LOOKING AT PERCEPTION OF CONTINUOUS TEMPO DRIFT - A NEW METHOD FOR ESTIMATING INTERNAL DRIFT AND JUST NOTICEABLE DIFFERENCE.

Sofia Dahl and Svante Granqvist

Department of Speech Music and Hearing
Kungl Tekniska Högskolan
sofia@speech.kth.se, svante@speech.kth.se

ABSTRACT

The method proposed here investigates if there is such a thing as an internal representation of a "steady tempo" - and whether this representation itself is free from tempo drift. The method uses a modification of the method for Parameter Estimation by Sequential Testing (PEST). Several click sequences are presented to the listener in each test and depending on the listener's response (correct or incorrect) the magnitude of the tempo drift is modified for the next presentation. The method does not rule out the possibility that the internal "clock" can have an inherent tempo drift. Such a drift would mean that some listeners could perceive an increasing tempo when, in fact, it is decreasing, and vice versa. Preliminary results indicate that some listeners tend to place their answers "biased" towards either increasing or decreasing tempo. The results also indicate that these listeners appear to be consistent in doing this. Thus we would like to propose a model of detectability of continuous (linear) tempo drift based on a person's *Internal Drift* (ID), which can be zero or biased in either direction. Surrounding this ID is an interval corresponding to twice the *Just Noticeable Difference* (JND).

1. INTRODUCTION

In many musical genres tempo changes are used intentionally to mark the structure of the piece. For instance, tempo tends to be increased in the beginning of a phrase but decreased towards phrase endings. However, even when the intention is to keep a steady tempo, measured timing data usually display both short-term and long-term variations. These unintended fluctuations and tempo changes include gradual changes of tempo over time, *tempo drifts*. In studies of finger tapping Madison reported linear drifts in both directions (increasing and decreasing) ranging between 0.05 and 0.3 % per interval [1]. Measured timing data from drummers also display drifts in tempo, when estimated as linear the drift was typically about 0.2 % per interval [2].

The measured tempo drifts in tapping and drumming data bring up interesting questions about how small gradually introduced tempo changes that can be perceived. It seems reasonably to assume that even very small continuous changes ought to be audible if enough time passes. However, this would require some kind of representation of the original tempo to compare the changing tempo with.

If there is such a thing as an internal representation of a "steady tempo", can we assume this representation to be isochronous? Madison has pointed out that isochrony is simply the special case where no monotonous drift is present and it does not need to have any special status [3]. This considered, there is a possibility that the internal reference itself may not be free from drift, and some earlier investigations have pointed in this direction. Vos et. al investigated perception of tempo drifts and found that subjects responses were biased by the the base tempo. For a base tempo of 240 *beats per minute* (BPM) subjects tended to respond that tempo was accelerating, while a base tempo of 60 BPM generated more responses that tempo was decelerating. For both these extreme tempi subjects' responses were biased *even with no tempo drift present*. For the intermediate tempo 120 BPM (nominal beat separation 500 ms), subjects perceived as many decelerations as accelerations and no bias was found [4].

As illustrated by the example above, the *physical tempo drift*, i.e. the physical lengthenings or shortenings of subsequent intervals in a stimulus, does not necessarily have the same direction as the perceived tempo drift. Such bias in the direction of tempo drift has a consequence on how the smallest perceivable drifts is measured.

In this paper we depart from the assumption that there is an internal reference that we use for comparison when judging tempo, without ruling out the possibility that this reference may have an inherent tempo drift. While we assume that there is a range of continuous tempo drifts that are too small to be perceived, we take into consideration that range may not be centered around isochronous physical tempo. We propose that this span of non-discriminable drift can be interpreted as twice the *Just Noticeable Difference* (JND), centered around an internal reference; the *Internal Drift* (ID).

2. METHOD

2.1. Model and definitions

If the perception of tempo drift is modelled as an internal drift of the subject and a band around this internal drift in which no tempo drift can be detected, the tempo drifts generated by our proposed method will converge to the positive or negative limits (just noticeable drifts) of this band. In the method the physical tempo drifts are generated by constant and continuous change of each *inter-onset interval* (IOI), that is, the interval between the onset of each click is consistently either increased or decreased by adding or removing a fixed number of samples.

The method uses a modification of Parameter Estimation by Sequential Testing (PEST), implemented in the custom made computer program "Beat". During a test session the program presents stimuli consisting of a number of clicks with either increasing tempo (successive shortening of IOIs) or decreasing tempo (successive lengthening of IOIs). After each completed playback of a stimulus the listener is asked to indicate the direction of the tempo drift by clicking one of the buttons "increasing" or "decreasing". Each clicking of a button is equivalent to a *response*. This procedure is repeated a number of times and the program adjusts the physical drift before each stimulus playback in order to arrive at the just noticeable drift. Two correct responses in a row are required to assume that the response was correct, while one incorrect response is considered an incorrect answer (see Figure 1). This causes the tempo drift to converge towards the 75 % correct responses level.

Each test session consists of several series running in parallel, each series consisting of several played back stimuli (normally between 20 and 60). Each test series is assigned a nominal tempo (in BPM), a stimulus duration (playback time in seconds), an initial drift (% of nominal initial IOI), an initial drift step, and a designation as either 'increasing' or 'decreasing'. Because of the designation, the correct response to a series designated 'decreasing', is always "decreasing" - also if the physical drift is increasing. In this way, the physical drift of a series can shift sign and cross the zero-line as the response curves converge around a perceived isochronous tempo.

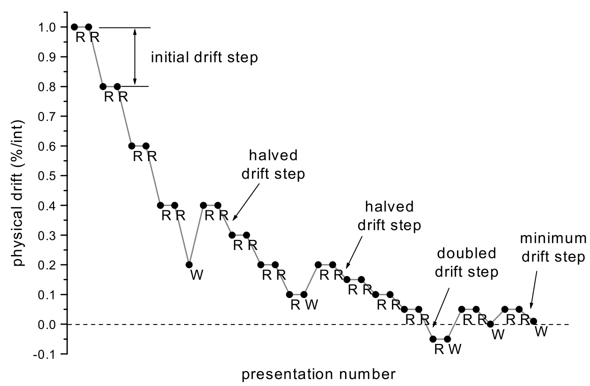

Figure 1: *Schematic illustration of a series and the associated responses from the subject. The first stimulus has a physical drift of 1 %/interval. After two correct responses (R) the physical drift is decreased by the initial drift step (0.2 %/int.). Later an incorrect response (W) is given and the drift is increased by the same drift step. After two correct responses following the incorrect response, the drift step is halved. The drift step never becomes smaller than the minimum drift step. Because of the designation of the series as 'decreasing' the response "decreasing" will always be considered correct, also if the physical drift should cross the zero-line.*

The Beat program runs on PCs with Windows and normal soundcards. The tempo (IOI) drift is achieved by inserting or removing a whole number of zero samples at the end of each click sound. The number of samples inserted/removed is kept exactly the same throughout each stimulus in order to achieve a smooth tempo drift.

3. LISTENING TEST

To test the perception of small tempo drifts and investigate whether listeners show ID $\neq 0$ a listening test was performed. Seven subjects, aged 25-57 years (mean 35.6 years) participated in the test. All subjects had received training in playing musical instruments or singing, their experiences ranging between 2 and 40 years.

The subjects were tested individually in three listening sessions of 40 minutes each. In order to expose the subjects to a mix of stimuli with both increasing and decreasing tempo drift, each session consisted of six parallel series; three designated 'increasing' and three designated 'decreasing'. The order in which series were to be played back was decided semi-randomly, so that the number of turning points in the six series were synchronized. In this way the amount of transitions from correct-incorrect or incorrect-correct answers for were approximately equal for each series throughout the test session. To reduce the risk of subjects being more inclined to use the right or the left button for their responses, the placement of the response buttons 'increasing' and 'decreasing' was randomized before each test session.

The subjects were instructed that they would hear a number of click sequences, 5 s long, with either increasing or decreasing tempo drift. For each presented stimulus the subject was to carefully pay attention to the direction of the drift and indicate this by clicking the corresponding button on the screen. All ten intervals in a stimulus had to be played before it was possible to respond. The listeners were not aware of the test design or that the method was adaptive. Each test session was interrupted after 40 minutes.

The nominal start tempo was 120 BPM, (nominal beat separation 500 ms), but the actual initial tempo that was used for each stimulus was randomized by the same amount as the total tempo change, from beginning to end, for the stimulus. Each stimulus consisted of ten intervals (~ 5 s) which were played after an initial silence with a randomized duration of 0.5 to 1 s. The maximum allowed drift step was 0.2 % of initial IOI and the minimum allowed drift step was set to 0.04 % of initial IOI. The sampling frequency used was 44100 Hz, which gives a resolution in the IOI drift at 120 BPM of approximately 0.0045 %.

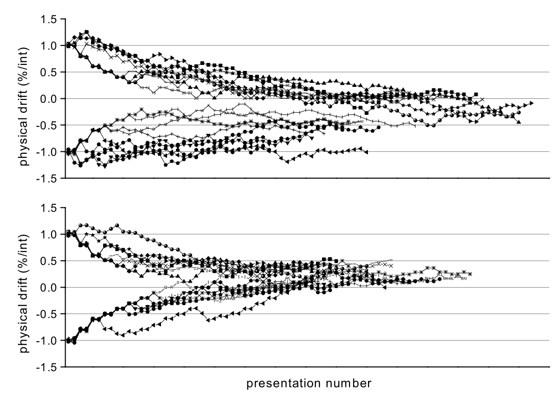

Figure 2: *Examples of series produced for two different subjects with different internal drift and different JND. Each panel presents 18 series originating from three test sessions. Each of these sessions consisted of three series with an initial physical drift of -1 %/interval (designated 'increasing'), and three series with an initial drift of 1 %/interval ('decreasing'). Both subjects display non-zero internal drifts. Subject GS (top) has an internal drift corresponding to a increasing tempo. By contrast, subject AA (bottom) has an internal drift corresponding to a decreasing tempo with a smaller JND than subject GS.*

4. RESULTS

Figure 2 shows all of the produced series for two of the subjects participating in the three test sessions. In each panel the nine series starting with a drift of 0.1 %/interval (designated 'decreasing'), and the nine series starting with a drift of -0.1 %/interval (designated 'increasing') are seen. Although the series originate from three different sessions the subjects are remarkable consistent in their judgements.

Many of the subjects showed bias in their judgements of the drifts. For instance, the 'decreasing' series for subject GS (top panel, Figure 2) all approach, or even cross, the zero-line for the physical drift, indicating that this subject has an internal drift corresponding to an increasing tempo. The span between 'decreasing' and 'increasing' series is rather large, indicating large JND. By contrast the produced series for subject AA (bottom panel) cluster closely together on the positive side of the physical drift, indicating an internal drift corresponding to a decreasing tempo and a smaller JND.

In the analysis of data the average endpoints across the 'increasing' and 'decreasing' series respectively, were taken as giving the span of non-discriminable drift centered around the internal drift (ID). The span was interpreted as twice the Just Noticeable Difference (JND).

4.1. Internal drifts

The results from the analysis of the Internal Drifts can be seen in Figure 3. The grand average ID, across subjects and sessions, was not significantly different from 0. Between subjects the IDs were significantly different ($F[6, 2.341] = 40.45, p < .001$, Greenhouse-Geisser compensated), and five of the subjects had an ID significantly different from 0. There were no significant differences between test sessions.

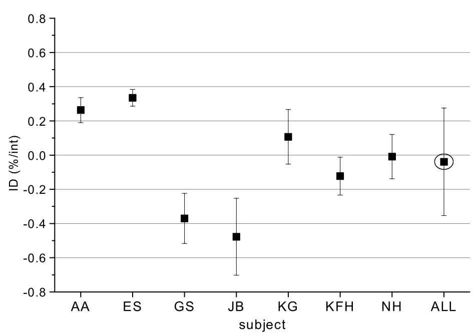

Figure 3: *The resulting internal drifts (ID) for the seven subjects. The vertical error bars show the standard deviation. The grand average across all subjects (encircled data point to the right) was -0.04 %/interval, which was found not to differ significantly from 0.*

4.2. Just Noticeable Differences

The JNDs relative to the found ID for the seven subjects can be seen in Figure 4. The mean JND, across subjects and sessions, was 0.27 %/interval, corresponding to 1.35 ms/interval. Like for

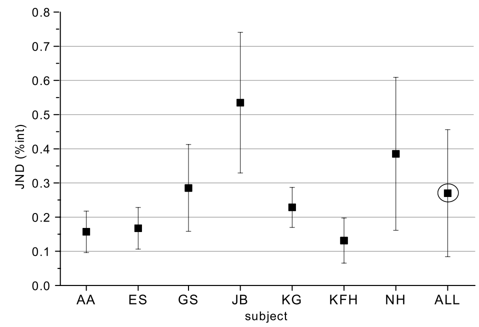

Figure 4: *The resulting just noticeable differences as related to the internal drift (see Figure 3) for the seven subjects. The standard deviation is shown by the vertical error bars. The grand average across all subjects (encircled) was 0.27 % per interval, equal to 1.35 ms/interval.*

the IDs the JNDs were significantly different between persons ($F[6, 1.962] = 11.79, p < .001$, Greenhouse-Geisser compensated).

5. DISCUSSION AND CONCLUSIONS

Although the subjects in our study clearly differ in their opinions of a "steady beat", with IDs covering a range of drift from about -0.5%/interval to +0.4%/interval, the grand average ID still ends up almost at zero (see Figure 3). We have, however, so far only investigated the tempo 120 BPM, the same tempo which Vos et. al found no bias for in their study [4]. Even so, the method appears to be successful in estimating the internal drift. Subjects proved to be very consistent in their responses, there were no significant differences in ID between series or test sessions. More tempi and more subjects are likely to help in shedding some light on whether the ID may be dependent on tempo.

The found values for the JNDs correspond to measured production data from drumming [2] and tapping [1]. However, when comparing to previously reported estimates of JND for drift and tempo detection (see surveys in [3] and [5]) they are surprisingly small. For a tempo with a nominal beat separation of 500 ms (120 BPM) these JNDs have been estimated to about 2 %. There may be several reasons for these discrepancies:

First, our model places the JND in relation to the ID of each subject. Considering that the found IDs proved to be non-isochronous for five of the seven subjects this makes comparisons to other studies somewhat hazardous. However, if our grand average JND is added to the largest ID for a subject, it still does not amount to more than 4 ms/interval (corresponding to 0.8 %/interval).

Second, many studies have concentrated on the detection threshold for stepwise tempo changes and perhaps not as often has continuous tempo drift been studied. However, a continuous tempo change of 0.2 %/int. over 10 intervals results in a total tempo change of 2 % over the whole sequence. It is possible that the accumulated tempo change over several intervals causes a detection, even when the change for each interval does not.

Third, many other studies have used pairwise comparisons or adaptive methods asking whether a change was heard or not. We chose the Forced Directional Response paradigm used by Vos et. al [4], asking the subjects to indicate the direction of drift. The fact that the subjects were informed about the presence of tempo drift, and that they were forced to make their best guess regarding the direction may have contributed to the lower JNDs.

Some experiments by Madison exemplify differences between methods. Using an adaptation of PEST, and asking whether drift had been detected (yes or no), several experiments showed that for 9 intervals at a tempo of 120 BPM, the JND was about 2 %/interval (or 10 ms/int.). In a following experiment the method was changed and subjects were instead asked to modify the tempo by pressing keys ('faster', 'slower') until they perceived it to be isochronous. The subjects could use up to 30 intervals to make their judgement before adjusting or confirming. The method also allowed the subjects modifications to cross the zero level, although this was rare. The results showed that the subjects' Point of Subjective Equality was about 0.2 % at a tempo of 120 BPM. The mean number of intervals used by the subjects across all trials varied from 8.1 to 16.8.

The number of intervals that subjects are allowed to hear also affects the JND. Drake and Botte [6], and Madison [3] have reported that detection thresholds decrease with increasing number of intervals presented. However, in many investigations of threshold for tempo detection, rather short stimuli durations have been used, usually between 4 and 6 intervals. In a musical context listener are also used to gradual changes occurring over more than a measure, and can be assumed to use more than 6 intervals to detect changes in tempo.

To conclude, the method seems to be successful in estimating the internal drift. There were differences between subjects, of which five displayed internal drifts significantly different from zero, either in the positive and negative direction. The found JNDs were small when compared to earlier studies and will need to be investigated further through more listening experiments.

6. ACKNOWLEDGEMENTS

The authors are indebted to all the subjects participating in the test, and to Joakim Westerlund for the statistic analysis.

7. REFERENCES

[1] Madison, G., "Variability in Isochronous Tapping: Higher Order Dependencies as a Function of Intertap Interval", Journal of Experimental Psychology: Human Perception and Performance, 27(2), 2001, 411-422.

[2] Dahl, S., "The Playing of an Accent - Preliminary Observations from Temporal and Kinematic Analysis of percussionists", Journal of New Music Research, 29(3), 2000, 225-234.

[3] Madison, G., "Functional Modelling of the Human Timing Mechanism", Published doctoral dissertation, Uppsala University, Sweden, 2001.

[4] Vos, P. G., van Assen, M., and Franek, M., "Perceived tempo change is dependent on base tempoand direction of change: Evidence for a generalized version of Schulze's (1978) internal beat model", Psychological Research Vol. 59, 1997, 240-247.

[5] Friberg, A. and Sundberg, J. "Time discrimination in a monotonic, isochronous sequence", J. Acoust. Soc. Am., 98(5), 1995, 2524-2531.

[6] Drake, C. and Botte, M-C., "Tempo sensitivity in auditory sequences: Evidence for a multiple-look model", Perception & Psychophysics, 54(3), 1993, 277-283.