On the robustness of time-shrinking

Remijn, Gerard Bastiaan
Unit of Experimental Psychology, Faculty of Social and Behavioural Sciences, Leiden University

Meulen, Gert van der
Unit of Experimental Psychology, Faculty of Social and Behavioural Sciences, Leiden University

Hoopen, Gert ten
Unit of Experimental Psychology, Faculty of Social and Behavioural Sciences, Leiden University

Nakajima, Yoshitaka
Department of Acoustic Design, Kyushu Institute of Design

他

https://hdl.handle.net/2324/1912120

On the robustness of time-shrinking

Ger Remijn,* Gert van der Meulen,* Gert ten Hoopen,* Yoshitaka Nakajima,**
Yorimoto Komori,** and Takayuki Sasaki***

*Unit of Experimental Psychology, Faculty of Social and Behavioural Sciences,
Leiden University,
PO Box 9555, 2300 RB Leiden, the Netherlands
**Department of Acoustic Design, Kyushu Institute of Design,
4-9-1, Shiobaru, Minami-ku, Fukuoka, 815-0032 Japan
***Miyagi Gakuin Women's College,
9-1-1, Sakuragaoka, Aoba-ku, Sendai, 981-0961 Japan

Time-shrinking is a well-established perceptual phenomenon by now. When two empty time intervals, marked by short sounds, are presented contiguously, the first interval can shrink the second one perceptually. This is almost always the case when the first interval is shorter than the second one, unless the difference gets greater than approximately 80 ms. The phenomenon is rather compelling, so it can be called an illusion of time perception. Our purpose in the present study is to show by three experiments how robust this illusion is. The first experiment showed that time-shrinking operates also when the last time interval is preceded by more than one interval (up to five at least). Moreover, the number of preceding intervals had no effect upon the amount of shrinking. The second and third experiment studied the effect of sound marker frequency on time-shrinking. It was found that the illusory phenomenon clearly appeared even when the sound marker frequencies differed by more than two octaves. However, the amount of shrinking appeared to be largest when frequencies were equal.

Keywords: Time perception, Auditory illusion, Temporal patterns, Time-shrinking

PACS number: 43. 66. Mk

1. INTRODUCTION

Auditory perception is obviously very important in our daily life. Somewhat less obvious is the fact that the precise timing within sound sequences is crucial. Sentences are very hard to understand if not spoken with a correct temporal pattern. Melodies with an improper rhythm and timing can hardly be recognized. Therefore, the study of the perception of time intervals with durations comparable to that of the constituents of speech (*e.g.* phonemes, morphemes) and of music (*e.g.* tones, tone groups) is important (*cf.* some recent studies from Taguti and Suzuki, 1995; Hirano, Nakajima, Ihle, and ten Hoopen, 1995; Kato and Tsuzaki, 1998; Yamada and Tsumura, 1998).

Traditionally, many students of time perception investigated the listener's ability to discriminate between two isolated durations, either empty or filled, in order to explore whether Weber's law held (*e.g.* Getty, 1975). Slightly more complex temporal patterns were studied by Israeli (1930), who used light flashes to delimit the empty time intervals. His main interest was whether neighboring intervals affect the perceived duration of a standard time interval. In Nakajima and ten Hoopen (1988) we took up this question again with regard to empty time intervals marked by short sound bursts. Moreover, we were interested in shorter time intervals than Israeli used. Quite unexpectedly, we

found a conspicuous context effect: when a short standard time interval (S) was immediately preceded by a shorter neighboring interval (N), the standard was considerably underestimated. We give one example: when an S of 120 ms was preceded by an N of 45 ms, the underestimation was about 50 ms. Because this illusory phenomenon had not been reported in the literature, we decided to perform a systematic series of experiments, which we shall shortly summarize.

Nakajima, ten Hoopen, and van der Wilk (1991) replicated the earlier finding: an S of 120 ms was underestimated considerably when N was smaller. Underestimations also appeared when S was 240 ms and N was 190 ms. No such conspicuous context effect could be observed when S was followed by N. In Nakajima *et al.*'s (1991) study, a small supplementary experiment was reported in which the listeners had to judge the duration of N instead of S. It was found that the underestimation of S was not accompanied by the same magnitude of overestimation of N, and this finding has been substantiated by Sasaki, Nakajima, and ten Hoopen (1998). Thus, the temporal context effect appears to be asymmetric, and we coined the term 'time-shrinking' for this illusion.

Ten Hoopen *et al.* (1993) showed that the illusion does not result from a difficulty in resolving the temporal structure, that is, N and S could be discerned. These authors also showed that listeners do not inadvertently judge the duration of N instead of that of S. Furthermore, they excluded forward masking as an explanation of the illusion, by varying the intensity levels of the sound markers within patterns. In Nakajima, ten Hoopen, Hilkhuysen, and Sasaki (1992) we found that time-shrinking is strongest when $50 < (S - N) < 70$ ms. However, in that study we applied only one N-value (50 ms). But ten Hoopen *et al.* (1993) showed this tendency to hold for N-values of 40, 80, and 160 ms as well, and maximum time-shrinking appeared in that study when $(S - N) = \sim 80$ ms. Recently, Nakajima *et al.* (in press) studied the apparent lawfulness that shrinking is at maximum if $(S - N) = \sim 80$ ms in a more systematic way. They applied a wide range of N from 40 to 480 ms, in steps of 40 ms (except for $N = 440$ ms). In almost all conditions where $(S - N)$ was 80 ms, shrinking of S was at maximum, though the amount was largest when $N \leqq 200$ ms (approximately 45 ms underestimation

of S).

Nakajima *et al.* (in press) also offered an explanation of the mechanism of time-shrinking. It is only possible to give the gist of it here in a simplified manner and for details one should consult that study. The core of the mechanism is that it takes a listener approximately 80 ms to mentally process an empty duration (Nakajima, 1987). Thus, our brain needs time to establish an amount of physical time. The sum of the physical time and the processing time results in psychological time. An example: if $N = 200$ ms, the psychological time is $200 + 80 = 280$ ms, and if $S = 400$ ms, its psychological time is $400 + 80 = 480$ ms. However, when $(S - N) \leqq 80$ ms, the perceptual system is able to detect quasi-equality between the psychological duration of N and the momentary psychological duration of S. For example, when $N = 200$ ms, like above, the psychological duration of N is 280 ms and stored in the brain. Let S be 250 ms. Processing of the S-duration starts immediately after the final sound marker. But when 30 ms of the 80 ms processing time have expired, the system detects at *that moment* that $250 + 30 = 280$ ms equals the stored psychological duration of N, and the system decides to stop processing. That is assimilation in perception, which means that events or objects that are not too different tend to be similarized. After the perceptual decision to stop further processing of S, it takes in this example another 20 ms to dampen processing (a kind of brain inertia). As a result, the psychological duration of S is $250 + 30 + 20 = 300$ ms, whereas it would have been $250 + 80 = 330$ ms when S were presented in isolation, as is done in the control conditions. Thus, the amount of time-shrinking is $330 - 300 = 30$ ms in this example.

Although the phenomenon of time-shrinking seems to be well established by now, one might critisize the fact that only very primitive patterns were used in all the above mentioned studies. The patterns never contained more than two intervals (N and S), and the patterns were always monotonous, that is, the frequencies of the sound markers within patterns never differed. Such patterns are of course not very ecological as compared to the complex patterning in speech and music. Furthermore, as we know from Bregman's (1990) and Handel's (1989) renowned textbooks, the way the listener organizes his/her auditory environment into events and streams (called "auditory scene analysis," Breg-

man, 1995) can be affected dramatically by changing and/or adding acoustic variables.

The motivation of this study is to investigate what happens to time-shrinking when the primitive temporal pattern we used in our previous studies is made more complex. The first experiment of the present study was set up to determine whether time-shrinking also operates in longer sequences, that is, in sequences in which S is preceded by more than one N. The second and third experiments were designed to investigate whether frequency differences between sound markers influence time-shrinking. It is a well established fact that tones with different frequencies, when alternated in a sequence, give rise to one coherent stream if the frequency difference is small, but to two streams when the difference is big (van Noorden, 1975). Even though the sequences used in streaming studies are longer than the short sequences applied in the present study, for safety we decided to design two separate experiments. Hence, the second experiment employed frequencies not too far apart (about 1.5 semitone), and the third experiment employed frequencies differing by more than two octaves.

2. EXPERIMENT 1

An important question about the robustness of time-shrinking has not been satisfactorily answered by now: What happens when the standard empty time interval (S), of which the duration has to be judged, is preceded by *more* than one neighbouring interval (N)? Does time-shrinking disappear because the mechanism only operates in two-interval sequences? Or does the amount of time-shrinking increase with increasing number of N, because more preceding neighbors exert a larger shrinking force? And if there is an effect of the number of neighbors, does it depend on the tempo of the sound sequence? In order to answer such questions we designed this experiment.

2.1 Method
2.1.1 Subjects
Four subjects with normal hearing participated, two of the authors (GtH, GvdM, both male), and one female and one male student. All of them were affiliated to Leiden University, the Netherlands.
2.1.2 Stimuli and design
The control conditions comprised of standard empty time intervals (S), followed 5 s later by a

comparison interval (C). The intervals were marked by sine tones of 3 kHz, lasting 5 ms. The sine tones were not envelope shaped, but started and stopped at zero crossing points. The intensity of the sound markers was such that a continuous 3 kHz tone amounted to 80 dBA as measured by a precision sound level meter (Brüel & Kjaer 2203), mounted with an artificial ear (Brüel & Kjaer 4152) and a microphone (Brüel & Kjaer 4144). The S-durations were 100, 125, 150, 200, and 250 ms. In the control conditions these durations were presented in isolation. In the experimental conditions, S was preceded contiguously by one to five neighboring intervals (N), which had half the duration of S (50, 62.5, 75, 100, and 125 ms respectively). This lead to 30 stimulus patterns (5 controls+25 experimental). For a different purpose, which we will not further discuss in this article, we added 10 stimulus patterns, nine in which the durations of N and S were equal and one isolated S-duration of 50 ms. Thus, in total we had 40 different stimulus patterns.
2.1.3 Apparatus and procedure
The duration of C in the 40 different patterns had to be adjusted to S in ascending (A) as well as in descending (D) series. Thus, a block contained 40×2 (A/D)=80 trials, which were preceded by eight warm-up trials. The block was replicated four times, in which the trials were randomized differently. Before beginning, the subjects had a practice session with 40 trials. The stimuli were generated by a Commodore Amiga 500 computer and were presented monaurally (right ear), via an amplifier (Nikko, type NA-500 II) through an AKG-headphone (type K140) in a sound attenuating booth. The task of the subject was to adjust the initial C to the subjective duration of S. The initial duration of C was clearly shorter or longer than S, and these durations were randomized. The adjustment was done by pressing a mouse-button on a 'shorten' and a 'lengthen' pane on the screen. The longer the button was pressed, the more the final sound marker of C was manipulated towards or away from the initial marker. Clicking a 'presentation' pane on the screen initiated the presentation of the stimulus pattern and the adjustable C. There was no limit to the number of presentations, that is, the subject could hear the pattern and adjust C as many times as he/she felt necessary. When the subject was satisfied with the match between S and

C, a 'finish' pane could be clicked at, after which the final value of C was registered by the computer as the point of subjective equality (PSE). The next trial was presented after pressing a button on the terminal.

2.2 Results and Discussion

The median PSEs of S, all based on 32 observations, were calculated and in Table 1 it can be seen that these median PSEs of the control conditions (0 neighbors) did not differ much from the points of objective equality (POEs) of S. Time-shrinking could be clearly observed when S was 100, 125, and 150 ms : the PSEs of S were smaller when there were preceding neighbors. The opposite effect, that is, systematic overestimation of S, was observed when S was 200 or 250 ms.

To test whether these underestimations and overestimations were significant, we performed five Friedman-tests, one at each level of S. All tests yielded significant differences at the 5%-level. Inspection by eye indicates that these significant effects were probably due to the difference between the control condition (0 neighbors) on the one hand and the experimental neighbor conditions on the other hand. To test whether there was an effect of the number of neighbors (1 through 5), we again performed five Friedman-tests, but now with exclusion of the 0-neighbor (control) condition. Except for the condition where S was 150 ms (and N was 75 ms), there was no significant difference with regard to the number of neighbors. In the $N/S = 75/150$ ms condition, the significant effect was caused by $N=2$, where the PSE of S was 104 ms. At present, we have no explanation for this exception. But in general, it seems safe to conclude that 1) time-shrinking also clearly occurs when there are more preceding neighbors, and 2) the extent of time-shrinking is not affected by the number of preceding neighbours. Thus, it seems that there is no accumulating effect of the preceding neighbors. However, we did only use preceding neighbors of equal duration in the series preceding S, and it is important to investigate in future research what happens to time-shrinking when the neighbors preceding S have different durations.

3. EXPERIMENT 2

In our previous studies, the empty intervals N and S (and C) were delimited by short sound markers of 5-10 ms, all of the same fundamental frequency (mostly 1 and 3 kHz). One might question what happens to time-shrinking if the sounds that delimit these empty intervals have different frequencies, but not so different that streaming would have arisen in longer sequences. For reasons of economic designing, we planned not to include control patterns in which the sound markers were equal, because we dispose of a huge data-base of underestimations of S which can serve as a base rate.

3.1 Method
3.1.1 Subjects

Four male subjects with normal hearing participated (two of the authors, GR, GtH, and two students of psychology at Leiden University). The ages of the subjects were 22, 50, 23, and 25 years.

3.1.2 Stimuli and design

The stimulus patterns $/N/S/$ had durations of $/80/80/$, $/80/100/$, $/80/120/$, $/80/140/$, $/80/160/$, $/80/180/$, $/80/200/$, and $/80/220/$ ms. The delimiting sound markers lasted 10 ms and started and stopped at zero crossing points. The markers approximated square waves of 1,000 Hz for the first marker, of 1,080 Hz for the second marker, and again 1,000 Hz for the third marker. The same eight stimulus patterns were also presented with swapped frequencies for the second and third markers : frequencies of 1,000 Hz, 1,000 Hz, and 1,080 Hz now marked the intervals N and S. The intensity level of the sound markers was, when presented continuously, 99.1 dBA for the 1,000 Hz markers and 99.3 dBA for the 1,080 Hz markers, measured as in Experiment 1. The control stimulus patterns comprised only of S, which had durations

Table 1 Median points of subjective equality (PSEs) in Experiment 1 as a function of the duration of the standard time interval (S), the duration of the preceding neighbor time interval (N), and the number of preceding neighbors. (PSEs based upon 32 observations ; all entries in ms).

N/S - pattern	Number of preceding neighbors					
	0	1	2	3	4	5
50/100	105	64	57	53	49	53
62.5/125	124	81	77	78	83	92
75/150	156	121	104	138	141	136
100/200	202	214	221	217	235	228
125/250	265	287	283	297	288	287

of 80, 100, 120, 140, 160, 180, 200 and 220 ms, delimited respectively by a 1,080 Hz tone as first marker and a 1,000 Hz tone as second marker (Order 1), and vice versa (Order 2). As a result there were 32 different stimulus patterns (16 controls and 16 experimental).

Both the experimental and the control patterns were presented in an ascending and a descending fashion, where the initial comparison (C) was either clearly shorter or longer than S. One block consisted of 64 randomized trials (16 control + 16 experimental patterns) × 2 (ascending/descending). These blocks were replicated eight times. The 64 trials in each block were randomized in a different order for each subject. The auditory stimuli were generated by a Commodore Amiga 500 + computer and presented on line via an amplifier (JVC AX11) and headphones (JVC HA D500) to the left ear. About 2.5 seconds after the start of S, C was presented.

3.1.3 Apparatus and procedure

The task of the subjects was to adjust the initial C to the perceived duration of S in the same way as described in Experiment 1. After two training sessions, covering one block of 64 adjustment trials, the subjects did the eight experimental blocks in 16 sessions, half a block of 32 adjustments per session.

3.3 Results and Discussion

A first inspection of the data revealed that two of the subjects (JB and GR) still showed considerable time-shrinking at the longest S-duration of 220 ms. Thus, although the difference between S (220 ms) and N (80 ms) was 140 ms, there was still no release from time-shrinking for these subjects, a fact we had never came across before. Therefore, we ran a small supplementary experiment with subjects JB and GR employing two longer S-durations, namely 240 ms and 260 ms. Procedures were exactly equal to those of the main experiment. It turned out that extending the experiment by these two extra steps on the S-duration scale was sufficient to get a release from time-shrinking for these two subjects.

A possible difference in the amount of shrinking between Order 1 and Order 2 was statistically analyzed. For each subject we calculated the differences between experimental and control PSEs at both orders. These differences were subjected to Wilcoxon-tests (two-tailed, 5%). It turned out that there was no difference between orders whatsoever.

Thus, for the sake of clear illustration of the course of time-shrinking, we took the median PSEs across these frequency-orders and plotted them as a function of S-duration in Fig. 1(subjects NP and GtH), and Fig. 2 (subjects JB and GR). Figure 1 shows
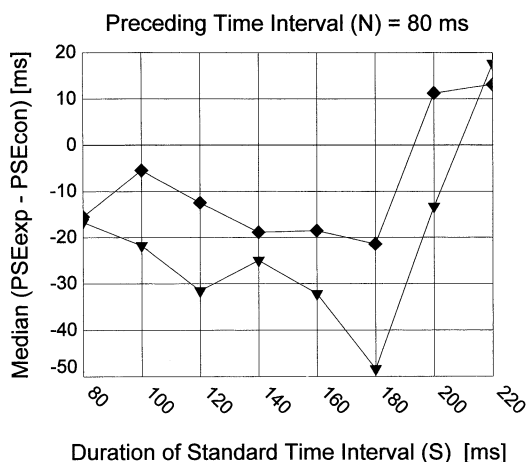

Fig. 1 The median amounts of overestimation ($PSE_{exp} - PSE_{con}$ in ms) as established in Experiment 2 for subjects GtH (diamonds) and NP (triangles) as a function of the duration of the standard time interval S (in ms), when preceded by a neighboring time interval N of 80 ms.
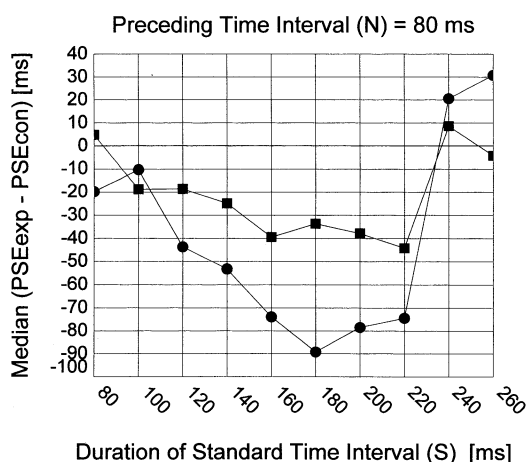

Fig. 2 The median amounts of overestimation ($PSE_{exp} - PSE_{con}$ in ms) as established in Experiment 2 for subjects JB (circles) and GR (squares) as a function of the duration of the standard time interval S (in ms), when preceded by a neighboring time interval N of 80 ms.

that the strongest time-shrinking occurred at an S of 180 ms for both subjects, that is, when $S - N = 100$ ms. Figure 2 shows that strongest time-shrinking for subject JB occurred when $S = 180$ ms. However, this subject still showed considerable shrinking when $S = 220$ ms and only after this S-duration there was a complete release of time-shrinking. The strongest time-shrinking for subject GR occurred when $S = 220$ ms. (All these amounts of underestimations were significant at 1% as tested by two-tailed Wilcoxon-tests.) Recall from the introduction that we stated that the time to process an empty duration was 80 ms on average. The present subjects, however, needed processing times ranging between 100 and 140 ms. This illustrates that individuals can vary considerably in the amount of processing time they need to establish durations.

4. EXPERIMENT 3

Experiment 2 showed that time-shrinking also clearly emerged when the marker frequencies differ. The maximum amount of underestimation of S was about 50 ms across subjects, and thus very comparable to the base-rate formed by our previous work (typically 40-50 ms). However, this frequency difference was only about 1.5 semitones, and one may wonder what occurs to the phenomenon of time-shrinking when the frequency difference is much larger. Hence, we planned to boost the frequency difference by more than two octaves, a difference which gives rise to streaming in longer sequences. This time we decided to include control trials, that is, patterns with markers of equal frequency. It is namely conceivable that the large variation of marker frequencies diminishes the amount of shrinking considerably as compared to the small variation in Experiment 2. If that is the case, it is desirable to dispose of an internal base rate.

4.1 Method
4.1.1 Subjects
Four male students of the Kyushu Institute of Design participated. They had normal hearing.
4.1.2 Stimuli and design
In the control conditions we employed three empty standard time intervals (S): 120, 200, and 280 ms. These intervals could be marked by 20 ms sine tones of 600 and 1,800 Hz, including rise and fall times of 3 ms. We used 600-600, 1,800-1,800, 600-1,800, and 1,800-600 Hz as marker types. Thus, there were 3 (S-durations)$\times$4 (marker types)$= 12$ different control patterns. In the experimental patterns, the preceding time interval (N) was always 120 ms. The duration of S, immediately following N, was 120, 200, and 280 ms. The $/N/S/$-patterns were also marked by 20 ms sine tones (including rise and fall times of 3 ms) and had frequencies of 600-600-600 Hz, 600-600-1,800 Hz, 600-1,800-600 Hz, 600-1,800-1,800 Hz, 1,800-600-600 Hz, 1,800-600-1,800 Hz, 1,800-1,800-600 Hz, and 1,800-1,800-1,800 Hz, respectively. Thus, in total there were 3 (N/S-patterns)$\times$8 (marker types)$= 24$ experimental patterns. The comparison intervals (C) were delimited by pulse-series tones of 500 Hz, of 8 ms duration, which were low-pass filtered at 3.5 kHz. The sound levels of the the 600, 1,800, and pulse series tone were 73.3, 72.5, and 73.0 dB(A) respectively as measured by a sound level meter (Brüel & Kjaer 2209) mounted with an artificial ear (Brüel & Kjaer 4152) and microphone (Brüel & Kjaer 4143).
4.1.3 Apparatus and procedure
The control (S) and experimental (N/S) standard stimulus patterns were recorded onto an opto-magnetic disk (Akai DD 1000). A Commodore (Amiga 500) computer triggered the opto-magnetic disk recorder to play back these patterns, and generated the comparison (C) sound stimuli. The time lapse between the last marker of the control and experimental stimulus patterns and the onset of the comparison pattern was about 3.5 s. Stimuli were routed via an amplifier (JVC AX-Z511) to the left shell of the headphones (AKG K141). The subjects were required to adjust the (subjective) duration of C to that of S by means of manipulating the computer mouse in the shorten and the lengthen pane on the computer screen. They were allowed to hear as many presentations as they felt necessary to be satisfied with the match between S and C. In half of the trials, the initial C sounded clearly longer (descending), and in the other half clearly shorter (ascending) than S. One block comprised of the 12 control and the 24 experimental patterns, all in an ascending and descending fashion, so there were $36 \times 2 = 72$ trials, which were randomly presented. The subjects did three blocks divided over 9 sessions which were taken individually in a sound attenuated booth. Each session started by two warm-up trials.

4.2 Results and Discussion

The first block was considered as training and was not included in the analysis. We calculated the median of the remaining 16 PSEs for each of the control and experimental conditions. Table 2 gives the median PSEs for the control conditions. It turned out that the listeners were quite able to approach the veridical empty durations of 120, 200, and 280 ms by their matches; the constant errors (CE) were small, and only in three out of the 12 control conditions, the PSE differed significantly from the POE. It should be emphasized that this capability of nearly veridical matching holds also for empty intervals marked by different frequencies (600–1,800, and 1,800–600 Hz).

Table 3 displays the median PSEs for the experimental /N/S/ - patterns. When S was 200 ms, thus S − N = 80 ms, huge underestimations of S arose. Six out of the eight underestimations were significant, according to sign-tests. When S was

280 ms, none of the median PSEs of S differed from that of the control condition. When S was 120 ms, two of the PSEs of S were significantly smaller than in the corresponding controls, but the extents of these underestimations (−4.5 ms and −8 ms) were far smaller than when S was 200 ms.

The main point, however, is that the data clearly show that time-shrinking also occurs when the marker frequencies differ. Experiment 2 already showed this to be the case for markers differing between 1,000 and 1,080 Hz (about 1.5 semitones), and Experiment 3 convincingly demonstrated that time-shrinking is that robust that it even emerges when marker frequencies differ by more than two octaves, a frequency interval that would give rise to streaming at these tempi, when there had been more tones in the sequence (Bregman, 1990; van Noorden, 1975). In other words, even though the pitches of the sound markers may have resided in potentially different streams, the mechanism of time

Table 2 Median points of subjective equality (PSEs) in Experiment 3 for the control conditions, in which the duration of a single empty standard time interval (S) had to be matched as for duration, as a function of marker frequency pattern.

Marker frequency pattern	Duration of standard time interval (ms)		
	120	200	280
600–600	121.5(1.5)	205.5(5.5)*	287.5(7.5)**
600–1,800	124.5(4.5)	201.5(1.5)	298.0(18.0)
1,800–600	117.5(−2.5)	206.5(6.5)	286.0(6.0)
1,800–1,800	115.5(−4.5)	197.0(−3.0)	299.5(19.5)**

The constant error (CE = PSE−POE ms) is given in parenthesis. Asterisks denote a significant CE (* = 5%, ** = 1%).

Table 3 Median points of subjective equality (PSEs) in Experiment 3 for the experimental conditions, in which the duration of a single empty standard time interval (S), preceded by a time interval of 120 ms, had to be matched as for duration, as a function of marker frequency pattern.

Marker frequency pattern	Duration of standard time interval (ms)		
	120	200	280
600–600–600	117(−4.5)**	142(−63.5)**	273(−14.5)
600–600–1,800	116.5(−8)	175.5(−26)	315(17)
600–1,800–600	114.5(−3)	168(−38.5)*	314(28)
600–1,800–1,800	107.5(−8)*	175.5(−21.5)**	279.5(−20)
1,800–600–600	116.5(−5)	174(−31.5)**	291.5(4)
1,800–600–1,800	114.5(−10)	178(−23.5)	316.5(18.5)
1,800–1,800–600	122.5(5)	147.5(−59)**	294(8)
1,800–1,800–1,800	113(−2.5)	143(−53.5)*	288(−11.5)

The difference between the experimental PSE and the control PSE in ms is given in parenthesis. Asterisks denote a significant difference (* = 5%, ** = 1%).

perception was not much affected by streaming. Evidence for this fact was presented already by ten Hoopen and Akerboom (1982), and ten Hoopen, van Meurs, and Akerboom (1982). By means of a stop-reaction time paradigm, these authors showed that the perceptual durations of empty time intervals between non-alternating tones and frequency-alternating tones hardly differed.

Nevertheless, it is worthwhile to inspect to what extent the amount of time-shrinking is affected by the marker pattern structure. Table 3 shows a general tendency that the significance of the underestimation of S was larger when the markers that delimited S, had the same frequency. We suspect that S stands out perceptually as a stronger Gestalt when its sound markers are equal, and is therefore temporally processed in a better way, also when S is subject to time-shrinking. To find support for this conviction, we inspected the amount of variability of the PSEs, expressed by the statistical index "semi-interquartile range" $((Q_3-Q_1)/2)$. Our data showed that this index was smallest, that is, matching accuracy highest, when the frequencies delimiting S were equal, and that this index was higher, that is, matching accuracy lower, when the frequencies differed. Our interpretation in terms of Gestalt-strength seems to be related to the classic psychophysical finding that discrimination between two empty time intervals deteriorates when the delimiting sound markers differ as regards spectrum (*e.g.* Divenyi and Danner, 1977).

5. CONCLUDING REMARKS

In our previous research, the stimulus patterns in which we invoked time-shrinking always comprised of two contiguous empty time intervals, separated by three short sound markers of equal frequency. The present results clearly show that the last empty time interval of the pattern was equally well shrunk perceptually when more preceding empty neighboring intervals were added. The phenomenon is also robust to considerable variation of sound marker frequency. Although the amount of time-shrinking was largest when the sound marker frequencies were identical, there was still a huge amount of underestimation of the last time interval when the markers were separated by more than two octaves.

These empirical results seem to shed light on a recurrent phenomenon in music and speech production: the phenomenon of stretching the last element in the sequence in time, even though "temporal rules" prescribe equal durations throughout the complete sequence. The elongation of the last element in the act of production is countered, as it were, by the perceptual act of time-shrinking. Although we did not systematically vary cross-cultural aspects, it should be clear from our present and previous studies that time-shrinking is a genuine perceptual mechanism: it is part of Dutch and Japanese perceptual "make up" equally well. In that sense one also might speak of robustness.

ACKNOWLEDGEMENTS

REFERENCES

Bregman, A. S. (1990). *Auditory Scene Analysis* (MIT Press, Cambridge, MA).
Bregman, A.S. (1995). "Constraints on computational models of auditory scene analysis, as derived from human perception," J. Acoust. Soc. Jpn. (E) **16**, 133–136.
Divenyi, P. L. and Danner, W. F. (1977). "Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra," Percept. Psychophys. **21**, 125–142.
Getty, D. J. (1975). "Discrimination of short temporal intervals: A comparison of two models," Percept. Psychophys. **18**, 1–8.
Handel, S. (1989). Listening: An Introduction to the Perception of Auditory Events (MIT Press, Cambridge, MA).

Hirano, K., Nakajima, Y., Ihle, A., and ten Hoopen, G. (1995). "On the perception of very short rhythm patterns," Proc. Spring Meet. Jpn. Soc. of Music Perception and Cognition, Fukuoka, 36–37.

Israeli, N. (1930). "Illusions in the perception of short time intervals," Arch. Psychol. **113**, 1–47.

Kato, H. and Tsuzaki, M. (1998). "Evidence for functional differences between rise and fall markers in discrimination of auditory filled durations," J. Acoust. Soc. Jpn. (E) **19**, 73–76.

Nakajima, Y. (1987). "A model of empty duration perception," Perception **16**, 485–520.

Nakajima, Y. and ten Hoopen, G. (1988). "The effect of preceding time intervals on duration perception," Proc. Autumn Meet. Acoust. Soc. Jpn., 381–382.

Nakajima, Y., ten Hoopen, G., and van der Wilk, R. G. H. (1991). "A new illusion of time perception," Music Percept. **8**, 431–448.

Nakajima, Y., ten Hoopen, G., Hilkhuysen, G., and Sasaki, T. (1992). "Time-shrinking: A discontinuity in the perception of auditory temporal patterns," Percept. Psychophys. **51**, 504–507.

Nakajima, Y., ten Hoopen, G., Sasaki, T., Yamamoto, K., Kadota, M., Simons, M., and Suetomi, D. (in press). "On the mechanism of time-shrinking," (Percept. Psychophys).

Sasaki, T., Nakajima, Y., and ten Hoopen, G. (1998). "Categorical rhythm perception as a result of unilateral assimilation in time-shrinking," Music Percept. **16**, 201–222.

Taguti, T. and Suzuki, N. (1995). "Percussive rhythm with claves in rumba —A measurement," J. Acoust. Soc. Jpn. (E) **16**, 37–39.

ten Hoopen, G. and Akerboom, S. (1982). "The perceived tempi of coherent and streaming tone sequences: II," Percept. Psychophys. **32**, 481–485.

ten Hoopen, G., van Meurs, G., and Akerboom, S.(1982). "The perceived tempi of coherent and streaming tone sequences," Percept. Psychophys. **31**, 256–260.

ten Hoopen, G., Hilkhuysen, G., Vis, G., Nakajima, Y., Yamauchi, F., and Sasaki, T. (1993). "A new illusion of time perception—II," Music Percept. **11**, 15–38.

van Noorden, L.P.A.S. (1975). "Temporal coherence in the perception of tone sequences," Unpublished PhD dissertation, Technische Hogeschool Eindhoven, the Netherlands.

Yamada, M. and Tsumura, T. (1998). "Do piano lessons improve basic temporal controllability of maintaining a uniform tempo?" J. Acoust. Soc. Jpn. (E) **19**, 121–131.