postnatal depression, mother-infant interactions, and child development: Prospects for screening and treatment. In J. Milgrom & A. W. Gemmill (Eds.), *Identifying perinatal depression and anxiety: Evidence-based practice in screening, psychosocial assessment, and managements* (pp. 139–164). John Wiley & Sons.
- Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development*, 27(4), 455–464.
- Nolan, F. (2006). Intonation. In B. Aarts & A. McMahon (Eds.), *Handbook of English linguistics* (pp. 385–405). Blackwell.
- Nordoff, P., & Robbins, C. (1977). *Creative music therapy: Individualized treatment for the handicapped child*. John Day Company.
- Odell-Miller, H. (1995). Why provide music therapy in the community for adults with mental health problems? *British Journal of Music Therapy*, 9(1), 4–10.
- Odell-Miller, H., Fachner, J., & Erkkila, J. (2018). Music therapy clinical practice and research for people with depression: Music, neuroscience and music therapy. In A. Zubala & V. Karkou (Eds.), *Arts therapies in the treatment of depression* (pp. 154–171). Routledge.
- Ogden, R. (2006). Phonetics and social action in agreements and disagreements. *Journal of Pragmatics*, 38(10), 1752–1775.
- Pavlicevic, M., Trevarthen, C., & Duncan, J. (1994). Improvisational music therapy and the rehabilitation of persons suffering from chronic schizophrenia. *Journal of Music Therapy*, 31(2), 86–104.
- Perilli, G. G. (1995). Subjective tempo in adults with and without psychiatric disorders. *Music Therapy Perspectives*, 13(2), 104–109.
- Ramus, F., Nespor, M., & Mehler, J. (2000). Correlates of linguistic rhythm in the speech signal. *Cognition*, 75(1), AD3–AD30.

- Rana, R., Latif, S., Gururajan, R., Gray, A., Mackenzie, G., Humphris, G., & Dunn, J. (2019). Automated screening for distress: A perspective for the future. *European Journal of Cancer Care*, 28(4), Article e13033.
- Robb, L. (1999). Emotional musicality in mother-infant vocal affect, and an acoustic study of postnatal depression. *Musicae Scientiae Special Issue*, 3, 123–154.
- Robledo, J. P., Hawkins, S., Cross, I., & Ogden, R. A. (2016). Pitch-interval analysis of “periodic” and “aperiodic” Question + Answer pairs. *Proceedings of Speech Prosody*, 2016, 1071–1075.
- Rolvsvjord, R., Gold, C., & Stige, B. (2005). Research rigour and therapeutic flexibility: Rationale for a therapy manual developed for a randomised controlled trial. *Nordic Journal of Music Therapy*, 14(1), 15–32.
- Saarikallio, S. (2011). Music as emotional self-regulation throughout adulthood. *Psychology of Music*, 39(3), 307–327.
- Santor, D. A., & Coyne, J. C. (2001). Examining symptom expression as a function of symptom severity: Item performance on the Hamilton Rating Scale for Depression. *Psychological Assessment*, 13(1), 127–139.
- Scherer, S., Stratou, G., & Morency, L. P. (2013, December). Audiovisual behavior descriptors for depression assessment. In *Proceedings of the 15th ACM International Conference on Multimodal Interaction* (pp. 135–140). ACM.
- Schober, M. F., & Spiro, N. (2014). Jazz improvisers’ shared understanding: A case study. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00808>
- Segrin, C. (2000). Social skills deficits associated with depression. *Clinical Psychology Review*, 20(3), 379–403.
- Senft, G. (2009). Phatic communion. In G. Senft, J.-O. Östman, & J. Verschueren (Eds.), *Culture and language use* (pp. 226–233). John Benjamins.
- Shannon, T. T. E., & Lan, S. S. (2016). Speech analysis and depression. In *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)* (pp. 1–4). IEEE.
- Sloboda, J. A. (1983). The communication of musical metre in piano performance. *The Quarterly Journal of Experimental Psychology*, 35(2), 377–396.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Blackwell.
- Spiro, N. (2010). Music and dementia: Observing effects and searching for underlying theories. *Aging & Mental Health*, 14(8), 891–899.
- Spiro, N., & Himberg, T. (2016). Analysing change in music therapy interactions of children with communication difficulties. *Philosophical Transactions of the Royal Society B*, 371(1693), 20150374. <https://doi.org/10.1098/rstb.2015.0374>
- Spiro, N., Tsiris, G., & Cripps, C. (2017). A systematic review of outcome measures in music therapy. *Music Therapy Perspectives*, 36(1), 67–78.
- Storm, S. (2013). *Research into the development of voice assessment in music therapy* [Unpublished PhD thesis, Institut for Kommunikation, Aalborg Universitet].
- Streeter, E. (2010). *Computer aided music therapy evaluation: Investigating and testing the Music Therapy Logbook Prototype 1 system* [Unpublished PhD thesis, University of York].
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7(1), 53–71.
- Torous, J., Bucci, S., Bell, I. H., Kessing, L. V., Faurholt-Jepsen, M., Whelan, P., Carvalho, A. F., Keshavan, M., Linardon, J., & Firth, J. (2021). The growing field of digital psychiatry: Current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, 20(3), 318–335.
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, 11(3), 188–195.
- Trevarthen, C., & Malloch, S. N. (2000). The dance of wellbeing: Defining the musical therapeutic effect. *Nordisk Tidsskrift for Musikkterapi*, 9(2), 3–17.
- Watt, R. J., & Ash, R. L. (1998). A psychological investigation of meaning in music. *Musicae Scientiae*, 2(1), 33–54.

- Wharton, T. (2009). *Pragmatics and non-verbal communication*. Cambridge University Press.
- Wigram, T. (2004). *Improvisation: Methods and techniques for music therapy clinicians, educators, and students*. Jessica Kingsley.
- Wigram, T., & Gold, C. (2006). Music therapy in the assessment and treatment of autistic spectrum disorder: Clinical application and research evidence. *Child: Care, Health and Development*, 32(5), 535–542.
- Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review*, 12(6), 957–968.
- Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, 20(1), 1–5.
- Yang, Y., Fairbairn, C., & Cohn, J. F. (2013). Detecting depression severity from vocal prosody. *IEEE Transactions on Affective Computing*, 4(2), 142–150.
- Zimmerman, M., Posternak, M. A., & Chelminski, I. (2005). Is it time to replace the Hamilton Depression Rating Scale as the primary outcome measure in treatment studies of depression? *Journal of Clinical Psychopharmacology*, 25(2), 105–110.