Timing in Piano Music
– a Model of Melody Lead

Johan Bjurling

KTH Computer Science and Communication

Master of Science Thesis
Stockholm, Sweden 2007
Timing in Piano Music
– a Model of Melody Lead

JOHAN BJURLING

Master’s Thesis in Music Acoustics (30 ECTS credits)
at the School of Media Technology
Royal Institute of Technology year 2007
Supervisor at CSC was Roberto Bresin
Examiner was Sten Ternström

TRITA-CSC-E 2007:115
ISRN-KTH/CSC/E--07/115--SE
ISSN-1653-5715
TIMING IN PIANO MUSIC – A MODEL OF MELODY LEAD

Abstract
In professional performances of piano music the melody voice tends to anticipate the accompaniment by 20-30 milliseconds. This may be due to dynamic differentiation. The melody voice is usually played louder than the accompaniment which may also cause the melody to lead, since the tone production in a grand piano is based on hammers being accelerated towards strings. A hard key strike results in a high hammer velocity whereas a soft key strike results in a low hammer velocity. Consequently, the hammers will arrive at the strings asynchronously if the keys are depressed with varying force. Presuming that the player strives at playing synchronously, this could easily be modeled and hopefully, provide more natural performances of piano music played by computers.

A model of this hypothesis, based on earlier studies, was implemented in pDM. The model was then tested in a listening test with experienced pianists. The findings indicate that the model simulates the typical asynchronous onsets found in piano music relatively well.
TIMING I PIANOMUSIK – EN MODELL AV MELODY LEAD

Sammanfattning

En modell av denna hypotes som baserar sig på tidigare studier implementerades i pDM. Modellen testades därefter i ett lyssnartest med erfarna pianister. Resultaten indikerar att modellen simulerar den för pianot typiska spridningen av ackordtoner relativt bra.
# Table of contents

1 Introduction ........................................................................................................................ - 1 -  
   1.1 Melody lead................................................................................................................. - 1 -  
   1.2 Outline of the report ................................................................................................. - 1 -  
2 Theory ................................................................................................................................ - 3 -  
   2.1 Tone production in the grand piano ............................................................................ - 3 -  
   2.2 Asynchrony in music................................................................................................... - 5 -  
   2.3 Perceptual aspects ....................................................................................................... - 6 -  
   2.4 The KTH rule system .................................................................................................. - 8 -  
   2.5 Summary ..................................................................................................................... - 9 -  
3 A melody lead model in pDM .......................................................................................... - 11 -  
   3.1 Development of the model ........................................................................................ - 11 -  
   3.2 Resulting performance .............................................................................................. - 15 -  
   3.3 Summary ................................................................................................................... - 21 -  
4 Listening test .................................................................................................................... - 23 -  
   4.1 Overview of the test .................................................................................................. - 23 -  
   4.2 Music examples ......................................................................................................... - 23 -  
   4.3 Participants ................................................................................................................ - 23 -  
   4.4 Procedure................................................................................................................... - 24 -  
   4.5 Results ....................................................................................................................... - 26 -  
   4.6 Discussion ................................................................................................................. - 33 -  
   4.7 Conclusions ............................................................................................................... - 35 -  
5 Discussion and conclusions.............................................................................................. - 37 -  
   5.1 Recommendations ..................................................................................................... - 37 -  
6 References ........................................................................................................................ - 39 -  
7 Appendix .......................................................................................................................... - 41 -  
   7.1 K-values for the performance rules ........................................................................... - 41 -
1 Introduction

1.1 Melody lead
In a grand piano the strings are excited by hammers which are in turn controlled by the pianist through the keys. When the pianist depresses a key, the impact is brought to the hammer through a mechanical arrangement and the hammer is accelerated towards the string. The velocity given to the hammer depends on the force imposed by the player’s finger. The harder the key is struck the higher the velocity and the larger the string amplitude and radiated sound level. The force also determines the duration from finger-key contact to hammer-string contact, typically between 20 and 200 milliseconds (Goebl, Bresin & Galembo, 2005). The harder the key is struck the sooner the hammer reaches the string.

Since skilled pianists strive to emphasize the melody note in a (homophonic) chord, that note (usually the highest) tends to be played 4 to 5 dB louder (Repp, 1996; Henderson, 1936) than the other notes within the same chord. Moreover, the melody usually precedes the rest of the chord by about 20-30 milliseconds. This phenomenon is called melody lead. The main tool for the player to draw attention to the melody is dynamic differentiation. However, since the melody is played louder it will also precede the other chord tones, presumed that all keys are struck simultaneously, since the velocities of the hammers are different. This makes the melody even more distinguishable in the musical context.

1.2 Outline of the report
The purpose of this thesis is to investigate if melody lead can be realistically modeled. In the next chapter the theoretical background to the thesis is covered. In chapter three the actual development of a model of melody lead is described. A listening test developed to evaluate this model is described in chapter four. In chapter five further studies are proposed and possible applications are discussed.
2 Theory

This chapter presents some theoretical concepts of importance for the thesis. It also serves as a summary of the literature that has been studied as starting point for the project. First the tone production in a grand piano is described. This is followed by a review of some earlier studies on asynchrony in music. In the next section, perceptual aspects of asynchrony in music are outlined. Finally, a brief overview of the KTH rule system for music performance is presented. It is included since a major part of the next chapter is dealing with pDM, which is a recent real-time implementation of the KTH rule system.

2.1 Tone production in the grand piano

The grand piano has remained more or less unchanged since the end of the nineteenth century. Even though innumerable studies have been made on the acoustics of the grand piano, its basic design is still the same it was a hundred years ago (Askenfelt & Jansson, 1990). This is at least partly due to reluctance among musicians, technicians and engineers (if you have played the piano for twenty years, you do not want your instrument to change).

2.1.1 The action of the grand piano

The mechanical function and timing properties of the grand piano have been described well by Askenfelt and Jansson in “From touch to string vibrations I: Timing in the grand piano action” (1990). The action of a grand piano is a delicate mechanical construction whose task is to transmit the force from the finger of the player to the hammer that eventually strikes the string. The main parts of the action are the key, the lever body, the hammer and the damper. The key is pivoted on approximately its midpoint. On the opposite side of the key (as seen from the player) the lever body is positioned (normally hidden by the chassis of the piano) upon the capstan screw. When the key is depressed, the lever body, with the hammer resting on it, is pushed into an upward motion and the damper is lifted from the string. Immediately before hammer-string impact, the hammer is let off from the jack and set into free flight towards the string. The hammer strikes the string setting it into vibration and then bounces back downwards. On its rebound it is being checked by the repetition lever making it possible to repeat the blow without restriking the key entirely (this is a mechanism only found in grand pianos, not in upright pianos). Figure 1 below displays the different parts of the piano action.

![Figure 1: The piano action (Askenfelt & Jansson, 1990).](image-url)
2.1.2 Temporal properties of the action

The timing of the tone production is largely affected by this mechanical setup. Striking the key hard, as in a forte attack, results in a short travel time (duration from the start of the key motion to hammer-string impact). A soft tone will require a much longer travel time. Furthermore, the key will reach its bottom position before hammer-string impact when depressed hard, but afterwards in a soft touch. These timing conditions can be adjusted by a piano-technician by setting the distances between the parts (e.g. let-off distance). A general goal for a piano technician is to make the temporal behavior of the action equal along the whole range of the keyboard (Goeb, Bresin & Galembo, 2005). This makes it necessary with different settings for different keys since the hammer masses vary along the keyboard with greater masses in the low compass and smaller masses higher up.

2.1.3 Pressed and struck touch

In playing the piano, two substantially different techniques are employed when striking the keys. Pressed touch (or legato touch) refers to playing with the finger resting on the surface of the key before the blow starts. Struck touch (or staccato touch) means the key is being hit from a distance above. To produce a very soft tone is only possible with a pressed touch while very loud tones only can be achieved with a struck touch (Goeb, Bresin & Galembo, 2005).

The temporal properties of force transmission from the finger, through the mechanical setup to the hammer, differ radically between pressed and struck touch. These properties were investigated by Goeb, Bresin & Galembo (2005) on grand pianos from three different manufacturers (Steinway, Yamaha, Bösendorfer).

Keystrokes were compared in pairs where the resulting intensities were nearly identical while the touch types were different. The key and hammer velocities were measured with accelerometers and the sound was recorded with a sound-level meter microphone. The results showed that in a pressed touch, the key and hammer are smoothly accelerated and the hammer reaches its maximum velocity immediately before string contact. In a struck touch the key is accelerated quite abruptly in the beginning. It then decelerates for a moment before a second acceleration phase gives the hammer its final speed. Even though the key acceleration starts with a sudden jerk when struck, the acceleration of the hammer maintains quite smooth but is delayed with several milliseconds. The overall process is 20-40 milliseconds shorter compared to a pressed touch, when the intensity is in the upper dynamic range. The measurements also indicated that the differences between pressed and struck touch are similar in different pianos.

One additional goal with their study was to extract functions describing the relationship between hammer velocity and travel time for the two types of touch. The study resulted in different curves for different touch conditions and different pianos. One of these functions was used in the melody lead model described in chapter four.

2.1.4 Tone quality

As mentioned above, the hammer is in free flight when it finally hits the string and thus one may argue that the type of touch should be of no importance for the tone quality of the produced tone (except in the extreme dynamic registers that only can be reached with the one or the other type of touch). Nevertheless, many pianists claim that the touch type affects the resulting timbre even at the same dynamic level. There is yet no consensus between pianists and experimenters on this, but one can of course expect that sound radiated from the action and from the finger bouncing on the key affects the pianist’s (and the listener’s) impression of the tone quality (Askenfelt & Jansson, 1991; Goeb, Bresin & Galembo, 2005).
2.2 Asynchrony in music

Notes indicated to be played simultaneously in a musical score are usually not played entirely synchronized in a performance. Deviations from the exactness in the score are included by the musicians both deliberately and unconsciously. Along with other expressive intentions such as changes in tempo, articulation, intensity, timbre et cetera, this is what makes music sound “alive” and beautiful. These are also the expressive means for the musician to communicate different emotions.

Rasch (1979) studied the timing in trio ensembles of various kinds and found that the melody instruments in string and woodwind ensembles tend to precede the other instruments slightly on the average, while no such tendencies were found in recorder ensembles (possibly because of the substantially different kind of music played by such ensembles). In all ensembles, the bass lay ahead of the middle voice on the average. The overall spread between instruments for nominally simultaneous notes varied between 30 and 50 milliseconds with larger spreads in the string ensemble (assumed by Rasch to be due to longer rise times in the amplitude envelopes for string tones).

2.2.1 Asynchrony in piano music

Special attention was given to asynchrony in piano music already in the 1930’s by Vernon (1936) and Henderson (1936).

Vernon (1936) studied the synchronization of chords in piano rolls recorded by world famous pianists. He found both melody leads and melody lags, as well as other types of asynchronous onsets such as upward and downwards rolls and classified these spreads as arisen by five different possible reasons. Chord spreads could be either accidental, due to tempo or tempo changes, brought in to facilitate perception, used on emphasized chords or lastly, applied to soften certain chords. These general explanatory principles were further subdivided within each category (e.g. a change of tempo could be an accelerando, a ritardando or an a tempo passage). These principles were then tested and support was found in the data for many of his assumptions. Especially, Vernon found greater spreads in slow movements than in fast.

When it comes to melody, Vernon concluded that it can be emphasized by playing it either early or late in a chord and that this strategy was employed to make the melody stand out in relief from the accompaniment (thus, falling into the category of spreads brought in to facilitate perception). When the melody occurred together with other chord tones in the right hand, it tended to lead. When the right hand played only melody notes, it tended to lag. Because of limitations in the material used in the study, Vernon could not investigate the dynamic properties of the recordings.

Henderson (1936) studied the performances of two pianists playing the choral section of Chopin’s “Nocturne No. 6 (G minor) Opus 15, No. 3”. Both pianists consistently played the melody 4 to 5 dB louder than the other voices and both showed a tendency to organize the intensity of the chords vertically. The soprano voice was dominating followed by the alto as the second loudest voice, the tenor as the third, and finally, the bass was the voice played with least intensity. In the right hand chords (soprano, alto, tenor) the melody tended to precede the other voices in almost all cases for both pianists. However, when the left hand bass notes were considered, these tended to lead the right hand for one pianist while they lagged the upper voices for the other.

---

1 Piano rolls were used in reproducing pianos in the early 20th century. By inserting a roll in such a piano the music could be played back automatically.
2 Bauer, Bachaus, Hofmann, Paderewski.
More recently melody lead in piano music has been studied by Palmer (1996), Repp (1996) and Goebl (2001). Two substantially different explanations of the melody lead phenomenon have been presented in these studies. Palmer, just like Vernon, concluded that melody lead is an expressive strategy, used by pianists, largely independent from other strategies such as tempo, articulation or dynamics. Repp (1996), on the other hand, presented the idea that melody lead is a consequence of the timing properties of the grand piano action, thus a velocity artifact correlated to dynamic differentiation. He studied ten graduate student pianists’ performances of three pieces and found correlations between dynamic differentiation and asynchrony within each hand. In his study, he also found different habits among different pianist regarding dependencies between hands. Five of his subjects had a tendency to lead with the left hand (bass). These subjects still led with the melody voice within the right hand however. The other five led all other voices with the right hand melody. This resembles the findings of Henderson described above. Repp’s data showed only weak correlations for a between-hand velocity artifact though.

Repp’s explanation was further tested by Goebl (2001) who studied asynchronies at the finger-key level by measuring tone onsets (hammer-string impact) and estimating the finger-key contact times through a timing correction function. Goebl found that finger-key asynchronies were reduced to about zero for the right hand while left hand notes tended to lead the right hand (just like some of the earlier findings). However, one pianist claimed to deliberately lead with the melody voice and his finger-key onsets indeed showed a melody lead of about 20 milliseconds. This deliberate lead still became substantially larger in the hammer-string domain.

### 2.3 Perceptual aspects

#### 2.3.1 Auditory scene analysis

The sounds that reach our ears in our daily lives emerge from many different sources. Still, we do not hear the sum of all these sounds as one single event. Instead the brain sorts them so that we can separate for instance a voice from a violin playing at the same time. Auditory scene analysis (ASA) (Bregman, 1990) is the psychoacoustical term for this process. As the term implies the brain decodes the scenery of sounds that is encoded in the signal that reaches the ears.

To be able to segregate different sound sources the brain utilizes grouping principles. These can be ordered in sequential principles working in the time domain, and spectral principles working in the frequency domain. The spectral analysis makes it possible for us to separate sounds of different timbres for instance. The sequential analysis is a fundamental prerequisite for speech recognition.

#### 2.3.2 Auditory stream segregation

One general principle for ASA is auditory stream segregation. The brain seems to organize sound events into separate auditory streams, such as a person talking or a car passing by. One side effect of this ability is something referred to as “the streaming effect”. When a rapid sequence of high and low tones are presented to a listener, this may be perceived as two separate sequences of tones presented simultaneously if the difference in pitch, timbre or spatial location is large enough. In music, this effect can be utilized by a composer. By separating one melodic line into different registers, it may be perceived as two separate lines. This is sometimes referred to as “virtual counterpoint” and an example of this is the “Alberti bass” commonly found in classical music.

---

1 Schumann’s “Träumerei”, Debussy’s “La fille aux chevaux de lin” and Chopin’s Prelude in D-flat major.
The perception of melody is related to auditory stream segregation since it is much easier to perceive if it falls into one prominent auditory stream. Thus, composers rarely let the melody wander around stochastically among the instruments of an ensemble, but rather devote it to one instrument (or instrumental group) of a timbre separate from the accompaniment. This is however not possible in piano music, where both melody and accompaniment is played by the same instrument. Thus, the spectral cues are quite vague since all tones are produced in a similar way (hammers striking strings) and there is a risk that the melody will not stand out in the musical context. Especially in simultaneous chords, the melody must be emphasized in some way to avoid being absorbed by the underlying voices. This makes temporal segregation important in the separation of different voices.

Another aspect of auditory scene analysis and melody (although of lesser importance for this study) is the melodic quality of a line. Coherence and structure is important for us to perceive a series of tones as a melody. Just as a composer can trick the ear and deliberately produce separate lines out of one, erroneous segregation (not intended by the composer or performer) can occur if the structural organization of notes is defective.

2.3.3 Masking
A perceptual aspect of importance for the study of melody lead is masking. The term denotes the perceptual cancellation or disruption of one sound by another. Masking can be both simultaneous and non simultaneous. A simultaneous noise added to a speech signal will make the speech unintelligible or imperceptible if the noise is loud enough.

Non simultaneous masking can be divided into forward and backward masking. Forward masking means the masker precedes the target sound making it imperceptible. Backward masking is the opposite, a sound heard before the masker is cancelled. Below, an experiment on masking in music is described.

2.3.4 Perceptual threshold of pitch jumps when masked by a drone
Rasch (1978) studied how asynchrony affects the perception of simultaneous notes in music. A test tone and a masker (a lower “drone” tone) were used as stimuli and presented to musical students. The task was to say whether the pitch of test tone was jumping up or down and a 75% score was considered to be the threshold for separate perception. The results showed that there was a much better detection of the target tone when the masker was slightly delayed (10-30 milliseconds). This way the masking effect was limited and the threshold for perceiving the target tone decreased with about 10 dB for each shift of 10 milliseconds (hence, the backward masking was diminished with larger shifts).

Another finding of the study was that simultaneous notes seemed to combine into one “sound object” when the harmonics of the target tone and masker coincided. Also this effect was largely reduced by a temporal shift. Furthermore, Rasch mentioned that unequal onsets are very difficult to hear even when the difference is as large as 30 milliseconds. Thus, the tones are perceived as simultaneous even though their onsets are temporally shifted. This fact is probably dependent on the amplitude envelopes of the individual tones though.

2.3.5 Perceived loudness affected by asynchrony?
Goebl & Parnicutt (2003) investigated the influence of asynchrony and intensity on the perceived loudness of a tone in a chord. They tested both single chords and sequences of chords in different intensity and asynchrony combinations. Their conclusion was that intensity is the main

---

4 Since there were only two alternatives, a 50% score would be reached by just guessing.
cue for the perceived loudness. The overall effect of asynchrony on loudness was quite small. However, when the target tone was soft compared to the others, anticipation helped it avoid being masked, resembling Rasch’s findings.

In the same study, an experiment was performed on a real music excerpt, Chopin’s “Ballade op. 38”. Again, intensity was the dominating cue. Small effects of asynchrony could be observed. Also a delay could increase the relative salience of the target voice when it was already louder.

In addition, in the above mentioned study by Vernon (1936), the limen for perceiving asynchronous chords was tested. He found that there was a significant difference in thresholds for musical versus non musical subjects. Contrary to Rasch’s more recent statement, he concluded that practicing musicians can perceive temporal shifts of about 20 ms or even lower. His test was made on piano tones, as opposed to Rasch’s that used computer generated tones and the difference in timbre and amplitude envelop may explain the different results.

2.4 The KTH rule system

2.4.1 Introduction

The KTH rule system is a system for music performance modeling developed at the department of Speech, Music and Hearing, KTH, Stockholm. The project was initiated in the early 1980's by Johan Sundberg and Lars Frydén. Since then the system has developed and been revised several times (Friberg, Bresin & Sundberg, 2006).

The system is based on different rules that take a musical score as input and manipulate the events defined in the score (tempo, note durations, articulations dynamics for instance). The output is an artificial performance of the music. Rules are defined for phrasing, timing, metrical patterns, articulation, tonal tension and intonation. The magnitude of each rule can be adjusted by the user to form different kinds of performances. Different amounts of the rules are necessary for different styles (e.g. Baroque, Jazz, and Romantic).

2.4.2 DM

The major platform for the development of the KTH rule system is the software Director Musices (DM) (Friberg, Bresin & Sundberg, 2006). In DM, a MIDI-file or a .mus-score\(^5\) can be opened and the performance rules can be adjusted, applied to the score and played back. The entire rule configuration can be saved as a rule palette\(^6\) and used on other scores as well. Thus, predefined rule configurations can be made for different styles or genres. One drawback with DM is however that the rules can not be changed when the music is being played back. Furthermore, to change the overall expression and go from a say, happy performance to a sad one, the user needs to change many rules which becomes quite awkward after a while. Both these problems are solved in pDM as described below.

2.4.3 pDM

pDM is a real time expressive sequencer developed by Anders Friberg and described in “pDM: An Expressive Sequencer with Real-Time Control of the KTH Music-Performance Rules” (Friberg, 2006). It is built in the graphical programming environment Pure Data (PD)\(^7\). In pDM a special file format (.pdm), which is a representation of a musical score with values for all performance rules embedded in the file, can be opened. It is preprocessed in DM and thus, one

---

\(^5\) .mus is the internal file format for saving musical scores in DM.

\(^6\) Rule palettes are saved in another file format (.pal).

\(^7\) http://www.puredata.org
has to open a score in DM first and save it as a pDM-file there. In the pDM-file, each note event is preceded by parameters for sound level, duration and tempo (given by the rules in DM). By multiplying these parameters with adjustable k-values in pDM the effect of each rule can be varied (amplified, set to zero, inverted). Since PD is a programming environment for real time purposes this can be done while the music is being played back. Consequently, the user can “conduct” the music by changing the k-values.

To make the steering of the playback more intuitive the k-values can be controlled indirectly through an interface. It is made up by a two-dimensional area where the user can point with the mouse and drag between different emotions (e.g. happy, sad). The input is then transformed to k-values corresponding to the desired emotional character. The four corners of the area represent four different emotions and k-values are defined for each emotion. Through linear interpolation of these values different shades of emotions can be achieved by dragging the mouse around.

2.5 Summary
Melody lead is not a phenomenon found exclusively in piano music, but rather a universal musical feature. The lead instrument is usually the “conducting” instrument in smaller ensembles and the others seem to follow, which may cause them to lag behind. In piano music, there is usually just one musician. Still the melody tends to lead the accompaniment. Because of the uniform tone production in pianos, these asynchronous onsets may aid the listener in detecting the melody.

The hypothesis suggested by Repp and Goebl, that melody lead occurs as a consequence of the mechanics of the grand piano along with dynamic differentiation, will be assumed to be valid. Even though melody lead may be employed as an expressive strategy, as suggested by Palmer (1989) and also partly found by Goebl (2001), the undoubted correlation between hammer velocity and travel time makes it a good starting point for modeling melody lead.
3 A melody lead model in pDM

3.1 Development of the model

In order to develop a model of melody lead in pDM described earlier, a method had to be found to make asynchronous onsets possible, since this was not yet implemented in the program. A thorough investigation of the program and its file-format preceded the actual modeling. pDM will be described in more detail in the following section.

3.1.1 The pDM file format

In the pDM file format, all note events are preceded by variables for sound level, articulation (duration), and tempo. These variables specify deviations from the nominal score and are therefore denoted delta sound level (DSL), delta articulation (DART) and delta tempo (DT). When these are read from the file they are all multiplied with k-values for the corresponding rules in pDM and added to form total delta values for sound level, articulation and tempo. The total delta sound level value is added to the nominal value for sound level and this defines the sound level of the next tone being played. The same applies for the duration and tempo (Friberg, 2006). An excerpt from a pDM file is shown in Figure 2 below.

```plaintext
923 BAR 2;
0 DT -0.04 -0.02 0 0.03 0.01 0.02 0.05 0 0.01 0.01 0.02 -0.01 0 0;
0 DSL -0.6 -0.6 0.1 0.6 0 0 0.2 -1.5 -0.5 -1.5 0.3;
0 DART -60 0 0 0 -462;
0 NOTE 65 1 0 2308;
0 DSL 0.4 0.7 0.4 0.7 0.1 0.3 0.2 -2.0 -0.2 -1.5 0.2;
0 DART -60 0 0 0 -923;
0 NOTE 41 5 0 4615;
```

Figure 2 An excerpt from a pDM file.

3.1.2 The sequencing of the pDM file

In pDM, a pDM file is opened through a simple user interface and then stored in an object called qlist. This object takes care of the sequencing of the file which is initiated when the user presses the play button. By passing a tempo variable to qlist the relative tempo of the sequencing can be changed. Tempo variables (total delta tempo) are constantly sent to qlist to account for expressive changes in tempo indicated in the file and manipulated by the user. The tempo variable is recalculated whenever a new note event is read from the pDM file. However, simultaneous note events are read simultaneously from the file and they are all affected similarly by the global tempo. Thus, asynchronous note onsets could not be modeled through the tempo variable sent to the sequencer object since asynchronous onsets require temporal control of every single note. Instead the solution was to add a delay object, in pDM called pipe, to the chain of processing so that all notes are assigned a specific delay, based on their sound level properties, before they are passed to the MIDI object and the sound board. Since the hypothesis assumes that a pianist will strive to play chords synchronously and that melody lead is an artifact caused by different hammer velocities and travel times, the most obvious solution was to estimate the travel time for each note, based on its sound level, and assign this value as the delay value.

3.1.3 Estimation of travel times

The travel time curves for the Bösendorfer SE290 computer-controlled piano presented by Goebl, Bresin and Galembo (2005) were chosen as functions for the estimation of the travel
times. These curves describe the relationship between the final hammer velocity (FHV), which is the velocity of the hammer just before its contact with the string, and travel time for the two types of touch (pressed and struck). The function for pressed touch was set to default and the program was made to switch to the function for struck touch whenever the property of staccato articulation was defined in the file (that is, when staccato articulation is indicated in the score by the composer). The temporal behavior between different pianos differs only marginally so it was assumed that the functions for the Bösendorfer could be used generally. To extract the travel times, the FHVs needed to be estimated in some way.

![Figure 3](image)

**Figure 3** Travel time curves for a Bösendorfer SE290 grand piano (Goebl, Bresin and Galembo, 2005).

### 3.1.4 From decibel to final hammer velocity

FHV depends on pitch and sound level. Equal FHV at different locations along the keyboard produces different sound levels. Data earlier collected from a Bösendorfer SE290 computer monitored grand piano (Goebl & Bresin, 2003) was used to extract FHV as a function of sound level for five selected keys, C1, G2, C4, C5, and G6. The five functions, one for each key, were estimated in a data file in Excel and implemented in pDM. Since the FHV-function for each key was only defined between the lowest and the highest sound level value in the data file, all sound
level variables were limited to lie within the specified range before they were passed to the FHV-function. To assign values to keys between the five sample keys, linear interpolation was used. Figure 4 displays the five functions. As can be seen in the diagram, the differences between the different keys become more accentuated in the upper dynamic range. The most effective key was C5 which reached a sound level of 109 dB with a FHV of 6.4 m/s. Generally, higher pitches produced louder sound level at equal hammer velocity.

**Velocity curves for Bösendorfer SE290**

![Velocity curves for Bösendorfer SE290](image)

**Figure 4** Final hammer velocity (FHV) as function of sound level for five different keys along the keyboard of a Bösendorfer SE290 grand piano.

The resulting performance of all keys and for different sound levels is given below. One can clearly see that the necessary FHV for a certain sound level decreases with higher pitches. The constant FHV for each sound level value below key 24 (C1) is a consequence of the fact that key 24 was the lowest measured key. Hence, it was impossible to interpolate values below that. The inconsistencies in the upper dynamic range (100 dB to 110 dB) are due to the dynamic limits of each key function. In the lower key range this results in all sound levels above 100 dB receiving the same FHV. In the upper key range the result is deflecting curves for the two highest sound levels.
Figure 5 Performance of all keys as reproduced by the model.

One problem in going from dB via FHV to travel time was the fact that in pDM, sound level values are defined as deviations from 0 dB, where 0 dB lies roughly in the middle of the dynamic range, corresponding to a MIDI velocity of 64 MIDI units. The dynamic range of the Bösendorfer SE290 ranged from 60 dB (after some outliers had been removed) to 110 dB in the data file, thus 85 dB was taken as mean sound level. In pDM the value 85 was added to the total delta sound level and the sum was used as argument for the functions translating sound level to FHV. This means for a loud tone an argument larger than 85 dB was used and vice versa for a soft tone.

The FHV, calculated for each note, was then used as argument for the travel time function to get the travel time for each note. Finally, the note was sent to the delay object (pipe) together with its travel time value.

3.1.5 Delayed offset
To avoid tones becoming too short the offset of each tone was delayed with the same delay value as the onset. This means also the offsets were asynchronous. These asynchronies were however assumed to be imperceptible.

3.1.6 Mathematical description
The model previously described can be summarized with a mathematical formula. The final hammer velocity is computed from sound level, $d$, through polynomial functions for different fixed keys $k_v$ (the coefficients $a_0$ to $a_6$ for each function are calculated from the data-file from the Bösendorfer SE290).

$$f_{h_{k_v}}(d) = a_{6k_v}d^6 + a_{5k_v}d^5 + a_{4k_v}d^4 + a_{3k_v}d^3 + a_{2k_v}d^2 + a_{1k_v}d + a_{0k_v}$$

*MIDI units span from 0 to 127.*
\[ K = [24, 43, 60, 72, 91] \]

To get the final hammer velocity for a certain key \( x \), the functions for the two enclosing keys are interpolated.

\[
f_{h_{x}}(d) = \left( (x - k_{n+1})/(k_{n} - k_{n+1}) \right) f_{h_{k_{n}}}(d) + \left( (x - k_{n})/(k_{n+1} - k_{n}) \right) f_{h_{k_{n+1}}}(d)
\]

\[ k_{n} < x \leq k_{n+1} \]

The result is then used as argument for one out of two functions, one for struck and one for pressed touch, computing travel time.

\[ tt_{st} = 58.39 f_{h_{x}}(d)^{-0.7377} \]

\[ tt_{pr} = 89.96 f_{h_{x}}(d)^{-0.595} \]

### 3.2 Resulting performance

#### 3.2.1 Relative delays in Schumann’s “Träumerei”

A pDM file of Schumann’s “Träumerei” was played back in pDM with a neutral rule setup (see appendix) and the model applied (and the melody set to be 5 dB louder than the other voices). All the travel times were recorded to a text file. By subtracting the travel times of the melody notes from the travel times of all other simultaneous notes the relative delays could be plotted. The result is displayed in Figure 6. The zero line represents the melody and the symbols represent the delay times (for simultaneous notes in the score) for the other four voices. Each voice was represented by a specific MIDI-channel ranging from 1, for the melody, to 5, for the bass. All voices lag the melody by 5 to 25 milliseconds except the bass which sometimes lags and sometimes anticipates the melody by 0 to 10 milliseconds. This tendency for the bass to lead is explained through the FHV-curves displayed earlier. The lower the pitch, the higher the FHV must be to produce a constant sound level. Moreover, higher FHV results in shorter travel time. Because of the distance in pitch between melody and bass the travel times for the bass in many cases became slightly shorter than those for the melody even though the melody was played 5 dB louder.
Figure 6 Delay times in Schumann’s “Träumerei” played with a neutral performance, the model applied and the melody 5dB louder than the accompaniment. The zero line corresponds to the melody. The symbols represent notes that are simultaneous with the melody in the score. Each MIDI-channel corresponds to a separate voice in the score.

The result resembles to a large extent Repp’s results for the five pianists playing “Träumerei” with bass leads. Figure 7 displays the average asynchrony profile for the first eight measures of the piece as played by the pianist’s that tended to lead with the left hand.

Figure 7 Average asynchrony profiles in the first eight measures of Schumann’s “Träumerei” as played by pianists that tend to lead with the bass (Repp, 1996). Vertical lines represent bars in the score.

Figure 8 displays approximately the same part of the piece as played by pDM with the model applied.
Delays in Schumann's Träumerei
(First 8 measures)

Figure 8 Delay times in the first eight measures of Schumann's “Träumerei” as played by the pDM model with the same settings as above.

“Träumerei” was also played back with a nominal performance and without dynamic differentiation. All voices were thus played with exactly the same sound level all the time. The model was still applied to the performance. The result is presented in Figure 9 below which depicts the effect of the varying FHV needed for different keys. Clearly, the melody lags all other voices (negative delays) by 0 to 25 milliseconds. This is explained by the FHV-functions (Figure 4) and the corresponding key properties (Figure 5).
3.2.2 Inverted melody lead

Of interest was also the result of an inversion of the model. As mentioned earlier, in pDM 0 dB is the default sound level and deviations from this are calculated to form a total delta sound level. 85 dB was then added to the total delta sound level and the sum was used as argument for the FHV-function. By multiplying the variable for total delta sound level with -1 (before adding 85 and sending the sum as argument to the FHV-function) the inverted effect of melody lead was achieved. This means instead of adding a couple of decibels to 85 dB for a hard key strike, they were subtracted. Hence, the higher the sound level for a note, the longer the travel time and vice versa. Once again “Träumerei” was played back with a neutral performance and the melody 5 dB louder than the accompaniment. The resulting delays are plotted against the melody in Figure 10 and it can be seen that the melody lags all other voices by 10 to 50 milliseconds. Once again this is explained by the key properties but here the differences are exaggerated by the inverted dynamic differentiation.

Figure 9 Delay times in Schumann’s “Träumerei” played with a nominal performance, the model applied and no dynamic differentiation between melody and accompaniment.
Inverted delays in Schumann’s Träumerei

Figure 10 Delay times in Schumann’s “Träumerei” when played back in pDM with a neutral performance, the melody 5 dB louder than the accompaniment and the model inverted.

The inter onset intervals $^9$ (IOI) of the melody voice were extracted for two nominal performances of “Träumerei”, one with the model applied and one without. The differences in IOI were then computed and plotted. The result is given in Figure 11 below. What can be seen is that the model imposes a random noise to the performance that represents the variability of the mechanical action.

---

$^9$ IOI is defined as the time duration from one note onset to the onset of the immediately following note.
IOI Difference between nominal performances of “Träumerei” with and without melody lead
(Only the melody voice considered)

Figure 11 Inter onset interval difference of the melody between two nominal performances of “Träumerei”, one with and one without the model applied.

### 3.2.3 Comparison with another model

This way of calculating the travel times was compared to a preliminary method proposed by Goebl. His solution is presented in the equation below.

\[
ml = \left( \left( \frac{fhv_n}{a} \right)^{1/b} \right) - \left( \left( \frac{fhv_1}{a} \right)^{1/b} \right)
\]

Where \( ml \) is melody lead in milliseconds, \( fhv_n \) is the final hammer velocity of the melody, \( fhv_1 \) the final hammer velocity of the accompaniment, \( a \), a constant of 2444.815927172549 and \( b \), a constant of -1.7363341578104656. The final hammer velocities were still computed with the FHV-functions in Figure 4 but instead of using the function shown in Figure 3 for travel time, the function showed below (extracted from Goebl’s suggested model) was used.

\[
n = \left( \frac{fhv}{a} \right)^{1/b}
\]

Figure 12 displays the delay values for the channel three (middle) voice in “Träumerei” when played back with the two different models. Filled triangles represent the present model (described under 4.1.3) and unfilled squares represent Goebl’s (previously described). As can be seen, the two methods produce almost identical results and the delays fall between 10 and 25 milliseconds.
Figure 12 Comparison of delays for channel 3 in Schumann's “Träumerei” between the present model of melody lead and a preliminary model proposed by Goebl.

3.3 Summary

A model of melody lead was implemented in pDM by including a delay object. Delay times were calculated in two steps. First the final hammer velocity for each note event was calculated. This was done by implementing FHV-functions for five keys and interpolating the values for pitches in between. The sound level and pitch of each note resulted in a FHV that was used in the second step to calculate the travel time. The travel time function for pressed touch, extracted by Goebl, Bresin and Galembo (2005), was used. Each note was sent to the delay object together with its corresponding travel time. This way the note events in the score basically represented the fingers of the pianist playing and the model simulated the temporal properties of the grand piano action.

The model was tested by playing back a score of Schumann’s “Träumerei” and the resulting delays relative to the melody were plotted. All voices but the bass lagged the melody by 5 to 25 milliseconds. The bass both lagged and anticipated the melody by 0 to 10 milliseconds.

The inverted effect of melody lead was achieved by multiplying the total delta sound level with -1. The resulting delays showed that this made the melody lag all other voices by 10 to 50 milliseconds.

Finally, the implementation was compared to another (preliminary) model proposed by Goebl. The resulting delays were very similar.
4 Listening test

In this chapter a listening test, that was developed in order to evaluate the model explained in the previous chapter, is described. The idea was to let highly trained pianists adjust the synchronization and tempo in different pieces of piano music.

4.1 Overview of the test

The test was developed in PD and consisted two parts of 5 examples each. By clicking buttons on a computer screen the participants were guided through the test. The music was played back through monitors.

Instructions were given to the participants on the screen. Before the test the instructions were also given orally to make sure the participants understood the task. They were able to ask questions if something was unclear. After each example the participants were instructed to save their settings by clicking a button. When they did, the values of the sliders were recorded to a text file.

An amateur pianist was asked to try out the test to see if something was difficult to understand and to eliminate possible problems. Some adjustments were made after this pilot test. The number of examples was reduced from 15 to 10 to speed up the test. When it later was performed with real test subjects it took about 30 minutes.

4.2 Music examples

Five different pieces of music were prepared: “Träumerei” and “Der Dichter Spricht” by Schumann, “Etude op. 10/3” by Chopin, “KV1” by Mozart and finally “Emigrantvisa”, a Swedish folksong arranged by Jan Johansson. One of these pieces (“KV1”) was already available as a .mus-file along with the DM program. A preprocessed file for “Träumerei” was given to the author by Anders Friberg. For the others, new files were created and preprocessed in DM.

In pDM, k-values for all the implemented performance rules were found that resulted in a relatively neutral performance of each piece. These values are listed in the appendix. In all examples, the melody was played back 5 dB stronger than the accompaniment to make the dynamic differentiation salient.

When the tests were performed the author had not yet found a way of implementing travel time functions for both pressed and struck touch (and switch between them as described in the previous chapter) in the program. Thus, the files were all played back with only the function for pressed touch implemented. This does not affect the resulting performance at any significance however, since all the chosen examples are to be played legato throughout (except for a short passage in the “Etude” and some chords in “Der Dichter spricht”).

4.3 Participants

All of the participants in the test had a long experience with piano music. Two were teachers at the Royal College of Music in Stockholm. Two students from the same school participated. One participant was a keyboard curator from the Stockholm Music Museum. A concerto pianist from Italy participated as well as an amateur pianist of unusually high standards. These were selected since they were expected to react more sensitively to subtle differences in piano music. All participants were male except for the Italian pianist. Their average age was 42. All participants were rewarded with a cinema ticket for taking part in the test.
4.4 Procedure

4.4.1 Part one – continuous synchronization lever

In the first part of the test the participants listened to the five pieces of music. For each example they were instructed to adjust the tempo and the synchronization to make it sound “as realistic as possible”. These exact words were chosen to prevent the participants from making erroneous associations. Another conceivable instruction could have been “as good as possible” but then maybe then the participants, in order to make it sound “perfect”, would strive at perfect synchronization (deadpan).

The musical examples were presented in random order. To control the two parameters, the participants used two physical levers on an input device. Visual feedback was also presented on the screen when the levers were moved. The tempo lever went from 0.5, which corresponds to 50 percent of the tempo specified in the pDM file, to 2, which corresponds to 200 percent of the specified tempo.

The lever for synchronization was a bit more complex. It ranged from -2 to 2. The sign of its value corresponded to normal (+) or inverted (-) melody lead. The absolute value of the synchronization lever corresponded to the amount of melody lead (or inverted melody lead) being added (implemented as percent of travel time). The user could therefore exaggerate the effect of both melody lead and inverted melody lead by up to 200 percent. Whenever the lever was set to its midpoint value 0, perfect synchronization was the result since the amount of melody lead, in percent of travel time, was 0 (all travel times were multiplied by 0).

To force the participants to really use their ears, the synchronization lever was randomly mirrored so that sometimes the upper value was -2 and the bottom value was 2 and sometimes vice versa. This was considered important since the function of the physical slider made each example start with the levers set the same way the preceding example ended. The participants did not know how the lever was configured however.
4.4.2 Part two – forced choice

In the second part of the test the same five pieces were played once again. This time the subjects once again adjusted the tempo with a lever but chose the synchronization with the mouse from three alternatives on the screen. Hence, it was a forced choice (ipsative measure). These alternatives were: deadpan (corresponding to perfect synchronization), 100 percent melody lead and 100 percent inverted melody lead. The musical examples were presented in random order and the three versions were randomly distributed among the alternatives. Once again the subjects were instructed to select a preferable tempo with the lever and choose the alternative for synchronization that they perceived as most realistic.
4.5 Results

4.5.1 Results of the first part of the test

All synchronization values from the first part of the test are plotted in Figure 15 for the five examples with different symbols for the seven participants. The examples are ordered with increasing average values (displayed in Figure 16). The highest value for synchronization was obtained for the Etude where the maximum value of 2 was reached. The lowest value was obtained for “Der Dichter spricht”, -0.91.
Synchronization
(First part)

Figure 15 All ratings for synchronization from the first part of the test. The examples are ordered with increasing average values.

In Figure 16 the average synchronization values and standard deviations for the five examples in the first part of the test are shown. The file receiving the lowest average value was Schumann’s “Der Dichter spricht”, with an average of 0.01. The highest average value for synchronization was obtained for “Emigrantvisa”, 0.49. None of the examples received a negative average value.

Average synchronization
(First part)

Figure 16 Average values with standard deviation bars for synchronization in the first part of the test. The examples are ordered with increasing average values.
Figure 17 All ratings for tempo from the first part of the test.

Figure 17 displays all tempo selections for the five examples and for the seven participants and Figure 18 displays the average values with standard deviation bars. Two outliers can be spotted in example three and four. It can be observed that the specified tempo for “Der Dichter spricht”, “KV1” and “Emigrantvisa” seemed to be suitable. The tempo for “Träumerei” was a bit too slow and the tempo for “Etude 10/3” a bit too fast in the pDM files. The average tempo for all examples and all participants in the first part of the test was 1.00.

Figure 18 Average values with standard deviation bars for tempo in the first part of the test.
In Figure 19 below, all synchronization values from the first part of the test are plotted against corresponding tempo. As can be seen from the trend line, there is a tendency for smaller synchronization values for higher tempos. The reader is reminded that a synchronization value close to zero means close to perfect synchronization (regarded as deadpan in the test).

The positive and negative values for synchronization are plotted separately against corresponding tempo in the diagram below and separate trend lines are included. The negative tilt of the trend line for positive synchronization values (normal melody lead) becomes steeper while the tilt of the trend line for negative synchronization (inverted melody lead) becomes flatter.
Figure 20 Positive and negative synchronization values plotted separately.

The outliner on the positive side in Figure 20 (rightmost value) was generated by p1 for “Träumerei”. Studying the tempo diagram in Figure 17 it can be seen that this value deviates quite much from the others. This value was therefore discarded and a correlation test was then made on the positive values for synchronization. A significant correlation was found.

<table>
<thead>
<tr>
<th>sync Pearson Correlation</th>
<th>sync</th>
<th>tempo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.760(**)</td>
<td>0.000</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>tempo Pearson Correlation</td>
<td>-0.760(**)</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

A correlation test was also made on the negative values for synchronization but no significant correlation was found.

<table>
<thead>
<tr>
<th>sync Pearson Correlation</th>
<th>sync</th>
<th>tempo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.136</td>
<td>0.628</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>tempo Pearson Correlation</td>
<td>-0.136</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>
4.5.2 Results of the second part of the test

In the second part of the test the alternatives for synchronization were discrete values: -1, 0 and +1, corresponding to inverted melody lead, deadpan and normal melody lead. Most of the times (20 out of 35) the participants selected the deadpan condition 0. Six times -1 was selected and nine times +1 was selected. The result is presented below in Figure 21. The examples are ordered with increasing number of positive ratings. “Der Dichter spricht” once again received the lowest overall rating. “KV1” on the other hand was the highest rated example with three positive ratings, three deadpans and one negative rating.

The participants still adjusted the tempo with the tempo lever and these values are therefore continuous. The average values for tempo assimilate those from the first part of the test very well as can be seen in Figure 22.
Figure 22  Average tempos with standard deviation bars for the second part of the test. The empty symbols represent the average values from the first part of the test.

In Figure 23 below the tempo values are plotted against corresponding synchronization for the second part of the test. Here the tilt of the trend line is positive with increasing tempo for increasing synchronization values, contrary to the result in first part. Also the spread of tempos increases with increasing synchronization values. The average tempo for all examples in the second part of the test and all participants was 1.02.

Figure 23  All synchronization values from the second part of the test against corresponding tempo.
A correlation test was made on the data from the second part of the test but no significant correlation was found as seen in the table below.

<table>
<thead>
<tr>
<th></th>
<th>sync</th>
<th>tempo</th>
</tr>
</thead>
<tbody>
<tr>
<td>sync</td>
<td>1,280</td>
<td>.104</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>tempo</td>
<td>.280</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.104</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
</tr>
</tbody>
</table>

### 4.6 Discussion

#### 4.6.1 Average synchronization

As seen in Figure 15 all examples in the first part of the test received a positive average value for synchronization, which means, on average the participants preferred positive synchronization values corresponding to normal melody lead. Negative synchronization never reached below -1 whereas positive synchronization reached above +1 for 6 cases. For the “Etude”, the maximum value of 2 for synchronization was reached with the tempo set to the minimum value of 50 percent of the tempo specified in the pDM file.

The file receiving the lowest average value in the first part of the test (close to zero) was “Der Dichter spricht”. One reason for this may be that the composition contains quite a long midsection of arpeggiated\(^\text{10}\) chords and long solo lines for the melody. In this part of the score it is very difficult to judge the synchronization because of the contrapuntal texture. If the participants moved the lever around during that section they could not really hear the effect of melody lead which could explain the ambiguous results. In the second part of the test “Der Dichter spricht” was the only example that received a negative average value which might be explained the same way.

The highest value for synchronization in the first part was achieved for “Emigrantvisa” which is arranged in a very homophonic way with a lot of chords in a steady but moderate pace. In this type of music it is probably easier to judge the synchronization. Furthermore, in such a piece the dullness of perfectly synchronous chords becomes more obvious. However, in the second part of the test the average synchronization for the piece was 0. One possible explanation to this is presented below.

In the second part of the test the results are generally more ambiguous. In most cases the participants chose the deadpan alternative of perfect synchronization. Many of them explained after the test that it was difficult to really hear any difference between the alternatives for some examples. This may have been due to fatigue since they had already listened to all the examples once. Another explanation to this might be that the possibility to exaggerate the effect in both directions was eliminated.

\(^\text{10}\) Broken chords played in a sequence.
Of all the examples in the second part of the test the participants selected deadpan or positive synchronization 29 times out of 35. The ratings from the first part of the test rarely exceeded +1 and never went beyond -1. Perhaps the participants chose a deadpan version instead of a positive or negative version when their preferred setting otherwise would have been somewhere in between. In order to test this hypothesis the synchronization values from the first part were mapped to discrete values in the following way. All values between -0.5 and +0.5 were quantized to 0. All values above +0.5 were quantized to 1 and all values below -0.5 were quantized to -1. The result is showed in Figure 24 (ordered as the result from the second part to facilitate comparison). For “Der Dichter spricht”, “Träumerei” and “Etude” the result is similar. “Emigrantvisa” had an overweight of positive ratings in the first part compared to the second whereas “KV1” had an overweight of negative ratings. Overall, the above presented explanation seems possible.

![Quantized result (First part)](image)

**Figure 24** Quantized result of the first part of the test.

### 4.6.2 Synchronization versus tempo

The trend lines in Figure 20 suggest that for positive synchronization values, the amount of tolerable melody lead increases with decreasing tempo, exemplified with the two extreme values for the Etude. When an outliner had been removed the correlation between synchronization and melody was highly significant. This is also consistent with Vernon’s finding that slow music is more asynchronous than fast music.

In the data from the second part of the test, the trend line for tempo against synchronization had a small positive tilt (Figure 23) contrary to the result of the first part. The correlation test showed no significance to this tilt however. It can also be seen in Figure 23 that the tempo spread was larger with positive synchronization. Maybe the participants were more flexible with the tempo when the model was functioning as intended.
4.7 Conclusions

A listening test was developed and tested on seven subjects of high musical training. The small amount of data makes it difficult to draw any certain conclusions. Nevertheless the result indicates that the model is useful. The way positive synchronization correlates with tempo assimilates reality. For negative synchronization values the relationship to tempo seems more arbitrary. Negative synchronization (inverted melody lead) on the other hand is a concept not found in piano music but invented for the purpose of this test and the arbitrary relation to tempo is thus logical. Another test without the option of negative synchronization would be an interesting sequel to this test.

The participants never knew how the synchronization lever was configured and this made some of them a bit frustrated. Maybe better results could be obtained with another type of test where the participants know in more detail what parameters they are adjusting. When the test was designed it was considered important however that the participants used their ears and not their knowledge.

In the second part of the test the participants found it difficult to hear the difference between the synchronization alternatives for some examples. This part of the test also showed more ambiguous results. Maybe better results could have been obtained with different fixed tempos and a continuous lever for synchronization.
5 Discussion and conclusions

5.2 The velocity artifact

The model of melody lead was evaluated in two ways. A comparison of the output of the model was made with Repp’s data which showed similarities. The delay times produced by the model were more modest however. Additionally, the model tends to produce bass leads similar to the ones found in one of Repp’s groups of pianists. Repp drew the conclusion however that leading with the left-hand is rather an expressive strategy than due to the velocity artifact. The result of bass leads in the model can be explained by the way pDM calculates sound levels. Henderson found that pianists tend to organize the chords vertically with increasing sound level for the top voices. This is probably not entirely implemented in pDM which may account for the bass leads.

Furthermore, the model was tested in a listening test, consisting of two parts, with skilled pianists. The results indicated that the model simulates melody lead quite well. Especially the first part of the test showed similarities with reality the way tempo correlated to positive synchronization. No such correlation was found for negative synchronization which seems natural since it is not modelled on reality.

By modelling melody lead this way there is a tendency for the bass to sometimes lead and sometimes lag. This is probably not totally consistent with the velocity artifact hypothesis but on the other hand not totally contrary to what is found in real performances either.

The analyzed data set was very small so it is difficult to draw any certain conclusions. It seems though that the velocity artefact hypothesis produces good results when applied in a straightforward process. Of course other variables are always included in a performance, such as motoric noise and deliberate asynchronies. Nevertheless, being able to reproduce asynchronies that sound realistic is the interesting result.

5.1 Recommendations

5.1.1 Further studies

In this Master’s project only a first attempt to model melody lead has been made. More thorough investigations must be performed in order to find a proper model. Especially, the relationship between the hands should be investigated.

Another test could be designed where a deadpan score is transformed into two different versions: one version with the present melody lead model applied, and another version with random variability within the same range. Test subjects would then decide which version sounds most realistic.

5.1.2 Applications

One possible application of a melody lead model is in sequencer programs where the quantizing of MIDI data from piano playing could be much improved. The playback of scores in notation programs such as Finale or Sibelius could also be made more realistic with such a model. Of course, the inclusion of a melody lead rule within the KTH system is another possible application.
6 References


## 7 Appendix

### 7.1 K-values for the performance rules

For the five music examples used in the test, k-values (used to steer the performance in pDM) for each performance rule were found that resulted in a relatively neutral performance. The resulting k-values are presented below. The k-value for tempo is excluded though since this was adjusted by the users dynamically. The k-value for sound level was set to 0 (no change from the sound level specified in the file) for all pieces except for “Emigrantvisa” which was boosted 7 dB to compensate for low sound level in the file. The rule setup for “Träumerei” was also used during the evaluation of the model (described in chapter three). These are the same performance parameters that are used in Director Musices.

### 7.1.1 Etude 10/3 – Chopin

<table>
<thead>
<tr>
<th>Rule</th>
<th>k tempo</th>
<th>k sound level</th>
<th>k articulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase arch 4</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 5</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 6</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 7</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 4</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 5</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 6</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Final ritardando</td>
<td>0,6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Punctuation</td>
<td>0,5</td>
<td>-</td>
<td>0,5</td>
</tr>
<tr>
<td>High loud</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Melodic charge</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Harmonic charge</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Duration contrast</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Inegales</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Double duration</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Repetition art</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Score legato art</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Score staccato art</td>
<td>-</td>
<td>-</td>
<td>0,4</td>
</tr>
<tr>
<td>Overall articulation</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

### 7.1.2 KV1 – Mozart

<table>
<thead>
<tr>
<th>Rule</th>
<th>k tempo</th>
<th>k sound level</th>
<th>k articulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase arch 4</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 5</td>
<td>1,5</td>
<td>1,5</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 6</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 7</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 4</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 5</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 6</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Final ritardando</td>
<td>0,5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Punctuation</td>
<td>1,2</td>
<td>-</td>
<td>1,2</td>
</tr>
<tr>
<td>High loud</td>
<td>-</td>
<td>0,5</td>
<td>-</td>
</tr>
<tr>
<td>Melodic charge</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Harmonic charge</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Duration contrast</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Inegales</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Double duration</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Repetition art</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Score legato art</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Score staccato art</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Overall articulation</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

### 7.1.2 Der Dichter spricht – Schumann

<table>
<thead>
<tr>
<th>Rule</th>
<th>k tempo</th>
<th>k sound level</th>
<th>k articulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase arch 4</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 5</td>
<td>1,5</td>
<td>1,5</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 6</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 7</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 4</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 5</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 6</td>
<td>0,5</td>
<td>0,5</td>
<td>-</td>
</tr>
<tr>
<td>Final ritardando</td>
<td>1,5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Punctuation</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>High loud</td>
<td>-</td>
<td>0,3</td>
<td>-</td>
</tr>
<tr>
<td>Melodic charge</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Harmonic charge</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Duration contrast</td>
<td>-0,9</td>
<td>-0,9</td>
<td>-</td>
</tr>
<tr>
<td>Inegales</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Double duration</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Repetition art</td>
<td>-</td>
<td>-</td>
<td>0,96</td>
</tr>
<tr>
<td>Score legato art</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Score staccato art</td>
<td>-</td>
<td>-</td>
<td>0,7</td>
</tr>
<tr>
<td>Overall articulation</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

### 7.1.2 Träumerei – Schumann

<table>
<thead>
<tr>
<th>Rule</th>
<th>k tempo</th>
<th>k sound level</th>
<th>k articulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase arch 4</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 5</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 6</td>
<td>0,5</td>
<td>0,5</td>
<td>-</td>
</tr>
<tr>
<td>Phrase arch 7</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 4</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 5</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Phrase ritardando 6</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Final ritardando</td>
<td>0,4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Punctuation</td>
<td>1,2</td>
<td>-</td>
<td>1,2</td>
</tr>
<tr>
<td>High loud</td>
<td>-</td>
<td>0,8</td>
<td>-</td>
</tr>
<tr>
<td>Melodic charge</td>
<td>0,2</td>
<td>0,2</td>
<td>-</td>
</tr>
<tr>
<td>Harmonic charge</td>
<td>0,2</td>
<td>0,2</td>
<td>-</td>
</tr>
<tr>
<td>Duration contrast</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Rule</td>
<td>k tempo</td>
<td>k sound level</td>
<td>k articulation</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------</td>
<td>---------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Phrase arch 4</td>
<td>1,2</td>
<td>1,2</td>
<td></td>
</tr>
<tr>
<td>Phrase arch 5</td>
<td>0,2</td>
<td>0,2</td>
<td></td>
</tr>
<tr>
<td>Phrase arch 6</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Phrase arch 7</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Phrase ritardando 4</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Phrase ritardando 5</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Phrase ritardando 6</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Final ritardando</td>
<td>0,5</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Punctuation</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>High loud</td>
<td>-</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Melodic charge</td>
<td>1,5</td>
<td>1,5</td>
<td></td>
</tr>
<tr>
<td>Harmonic charge</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Duration contrast</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Inegales</td>
<td>0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Double duration</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Repetition art</td>
<td>-</td>
<td>-</td>
<td>1,2</td>
</tr>
<tr>
<td>Score legato art</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Score staccato art</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Overall articulation</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

7.1.2 Emigrantvisa – Jan Johansson