Perception and judgement of whispered vocalisations

Jasmin Cirillo [1] & Dietmar Todt

(Institute of Biology/Behavioural Biology, Free University of Berlin, Haderslebener Str. 9, *12163 Berlin, Germany)*

(Accepted: 27 January 2005)

Summary

Whispering is regarded as a close-contact vocalisation which, in structural terms, clearly differs from normal (phonated) speech. Here, we present the first experimental evidence for specific functional differences that additionally exist between these two forms of human vocal communication. Such evidence was collected by an inquiry into the perception and also the social judgement of whispered vocalisations. Subjects were young adults (mainly students; n=200) who were exposed to auditory stimuli which, for exclusion of verbal effects, were given in artificial vocal patterns only. To test for possible social effects, our stimuli (whispered phrases or, for control, phonated phrases) simulated exposures to three socially different situations: 'monologues', 'dialogues', and 'dialogues including laughter'. Evaluation of self-report data collected after each stimulation revealed that only the whispered stimuli received significant numbers of socially negative judgements or votes, that documented 'feelings of social segregation'. Such judgements were most frequent after exposures to 'dialogues including laughter', but less frequent after 'monologues'. Taken together our study suggests that whispering is not just a mere close-contact vocalisation, but a vocal expression with specific social side-effects. To explain these effects we presume that whispering signals primarily a strong affiliation of its users, and that such role may make other listeners feel to be excluded from a whispered interaction.

*Keywords:*whispering, unvoiced speech, nonverbal communication, affiliation and bonding, ingroup-outgroup effects

[1] Corresponding author: email address: nimsaj@zedat.fu-berlin.de

Introduction

Whispering is wide-spread across human cultures and - like laughing, crying and screaming - regarded as an universal form of vocalisation (Jensen, 1958; Bachorowski & Owren, 2001; Cirillo, 2001). However, whereas crying and screaming (Newman, 1985; Todt, 1988) and also some vocal precursors of laughing (Provine, 1996; Vettin et al., 1999; Kipper & Todt, 2002) can occur in non-human primates as well, whispering seems to be an uniquely human accomplishment (Pinker, 1995; Weissenborn & Höhle, 2000). Whispering, however is not a regular component of human communication, and is applied rather seldomly. There are hints suggesting that several factors, such as correlates of its production and perception or even social convention, may account for its rare occurrence (Cirillo, 2003). One of the aims of our study was to further clarify these factors and also elucidate some seemingly strange properties related to the social role of whispered vocalisations.

The majority of prior research on whispering concentrated on either its production or the physiology of its perception, and uncovered clear structural differences between this vocalisation (synonymous term: unvoiced speech) and normal speech (synonymous term: phonated speech). In contrast to phonated speech, whispering is produced by preventing the vocal fold from vibrating (Mansell, 1973; Tsunoda et al., 1997). Thus, the passing air does not generate any fundamental frequency, but just a turbulent noise. During several subsequent processes this 'noise' is used to generate vowels and consonants (Monson & Zemlin, 1984; Tsunoda et al., 1994). Therefore and also because parts of some formants are raised in pitch (Kallail & Emanuel, 1984; Traunmuller & Eriksson, 2000), whispered vocalisations can clearly be distinguished from patterns uttered in the normal phonated manner (Tartter, 1991; Tsunoda et al., 1997; Higashikawa & Minifie, 1999).

Until recently, communicative aspects of whispering and aspects of its social role remained neglected issues. Some early studies were merely descriptive (Miller, 1934; Panconcelli- Calzia, 1955), and the first experimental investigations were conducted only decades later (Hultsch et al., 1992; Tartter & Braun, 1994; Cirillo, 2001). These experiments showed that whispered vocalisations can be used as an expedient tool to mediate basic emotional information (Cirillo & Todt, 2002). Newly the latter finding was supplemented by data of a general inquiry substantiating interesting relationships between a judgement of whispered speech and the quality of a social situation (Cirillo,

2003). In addition, it was documented that whispering is not an appropriate substitute of normal speech. Due to their low sound pressure level and some other structural features (see para above), whispered vocalisations do not reach very far and also can easily be masked by environmental noise (Cirillo, 2004). Therefore they are expedient signals only if applied in close contact to a given addressee.

There is evidence that also many social birds and mammals may make use of soft calls or utter their normal call types with a lowered amplitude if addressing them to a nearby partner (Todt & Naguib, 2000). This has been reported e.g. for the greeting duets of many tropical songbirds (Todt, 1970; Todt & Fiebelkorn, 1980; Farabough, 1982; Levin, 1996) or for the so-called 'grunts' documented for different species of primates (Snowdon & Hodun, 1981; Cheney et al., 1996; Silk et al., 1996; Palombit et al., 1999). Most studies on these matters showed that applications of soft vocalisations were usually confined to interactions among specifically affiliated individuals (Hultsch & Todt, 1984).

When designing the study reported here, we investigated a similar aspect and hypothesised that a performance of whispered vocalisations could signal affiliation between interactors as well. Our hypothesis was supported by the results of both a pilot study and a general inquiry. The pilot study documented that many cases of whispering among partners were accompanied by particular visible signals, such as slightly lifted shoulders connected with a subtle head turn towards an addressee and sometimes also a raised hand shielding the vocalisation from other listeners. The general inquiry, in contrast, suggested a means of distinguishing between private and public uses of whispering. Whereas private use seemed to have positive effects, e.g. mediating tenderness or even supporting the bond of mates, its public application seemed to have several negative consequences, e.g. inducing mistrust or diverting the attention of listeners who were not addressed by the vocalisation (Cirillo, 2004).

To further clarify such social effects of whispering we conducted an inquiry into the mechanisms of its perception and judgement. The study was designed to investigate experimentally whether and when an exposure to whispered vocalisations would induce judgements of being socially segregated or excluded from a given conversation. Furthermore, we wanted to test whether such judgements could be caused by the nonverbal properties of whispering alone. To clarify these issues, adult participants were presented

with different auditory stimuli, and then asked to individually judge them. Stimuli were either whispered or phonated phrases (control) that, for exclusion of verbal effects, were given as vocal patterns which participants could not decode. Additionally, our stimuli simulated exposures to three socially different situations: 'monologues', 'dialogues', and 'dialogues including laughter'. We expected that our experiments would uncover how participants judged the stimuli, and that socially negative judgements could be more probable if participants were not listening to the voice of a single person only, but hearing a simulated conversation or a conversation including a phrase of laughter. This expectation was derived from findings suggesting that laughter itself is a social signal (Bachorowski & Owren, 2001; Cirillo, 2001; Kipper & Todt, 2002), and thus could give rise to possible social effects of whispered stimuli.

Methods

Whispered vocalisations are distinctly (structurally) different from phonated ones. In contrast to normal speech, for instance, they are not based on a fundamental frequency (Tsunoda et al., 1997). Therefore, it was easy to reliably identify and classify vocal productions and vocal patterns serving as stimuli. However, all vocal patterns were nevertheless checked by ear and, if remaining unclear, for control also on the basis of frequency spectrograms (Figure 1).

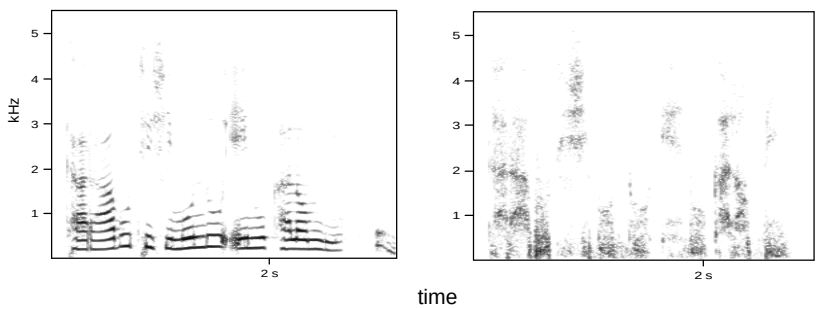

Figure 1. Frequency spectrograms of two stimulus patterns that both represent the same vocal sequence (here: *'karabu tishimundu enu kalombi').* Left: stimulus pattern uttered in a normal phonated manner. Right: whispered stimulus pattern

Subjects and stimulus material

Participants of the study were students of biology (n = 202) with a mean age of 25 (+/- 4) years and a symmetrical gender distribution. None of them showed any hearing deficits.

To prepare the experiments, six individuals (three females, three males) who were not involved in the experiments otherwise, were asked to generate a pool of stimulus material. Each individual contributed to the pool by producing six different samples of stimuli: One sample simulated a 'monologue', another one a 'dialogue', and a third one a 'dialogue containing a phrase of laughter', each given in two renditions: a whispered version and, for control, a phonated version as well. Every stimulus had a length of 20s and was produced in an 'artificial language' (Fig.1). That is, all stimuli were composed of individually invented sound patterns; thus none of our participants could decode any verbal meaning from the material. Samples of different speakers differed in their vocal composition or the quality of the used 'artificial language', respectively. Upon production, all stimuli were sampled by a SONY MZ-R2 minidisk recorder equipped with a Sennheiser Me80 microphone, then marked by a secret label for each speaker and finally incorporated in our pool of test material. 'Dialogues' were sampled by dual-channel recording. This material served as stimuli in two different experiments.

Experiments

In *experiment I* participants (n=104) were presented with a simulated 'monologue' and a simulated 'dialogue', each given in a whispered version and, for control, a phonated version as well. In *experiment II* other participants (n=98) who did not take part in the first one and who also remained uninformed about it, were exposed to whispered and phonated versions of two different simulated 'dialogues': One had the same quality as during *experiment I*, but the other dialogue contained a phrase of laughter.

Aside from such differences, the performance of the two experiments was always alike. That is, before a given test session, a set of four stimuli produced by the same speaker was chosen randomly from the pool of test material. Then, the groups of participants were subdivided into subgroups each comprising about 12 students. Subgroups were invited into a sound-protected room, where they were placed in a manner that allowed us to test them individually, and then exposed to a succession of 4 different auditory stimuli.

Stimuli were presented via headphones, thus subjects could hear each stimulus simultaneously. During test sessions each subject was presented with four stimuli in total (see above). Within a given session, we did not use material received from different speakers. Across different test sessions, however, this material as well as the succession of stimuli was changed randomly.

For the experimental playback of stimuli, we used a minidisk recorder again that here was connected parallel to a set of 14 Sennheiser headphones (12 for subjects + 2 for the experimenters). This allowed for synchronous tests without loss of sound quality. To simulate an almost 'natural' stimulation, stimulus amplitude was kept between 65 dB and 40 dB (1m apart from headphone membrane; see Cirillo, 2003). In addition, stimulus length had been standardised to 20s (see above). According to other studies, this duration was optimal: i.e., long enough to allow subjects to become familiar with a given speaker's voice and to extract sufficient information for a judgement, but at the same time also not so long as to induce stress or habituation (Wiedenmann & Todt, 1992; Grahe & Bernieri, 1999).

Data sampling and analyses

Immediately after each playback, subjects were asked to judge the stimulus by using written questionnaires with bipolar adjective scales of 13 items that in pilot-experiments had proven to be well suited for this purpose. Besides items that, in terms of the studied objective, were 'neutral' and served as simple fillers, the list included one particular pair of items that was considered to be relevant here. It concerned the alternative that the stimuli were estimated as either 'socially segregating' or 'not socially segregating'. Each pair of items allowed to choose among five votes: two votes for each item [e.g.: '-2' (or '+2') representing a strong vote, or '-1' (or '+1') representing a weak vote], or the option '0' indicating a neutral vote or just 'no decision'. With this structure, our questionnaires were in line with criteria outlined by Teske (1989).

In a first step of data analysis, we evaluated the distribution of ratings or votes, respectively, and also the loading of items. In a second step, we compared the data across different subjects and test sessions and then analysed them according to the stimuli. To achieve a ready detection of a particular judgement, we simplified the analytical procedure in one line of analysis. That is, we did not consider differences in item loading, but pooled all votes

addressed to a specific item (here: 'socially segregating'). Then, we compared the received number to another number received by pooling both all votes addressed to the opposite item (here: 'not socially segregating') and all neutral votes. This was done by the judgement coefficient C_J. Values of C_J were calculated according to:

$$C_J = (r - s) / (r + s)$$

s = number of subjects who voted for a specific item (here: 'socially segregating'); r = number of all remaining subjects (i.e., voting for the opposite item or for neutral). Statistical significance of differences between calculated values was tested by a two-tailed ANOVA, and accepted at a level of $p< 0.05$.

Results

Analyses of data on speech perception collected by the item pair 'socially segregating' / 'not socially segregating', and thus concerning judgements of auditory stimuli which simulated different social challenges, revealed clear differences between whispered and phonated vocalisations.

Experiment I provided two results (Fig. 2). First, in contrast to phonated stimuli, whispered stimuli received relatively large number of votes indicating a socially negative judgements or feelings of social segregation. These differences between the judgement of stimuli were highly significant [$F (1, 7) = 307.23$; $p< .0001$]. Second, exposure to a whispered 'dialogue' induced a larger number of votes that indicated feelings of social segregation than exposure to a whispered 'monologue'. The difference between the judgements of these stimulus qualities was also significant [$F (1, 7) = 307.23$; $p< .0001$]. A similar result was not found for the judgements of phonated stimulus qualities [$F (2, 14) = 2.37$; $p = .13$].

Experiment II provided two further results (Fig. 3 & 4). First, implementation of unvoiced laughter into a whispered dialogue had a strong effect on the composition of self-report data. That is, the number of judgements documenting increased 'outgroup'-feelings among participants was remarkably high for whispered stimuli, but not for phonated stimulus versions. Second, this effect was more prominent for the stimulus 'dialogue plus laughter' than for the stimulus 'dialogue' alone. In addition to these clear finding we

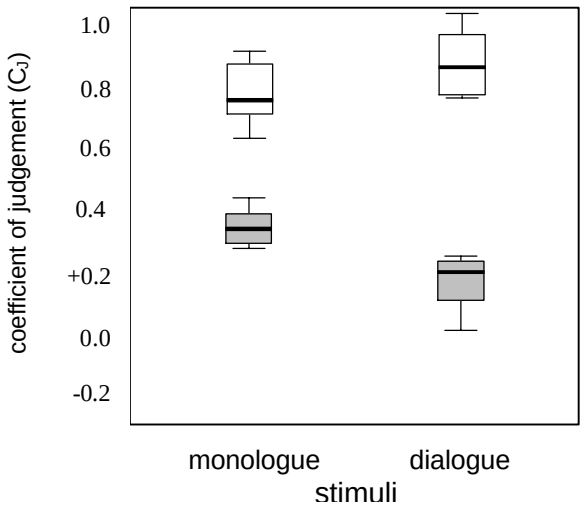

Figure 2. Judgement data received after auditory exposure to a simulated monologue (left) or a dialogue (right). Stimuli were given either in a whispered (unvoiced) version (bottom) or a phonated (voiced) version (top). Boxplots refer to values of the judgement coefficient C_J calculated for a socially relevant item. Note: The more C_J - values differ from '1.0', the stronger a judgement indicating feelings of 'social segregation' (see text).

conducted a further analysis. Here, we investigated the frequency distribution of ratings in more detail and without pooling them, as done by the calculation of C_J. This procedure revealed that the number of subjects who contributed to a high loading of the item 'socially segregating' and who also provided particularly 'strong' votes, was clearly larger for the stimulus 'dialogue plus laughter' than for the stimulus 'dialogue' alone. Again this effect occurred only in the experiments with whispered stimuli [$F(2, 14) = 24.76$; $p < .0001$] but not in those ones with phonated stimuli [$F(2, 14) = 1.86$; $p = .21$] (Fig. 4).

Discussion

Our inquiry into the perception and judgement of whispered stimuli provided the first experimental evidence for specific social effects of this uniquely human vocalisation. In addition, or study showed that these effects could be induced by the acoustic properties of whispered patterns alone, and that so-

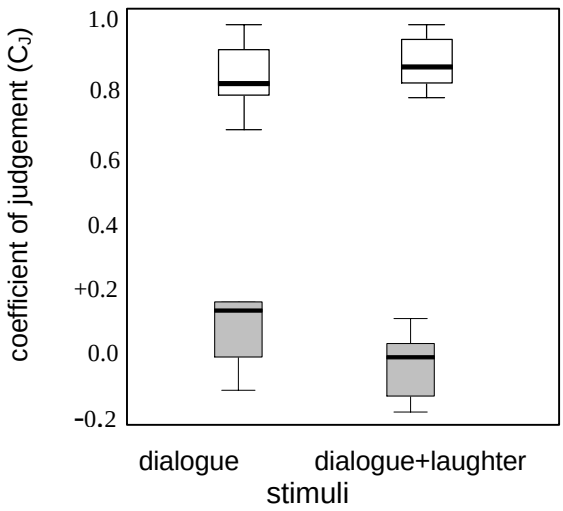

Figure 3. Judgement data received after auditory exposure to a simulated dialogue (left) or a dialogue that included a phrase of laughter (right). Stimuli were given either in a whisper (unvoiced) version (bottom) or a phonated (voiced) version (top). Boxplots refer to values of the judgement coefficient C_J calculated for a socially relevant item (see Figure 3 and text).

cially negative judgements, e.g. votes documenting 'feelings of social segregation', were more frequent if participants had been presented with a simulated conversation than just the voice of a single person. These results bear a number of implications that briefly will be discussed in the following.

Nonverbal aspects of whispering

Nonverbal signals play a crucial role in human social life, and their perception is often used to detect emotional information (Darwin, 1872; review in Cirillo & Todt, 2002), or to make personality judgements about other people (Kramer, 1963; Anisfeld, 1972; Aronovitch, 1976). To avoide such voice-related ‘checks’, the stimulus material applied in our study originated from six different speakers without emotional or exceptional voice qualities and was varied across trials. Within a given trial, however, raters were always exposed to stimuli produced by the same speaker. Therefore, and also because vocal patterns from which raters could not decode any verbal message, our results and especially the differences found between judgements of whispered and phonated stimuli can be

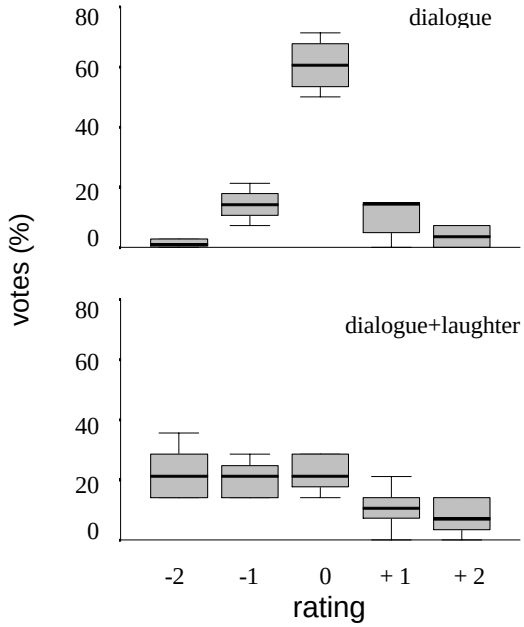

Figure 4. Boxplots documenting the distribution of ratings collected by bipolar adjective scales for a selected pair of items. Ratings were sampled after auditory exposure to either a whispered dialogue (top) or a whispered dialogue that included a phrase of whispered (unvoiced) laughter (bottom). Note: Negative numbers refer to ratings of 'socially segregating' (-2 > -1), whereas positive numbers refer to ratings of 'not socially segregating' (+2 > +1). Data given at '0' refer to neutral ratings or the vote 'no decision'.

explained as genuine consequences of the acoustic differences between these two types of vocalisation.

Besides such prosodic or nonverbal features, also some social characteristics of the test regime turned out to have affected our experimental subjects. For instance, stimuli simulating a conversation of two people (dialogues) were more effective in inducing 'feelings of being socially segregated' than stimuli simulating just a single voice (monologues). We explain these results as side-effects of a more basic role of whispering, which had been predicted by our hypothesis (see 'Introduction', para 5). Accordingly, whispering signals primarily indicate a strong affiliation between users. A similar social function has been reported for the vocal component of laughter: Performed among partners or allies, laughter has clearly positive effects, whereas negative effects can be observed across different groups (Alexander, 1986). These

properties are closely related to the 'contagious effect' (Provine, 1992; 1996) which, in the case of laughter, can be so strong that people who do not join in can be regarded as 'outsiders' (Todt, 1997). As documented by Cirillo (2004), a contagious effect can be induced by whispering, too. In addition, whispering and laughter can share even more properties. This is clear especially for the case of unvoiced laughter which, as reported by Bachorowski and Owren (2001), can induce negative feelings in recipients. Such effects were now confirmed by our results showing that a combination of unvoiced laughter and whispered vocalisation raised the number of subjects who reported feelings of being socially segregated.

How to explain the observed judgements?

By concentrating on data collected specifically with respect to the items 'socially segregating' or 'not socially segregating', our study showed that only whispered stimuli, but not any phonated pattern, induced socially negative votes. The votes documented that a large number of participants neither felt personally addressed by whispered vocalisations nor included in the heard conversations. Currently, there are a few aspects which, at least indirectly, contribute to an explanation of these judgements. Such aspects can be derived, for instance, from findings of a general inquiry suggesting that people who are not addressed may nevertheless listen to the whispering and, e.g. by trying to understand its meaning, may feel disturbed or even socially uncomfortable with the situation (Cirillo, 2004). Additional aspects can be derived from studies that interpreted the social role of whispering by presuming that it serves to communicate secret messages (Hultsch et al., 1992; Eckert & Laver, 1994; Cirillo, 2001) and that such are predominantly addressed to a nearby friend or a partner of the whisperer (Cirillo, 2003).

For a more direct explanation of the observed judgements we like to point, however, to another perspective. That is, we suggest classifying whispering as a vocal display which, upon both performance and perception, is distinguished from normal speech and also could signal a close affiliation of interactors. Such view would predict that whispering primarily serves private forms of communication, and that its public use can thus socially challenge all listeners who have a problem with its address or may regard the use as an impolite habit. This explanation incorporates the aspects listed above and is also in line with several criteria that originally were postulated for the

process of signal generation by ritualisation (Huxley, 1966). Ritualisation is wide-spread across animals (for review see Baerends, 1975) and humans as well (Koenig, 1970; Eibl-Eibesfeldt, 1981). In either case ritualisation is linked with distinct changes in signal structure and performance as well as with a specific functional differences between a 'ritual' and its original (review in Todt, 2003).

If these characteristics are applied to the comparison between normal and whispered speech, the following reasons support the view that 1the latter could be interpreted as a ritualised class of the former. First, there are clear structural differences between the two vocal patterns, e.g. whispering does not built on any fundamental frequency, but just a turbulent noise (Mansell, 1973; Tsunoda et al., 1997). Therefore and also because parts of some formants are raised in pitch (Kallail & Emanuel, 1985; Traunmuller & Eriksson, 2000), whispered vocalisations can clearly be distinguished from patterns uttered in the normal phonated manner (Tartter, 1991; Tsunoda et al., 1997; Higashikawa & Minifie, 1999). Second, whispering is not an appropriate substitute of normal speech, because it, e.g. due to a low sound pressure level, does not reach very far and also easily can be masked by environmental noise (Cirillo, 2004). Therefore its application requires a close contact between interacting individuals. Then, whispering can be accompanied by visible signals, and it shares this characteristic with some social rituals of animals, e.g. the duet displays which signal strong affiliation or even stable bonds between mated tropical birds (Todt et al., 1981; Hultsch & Todt, 1984). Further arguments finally can be derived from an other study which showed that whispering develops only after a good competence in the use of normal speech has been achieved, and that its production requires more effort than that of phonated patterns (Cirillo, unpublished data). The evidence that whispered vocalisations are more arduous than phonated ones, invites to ask whether whispering could be regarded as a 'costly' display, and thus could be considered as a case of honest signalling (Grafen, 1990; Zahavi & Zahavi, 1997).

Biological analogues

To date, whispering is regarded as an uniquely human accomplishment. The outcome of our study suggests, however, that a search for biologically analogous phenomena may nevertheless be both interesting and promising. First,

our study showed that certain effects of whispering were not related to any verbal quality, but induced by acoustic signal properties, and such properties play role in animals, too. Second, explaining whispering as a vocal display which serves particular social functions could stimulate questions about how animals would deal with similar functions.

There is evidence already that several social birds and mammals may make use of soft calls or utter their normal call types with a lowered amplitude if addressing a nearby partner (Todt & Naguib, 2000). Such was documented, for example, for the greeting duets of many tropical songbirds (Todt, 1970; Todt & Fiebelkorn, 1980; Farabough, 1982; Levin, 1996) and also the so-called 'grunts' described for different species of primates (Snowdon & Hodun, 1981; Cheney et al., 1996; Silk et al., 1996; Palombit et al., 1999). Taking the results of our whispering study as a reference, it would be stimulating to now focus the attention of investigators on questions such as: Do individuals who address a nearby mate or ally only reduce the vocal amplitude of their signals or also modify them otherwise? Do other conspecifics observe the interaction of mated individuals? If so, how do they respond to such a communication? Are there hints, for example, that observers or listeners could dislike exposures to such interactions? Currently, there is evidence only for duet performances which mated birds themselves address to other conspecifics, namely territorial neighbours (review in Todt & Naguib, 2000). In contrast to the soft duets performed during the greeting ceremonies of partners, however, these territorial displays are orchestrated by a clearly raised vocal amplitude and additionally exhibited at exposed positions, often accompanied by conspicuous movements or other visible signals (Todt et al., 1981; Hultsch & Todt, 1984; Levin, 1996).

Acknowledgements

We thank Roger Mundry for statistical advice and Henrike Hultsch for valuable comments on an earlier draft of the manuscript. Many thanks also to Nadine Moeller for her assistance during the experiments. J.C. appreciated a fellowship of the Fazit-Foundation.

References

Alexander, R.D. (1986). Ostracism and indirect reciprocity: the reproductive significance of humor. — In: Ostracism, a Social and Biological Phenomenon (M. Gruter. & R.D. Masters, eds.). Elsevier, New York, p. 105-122.

Anisfeld, M. (1972). Social perception of speech. — In: Current Trends in Linguistics, 12. (T.A. Sebeok, ed.), Mouton, The Hague, p. 314-327.

Aronovitch, C.D. (1976). The voice of personality: stereotyped judgements and their relation to voice quality and sex of speaker. — J. Soc. Psychol. 99, p. 207-220.

Bachorowski, J.A.& Owren, M.J. (2001). Voiced but not unvoiced laughter readily elicits positive affect. — Psychol. Science 12, p. 252-257.

Baerends, G.P. (1975) An Evaluation of the Conflict Hypothesis as an Explanatory Principle for the Evolution of Displays. — In: Function and Evolution in Behaviour. (G. Baerends, C. Beer & A. Manning, eds.), Clarendon Press, Oxford, p.83-97.

Cheney, D.L, Seyfarth R.M. & Palombit R.A. (1996). The function and underlying mechanisms of baboon 'contact' barks. — Anim. Behav. 52, p.507-518.

Cirillo, J. (2001). Social aspects of whispered vocalisations. — Advances in Ethology 36, p. 137.

Cirillo, J.(2003). Psychobiological aspects of whispered speech. — PhD Thesis, FU Berlin.

Cirillo, J.(2004). Whispering: communication by unvoiced speech. — Ann. Braz. Acad. Sciences 76, p. 413-423.

Cirillo,J. & Todt D. (2002). Decoding whispered vocalizations: relationships between social and emotional variables. — Proc. Internat. Conf. Neural Inform. Processing, Singapore, p. 1559-1563.

Darwin, C. (1872). The expressions of the emotions in man and animals. — Philosophical Library, New York.

Eckert, H. & Laver, J. (1994). Menschen und ihre Stimmen: Aspekte der vokalen Kommunikation. — Beltz, Flensburg.

Eibl-Eibesfeldt, I. (1981) Stammesgeschichtliche und kulturelle Ritualisierung. — Nova Acta Leopoldina 54, p. 722-729.

Farabaugh, S.M. (1982). The ecological and social significance of dueting. — In: Acoustic communication in birds (D.E. Kroodsma & E.H. Miller, eds.). Academic Press, New York, p. 85-124.

Grafen, A. (1990). Biological signals as handicaps. — J. Theor. Biol. 144, p. 473-516.

Grahe, J.E. & Bernieri, F.J. (1999). The importance of nonverbal cues in judging rapport. — J. Nonverb. Behav. 4, p. 253-268

Higashikawa, M. & Minifie, F. D. (1999). Acoustical- perceptual correlates of whisper pitch in synthetically generated vowels. — J. Speech. Lang. Hear. Res. 42, p. 583- 591.

Hultsch, H. & Todt, D. (1984). Spatial proximity between allies: a territorial signal tested in the monogamous singer *Cossypha heuglini*. — Behaviour 91, p. 286-293.

Hultsch,H., Todt,D. & Zühlke, K. (1992). Einsatz und soziale Interpretation geflüsterter Signale, Umwelt und Verhalten. — K. Pawlik, & K. H. Stapf (eds.), Huber Verlag, Bern, Göttingen, Toronto, p. 391-406.

Huxley, J. (1966) A discussion of ritualisation of behaviour in animals and man. — Phil. Trans. Royal Society London 251, p. 247-256.

Jensen, M.K. (1958). Recognition of word tones in whispered speech. — Word 14, p. 187-196.

Kallail, K.J. and F.W. Emanuel (1984). An acoustic comparison of isolated whispered and phonated vowel samples. — J. Phonetics 12, p. 175-186.

Kipper, S. & Todt, D. (2002). The use of vocal signals in the social play of barbary macaques. — Primates 43, p. 3-17.

Koenig, O. (1970). Kultur und Verhaltensforschung. — Deutscher Taschenbuch-Verlag, München.

Kramer, E. (1963). Judgement of personal characteristic and emotions from nonverbal properties of speech. — Psychol. Bul. 60, p. 408-420.

Levin, R.N. (1996). Song behaviour and reproductive strategies in a duetting wren (*Thyrothorus nigicapillus*). — Anim. Behav. 52, p. 1107-1117.

Mansell P. (1973). An experimental investigation of articulatory reorganisation in whispered speech. — Working Papers, Institut f. Phonetik und sprachl. Kommunikation 2, München, p. 201- 253.

Miller, C.B. (1934). An experimental phonetic investigation of whispered conversations, considered from a linguistic point of view. — PhD Thesis, Univ. Hamburg.

Monson P. & Zemlin W. R. (1984). Qualitativ study of whisper. — F. Phonetica 36, p. 53-65.

Newman, J.D. (1985). The infant cry of primates: an evolutionary perspective. — In: Infant Crying (B.M. Lester & C.F.Z. Boukydis,eds.). Plenum Press, New York.

Palombit R.A., Cheney D.L. & Seyfarth R.M. (1999). Male grunts as mediators of social interaction with females in wild chacma baboons (*Papio cynocephalus ursinus*). — Behaviour 136, p. 221-242.

Panconelli-Calzia G. (1955). Das Flüstern in seiner physiopathologischen und linguistischen Bedeutung. — Lingua 4, p. 369-378.

Pinker, S. (1995). The language instinct (2. ed.). — Harper Perennial, New York.

Provine, R.R. (1992). Contagious laughter: laughter is a sufficient stimulus for laughs and smiles. — Bull. Psychonom. Soc. 30, p. 1-4.

Provine, R.R. (1996). Laughter. — Am. Scientist, p. 121- 128.

Silk, J.B.; Cheney, D.L. & Seyfarth, R.M. (1996). The form and function of post-conflict interactions between female baboons. — Animal Behaviour 52, p.259-268.

Snowdun, C.T. & Hodun, A. (1981). Acoustic adaptations in pygmy marmorset contact calls: Locational cues vary with distances between conspecifics. — Behav. Ecol. Sociobiol. 9, p. 295-300.

Tartter V.C. (1991). Identifiability of vowels and speakers from whispered syllables. — Percept. Psychophys. 49, p. 365-372.

Tartter, V.C. & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. — J. Acoustic. Soc.Am. 96, p. 2101-2107.

Teske, K. (1989). Zur Beurteilung frühkindlicher Schrei-Serien durch erwachsene Zuhörer: Datenerhebung mittels Fragebögen. — PhD Thesis, FU Berlin.

Todt, D. (1970). Die antiphonen Paargesänge des ostafrikanischen Grassängers Cisticola hunteri prinioides Neumann. — J. Ornithol. 111, p.332-356.

Todt, D. (1988). Serial Calling as a Mediator of Interaction Processes: Crying in Primates. — In: Primate Vocal Communication (D. Todt, P. Goedeking & D. Symmes, eds.), Springer-Verlag, Berlin, Heidelberg, p. 88-107.

Todt, D. (1997). The codex of social play: a system of rules that can prevent ingroup violence. — Schriften Rechtspsychol. 3, p. 305-315.

Todt, D. (2003). Rituale als Spielmaterial einer denkbaren Evolution von ästhetischer Erfahrung. — In: Dimensionen ästhetischer Erfahrung (J. Küpper & C. Menke, eds.). Suhrkamp Taschenbuch Verlag, Frankfurt, p. 308-332.

Todt, D. & Fiebelkorn, A. (1980). Display, timing and function of wing movements accompanying duets of *Cichladusa guttata.* — Behaviour 72, p. 82–106.

Todt, D. & Naguib, M. (2000) Vocal interactions in birds. — Adv. Study Behav. 29, p.247-296.

Todt, D. , Hultsch, H. & F. Duvall (1981) Behavioural significance and social function of vocal and non-vocal displays in monogamous duet-singers. — Zoolog. Beiträge 27, p. 422-448.

Traunmuller, H. & Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women and children. — J. Acousic. Soc. Am. 107, p. 3438-3451.

Tsunoda, K., Niimi S. & Hirose, H. (1994). The roles of the posterior thyrophryngeus muscles in whispered speech. — Folia Phoniatr. Logop. 46, p. 139-151.

Tsunoda, K. , Otha Y., Soda Y., Niimi S. & Hirose, H. (1997). Laryngeal adjustment in whispering magnetic resonance imaging. — Ann. Otol. Rhinol. Laryngol. 106, p. 41-43.

Vettin, J., Skirl, J. & Todt, D. (1999). From social play to tickling and laughter: a comparative approach. — Adv. Ethol. 34, p. 37.

Weissenborn, J. & Höhle, B. (2000). Approaches to bootstrapping: phonological, lexical, syntactic and neurophysiological aspects of early language acquisition. — John Benjamins Publishing Company, Amsterdam, Philadelphia.

Wiedenmann, G. & Todt, D. (1992). Discrete responses of adult subjects during exposure to infant cry sequences. — Early Child Develop. Care 65, p. 179-188.

Zahavi, A. & Zahavi, A. (1997). The Handicap Principle. — University Press, Oxford.