

Theoretical feasibility of suppressing offensive sports chants by means of delayed feedback of sound

Sander J. van Wijngaarden^{a)} and Johan A. van Balken
TNO Defence, Security & Safety, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands

(Received 3 July 2006; revised 24 April 2007; accepted 24 April 2007)

A novel approach for disrupting offensive chants at sporting events is proposed, based on attacking synchronization between individuals. Since timing is crucial for coordination between chanters, disruption of timing is expected to be effective against undesired chants. Delayed auditory feedback is known to disrupt timing in individual sound production. It may be expected to have similar effects on groups of chanters. To test this hypothesis, a controlled laboratory study was carried out. This showed that the timing of individuals joining in with sports chants can indeed be severely disrupted by also presenting an artificially delayed version of this chant (distracter). This effect is reduced as an individual is given more cues (direction, fidelity) to differentiate between original chant and distracter. However, informal field trials showed that it may be hard to exploit the perceptual effects discussed here for countering offending sports chants in a real-life setting, particularly due to feedback distortion at the required high levels. © 2007 Acoustical Society of America.

[DOI: 10.1121/1.2740054]

PACS number(s): 43.66.Lj, 43.75.Rs [DOS]

Pages: 436–445

I. INTRODUCTION

Chanting is a common phenomenon at sporting events. Many sports chants reinforce a positive atmosphere (Armstrong and Young, 1999), but others can be highly offensive. Attempts are often made to disrupt offensive chants, usually by means of masking sounds such as loud music (McRobb Calder, 2002). Earlier, we have proposed an alternative approach, which has the potential of being effective at lower sound levels (Van Wijngaarden and Van Leeuwen, 2006; Van Wijngaarden and Van Balken, 2006). This approach is based on attacking synchronization between individuals. Since timing is crucial for coordination between musicians as well as in speech (e.g., Jungers *et al.*, 2002), disruption of timing is expected to be effective against chants. A mechanism known to affect an individual's timing of speech and music is delayed auditory feedback.

When a person's speech is fed back to this person at a certain delay, for instance by means of headphones, then effects on speech production are noticed (Lee, 1950; Yates, 1963). Depending mostly on the delay time, delayed auditory feedback may induce a reduction of the speaking rate, skipping or repeating of words and syllables, or even total confusion. Speech production is influenced by one's perception of one's own speech; introducing time delays disrupts this natural feedback process. Similarly, timing of music performances is affected by delayed feedback (Pfordresher and Palmer, 2002).

Delayed auditory feedback is primarily known to lay audiences (if at all) as a party trick sometimes used in television game shows, but is also used in a more serious professional context, such as in stuttering therapy (Kuniszky-Jözkowiak, 1996). Under certain delayed feedback

conditions, many stutterers are able to speak more fluently. Applications are also found in clinical audiology (Tye-Marry, 1992). Delayed auditory feedback can be used to determine whether children who have been given cochlear implants are able to decode information that is becoming available through these implants; they will then show the "normal" response to delayed auditory feedback, such as reduction of the speaking rate through prolonging of speech sounds.

In groups, delayed feedback may have slightly more complex effects. Presumably, each individual in a group producing synchronous speech or music does not only use his own speech as a feedback signal, but the joint speech signal from the entire group; otherwise accurate synchronization would be impossible to achieve. Also, the combined contributions of other group members will normally have a much higher sound level than just one's own speech, with the possible exception of small groups of just a few people. In other words, timing performance of each member of the group affects all other members of the group. For a group to maintain an overall acceptable synchrony and rhythm, a certain process of social self-organization is needed. Introducing delays between different contributors to the overall sound may have surprising consequences. Pairs of musicians, clapping rhythms together while listening to a delayed version of each other's sound, were found to show a deceleration of the clapping rate that seems consistent with classical effects of auditory feedback (Chafe *et al.*, 2004). However, at relatively short time delays, a surprising increase in clapping rate was observed.

Rhythmic applause by audiences of music performances has been studied using models taken from statistical physics (Néda *et al.*, 2000). After a well-received musical performance, a period of a few seconds of incoherent applause is often followed by a phase of accurately synchronized clapping. The transition between these two phases can be understood by modeling the audience as a set of globally coupled

^{a)}Electronic mail: sander.vanwijngaarden@tno.nl

oscillators. The condition for synchronization is that the dispersion (statistical spread in timing) between group members must be smaller than a certain critical value. Hence, one would expect disruption of coordination between chanters (due to delayed feedback channels) to be potentially effective in undermining synchronization, by increasing dispersion between group members. This can only be true if timing of individual chanters is, in fact, affected by delayed feedback from other chanters. The purpose of the current study is to test this hypothesis.

Whether or not a suppression technique based on delayed sound feedback can be effective in practice will also depend on other (largely nonacoustic) factors. Practice has proven sports chants to be remarkably resistant against suppression by external influences. Sociologists view sports chants as a collective expression of social-cultural identity (Armstrong and Young, 1999). Comparisons to primitive rituals to prove one's masculinity, or even ritual warfare, have even been drawn (Bromberger, 1993). In all, it seems safe to say that very potent measures are needed to successfully suppress a chant once sports fans have set their hearts on singing it. Our approach, which is essentially based on subconscious, involuntary perceptual mechanisms, is appealing since chanters may be expected to find it hard to devise counterstrategies.

Two laboratory experiments were carried out in order to determine the conditions under which delayed feedback affects individuals in a group (Secs. II and III). These experiments place participants, one by one, in a group of simulated fellow chanters (based on digitally manipulated recordings). The participant tries to chant along with this group. Since the group is simulated, any arbitrary delay can be added to the feedback signal presented to the participant.

Although these laboratory experiments do not recreate the dynamic, interactive group process observed in real life, it does give insight into the behavior of individuals in reaction to delayed contributions from other chanters. It allows us to test the formulated main hypothesis: Can chanting individuals be disrupted by delayed feedback of other chanters' voices? Additionally, two less formal field trials took place, which are briefly presented in Sec. IV.

II. EXPERIMENT 1: EFFECT OF DELAY AND DISTRACTER SIGNAL LEVEL

A. Goal of the experiment

The primary goal of experiment 1 was to test the hypothesis that timing of individual chanters can be adversely affected by presenting them with delayed versions of the contributions by other chanters. Also, optimum time delays and signal-to-noise ratios were investigated. Subjects were presented with "target" and "distracter" chanters, and instructed to synchronize with the target chanters, while ignoring the distracter chanters. Sound level and delay of the distracters was varied.

Additionally, the question whether or not individual chanters experience distracter chanters to be disturbing was investigated.

B. Method

1. Participants

Eight paid volunteers, all male students with an interest in sports, took part in this experiment. They varied in age between 19 and 33 years (mean 23 years). They completed a questionnaire specifically inquiring about hearing impairments, speech impairments, and linguistic disorders. None of the subjects reported abnormal speech or hearing.

2. Stimuli

Participants were instructed to sing along with four different chants (see the Appendix). These chants were derived from relatively well-known Dutch football chants, and would generally not be considered offensive. The target signal consisted of high-quality recordings of a (synchronized) group of seven chanters (six males, one female). The distracter signal was another recording of the same seven chanters; because of between-session differences, subtle differences between target and distracter signals existed. The target signal was presented with zero delay, and at a fixed signal level of 80 dB(A) at the position of the participant's head.

Depending on the test conditions, the distracter signal was presented with time delays ranging from 0 to 900 ms, and signal levels ranging from 6 dB below the target signal level to 15 dB above the target signal level. In other words, the target-to-distracter ratio ranged from +6 to -15 dB.

3. Apparatus and procedure

Participants were placed in an anechoic room, in a standing position. The target signal as well as the distracter signal was presented from a loudspeaker at a distance of 1.40 m straight in front of the participant, at the approximate height of the participant's ears. Subjects wore a lightweight boom-microphone which was used to record their own chanting during the experiment.

Before starting the actual experiment, subjects were given several practice runs to get acquainted with equipment and procedures. The experiment was divided into 3 min runs of singing the same chant, structured in the following way:

- (1) 12 s of singing along without disruption (no delay of the distracter signal).
- (2) 45 s during which either the signal level or the delay of the chanter was increased (depending of the current test condition).
- (3) This whole pattern (12 plus 45 s) was repeated three times, leading to a total of nearly 3 min.

Two different types of test conditions were presented: runs during which the delay was constant and the distracter level increased, versus runs during which the delay was increased at a fixed level of the distracter. The conditions included in experiment 1 are given in Table I.

It appears likely that the effect of delayed feedback will not only depend on the delay and distracter sound level at the time of measuring the effect, but also on the way these parameters varied in the moments before measurement. This should become apparent by investigating differences be-

TABLE I. Conditions included in experiment 1. Each condition was repeated three times per subject. The presentation order was counterbalanced across participants and repetitions, in order to cancel out any effects of presentation order on the mean results.

	Chant	Delay (ms)	Target-to-distracter ratio (dB)	Number of conditions
Baseline	1-4	0	-	4
Fixed delay	1-4	200	-15 to 3	4
Fixed delay	1-4	300	-15 to 3	4
Fixed delay	1-4	400	-15 to 3	4
Fixed delay	1-4	600	-15 to 3	4
Fixed delay	1-4	800	-15 to 3	4
Fixed target-to-distracter ratio	1-4	0-900	6	4
Fixed target-to-distracter ratio	1-4	0-900	0	4
Fixed target-to-distracter ratio	1-4	0-900	-6	4
Fixed target-to-distracter ratio	1-4	0-900	-12	4

tween the two types of conditions (fixed delay and increasing distracter level, versus increasing delay and fixed distracter level).

Subjects were instructed, and frequently reminded, that their objective was to maintain synchrony with the original target chant. A small financial bonus was promised (and given) to subjects who were able to synchronize well, in order to provide additional motivation to maintain synchrony.

4. Data analysis

Two different measures were extracted from each condition: a subjective rating of the degree of disruption, and an objective synchronization measurement. The subjective rating, on a 1-5 scale, was given by each participant upon completion of each 3 min run. The subjective rating reflects the overall judgment of the entire run, some parts of which may have been more disturbing than others.

The synchronization measure was derived from the boom microphone which recorded the participant's chanting. This measure was derived using a correlation technique based on the amplitude envelope of the signal, which was calculated from the recorded audio and resampled to 200 Hz. This was done for each tested condition, but also for a normalization condition in which the participant was allowed to synchronize as accurately as possible, without any disruption. The two envelopes (test condition and optimally synchronized condition) were compared by calculating a correlation coefficient for every segment of 2 s. A correlation coefficient of 1 indicates perfect synchronization of the segment, a value close to 0 indicates no (time) relation between the segments. Hence, each minute of recorded signal resulted in 30 measures reflecting the degree of synchronization at that particular moment. These measures were used for further analysis in three ways.

(1) Calculating the proportion of accurately timed segments.

In order to decide whether a 2 s segment was accurately

timed (synchronous) or not, a threshold on the measured correlation coefficients of 0.70 was used. This threshold was chosen somewhat arbitrarily; it corresponds to the 75th percentile when correlating different instances of the same, undisturbed but somewhat sloppily timed chant. Other choices for the threshold between 0.50 and 0.90 were verified to produce similar results.

(2) *Calculating the required time shift to restore synchronization.* Another measure of synchronization is how much adjustment (time shift) the chant contribution by a participant would need in order to be optimally synchronous with the target. If a participant's timing is ahead or behind relative to the target chant, then the correlation with the target will be reduced; it can be calculated how much time shift is needed to reach an optimal correlation.

(3) *Direct statistical analysis on mean correlation coefficients, after normalization.* Methods (1) and (2) are useful to gain qualitative insight into the data, but do not easily allow statistical analysis. For this reason, the correlation coefficients measured from the participant's chant contributions were also considered more directly, i.e., by calculating average correlation coefficient values and associated standard errors for each delay and target-to-distracter ratio. Before averaging, an additional normalization was applied to the correlation coefficients. Even if a chant is repeated several times under perfect conditions, the correlation between successive instances of the chant will always be less than optimal, i.e., smaller than 1. The reason for this is that perfect timing is nearly impossible to achieve. However, since difficulty of timing may vary between chants, it is desirable to account for this natural variability. This is done by dividing all correlation coefficients by the average correlation between two undisrupted instances of the same chant.

C. Results

The results from experiment 1 are presented in three sections. First, we show the general pattern of responses within typical test runs. This section serves to create a qualitative impression of the way that chants become disrupted by the distracter signal. The second section contains an analysis of objective synchronization data for each condition, averaged across subjects. This gives an impression of the magnitude of the disruptive effects, as a function of delay and distracter sound level. The third section presents subjective disruption ratings, showing the degree to which participants are consciously aware of (and suffering from) disruption attempts.

1. Synchronization within a test run

Figure 1 shows the average proportion of appropriately timed (synchronous) 2 s segments, as a function of time, averaged over a series of typical test runs. In this example, the target-to-distracter ratio was fixed at -12 dB while the delay increased gradually from 0 to approximately 900 ms.

During the first 12 s the target and distracter were still synchronous. This leads to a relatively high proportion of accurately timed segments, although some initialization ef-

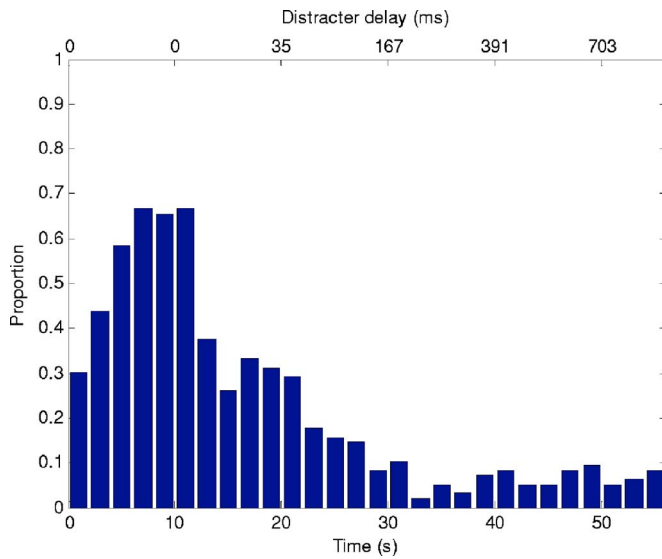


FIG. 1. (Color online) Proportion of accurately timed chant segments, as a function of position within a test run ($N=96$; 8 participants, 4 chants, 3 repetitions). The target-to-distracter ratio for this condition was -12 dB. The delay of the distracter was 0 ms during the first 12 s, then increased up to approximately 900 ms by slowing the distracter down at a rate of 0.5%/s.

fects appear to be present in the first 6 s. Once an (initially small) delay of the distracter is introduced, the proportion of accurately timed segments decreases rapidly.

Figure 2 is based on the same condition as Fig. 1, but this time the required time shift to restore optimal synchronization is shown.

Figure 2 shows some erratic data points at any moment during the test run, including during the undisrupted beginning of the run. However, as a delay is introduced to the distracter chant, the data points appear to fall apart into two distinct groups: one following the target chant (around a delay of 0 ms), and one following the distracter chant, at an

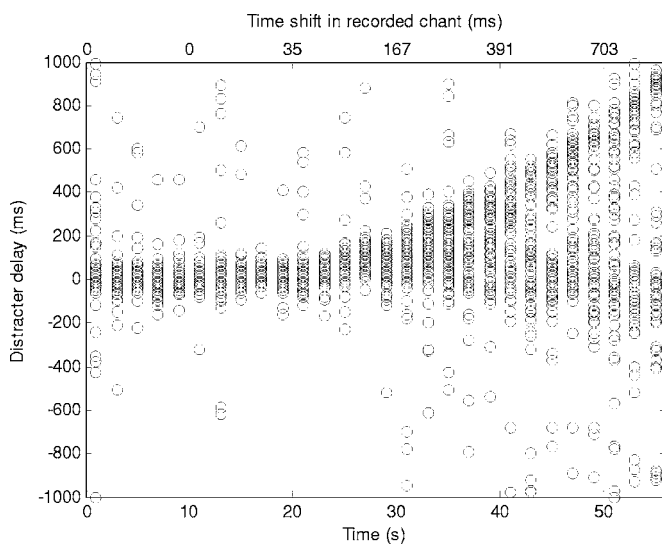


FIG. 2. Time shift (delay) of the participant's recorded chant for which timing correlates optimally to the target chant. Data from each participant, chant, and repetition are shown as separate data points ($N=96$; same condition as Fig. 1).

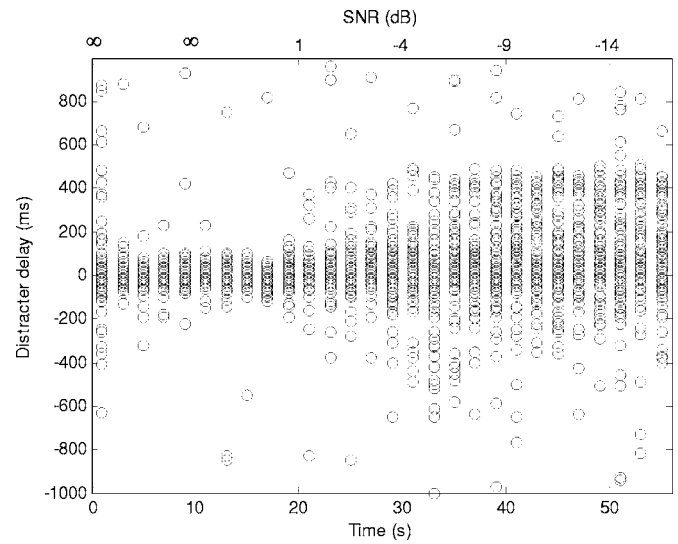


FIG. 3. Time shift (delay) of the participant's recorded chant for which timing correlates optimally to the target chant, at a fixed distracter delay of 400 ms and an increasing distracter loudness ($N=96$; target-to-distracter ratio decreasing at 0.5 dB/s).

increasing delay. These are not specifically between-participant differences; all participants showed both types of behavior.

A different pattern is observed for conditions where the distracter delay was fixed and the distracter level was increased gradually; see Fig. 3 for a typical example. In this case, the distracter was introduced after 12 s at a delay of 400 ms. The target-to-distracter ratio was brought from infinity to $+3$ dB in about 1 s, and decreased gradually (0.5 dB/s) after that. In such cases (as the example shown in Fig. 3), data points are not only clustered around the target and distracter delay (at, respectively, 0 and 400 ms), but distributed more or less evenly between those values. This indicates a greater degree of uncertainty on the part of the participant; instead of strictly following the timing of either the distracter or the target, participants appear to be switching back and forth. Hence, gradually introducing delay leads to more predictable behavior on the part of the participants; if they are at all affected by the distracter, they are likely to attempt to synchronize to it.

2. Synchronization results for each tested condition

Mean normalized correlation coefficients of all conditions with fixed target-to-distracter ratios are given in Fig. 4.

Figure 4 shows that the correlation increases with target-to-distracter ratio. In other words, distracters are more disruptive if their level is higher. There appears to be a ceiling effect toward a target-to-distracter ratio of 0 and $+6$ dB; in order for the distracter to be disruptive, its level must be higher than the target sound level. Also, the correlation decreases (the distracter is more disruptive) as the delay increases.

Figure 5 gives similar results as shown in Fig. 4, but this time for conditions with a fixed distracter delay. Individual delayed auditory feedback, as used in classic experiments mentioned in Sec. I has an optimum effect at a delay time around 300 ms. There does not appear to be such a clear

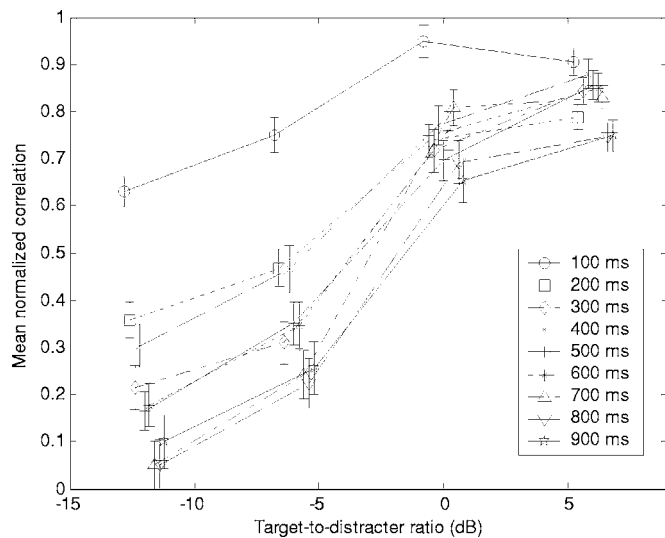


FIG. 4. Mean normalized correlation coefficient (correlation during test run divided by correlation without distracter) at fixed target-to-distracter ratios of -12 , -6 , 0 , and 6 dB, averaged around instantaneous delays as indicated in the legend, ± 50 ms. The error bars indicate the standard error ($N=96$).

optimum for the delay time in this case, although a local minimum in correlation may be present around 300–400 ms. In general however, and especially at lower target-to-distracter ratios, longer delays tend to be more disruptive.

A four-way analysis of variance (ANOVA; delay, target-to-distracter ratio, chant, repetition) was carried out on all data of Fig. 4 as well as Fig. 5. Both ANOVAs lead to the same results: significant ($p < 0.001$) main effects for delay and target-to-distracter ratio, and also significant main effects ($p < 0.05$) for chant and repetition. The significant effect for the type of chant can be understood by considering the difference in timing difficulty between chants. Different repetitions lead to different results because of learning and initialization effects: The first repetition shows consistently lower correlations, which can be understood by assuming that it takes subjects some time to get adjusted to the rhythm of the

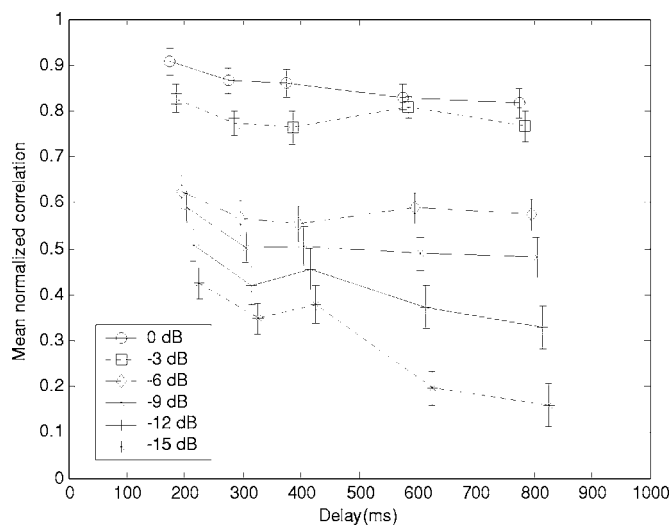


FIG. 5. Mean normalized correlation coefficient at fixed distracter delays of 200, 300, 400, 600, and 800 ms, averaged around instantaneous target-to-distracter ratios as indicated in the legend, ± 1 dB. The error bars indicate the standard error ($N=96$).

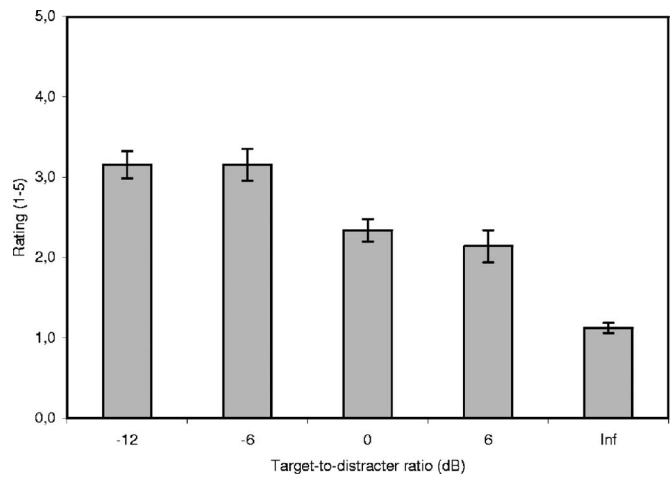


FIG. 6. Mean subjective degree of disruption experienced by participants, as a function of target-to-distracter ratio, rated on a scale from 1 (no disruption) to 5 (absolute disruption). Results are averaged across eight participants and four chants ($N=32$). The error bars indicate the standard error.

chant. Because the experiment was counterbalanced across repetitions and participants, the influence of this learning effect on the mean results for the other variables will be averaged out.

There is also a significant interaction ($p < 0.001$) between delay and target-to-distracter ratio. This can be understood by inspecting Figs. 4 and 5: longer delays are more effective, but this effect also appears to be enhanced by distracter level. Another significant interaction ($p < 0.05$) occurs between delay and chant. Chants that are more difficult to time are more profoundly affected by longer delays.

When comparing Figs. 4 and 5, it is noticed that any combination of delay and target-to-distracter ratio results in higher correlations in Fig. 5 than in Fig. 4. It appears that a gradually increasing delay is more detrimental to proper synchronization than increasing the distracter level at a fixed delay.

3. Subjective disruption ratings

Average disruption ratings are given in Figs. 6 (fixed target-to-distracter ratio) and 7 (fixed delay).

The subjective experience of disruption follows the same patterns as the objective synchronization data: Longer delays lead to more disruption, without a clear optimum delay time. Upon introduction of delayed feedback, the disruption rating increases from 1 (indicating no disruption) to an average of around 3, indicating moderate disruption. It should be kept in mind that this is an average rating, pooled across various (easy and hard) conditions. The hardest conditions received individual ratings of 4 or 5 by all participants.

D. Discussion

The results from experiment 1 show that timing and synchronization of an individual chanter is indeed adversely affected by a delayed contribution. Individual chanters also experience disruption subjectively. Longer delays tend to be more effective in disturbing synchronization; Delayed contri-

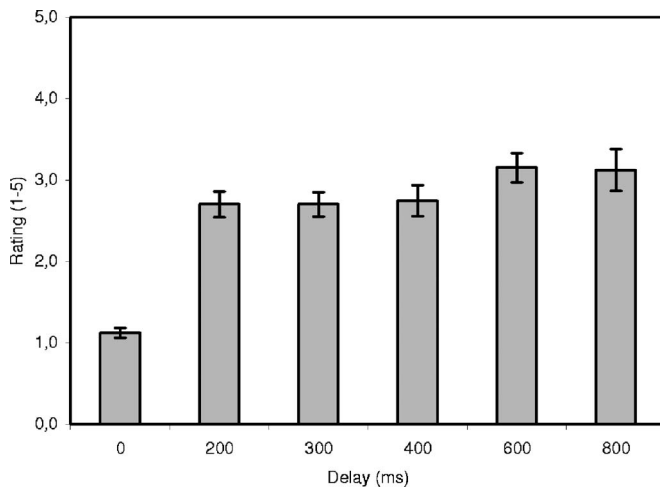


FIG. 7. Mean subjective degree of disruption experienced by participants, as a function of distracter delay, rated on a scale from 1 (no disruption) to 5 (absolute disruption). Results are averaged across eight participants and four chants ($N=32$). The error bars indicate the standard error.

butions must also have a sufficiently high level (higher than the original chant) to be disruptive. Timing is affected differently by a distracter that gradually increases in sound level, than by a distracter of which the delay is gradually increased. In the latter case, participants are likely to synchronize to the distracter signal instead of the original chant. If the delay is not introduced gradually (but instead the level increases gradually) the timing is more erratic and hard to predict. Overall, the effect of a gradually increasing delay is greater than the effect of a fixed delay at a gradually increasing distracter level.

III. EXPERIMENT 2: INFLUENCE OF DIRECTION AND SIGNAL FIDELITY

A. Goals of the experiment

In experiment 1 subjects were given very few cues to distinguish between target and distracter signals. In fact, since the direction and signal quality were the same, the main cue was the fact that the distracter was introduced later, after adaptation to the original chant.

If feedback signals are to be used for chant suppression in real stadiums, then there will always be additional cues for distinguishing between the original chant (sung “live” in a stadium) and artificial feedback signals. The effects of two such cues, direction of the sound sources and signal fidelity, were investigated in experiment 2.

B. Method

1. Participants

A group of eight male participants, none of whom also participated in experiment 1, took part. The participant group was matched to experiment 1 in terms of background and age. Scores on a set of nine reference conditions (see Sec. III B 3) that were identical between experiments 1 and 2 showed no significant between-group differences in the participant groups.

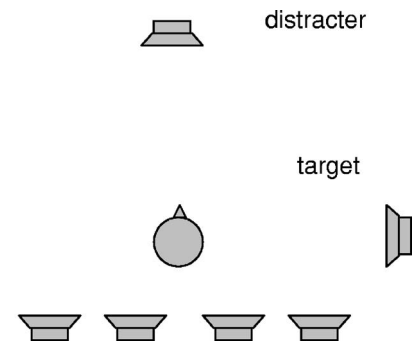


FIG. 8. Schematic representation of loudspeaker locations for presentation of the target chant in experiment 2.

2. Stimuli

The same recordings were used as in experiment 1. Additionally, a “low fidelity” version of the recorded chants was created by digitally introducing bandwidth limiting and non-linear distortion components similar to the characteristics of a specific horn-type loudspeaker (Philips LBC 3403) sometimes used in open-air public address systems. Calculated in octave bands, the frequency transfer is nearly flat in the 500, 1000, and 2000 Hz bands, and approximately 8 dB lower in the 4000 Hz band. Other octave bands do not contribute to the useful sound (i.e., produce no sound at all, or only distortion components).

3. Apparatus and procedure

In experiment 1, the target chant as well as the distracter chant were always presented using a loudspeaker in front of the participant. In experiment 2 (described in this section), the distracter chant was still produced by that same loudspeaker, straight in front of the participant. However, the way that the target chant was presented was differed between conditions. A group of seven loudspeakers was used to create various versions of the target chant (see Fig. 8).

The loudspeakers configuration of Fig. 8 was used in three different ways: (1) the front speaker was used for target as well as distracter (replicating experiment 1); (2) the loudspeaker off to the right-hand side of the participant was used to present the target chant; and (3) a different recording of a single chanter was presented over each of the seven loudspeakers, surrounding the participant with voices of simulated fellow chanters.

Condition (3) comes closest to what one might expect in a real-life setting. Condition (1) replicates experiment 1, albeit for just a subset of the conditions tests in experiment 1. The overlap between experiments 1 and 2 consists of a group of nine “reference conditions,” which allow for comparison between the participant groups. Four conditions featured a fixed target-to-distracter ratio (6, 0, -6, and -12 dB) at a varying delay. For five conditions, the delay was fixed (200, 300, 400, 600, and 800 ms), and the target-to-distracter ratio varied. In all cases, only chant 1 was used. Post-hoc tests did not reveal statistical differences between the participant groups in these nine reference conditions.

The conditions included in experiment 2 are given in Table II.

TABLE II. Conditions included in experiment 2. Each condition was repeated three times per subject. The presentation order was counterbalanced across participants and repetitions, in order to cancel out any effects of presentation order on the mean results.

	Chant	Delay (ms)	Target-to-distracter ratio (dB)	Target loudspeaker	Distracter fidelity	Number of conditions
Baseline	1	0	-	Front	High	1
Reference conditions Fixed Delay	1	200, 300, 400, 600, 800	-12 to 3	Front	High	5
Reference conditions Fixed target-to-distracter ratio	1	0-900	6, 0, -6, -12	Front	High	4
Low fidelity Fixed delay	1	200, 300, 400, 600, 800	-12 to 3	Front	Low	5
Low fidelity Fixed target-to-distracter ratio	1	0-900	6, 0, -6, -12	Front	Low	4
Side loudspeaker Fixed delay	1	200, 300, 400, 600, 800	-15 to 3	Side (90°)	High	5
Side loudspeaker Fixed target-to-distracter ratio	1	0-900	6, 0, -6, -12	Side (90°)	High	4
All loudspeakers Fixed delay	1	200, 300, 400, 600, 800	-15 to 3	All	High	5
All loudspeakers Fixed target-to-distracter ratio	1	0-900	6, 0, -6, -12	All	High	4

C. Results

1. Effects of signal quality

Correlation results for the different signal quality conditions (low fidelity and high fidelity) are given in Fig. 9. Results were subjected to ANOVA analysis, similar to the results of experiment 1.

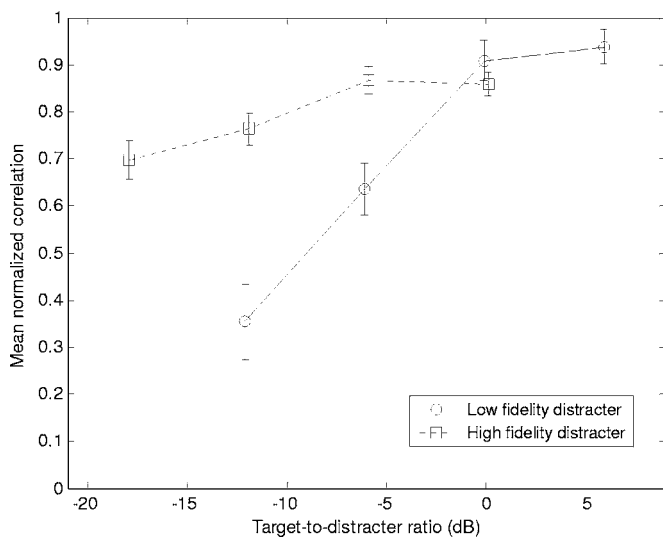


FIG. 9. Mean normalized correlation coefficient at fixed target-to-distracter ratios of -18, -12, -6, 0, and 6 dB, averaged across instantaneous delays between 200 and 800 ms. The error bars indicate the standard error ($N = 24$).

Figure 9 shows that the distracter signal is much less disruptive if it can be distinguished from the target signal by differences in signal quality, although there is a significant effect of target-to-distracter ratio on synchronization ($p < 0.05$) even in the low-fidelity condition. The distorted signal leads to a correlation value at -18 dB that is similar to the correlation for the undegraded signal at 6 dB. In other words, the degraded distracter has to have 12 dB higher level to be equally effective. In reality, the difference may be smaller; ambient stadium acoustics, as well as interactions between speech signals from many chanters, will reduce the perceived quality of all chant signals compared to the “hifi” signal used in the laboratory experiment. The difference between target and distracter in the laboratory experiment is artificially large. However, the fact that signal quality cues may play an important role is clearly shown by this experiment.

2. Effects of sound direction

Conditions at various (fixed) target-to-distracter ratios were again considered, now with the target chant originating from different sources (in front, to the side, and surrounding the participant). Results are given in Fig. 10.

Again, synchronization is significantly affected by raising the distracter signal level. Clearly, directional cues derived from the target signal are helpful in distinguishing between target and distracter. The “group” condition with seven loudspeakers comes closest to the situation that one would

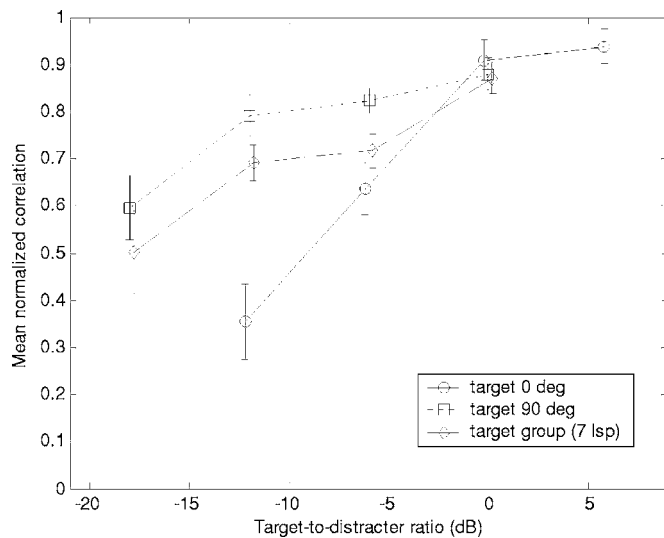


FIG. 10. Mean normalized correlation coefficient at fixed target-to-distracter ratios of -18 , -12 , -6 , 0 , and 6 dB, averaged across instantaneous delays between 200 and 800 ms. The error bars indicate the standard error ($N = 24$).

find in a true stadium setting, simulating a row of chanters behind the participant, as well as chanters to the side and in front. This condition allows for a less effective disruption by a delayed distracter than the conditions tested in experiment 1.

D. Discussion

Whereas experiment 1 showed that delayed feedback can be an effective disruptor of synchronization, experiment 2 makes it clear that real-life limitations, such as directional restraints and limits on signal quality, have a degrading effect on the effectiveness of the proposed technique of undermining sports chants. The extent to which the effect would be reduced in a real stadium setting cannot be determined from this experiment; the stylized laboratory experiments are too artificial compared to the situation in a stadium. Also, it is unclear how various cues differentiating between target and distracter would combine into an overall effect. However, the fact that directional cues and signal quality cues will play a noticeable role is adequately demonstrated.

IV. FIELD TRIALS

A. Need for field experiments

The above-described laboratory experiments have shown that chant timing may be disrupted by playing back delayed versions of a chant, and that the effectiveness of this approach depends on (at least) delay time, signal level, signal quality, and sound direction. However, the question whether a system based on delayed feedback can be effective against real-life chants, in real stadiums using large-scale sound systems, remains unanswered. Next to perceptual and (electro-) acoustic considerations, psychological aspects (such as adaptive group behavior) will determine whether this is the case. Only large-scale field testing will provide a definitive answer.

Translation of the feedback approach into a real system will involve an electro-acoustic setup, involving directional microphones (for picking up the original chant signal from the crowd) and directional loudspeakers. The sound cannot be fed back to the exact microphone position, or feedback distortion (howling, whistling) will occur. The simplest way to tackle this complication is to use multiple feedback loops, and to point each loudspeaker away from its own microphone. Each loudspeaker targets a separate section of the crowd, using the source signal from another section. If this works as planned, then each section may be presented with a different version of the distracter chant. However, since all of these feedback loops are placed in a common acoustic space, feedback distortion is still an issue; this is partly due to surface reflections and partly due to technical limitations to the directivity of microphones and loudspeakers. In practice, this means that the feedback signal cannot have a higher level than a certain maximum, which may or may not be high enough for the distracter chant to be sufficiently disruptive.

B. Results from preliminary field trials

While full-scale validation in field experiments was beyond of the scope of the study reported here, two preliminary field trials were organized to gain insight into complications that arise in practice. Primarily, these trials aimed to answer the question whether simple, commonly used sound systems can be used to produce a sufficiently disruptive feedback signal in the environment of a real stadium. Limits imposed on the distracter level due to feedback distortion were expected to be a major issue.

The first trial used a group of 350 participants, divided into three sections (associated with three feedback loops). Sound from the first section was presented to the second (adjacent) section, sound from the second section to the third, and sound from the third section to the first section. Various delay times (including fluctuating delays) were tested, as well as various loudspeaker setups. A single Bose 802b loudspeaker was used for each feedback loop; hence one loudspeaker was used for each section. Various microphone types (ranging from omnidirectional to highly directional) were tested. The participants (students, averaging 17 years of age) were given freedom to select chants at will.

The second trial was similar in setup, but with a group of only 115 participants (students, average age 19), divided into only two sections. In this case, two loudspeakers were used for each section (one in front of the crowd, and one behind the crowd). Also, a greater degree of structure was imposed on the trial (fixed chant lyrics, synchronized start of the chant, each participant standing exactly at a fixed position).

Both trials led to the same conclusions. It was found that the timing of subjects could be affected to such a degree that the intelligibility of the chant as a whole was subjectively somewhat reduced, according to the participants' own interview responses. However, participants could not be confused into discontinuing their chanting, and could not be split up into distinct groups sticking to different timing standards. A probe microphone, placed at ear-height in the middle of the

crowd, recorded A-weighted sound pressure levels ranging between 96 and 103 dB during chanting. At best, the level of the feedback signal in the middle of the crowd was high enough to achieve a target-to-distracter ratio of around 0 dB; higher levels resulted in such a degree of feedback distortion that this itself becomes a strong cue to differentiate between target and distracter sound.

Depending mostly on the loudspeaker configuration, up to 32% of participants in the 115 participant trial reported to experience disruption (indicated, according to instructions given beforehand, by hand-raising upon finishing the chant). These were consistently the subjects closest to the loudspeakers.

C. Conclusions from the preliminary field trials

The simple setup tested in the preliminary field trials will not be sufficiently effective to suppress chants in practice. The maximum levels of feedback sound that can be achieved are simply not high enough. Based on the laboratory results, the sound level of the feedback level should exceed the level of the original chant. This cannot be achieved with the type of sound system (in particular, the type of loudspeaker) used in the field trials. A more complex system, featuring highly directional loudspeakers and so-called electronic feedback destroyers, may result in higher maximum levels of the distracter signal. Whether this would allow for a more effective disruption of unwanted chants remains to be proven.

V. GENERAL DISCUSSION AND CONCLUSIONS

We showed that the timing of individuals joining in with sports chants can be severely disrupted by also presenting an artificially delayed version of this chant (called distracter): sometimes individuals synchronize to the distracter, but in other cases their timing is poorly correlated with the original chant as well as the distracter. If the distracter sound level is sufficiently high, and the delay is long enough, synchronization is affected. In addition to this (objective) effect, subjects also subjectively find the chants significantly harder to sustain. However, these effects are considerably reduced if the individual is given more cues (direction, fidelity) to differentiate between original chant and distracter.

Delayed auditory feedback is classically found to be optimally disruptive at delay times around 300 ms, and less effective at longer delays. Although our data suggest a local optimum around 300–400 ms in some cases, increasing the delay beyond 400 ms makes the feedback signal even more disruptive. In the case described here, the production-perception feedback loop is not actually closed within the individual, since the confusing, delayed signal stems from an external source. The difference in optimum delay (or rather, the lack thereof in the investigated range of delays) can be seen as an indication of differences in processing during individual sound production versus sound production in a group setting.

Although the laboratory experiments did yield sufficient proof of the perceptual principles exploited by the proposed chant disruption approach, the system tested in a set of pre-

liminary field trials was not sufficiently effective. At least one reason for this could be identified: The maximum sound level of the distracter signal that could be achieved without unacceptable feedback distortion (i.e., feedback distortion becoming an additional, strong cue to differentiate between target and distracter) was not high enough. More advanced sound systems are likely to allow higher levels of feedback sound. Whether this will be enough to effectively suppress chants will also depend on other factors that have not yet been studied, such as (mostly nonacoustic) interactions between individuals in groups, and their effectiveness in devising—yet unidentified—counter-strategies.

APPENDIX: LYRICS OF FOOTBALL CHANTS USED IN EXPERIMENTS 1 AND 2 (IN DUTCH)

Chant 1

Hij is een onbenul
Hij is een onbenul
Hij is een onbenul
Hij is een onbenul

Chant 2

Alles of niets olé olé
Alles of niets olé olé
alles of niets
alles of niets
Alles of niets olé olé

Chant 3

Come on PSV
 (repeated continuously)

Chant 4

Heerenveen maak die goal
Maak die goal
Maak die goaloal

- Armstrong, G., and Young, M. (1999). "Fanatical football chants: Creating and controlling the carnival." *Sport in Society* 2(3), 173–211.
- Bromberger, C. (1993). "Fireworks and the ass," in *The Passion and Fashion: Football Fandom in the New Europe*, edited by S. Redhead (Aldershot).
- Chafe, C., Gurevich, M., Leslie, G., and Tyan, S. (2004). "Effect of time delay on ensemble accuracy," *Proceedings of the International Symposium on Musical Acoustics (ISMA2004)*, Nara, Japan.
- Jungers, M. K., Palmer, C., and Speer, S. R. (2002). "Time after time: The coordinating influence of tempo in music and speech," *Cognitive Processing* 1-2, 21–35.
- Kunizyk-Jözkowiak W. (1996). "A comparison of speech envelopes of stutterers and non-stutterers," *J. Acoust. Soc. Am.* 100(2), 1105–1110.
- Lee, B. S. (1950). "Effects of delayed speech feedback," *J. Acoust. Soc. Am.* 22(6), 824–826.
- McRobb Calder, R. (2002). "Method of crowd control," UK Patent application GB 2 368 958 A.
- Néda, Z., Ravasz, A., Brechet, Y., Vicsek, T., and Barabási, A-L. (2000). "Self-organizing processes: The sound of many hands clapping," *Nature (London)*, 403, 849–850.
- Pfordresher, P. Q., and Palmer, C. (2002). "Effects of delayed auditory feedback on timing of music performance," *Psychol. Res.* 16, 71–79.

- Tye-Murray, N. (1992). "Young cochlear implant users' response to delayed auditory feedback," *J. Acoust. Soc. Am.* **91**(6), 3483–3486.
- Van Wijngaarden, S. J., and Van Leeuwen, D. A. (2006). "Method and means for counteracting undesired sound utterances of a group," International Patent Application PCT/NL2005/000668.
- Van Wijngaarden, S. J., and Van Balken, J. S. (2006). "Suppression of sports chants through delayed feedback of sound," *J. Acoust. Soc. Am.* **119**, 3332.
- Yates, A. J. (1963). "Delayed auditory feedback," *Psychol. Bull.* **60**(3), 213–232.