CHAPTER 10

MUSIC PERFORMANCE ANALYSIS

Music is a performing art. While the differentiation between the score (or the underlying musical ideas) and its performance is hard in the case of popular music, this is not the case with classical western music. Here, it requires a performer or a group of performers who "self-consciously enacts music for an audience" [387]. The performers render the composer's work, a score containing musical ideas and performance instructions, into a physical realization.

10.1 Musical Communication

The communication between composer and listener can be visualized as a chain of musical communication derived from Kendall and Carterette as shown in Fig. 10.1 [388]. No direct communication takes place between composer and listener. Instead, the composer translates his musical ideas into a score which is analyzed by the performer to devise a performance concept or plan and finally to render the acoustic realization — the actual music performance — which is subsequently perceived by the listener. Each of the communication stages allows or even enforces interpretation, modification, addition, and dismissal of information.

10.1.1 Score

A musical *score* standing in the tradition of western music history always contains information on pitch and (relative) duration of each note as well as instructions on musical

An Introduction to Audio Content Analysis: Applications in Signal Processing and Music Informatics,
First Edition. Alexander Lerch.

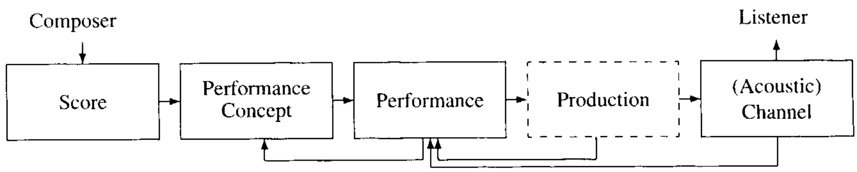

Figure 10.1 Chain of musical communication

dynamics (compare Sects. 4.2, 5.2, and 6.2). Additional instructions, for instance, on character, quality, or specific ways to perform may also be found in the score. Some of the contained information is available only implicitly (for example, information on the musical structure) or might be ambiguous or hidden, complicating its description and quantification as pointed out by many musicologists and music psychologists [319, 389–391].

All this information is subject to the performers' interpretation — they detect and evaluate implicit information, try to understand and explain performance instructions, identify ways to convey their understanding of musical ideas to the listener, and transform the discrete score representation of pitch, duration, and dynamics to continuous scales.

It can be observed that modern scores tend to be more explicit in terms of performance instructions than historic scores, indicating the composers' intention to eliminate the unspecified or ambiguous information in the score [389]. This may be due to the increasing awareness of the fact that scores often take into account performance rules that may seem "natural" at the time of composition but may change over decades and centuries, possibly leading to "unintended" performances.

Although the literature on musical performance sometimes conveys the impression that imprecision and restriction of the score representation is undesirable, there can be no doubt that there is no true, absolute, or optimal interpretation. Music is a living art and constant re-interpretation is the artistic breath giving music life.

10.1.2 Music Performance

The *music performance* is the acoustic realization of the score. Several authors defined the expressive parts of a performance as the deviations from a reference performance. Seashore saw a "neutral," mechanical score rendition as reference [45]. Other authors defined such a reference performance as a performance which is *perceived* as mechanic (which may not be necessarily a mechanical performance [191]) or as a performance with "perfectly normative rubato (and the equivalent on all other relevant expressive parameters)" [392] which is a performance matching all standard or default expectations of the (average) listener.

Every performance requires a concept or plan created by either a rigorous or a rather intuitive and unsystematic analysis of the score (for instance, in the case of sight-reading). The performance plan is a mental representation of the music, an abstract list of actions that may be realized in an indefinite number of ways and is specified only relative to the context [192, 393]. The performance plan is so closely related to the performance itself that it does not always make sense to treat them separately, and the following paragraphs will not always differentiate between the plan and the performance itself.

A music performance is highly individual in both its production and its perception. Still, a list of parameters that the performance may depend on can be compiled. The number of influencing parameters on the performance (and the performance plan) itself is probably infinite; nevertheless, the following list attempts to describe the main influences which may explicitly or implicitly influence a musical performance. This list has been inspired by numerous texts on music performance [191, 387, 389, 390, 393–399]:

- *General interpretative rules*: These are rules, conventions, or norms which every performance follows because it would be perceived as uncommon or even unnatural otherwise.
- *Performance plan and expressive strategy*: A concept of interpretation as a list of actions that may be influenced by
  - *the interpretation of musical structure* or shape, e.g., the question of how to successfully convey melody, phrases, etc. to the listener,
  - *the addition of unexpectedness* by deviation from expected conventions or rules,
  - *the stylistic and cultural context and rules* possibly changing over time, varying between countries or following "performance fashions" [396], depending on both the historic context (the time the piece of music was composed or premiered) as well as on the context at the time of the performance. This includes instruments or instrument characteristics, tuning frequencies and temperaments, and specific performance styles with respect to articulation, ornamentation, vibrato styles, tempo, and *rubato*,
  - *the musical mood and emotional expression* the performer plans to convey to the listener, and
  - *the performance context* such as the expected audience, the style and performance plan of other performances and pieces in the concert program.
- *Performers' personal, social, and cultural background*: A broad category including, e.g., previous performing and general experiences, teachers and mentors, attitude, manners and mannerisms, etc.
- *Physical influences*: The auditory and motorical or — more generally — physical and cognitive abilities of the performer may lead to forced or unintended deviations from the performance plan. This covers general human limitations such as the motoric precision in timing as well as attributes of the musical instrument that impose limitations on, e.g., fingering and breathing.
- *Rehearsal*: The rehearsal phase allows direct feedback on the performance plan and may also train some specific motorical abilities of the performer.
- *Immediate influences*: Influences which may change the performance at the time of performance and may lead to a deviation from the performance concept such as
  - *runtime feedback control*, the feedback that the performer directly receives consisting of auditory, visual, tactile, and other cues [400]; examples include the instrument's sound and reaction, the performance of co-performers, the acoustics of the environment, and the reaction of the audience,
  - *external influences* not directly related to the performance such as humidity, temperature, distractions, and
  - *"internal" influences* such as the emotional and physical state of the performers (stress, stage fright, fatigue, illness, etc.).

10.1.3 Production

Recorded performances can differ significantly from the live performance, even in the case of so-called live recordings [397, 401]. The reason is that persons other than the performers themselves, for instance, the producer, sound engineer, and editor will influence the final result during the production stage. Furthermore, mechanical and technological restrictions enforce differences between an original and reproduced performance but also open up new possibilities of improving a recorded performance during the post-production process. To give an example, it is established recording practice to not only record several complete performances and finally choose the "best," but instead to record several so-called *takes* of passages of the musical piece. The recording process can also involve repeated listening to the recorded takes and discussions on the performance with influence on the following performances. Afterward, it is decided which parts of these takes will finally be used on the published CD, and these will be edited in a way that the edit points are inaudible. Having analyzed seven productions of Beethoven's 9th Symphony, Weinzierl and Franke found between 50 and 250 cuts between different takes in each production; the average number of edits increased with the technical evolution [402]. Modern software allows the editing of audio signals at nearly any score position.

Microphones and their positioning as well as signal processing applied by the sound and mastering engineers may impact the loudness, the timbre, the reverberation, and other parameters of the recording. These "interventions" can also vary over time to artificially increase or decrease acoustical or performance-based effects (e.g., increase the loudness of a specific instrument for its solo part). Maempel et al. give an overview on processing options and typical objectives in the post-production context [403].

The musician's and the producer team's influences are not distinguishable on the final product, for example, the CD [404]. It is common practice to refer to the resulting recording as music performance; this seems to be a valid approach as the artist usually states his final agreement with the recording.

10.1.4 Recipient

The listener, as the receiving end point of the chain of musical communication, subjectively interprets the music. He listens to a performance and conceives musical ideas and other information conveyed by the performance. As Lundin points out, the kinds of possible affective reactions of listeners are practically limitless [405] and allow a multiple of different research angles.

10.2 Music Performance Analysis

Music Performance Analysis (MPA) aims at studying the performance of a musical score rather than the musical score itself. It deals with the observation, extraction, description, interpretation, and modeling of music performance parameters as well as the analysis of attributes and characteristics of the generation and perception of music performance.

Different areas of research contribute to the field of MPA, including musicology, (music) psychology and engineering. A valuable introduction to the research field is given by Clarke [406]. Articles providing extensive overviews have been compiled, for instance, by Gabrielsson [192], Palmer [390] and Goebl et al. [407].

Three basic directions can be identified in the field of systematic performance analysis, namely the study of

- the *performance* itself, i.e., the identification of common and individual characteristics in the performance data, general performance rules, or differences between individual performances,
- the generation or *production of a performance*, i.e., the understanding of the underlying principles of performance plans, the relation of the performers' intention to objective performance parameters, and the performers' motoric and memory skills, and
- the *reception of a performance*, i.e., the investigation of how performances or the variation of specific parameters are perceived by a listener and how he is affected.

MPA could on the one hand lead to more explicit formulations of the different (objective) performance characteristics in the practice of music teaching or enable the development of teaching assisting systems giving the student direct and objective feedback on the performance parameters. On the other hand, it could assist the implementation of performance models which generate computer renditions of human-like music performances. MPA also allows us to gain valuable insights for the research fields music psychology, music aesthetics, and music history.

As Clarke points out, "musical analysis is not an exact science and cannot be relied upon to provide an unequivocal basis for distinguishing between errors and intentions" [406], emphasizing the challenge of meaningful interpretation of extracted performance data. A related difficulty that MPA has to deal with is to distinguish between inherent performance attributes and individual performance attributes. In the context of musical accents, Parncutt [191] distinguishes between *immanent accents* which are assumed to be apparent from the score (structural, harmonic, melodic, metrical, dynamic, instrumental) and *performed accents* "added" to the score by the performer. This approach may be applied to nearly all extracted parameters, and in the general case it might not be possible to distinguish score-inherent and performer-induced characteristics.

The interpretation of the meaning of parameters derived from performance data is a difficult task. In the end, final conclusions can only be drawn by taking into account subjective judgments. The methodology and questionnaire or rating scale for such subjective tests and how they relate to performances, however, has only begun to evolve to systematic approaches during the last decade [408]. The problem of extracting relevant characteristics is apparent in the design of systems intended to automatically generate music performances from a score. Clarke notes (in the context of parameters possibly influencing performances): "Whatever the attitude and strategy of different performers to this wealth of influence, it is clear that a theory of performance which is presented as a set of rules relating structure to expression is too abstract and cerebral, and that the reality is far more practical, tangible and indeed messy" [396, p. 66].

10.2.1 Analysis Data

10.2.1.1 Data Acquisition

The acquisition of empirical data is one of the crucial points in systematic MPA. Among the various methods that have been proposed and used to acquire data, two general approaches can be identified: monitoring performances (or performance parameters) by mechanical or technical devices, or extracting the parameters from an audio recording of the performance. Both concepts have inherent advantages and disadvantages.

The monitoring approach usually provides accurate and detailed results since the measurement devices can track the performance parameters more or less directly, but the analysis is exclusively restricted to specific performances which were produced under special conditions and with the specific performers that were available.

The direct extraction of performance parameters from the audio signal — as opposed to from the instrument with sensors — is difficult and most definitely results in less accurate data. This is true for both the manual annotation of audio (such as marking onset times) and the fully automated extraction of data. Additionally, some parameters of interest may be even impossible to extract from the audio such as information on piano pedaling or note-off times. Other parameters of interest such as the performers' movements are obviously not extractable from the audio at all.

The advantage of extracting parameters directly from the audio signal is the possibility to analyze an enormous and continuously growing heritage of recordings, including outstanding and legendary performances recorded throughout the last century and until now. Hence, audio-based approaches allow to widen the empirical basis considerably with respect to the amount of available sources and their significance.

To extract the tempo curve from an audio recording, the usual approach is to either tap along with the performance [409, 410] or to manually annotate the onset times in a wave editor/display or a similar application [184, 411–419]. Both approaches have also been automated or partly automated by the use of automatic beat tracking systems (see Sect. 6.5) — followed by manual correction of beat times — [306, 420–423] or more recently by audio-to-score alignment algorithms using MIDI data as additional input [238, 269, 273] (compare Sect. 7.3.2). The main difference between tap-along and beat-tracking approaches as compared to manual onset time annotation and alignment systems is that in the former case the resulting tempo curve resolution is on the beat level, meaning that between-beat timing variations cannot be analyzed, while the latter usually takes into account each single onset time, whether this note lies on a beat or not.

Piano or Keyboard Performance

The introduction of mechanical pianos at the end of the 19th century made the acquisition of objective performance data possible through piano rolls. For example, Hartmann presented an early analysis of tempo and timing of two piano performances based on their piano rolls [424]. There are also later approaches to the analysis of performance data from piano rolls [425].

Other historic approaches used proprietary sensors that were built to extract performance data. The most prominent example is the *Iowa Piano Camera* that was used by Seashore and his team at the University of Iowa in the 1930s [45]. For each piano key, this "camera" recorded onset and note-off times and hammer velocity by optical means. Another example of a proprietary system is Shaffer's Bechstein grand piano using photo cells to detect hammer movements [187].

The introduction of the MIDI specification in the 1980s [3] resulted in an increasing number of electronic instruments and MIDI sequencers as well as compatible computer hardware and software solutions and opened up new possibilities to measure, store, and analyze pianists' performance data. Partly, music performance research has been done with the help of electronic instruments such as synthesizer keyboards and electronic pianos [426–428], but the majority concentrated on using acoustic instruments with built-in sensors which automatically output MIDI (or similar) data such as the Yamaha Disklavier product series or Bösendorfer grand pianos with the so-called *SE-System* [180, 332, 334, 429–440].

Other Instruments or Instrumentations

Most non-piano instruments represented in the literature on music performance are monophonic, meaning that two or more notes can never occur simultaneously. In this case, common approaches to fundamental frequency detection are robust enough to extract the variation of pitch over time (compare Sect. 5.3). Proprietary as well as commercially available systems have been applied to the task of pitch extraction for MPA [45, 115, 324, 441–447]. Seashore invented the "Tonoscope" for the pitch analysis of monophonic signals [4]. It consists of a rotating drum covered with a paper containing small dots, each representing a certain frequency. The input signal is — by the means of a light-emitting gas tube — projected on the rotating paper. If the input frequency matches one of the frequencies a dot represents, this line of dots will stand still for the observer and gives a clear indication of the frequency. The "Melograph" appears to be basically of a similar design [444]. Other studies work with spectrogram visualizations, use commercially available software solutions for the detection of monophonic pitches, or implemented their own software algorithms for the pitch detection.

The majority of these systems are not able to extract onset times with sufficient accuracy, so tempo and timing information is either not analyzed or is extracted by manual annotation. However, to name two counter-examples, Kendall compared timing and dynamics of monophonic melodies performed on piano, clarinet, oboe, violin, and trumpet [388], and Ramirez et al. used automatically extracted timing data for the identification of performers of violin recordings [447].

The tempo and timing data for other, non-monophonic signals has usually been extracted by tapping along [410] or by manually setting onset time labels [186], [448], [449]. Clynes did not analyze the tempo on a beat or onset level but measured the overall duration of single movements [450].

Lerch analyzed a set of string quartet performances of a movement of a late Beethoven quartet with respect to tempo, timing, and timbre by utilizing an automated system accompanied by manual correction [274].

10.2.1.2 Instrumentation

The majority of musical performance research focuses on the piano as the instrument of main interest. One of the obvious reasons is that the piano is a very common instrument with a large (solo) repertoire, but there are more reasons that make the piano an appealing choice. The tones produced by a piano have a percussive character that makes this instrument far more suitable for accurate timing analysis than, for instance, string instruments. Its mechanics make it possible to measure data with sensors less intrusive than on other instruments that offer a more direct interaction between performer and sound production. Furthermore, the pianist is in some ways more restricted than other instrumentalists; he is limited to fixed (and equally tempered) pitch frequencies which rules out expressive intonation and other performance specifics such as vibrato. He also has little influence on the timbre of a played note, and after hitting a key, he is not able to control any of the typical note parameters such as pitch, loudness, or timbre except its duration. From a technical point of view, these restrictions seem to make the piano a rather unattractive instrument with limited degrees of freedom, but even with these limitations, piano performances are an integral part of western cultural life, meaning that the mentioned restrictions do not really impede the communication of musical expression between pianist and audience. The reduction of possible parameter dimensions is, however, beneficial in performance research because it keeps the measurement data set smaller. Last but not least, the (commercial) availability of electronic and acoustic instruments using MIDI as a universal communication protocol simplified the performance data acquisition significantly since custom-built solutions were no longer necessary. While the recording of MIDI data from other non-keyboard instruments is at least partly possible, the fact that MIDI is a keyboard-focused protocol results in limited usefulness in many cases.

Despite the good reasons for the usage of piano as the main instrument for performance analysis, it has not yet been conclusively shown that the insights gained from piano performance analysis can be applied to performances with other instruments and ensembles (although the few studies done on other instruments indicate that this might at least partly be the case).

Other solo instruments include the singing voice [45, 115, 441, 445], string instruments such as violin, viola, and violoncello [45, 324, 388, 443, 444, 446, 447], wind instruments such as flute, clarinet, oboe, and trumpet [388, 442, 444], organ [448, 449], and percussion instruments [451].

There exist also some publications on chamber music performance [186, 274, 410, 450].

10.2.1.3 Variety and Significance of Input Data

With respect to the question if and how reliably conclusions can be drawn from the extracted data, it is important to verify how and from whose performance this data has been generated.

For example, it could be argued that performance data gathered under "laboratory conditions" is insignificant per se due to the unnatural recording environment; however, these special conditions are also given for many (studio) recording sessions which resulted in recordings that are in fact perceived as convincing performances by the listeners.

Still, when the data is acquired under such laboratory conditions, it implies that the number and possibly the skill of the available performers might be limited. For example, research had partly been done on student performances [180, 332, 413, 429–431, 433, 434, 439, 443]. This fact by itself is not too remarkable, but it nevertheless emphasizes the question if and how research methods and conclusions take into account the possible discrepancies between the performances of student pianists (or just *available* pianists) and the performances of professional (and *famous*) pianists. Under the assumption that fame is related to higher professional skills of the performer this could be a noteworthy criterion.

Due to the difficulties of acquiring large sets of performance data described above, the number of performers per study is usually small. The majority of research in the presented paper database has been done with a number of 5 or less performers per publication [186, 187, 194, 324, 388, 410, 411, 420, 423, 424, 427, 428, 433, 435–438, 440, 441, 445, 449, 451–454] or 6–10 performers [180, 421, 422, 429–432, 439, 455]. Some studies evaluate a larger number of 15–25 performers [184, 274, 412, 413, 434, 444] and an analysis of an outstanding number of 108 performers (115 performances) has been presented by Repp in the late 1990s [414–416].

This raises the question if and how insights gained from a small group of performers can be extrapolated to allow general assumptions on performances.

10.2.1.4 Extracted Parameters

The basic categories of information extractable from audio signals have been introduced earlier as temporal, tonal, intensity-related, and timbral (compare Sect. 1.1).

The variation of tempo and timing is one of the most thoroughly researched aspects in MPA. The analysis of the articulation is in most cases restricted to keyboard performances captured in MIDI format. Articulation is then simply interpreted as a measure of performed note overlap or note duration with respect to the note duration as given by the score.

In order to analyze the musical dynamics in a performance, the level or loudness is extracted using sound level or psycho-acoustically motivated loudness measurements as presented in Chap. 4. Strictly speaking, such measurements do not correspond directly to musical dynamics as these would depend on the musical context, on the instrument or instrumentation, and on the timbre. Nevertheless, intensity and loudness measurements seem to provide a reasonable approximation to dynamics [84, 456].

Pitch-related performance parameters such as vibrato and intonation can be analyzed by extracting the fundamental frequency variation from the audio signal. Due to technological restrictions of current analysis systems for polyphonic music, this usually has been limited to monophonic input signals.

The analysis of timbre deviations in performances is probably one of the least-researched parameters in MPA. This may be on the one hand due to the multi-dimensional nature of timbre (see Sect. 3.3), on the other hand because it is assumed to be of least importance and partly of high correlation with dynamics. One study on the timbre variation of string quartet performances led to inconclusive results [274].

10.2.2 Research Results

10.2.2.1 Performance

Many studies focus on a rather descriptive approach to performance analysis by just analyzing extracted data such as the tempo curve [45, 184, 187, 411, 412, 414, 424, 426, 457] or the loudness/energy curve [45, 415, 429] to identify attributes of the extracted parameters between different performances and performers.

The relation of musical structure (melodic, metric, rhythmic, harmonic, etc.) or the musical gestalt to tempo and loudness deviations has been intensely researched [187, 188, 387, 390, 410, 415, 424, 429, 440, 453, 458, 459]. Most authors agree on the close relationship between musical structure such as musical phrases or accents and performance deviations mainly in tempo and timing. In particular, larger tempo changes seem to be most common at phrase boundaries. There is a general tendency to apply *ritardandi* or note lengthening at the end of a phrase and moments of musical tension [184, 274, 412, 414, 426].

There are no conclusive results on the coupling of timing with dynamic patters [415, 429].

Desain et al. and Repp report on the influence of overall tempo on expressive timing strategies [427, 460]. They find that the concept of relational invariance cannot be simply applied to expressive timing at different tempi, a result similar to Windsor's, who analyzed tempo-dependent grace note timing [435]. The overall tempo might also influence overall loudness, an effect possibly linked to the increasing amplitude of pianists' vertical finger movements toward higher tempi [461].

Goebl investigated the relationship of the composer's tempo indications (*andante*, *allegro*, etc.) with the "real" tempo and was not able to separate different tempo classes sufficiently well with the tempo extracted from the performance [194]. The number of note events per minute, however, seemed to be easier to map to the tempo indications.

Studies on the timing of pedaling in piano performance indicate some relationship between pedal timing and overall tempo [428, 431].

In the context of keyboard instruments the articulation, or the amount of key (non-) overlap, has been studied [424, 426, 432, 433, 448, 449, 462]. In summary, key overlap times for *legato* articulation seem to decrease with increasing IOIs.

Studies of the accuracy of timing synchronization of two and more performers showed that performers are highly capable of synchronizing onset times even when modulating the tempo [186, 187]. Other publications deal with the timing synchronicity between both hands or between the melody and the accompaniment in piano music [187, 424]. In many cases of piano performance, a lead of the melody before accompanying voices can be observed [426], but whether this represents a performance concept or a consequence of the higher velocity of the melody tones is subject of discussion [180, 434].

The evaluation of the consistency of repeated performances of the same performers has shown their ability to reproduce a rendition quite exactly in terms of timing [45, 187], dynamics [429], and pedal timing [428]. This seems to be the case for performances spaced by several years as well [410, 415]).

Performance data from student and professional performances has been compared in [426] and [413]. While individual differences tended to be more pronounced among the professionals, both groups seemed to share the same general performance concepts.

Repp investigated the (statistical) relationships between the extracted performance data and sociocultural variables such as the artists' gender, nationality, year of birth, and recording date but, although the correlation was sometimes significant, pointed out that these results should be regarded with caution and that individual differences are likely to outweigh any sociocultural correlation [414, 415]. In a similar study with a smaller data set, Lerch found no significant relationships [274].

Walker showed that instrumental timbre may influence several performance parameters such as timing, articulation, and dynamics [442].

The analysis of vocal performances focuses frequently on the evaluation of vibrato rates and depth and the change or stability of pitch over time [45, 115, 443, 445] or other intonation characteristics of the performance [441, 444]. Fletcher analyzed the vibrato (and other acoustical features) of flute players [116].

Statistical and machine learning approaches have been tested to use the extracted tempo and loudness information for the purpose of classification, structuring the data, or extracting general rules from the data. Dovey tried to extract general as well as individual rules from two of Rachmaninov's piano roll recordings by using a logic programming approach [425]. Supervised learners can be used to assign representations of the extracted performance data to the corresponding artists with promising results [421, 422, 436, 454]. Other machine learning methods have been used to identify general performance rules [417–419, 437, 438] and to determine individual differences between artists [420].

10.2.2.2 Performer

While the publications listed above deal mainly with the analysis of the performance itself, the second area of MPA tries to determine the capabilities and goals of performers.

Repp analyzed the type of errors (i.e., pitch deviations from score) pianists make during a performance and checked if and how severe they were perceived by listeners, coming to the conclusion that the errors concentrated in less important parts of the score and thus were hard to recognize [430].

The relationship between the performers' intentions and the parameters extracted from performances has been studied in various ways. Palmer found good correspondence between notated intentions with respect to melody and phrasing and the extracted timing parameters [426]. Also, systematic relationships between the intended emotionality of the performance and the performance data (that is, representations of loudness and timing) can be detected [323–325, 454].

Other studies investigate the importance of the feedback of the music instrument to the performer [393]; there have been studies reporting on the effect of deprivation of auditory feedback [439, 455], investigated the performers' reaction to delayed or changed auditory feedback [463–465], or evaluated the role of tactile feedback in a piano performance [466].

Publications on the nature of memorization and learning of a musical piece (or its performance) tried to identify differences between novice and expert performers [467], to learn more on the nature of performance memory itself [468–470], and to find out more on the relation between a real and a virtual, imagined performance [439].

10.2.2.3 Recipient

It is the listener of a music performance who ultimately consumes, interprets, and probably judges it. Overall judgment ratings of performance data have been investigated in various studies. In an early publication, Repp reported some significant relations of ratings to measured timing patterns [412], while in a later study he had to conclude that "the aesthetic impression of the original recordings rested primarily on aspects other than those measured (such as texture, tone, or aspects of timing and dynamics (...))" [416]. Timmers did a similarity rating experiment and concluded that performances are judged in other ways than generally used to represent performance data [306]. In a different study, she let listeners rate the goodness of fit of two parts of different performance pairs [471]. Kendall investigated the communication of three levels of expressiveness: without expression, with appropriate expression, and with exaggerated expression [388]. Listeners were in many cases able to identify these three levels. Thompson et al. investigated the variation of listener ratings for a performance over time and found that the listening time to reach a decision was typically in the short range of 15–20 s [472]. Weinzierl and Maempel investigated how much of the listener's impression of the performance can be explained by common acoustical features [408].

Juslin detected relationships between moods and both tempo and loudness cues [323], and Kantor reported on associations of such cues and emotional reactivity [331]. Similar conclusions have been drawn from studying the time-variant emotional valence or the arousal and its relationship with performance data [332, 333]. Timmers found strong correlation between the dynamics and listener's judgments of emotionality and very good communication of emotional activity between performer and listener [334, 473]. In another study, she examined the influence of recording age and reproduction quality, observing that judgments of age and quality changed strongly with the recording date, in contrast to the perceived emotion which were mostly independent of the recording date; the communication of emotional valence tended to be more restrained for old recordings [474]. Husain varied the tempo and the mode (major, minor) of a performance and found indications that tempo modifications had an effect on arousal and mode modifications on mood [335]. Krumhansl evaluated the influences on timing and loudness variations on judgments of musical tension and found a close relationship of musical structure with both the listeners' musical tension rating and the performance data [458].

has been studied by Dixon, who found listeners to prefer smoothed beat sequences over the performed ones [423].

Lapidaki investigated the dependency of the initial tempo of a performance on the preferred tempo of a musical piece [475]; he found a general dependency although he also identified a group of listeners that were able to come to very consistent tempo preferences. Repp found systematic deviations between the tapping of listeners and metronomical time of music events, a result that seems to correspond well with the performers' inability to render a performance mechanically [476]. Aarden reported dependencies between tempo and "melodic expectancy" [477].