

# 6



## GENERAL DISCUSSION

MY INVESTIGATION has uncovered several properties of whispered speech that now allow specifying the communicative role of this verbal accomplishment. My results on this matter pertain (bear) a number of implications which merit to be discussed in more detail, and thus will be treated in the following discourse. A main emphasis will be on 'social aspects' and on 'psychobiological aspects' of whispering. In order to highlight some new relationships between the results, the treatment of these aspects will meet an order which, as compared to the succession of chapters in the preceding part; however, it will be reversed. In this framework, I will first review significant findings of other investigators and only then discuss how my results relate to this knowledge.

### Psychobiological Aspects

The term 'psychobiological' is used here in the sense of Immelmann et al. (1988) who applied it to all matters that are fundamental to behaviour, but addressed it particularly to the mechanisms of behaviour, or intrinsic variables, respectively, which underlie the performance of behaviour. In my study, two of such variables, namely 'vigilance' and 'emotion', were investigated in relation to whispered speech. In the subsequent part of this chapter, a short introduction to these two variables will serve as a framework for a discussion of my results.

## Vigilance

Vigilance is a major topic of both psychological and biological research. Its study began about 50 years ago and was concentrated, then, on attentive behaviours of human-beings. Only many years later, it was addressed also to comparable behaviours of non-human organisms (Keverne et al. 1978). In free-ranging animals, namely birds and mammals, vigilance has a high survival value, and this explains why so many animals invest an immense amount of time in being vigilant or watchful and alert, respectively (Caine & Marra 1988; Tomasello et al. 1998). From a biological perspective, vigilance supports an effective detection of risks (e.g. predators) or chances (e.g. prey), and from a social perspective, it can help individuals to better cope with competitors or to improve their contacts to cooperative partners (Warm & Jerison 1984; Davies & Parasuraman 1982). With reference to the latter aspect, it was suggested that vigilance has an adaptive value which contributed to the evolution of cognitive processes and especially of so-called 'social intelligence' (Todt & Brumm 2001; Bachar et al. 1993).

In psychological research, vigilance indicates a state of being wide-aware and open for several different perceptions (Grillon et al. 1997). Some authors treat this as a kind of 'sustained attention' which they, however, at the same time distinguish from a 'focussed attention' that occurs after a distinct matter was identified and found to be interesting enough for a more detailed exploration (Hawk et al. 1992). Although both, vigilance and attention mainly concern the afferent side of an intrinsic data processing, their execution can obviously influence also variables which, e.g. like arousal, are related to the efferent side of such processing. There is evidence that a vigilance-related arousal increase can generate positive feelings which then can operate as a positive reinforcer and finally generate a sort of 'conditioned vigilance' (Todt 1986).

The main portion of knowledge on vigilance comes from studies of *visual* vigilance (Keverne et al. 1978; Hunter & Skinner 1998), whereas auditory vigilance remained a relative neglected issue. Basically very practical reasons have contributed to this situation. In the case of visual vigilance, investigators often have access to conspicuous behavioural cues which are related to its performance. Head and eye movements, for example, are commonly accepted as expedient indicators of visual

vigilance, and their rates or durations can be taken even as measures of its intensity (for citations see: Todt & Brumm 2001). The few studies which, nevertheless, successfully investigated properties of auditory vigilance, achieved this aim by specific methodological manoeuvres, for instance, by using the so-called 'startle response' as an empirical measure (for citations see Hoffmann-Kuhnt 2003).

Auditory vigilance plays an important role in the live of blind people who are believed to be hypervigilant because of their constant auditory monitoring of the environment. However, when being exposed to specific acoustical stimuli, these subjects did not show an increase in startle responsivity compared to sighted individuals (Bachar et al. 1993). Studies with sighted subjects tested in the dark resulted in similar findings (Grillon et al. 1997). The experiments suggested that darkness or deprivation from visual information, respectively, did not enhance the processing of auditory information. On the other hand, if such deprivation was based on an individual's self-control, i.e. if subjects voluntarily closed their eyes, the outcome was different: here, the latency of the startle response was clearly shorter with eyes closed as compared to eyes open (Hawk et al 1992).

My own results on the effect of darkness can be taken as another contribution to these issues. As shown, the decoding of verbal stimuli heard in the dark was a bit more successful than that one achieved under normal light conditions. However, since the difference was statistically not significant, my findings remained in line with the literature (Grillon et al. 1997). In addition, my results did not confirm a raised vigilance level, in this case. To recall, due to my hypothesis I had expected that whispered stimuli increase the vigilance of subjects and hence would be decoded better than phonated stimuli. The fact, that their decoding was even poorer than that of the latter, was explained as a possible consequence of methodological factors. For example, the stimulus amplitude applied here was 25 dB only, and this could have been just too low for a facilitation of such an effect (see also Figure 2.4). I anticipate that an experimental design which would have substituted an exposure to darkness by a voluntary eye-closing of subjects could have improved the effects (see e.g. Hawk et al 1992). This expectation is based on Smith's vigilance theory which suggested the possibility that 'typical

experimental subjects differ not so much in their ability to maintain attention as in their willingness to do so' (Smith 1966).

In contrast to the effects of 'whispering in the dark', other experimental effects provided a clear support for the vigilance-hypothesis (H2: whispering can raise the level of auditory vigilance). Such support can be deduced from my results on an interference between auditory vigilance and visual attention. To recall, auditory exposure to whispered stimuli, but not to phonated stimuli, reduced the success of visual learning task, or the rate of correctly memorized visual stimuli, respectively. I have explained this finding as a consequence of an auditory vigilance which here was strong enough to distract the attention of subjects away from the visual stimuli. Currently, however, this view can not be regarded as a final interpretation. It should be clarified, for example, which part of the chain of processes mediating between stimulus recognition and memory retrieval was specifically affected by the experiment. Based on Miller's theory (1956), I assume that the short-term memory could play a crucial role, in this respect.

The results of my experiment on an interference of auditory and visual attention invite to be compared to the results of dual-task experiments conducted and interpreted by other investigators. Findings of some early studies on this matter suggested that stimuli presented in the same modality can interfere more than stimuli presented in different modalities (Treisman & Davies 1973; Pashler & O'Brien 1993). Such findings have been explained by referring to a hypothesis which predicted that the resources in the intrinsic data processing are especially limited, if data related to the same modality have to be processed (Norman & Bobrow 1975). More recently, however, the view of such modality-specific attentional resources (Warm & Jerison 1984), has been challenged by behavioural data which demonstrated the existence of extensive cross-modal attention-links between audition and vision across a wide range of dual-tasks situations (Driver & Spence 1994; Eimer 1999). For example, Spence and Driver (1997) reported a decrement of 8% in visual spatial discrimination when participants simultaneously had to additionally monitor auditory stimuli. Other studies have provided similar results suggesting that corresponding attentional resources are used to process auditory and visual information, and that such information processing is competing for the same pool of attentional resources (Horswill & McKenna 1999; Spence & Read 2003). The results

of my experiment are in good accordance to this actual view. Nevertheless, there is a further aspect that merits consideration.

This aspect concerns the question, how whispered speech which participants could not decode could have attracted their attention. Currently, I like to explain this question by the predictions of the vigilance-hypothesis cited above. The predictions imply that even a not-decodable expression of whispering should be more effective than just a background-noise, comparable e.g. to music. Listening to music, at least when played at a moderate volume, does namely not impair a performance of visual tasks (Beh & Hirst 1997; Strayer & Johnston 2001). This effect has been explained by assuming that listening to music does not require as many attentional resources as actively listening or conversing (Brodsky 2002).

## Emotion

The term *emotion* labels a category of intrinsic states that - such as anger, fear, sadness or joy - often are mirrored in the display behaviour of individuals, e.g. by modifying particular parameters of facial or vocal expressions (Sobin & Alpert 1999). Besides in humans, emotions can be found in other mammals and birds, too, and they play a significant role in the social life of these organisms. Here, however, the term *emotion* is usually replaced by the term *affect* (Scherer & Kappas 1988). From an evolutionary perspective, affects have two adaptive values: First, they are components that supplement other intrinsic components, e.g. motivations, and thereby contribute to an efficient control of complex behavioural interactions (Jürgens 1988). Whereas motivations facilitate or induce the execution of particular target-directed motor activities, affects are related to the perceptible side of behavioural control (Todt 1986). According to Scherer (Scherer & Kappas 1988) affects are *feelings* that serve as intrinsic *filter*-components and shape an individual's evaluation of perceived stimuli.

The other adaptive value of affects concerns their social role, and these benefits from signals or displays which encode affectional or emotional information via the following three aspects: (a) Group members are often attracted by such displays or at least direct their attention to the signaller (Todt & Brumm 2001). A similar function has been reported also for signals of so-called *learned helplessness* (review in Todt 1988). (b) Group members who correctly decode the information can mutually better predict the potential replies of other group mates. (c) Expressions of special affects or emotions can finally lead to social facilitation and e.g. synchronize the intrinsic state of group mates (Todt 1986). In humans, a well-known example of the latter aspect can be found in the contagious role of vocal laughter (Kipper & Todt 2003). As shown in chapter 4, there is evidence now that a perception of whispering can be contagious, too.

Several investigators have studied the prosodic properties of spoken language that mediate emotions (Scherer & Kappas 1988; Hultsch et al. 1992; Tartter & Braun 1994; Sobin & Alpert 1999). However, the study reported here has provided the first comprehensive comparative approach to this issue, which also included the role of whispered speech. Some exceptions aside (see e.g. 'joy'), the portion of correctly decoded emotional expressions was remarkably large and exceeded that one reported for other studies. I interpret this effect as a consequence of our method which was found to be very expedient already in other investigations (Wiedenmann & Todt 1992). The method allowed subjects to first extract a reference for discrimination, and then to decide among a few (here: five) alternatives, only. Based on these methodological aspects and also on my results, I conclude that the decoding of emotional expressions may be more likely a discrimination process than a process of instant identification.

Besides the role which whispering can play in mediating prosodically encoded emotions, also its role in inducing specific feelings of people merits more consideration than it received in the past. In my study, such feelings have been documented for both addressees and non-addressees of whispering, or co-listeners, respectively. And in most cases, these feelings were not at all neutral, but rather positive or negative, depending on how a given person judged a perceived whispering from a social perspective. The best explanation of how such feelings are generated is provided by Scherer's 'Model of Vocal

Affect Expression' (Scherer & Kappas 1986). The basic assumption of the model is that the evaluation of any stimulus complex which produces an emotional state consists of a sequential process of five stimulus evaluation checks. Here, the differentiation of an emotion is determined by the specific combination of outcomes of these checks. For a processing of socially relevant information, this model would need, however, a few supplementary components. I assume that suggestions of such components can be made, if further experiments with whispered speech are concentrated on this issue.

## Social Aspects

A major part of my study on the communicative role of whispered speech was concentrated on social aspects, and has centred particularly the so-called 'ingroup/outgroup' paradigm (Billig & Tajfel 1973). Such social matters cover a crucial part of our daily life, where they can occur in a variety of different forms. An elucidation of such matters is prosecuted not only from the field of sociopsychology, but also from the side of behavioural biology. The latter discipline investigates the fundamental components of social accomplishments, and thereby contributes to a better understanding of the often highly advanced social achievements of humans (review in Todt 1987). In order to incorporate knowledge accumulated by these interdisciplinary ventures, the subsequent section of this chapter will treat some relevant results of behavioural research first, and only after this focus on issues provided by sociopsychological research. Treatments of either matter will serve as a framework for a discussion of my results on whispered speech.

### The behavioural perspective

The 'ingroup/outgroup' paradigm of humans can be compared to a many similar social phenomena of animals. The most wide-spread form has been described as a 'closed society'. Closed societies can be found in almost all taxa; that is, in mammals as well as in birds and even in non-vertebrates. Honeybees and ants, for example, which often

develop extremely differentiated societies, use particular pheromones to distinguish members of their own group from strangers, i.e. members of a foreign group. Strangers are treated as competitors or even as 'enemies', and thus repelled or even seriously attacked. Among familiar group members, or within a given group, respectively, mutual cooperation is a predominant trait, and in some societies, individuals even may care for one another (Krause & Ruxton 2002).

In birds, the most common society is a pair. Here, a male and a female usually join to cooperate in founding a territory and in raising their offspring. In the tropics, many of such pairs establish persistent bonds that sometimes can be maintained life-long. Persistent bonds are linked with specific rituals that basically include pair-specific vocalisations known as antiphonal duets. Performing the duets is often a sophisticated manoeuvre which requires competence and skill, and thus can be developed and precisely executed only by mates who stood together for an extended amount of time. The duets have two functions: First, they help mates to keep contact and to strengthen their bonds, e.g. by serving as greeting ceremonies. Second, they help mates to defend their territory and to successfully repel co specific intruders (Todt & Naguib 2000). In other words, and in the terminology of the 'ingroup/outgroup'-paradigm, dueting is an 'ingroup'-ritual which, at the same time also serves distinct 'outgroup'-functions (Todt 2003).

Based on this characterization, one could ask whether the whispering of humans could be interpreted as a ritual which has evolved analogously to dueting. Although there are many clear-cut differences between the two classes of displays, two aspects suggest an answer like 'may be'. One aspect refers to the fact that such duet systems are not constraint to birds only. And, remarkably enough, they occur also in those species of non-human primates, that, such as Gibbons and Siamangs, form stable pair bonds, too. The other aspect is related to my own results. As substantiated in chapter 5.2, my findings suggest that whispering can be regarded as an 'ingroup'-ritual which has evolved from normal speech. Currently, like I idea that this ritual was developed to enrich the repertoire of signals which mates or partners can use to communicate their private messages. On the other hand, however, I do not see a parallel between the 'outgroup'-functions of duets and the problems that whispering could cause in the public.



Rather, I like to interpret the latter effect as a consequence of a functionally incorrect use of the ritual (see chapter 5.2, section 'Conclusions').

An aspect that complements the rational described above by an interesting facet can be derived from the social life of non-human primates. A few exceptions aside, these primates form large societies in which the formation of subgroups or coalitions, respectively, is a common strategy. The composition of such alliances can vary and, depending on their particular purpose, include e.g. close relatives or individuals who just know each other very well, but usually profit from assisting one another. The most significant signal strategy to exhibit a given alliance is the performance of close body contacts and extended episodes of mutual grooming (Smuts et al. 1987). In addition, however, there is a further factor which is not specifically communicated, but related to a property termed 'social intelligence'. According to recent studies, this intelligence is connected to vigilance and a persistent mutual observation of group members, and also to an ability to extract and make use of socially relevant information (de Waal 1982; Todt 1997). Given this property one can expect that, provided the primates would develop and use a specific 'ingroup'-ritual, its performance should affect other group members similarly to what in my thesis was documented for a performance of 'whispering in the public'.

## The sociopsychological perspective

Sociopsychology was established basically by the work of Henry Tajfel and John Turner who successfully combined knowledge provided by behavioural research (see 6.2.1) with theories provided by sociology (Bernstein 1981), but additionally developed some new concepts of human social accomplishments. Their most influential concept was the 'Theory of Social Identity' (Tajfel & Turner 1979). This theory postulates four processes as crucial for matters of social integration or segregation, respectively, and predicts that an individual who is exposed to a socially new environment will execute the following accomplishments: (a) She/he will first observe and then structure her/his environment into subgroups. (b) Subsequently she/he will define her/his personal identity in the society, what includes that she/he will join one of the subgroups, at least mentally. (c)

After this, she/he will search for 'stereotypes' which exhibit the contrast between subgroups and thereby allow distinguishing them. (d) Finally, she/he will rank the subgroups and positively appraise her/his own group, but negatively judge the other subgroups. Some terminological aspects aside (Sumner 1906), this concept is now considered as the theoretical foundation of the 'ingroup-outgroup'-paradigm.

In my study on social aspects of whispering, the 'ingroup-outgroup'-concept had served as a framework. And, as shown in chapter 3 especially, this concept allowed an expedient explanation of my results. From the perspective of the theory sketched above (Tajfel & Turner 1979), the whispering-related splitting of social feelings, which I observed among the participants of my experiments, can be seen in new light. That is, such splitting could be interpreted as a reflection or even a consequence of processes of social identification and social segregation, which in a given number of people were initiated by the signal properties of a whispering voice.

There is evidence from other investigations that people make personality judgments about other people based on vocal cues. For example, raters consistently judged certain voices as reflecting various personality characteristics, regardless of whether or not the person whose voice was rated possessed these characteristics (Aronovitch 1976). In particular, pitch and loudness turned out as cues upon which judgments of a speaker's personality were found to be based, but the manner in which these vocal characteristics were employed differed for male and female speakers (Aronovitch 1976). For two reasons, however, such voice-related personality checks could not have played a role in my experiments. First, because I excluded speakers with exceptional voice qualities, and second, because the stimulus material used in my study originated from six different speakers and was varied across trials. Thus, the described results could not be caused by individual voice parameters, but by the quality of whispering only.

Besides whispering, there is an other universal display of humans, that merits to be considered in the framework of the 'ingroup-outgroup'-concept. This display is human laughter (Kipper & Todt 2003; Vettin 2003). Performed within a given group, it has clearly positive effects, whereas negative effects can be observed across groups

(Alexander 1986). These properties are closely related to the so-called 'contagious effect' (Provine 1992; 1996). As documented in chapter 2, this effect is shared by whispering, too (see also Hultsch et al. 1992). In terms of laughter, however, the effect can be so strong that people who do not join-in can be regarded as 'outsiders' (Todt 1997). Interestingly, there is a structural relationship between whispering and laughter. It becomes clear, if people are exposed to unvoiced laughter. Such exposures can induce negative feelings in addressees (Bacharowski & Owren 2001). Now, this effect was documented also by my experiments with unvoiced laughter inserted into a sequence of undecodable whispered words. As shown (chapter 2.2), such stimulus combinations increased the effects of whispering and raised the distinctness of 'outgroup'-feelings in a large number of people.

In daily life, whispering is regarded to signal secrecy and confidentiality (Eckert & Laver 1994). The results of my study do not contradict this view, but rather specify it. Such is true also for the outcome of my study on the ontogeny of whispering. As shown, whispering can be induced already at an age of about two years. Its normal interactive use, however, develops only later; that is at an age of four to five years, when children have developed also a concept of secrecy and privacy (Meares & Orly 1988; Watson & Valtin 1997). Since such relationships are extremely interesting, I would like to complement the approach described and discussed above, by a further clarification of developmental accomplishments.