The effects of verbal and nonverbal elements in persuasive communication: Findings from two multi-method experiments

NIKOLAUS JACKOB, THOMAS ROESSING and THOMAS PETERSEN
E-Mail: nikolaus.jackob@uni-mainz.de; roessing@uni-mainz.de; tpetersen@ifd-allensbach.de

Abstract
This article addresses the relationship between content, voice, and body language in persuasive communication and the contribution of these three elements of persuasive performances to its overall persuasiveness. Findings are presented from two separate laboratory experiments. In the first experiment three versions of a video displaying a speech were shown to three different groups of participants: (1) without vocal emphasis and without gestures of the speaker, (2) with vocal emphasis but without gestures, (3) with vocal emphasis and gestures. Audio tracks of the first two experimental conditions were later used in the second experiment to analyze the effects of vocal emphasis when no visual cues are present. Measurement included a questionnaire as well as Real Time Response-measurement (RTR). It was found that content dominates the effect of the speech; emphasis and gestures, however, improved the perception of some features of the speech, such as liveliness and power. Audio-only versions yielded similar results but were rated more favorably in general.

Keywords: Persuasion, Rhetoric, Nonverbal Communication, Experiment, Real Time Response-Measurement.

Introduction
There are numerous myths circulating amongst both academic researchers and communication experts about the effects of vocal emphasis, facial expression and gestures in persuasive communication, as there is no shortage in ingenious, appealing answers to the question of how communicators can achieve the greatest effect on their public. Political advisors and rhetoric coaches alike boast a myriad of supposedly tried and tested recipes for creating the perfect mix of verbal and non-verbal stimuli; vocal emphasis and verbalization, facial expression and gestures. The myth of the so-called “communication pyramid” is a perfect example of
such an attempt to reduce the effects of persuasive communication to a single rule of thumb: 55 percent of the effect of a speech results from the body language, 38 percent from the voice, and just 7 percent from the content of a speech (Maurer and Reinemann, 2007, p. 320). This formula, a myth proliferated for decades, can be found everywhere, from handbooks, teach yourself manuals and handouts at political, advertising and public relations seminars to training courses for sales representatives. Although in no way proving its authenticity, the fact that this myth has remained in such wide circulation for so long does reveal a strong desire among communication experts and researchers alike for a simple, practical formula that can be applied to central issues such as: In what ways do presentations affect audiences? What is the most effective mode of presentation? And, more specifically: Which part of persuasive communication plays the most significant role in dictating its effect?

Theoretical and empirical background

Existing theory and research

These questions are not new. The special role of the voice and body language in persuasive communication was taken up in works as early as Aristotle’s *Rhetoric* (Ar. rhet. III,1,4–5; see also Knapp and Hall 2005), although Aristotle expressed regrets about the manipulative and theatrical nature inherent in the use of such techniques. It was also clear to Cicero, that along with vocal elements, gestures and facial expressions also had important roles to play in communication: He believed, for example, that facial expression next to the voice has the greatest effect (Cic. or. 60). During the 20th century many new techniques were developed to study the effects of auditory and visual elements in persuasive communication. Even a cursory review of the relevant research literature reveals that the observations made by Aristotle and Cicero have lost none of their relevance. Most research has concentrated on the question of how persuasive appeals can be presented most effectively (Stiff and Mongeau, 2003, p. 190). Hovland and his colleagues, for example, ascertained that (alongside the characteristics of the audience and the message of the speech): “observable characteristics of the perceived source of communication” are, above all, crucial in determining the effect of persuasive communication (Hovland, Janis, and Kelley, 1953, p. 11). These observable elements fall into the category of “nonverbal” stimuli, a group made up of kinetic (body language), vocal (mode of speech and vocal emphasis), spatial (proximity and surroundings) and haptic (touch) stimuli (Burgoon, Dunbar, and Segrin, 2002; see also Knapp and Hall, 2005; Manusov and Patterson, 2006; Bente and Krämer, 2008), a
group which has been extended to include chronenical (timing and context) and iconic (for example, symbols, pictures) stimuli (Cronkhite and Liska, 1980, p. 109).

The vast range of stimuli involved in persuasive communication and the interplay between them makes effects research in this area a very complex undertaking. As a result, a wide variety of approaches has been developed and some very different types of study employed in what has become a very heterogeneous research field (see for an overview e.g. Burgoon, Birk, and Pfau, 1990; see also O’Keefe, 2002; Perloff, 2003; Knapp and Hall, 2005; Manusov and Patterson, 2006). This fact is reflected in the following summary of relevant findings, which brings together studies from such diverse research backgrounds as propaganda research, media effects research, persuasion research and political (election) communication research.

It is a more or less accepted fact that audience perception differs according to the mode of presentation. For example, early propaganda and media effects research approaches found differences in the effectiveness of video, audio and textual modes of presentation (Wilke, 1934; Cantril and Allport, 1935; Knower, 1935; 1936). Audiovisual presentations provide audiences with more (visual) information on the basis of which to draw conclusions — in particular, conclusions about the presenter (e.g. Frey, 1999; Hassin and Trope, 2000; Kepplinger, 2010a) — and therefore have different effects. Research shows that audiovisually accentuated presentations are apt to grab an audience’s attention and focus that attention on the presenter’s characteristics while at the same time tending to divert attention away from the content of a presentation (Bettinghaus, 1973, p. 119; Stiff and Mongeau, 2003, pp. 190–195). There is substantial research evidence for the fact that gestures, facial expressions, and certain forms of vocal emphasis contribute to audience impressions that presentations without non-verbal components do not, such as the nervousness, attractiveness or openness of the speaker; aspects that are not normally considered important for rational, intellectual information processing and, as such, are considered more likely to stimulate emotional reactions (Wagner and Brandstaetter, 1980; Kepplinger, 1987; Kepplinger et al., 1994; Kepplinger and Maurer, 1999; Mauer and Reinemann, 2007; Kepplinger, 2010a). Nonverbal behaviors such as facial expressions and gestures play a role in affective experience and at the same time convey non-semantic information about the personality or personal background of an individual, about its cultural context and its emotional state (e.g. Krauss, Chen, and Chawla, 1996; Burgoon et al., 1990; Marsh, Hart-O’Rourke, and Julka, 1997; Hassin and Trope, 2000; Bente and Krämer, 2008).
Studies on source credibility and/or persuasion respectively have shown that the way in which communication affects an audience also depends on the composition of audio and visual stimuli. Related research produced a great deal of (partly heterogeneous) findings (Burgoon et al., 1990, pp. 146–152; see also Mehrabian and Williams, 1969; Pearce and Conklin, 1971; Pearce and Brommel, 1972; Rauscher, Krauss, and Chen, 1996; Marsh et al., 1997, pp. 563–564; Krauss, 1998; Krauss and Hadar, 1999; Bavelas, 2000; Tusing and Dillard, 2000; Hosman, 2002; Perloff, 2003, pp. 198–201). The following cursory summary refers only to findings relevant for the research context at hand: With respect to the vocal characteristics of a presentation it has been demonstrated that its persuasiveness varies according to the speed the presenter speaks, the pitch variety of the voice, as well as whether or not the content is put across smoothly, fluently, and coherently: fluent speech is more credible than nonfluent speech, greater pitch variety is likely to improve character and competence judgments, rapid speaking seems to enhance competency based assessments. Nonfluencies, hesitations, and monotonous speech, on the other hand, are likely to undermine the persuasiveness of a presentation. In sum, it seems that increased vocal intensity in the form of louder amplitude, greater intonation, greater fluency, and faster tempo are likely to facilitate source credibility and overall persuasiveness. With respect to body language there is some evidence that facial activity and the adequate use of gestures increase perceptions of the credibility and the persuasiveness of a speech. The prevailing view in research is that certain gestures enhance communication by conveying additional information that amplifies and modulates information conveyed in the speech channel. Some authors argue that eliminating gestures leads to a less vivid speech (e.g. Rauscher, Krauss, and Chen, 1996; Krauss, 1998; Krauss and Hadar, 1999): It has been demonstrated, for example, that speakers who are not allowed to gesture have more difficulties in producing speech — their presentations are perceived as less fluent and in consequence as less credible and/or persuasive. Latest research suggests that recipients judge integrated verbal-plus-nonverbal presentations as shorter, clearer, and probably more effective. In sum, there are numerous studies demonstrating that differences in gestures and facial expressions can influence how communicators are perceived (e.g. credibility) as well as the persuasiveness of their communication (Burgoon et al., 1990, p. 164).

However, with regard to modern dual processing models such as the elaboration likelihood model (ELM) or the heuristic systematic model (HSM) (see e.g. Chaiken, 1980; Chaiken and Eagly, 1983; Petty and Cacioppo, 1986; Booth-Butterfield and Welbourne, 2002; Todorov et al.,
2002; Schenk, 2002), it has been argued that the effects of nonverbal cues are limited by the role of audience-involvement in the process of message perception and procession: “For cognitive rather than affective tasks, the verbal channel may be more important (...). Thus, for many persuasion tasks, nonverbal behavior could have a minimal impact. It seems likely that nonverbal behavior will have more influence on attitudes when the topic is more emotionally involving than when the social and emotional importance is minimal” (Marsh et al., 1997, pp. 563–564).

For the study at hand, the framework of the dual processing theories is necessary in order to properly evaluate the methodical setup (e.g., involvement of the experiments’ participants) and the research findings. According to the ELM and the HSM there are two different kinds of persuasion processes, depending on the degree of ego-involvement: Highly involved individuals, on the one hand, engage in systematic thinking, they elaborate the arguments carefully and focus on the quality of communication content; nonverbal signals are less important as factors facilitating persuasion (central route of persuasion). Indifferent or uninterested recipients, on the other hand, tend to process information superficially, the quality of arguments is less important, they base their judgments rather on source characteristics (e.g. the speaker’s attractiveness) or characteristics of the source’s nonverbal performance (e.g. facial expressions or body language) on this, peripheral route, nonverbal signals may be important sources of persuasion effects.

Several variables are likely to have an impact on the individual degree of elaboration: e.g. the relevance of presentation’s topic, personality characteristics of the recipients (such as their need for cognition), the presence of potentially distracting factors, the amount of relevant background knowledge among the auditory, and the prevalence of counterattitudinal arguments. Or as stated by O’Keefe (2008: p. 1476): “So, for example, when a topic is personally relevant to a receiver, the receiver is generally predisposed to engage in careful thinking, the receiver has extensive relevant background knowledge, and the receiver is undistracted, elaboration will presumably be high – and the central route of persuasion will be engaged”. Studies focusing on the biological backgrounds of human information processing show that these different paths are probably rooted in different cerebral structures: It has been argued that linguistic-verbal communication is associated with left hemisphere functioning whereas nonverbal communication is associated with the right hemisphere or both hemispheres, respectively (Buck and VanLear, 2002: pp. 524; 533; 536).
Problems and unaddressed questions of related research

The potentially endless range of ways in which persuasive stimuli can combine in persuasive communication makes it very difficult for researchers to develop studies that provide generalizable findings. Almost all empirical studies related to the role of nonverbal components in communication suffer from serious methodological problems (Burgoon et al., 1990, p. 141; Marsh et al., 1997, pp. 563–564): It has been criticized, for example, that studies on nonverbal persuasion mostly are piecemeal empirical investigations and researchers usually did not link the pieces together – most experimental studies typically manipulate a single nonverbal cue, such as the speed of a speech, examining the separate effects of this single cue instead of investigating the effects of behavioral composites or of multiple behaviors. Furthermore, few studies to date have been able to provide insight into “(...) the relative potency of verbal versus nonverbal cues (...)” (Burgoon et al., 2002, pp. 462–463) and investigated the interplay between the different variables: “(...) the examination of isolated cues prevented relative comparisons among cues as to their importance. We do not know, for example, whether speaking tempo or gestural animation contributes more to persuasive impact (...))” (Burgoon et al., 1990, p. 141) While research has ascertained many of the effects of various modes of presentation, it remains completely unclear whether certain modes of presentation exert stronger effects (Stiff and Mongeau, 2003, p. 190; cf. also Andreoli and Worcel, 1978; Chaiken and Eagly, 1983), as suggested in the so-called communication-pyramid. In fact, no universal hierarchy of effectiveness, attributing stronger effects to certain forms of presentation, has, so far, been confirmed scientifically (Maurer and Reinemann, 2007). And finally, it is criticized that many studies suffer from comparatively “artificial manipulations” (Burgoon et al., 1990, p. 141) under quite artificial circumstances “in the laboratory” (see e.g. Petersen 2006, p. 50) with consequences for the external validity of these studies.

Thus, on the whole, the discussion here shows that there are still some very important questions to be answered. Although, for example, it has been clearly established that the effect of persuasive messages is dependent on the mode of presentation, it remains, nevertheless, unclear how great a contribution verbal and non-verbal elements make to the overall effect of communication in relation to other elements. Furthermore, it is certain that not only visual stimuli but also auditory stimuli, e.g. the speed and flow of a presentation, influence its persuasiveness. However, hardly any studies have investigated whether combined variations in speed and vocal emphasis benefit the persuasive power of messages presented. Given that “(...) nonverbal cues may not be particularly mean-
The effects of verbal and nonverbal elements

It is necessary to design studies manipulating a "(…) a variety of nonverbal cues simultaneously (…)" which "(…) may have a greater impact than manipulating a single cue in isolation" (Marsh et al., 1997, p. 564). There are also very few studies that systematically investigate the relative effectiveness of the types of presentation that are used on a *day-to-day basis* in practice (presentation forms, as such, which are vitally important for establishing the external validity of research findings). It is consistently assumed that a message presented using a variety of vocal emphasis, gestures, and facial expressions will have the strongest possible persuasive effect on an audience and be superior relative to presentation forms that do not employ nonverbal support even though there has been not much rigorous academic discussion of such assumptions in a realistic context on an adequate empirical basis (cf. for example, Andreoli and Worcel, 1978; Burgoon et al., 1990; Maurer and Reinemann, 2007).

It is these unaddressed issues that were taken up in a series of experiments conducted by the authors: The *first experiment* centered on the question of whether a speech including body language as a mode of communication via facial expressions and gestures in combination with vocal emphasis is more persuasive than a speech with vocal emphasis which does not contain much facial expressions or gestures. The experiment also investigated whether these two versions were both more persuasive than a third speech which was comparatively free of facial expressions and gestures and vocal emphasis. The study investigated not only which of the three presentations was more persuasive, but also whether the differences in presentation techniques employed led to different perceptions of persuasiveness. This was done, for example, by analyzing effects profiles to establish whether there were any systematic differences between the speech presented with vocal emphasis and supported by gestures and facial expressions and the speeches presented without particular vocal emphasis and/or support from facial expressions and gestures. A *second, follow-up experiment* investigated whether the differences in the strengths of effects and effect profiles ascertained for the different forms of presentation in the first experiment were replicated if the stimulus was presented in an audio only and not (as was the case in the first experiment) an audiovisual format. Thus, the key goal of the second experiment was to establish whether a message presented with vocal emphasis was more persuasive than one presented without vocal emphasis and whether different patterns were evident in the effects profiles of the two versions of the speech. Furthermore, the findings of the two experiments were analyzed to establish whether the audio versions of the speech in the second experiment displayed any significant differ-
ences to the video versions in the first experiment in relation to the issues under investigation. By choosing this study design the authors try to meet some of the methodological requirements which were pointed out by Burgoon et al. (1990) and Marsh et al. (1997): the study focuses on a realistic combination of nonverbal stimuli, specifically on the variation of vocal characteristics, facial expressions, and bodily gestures simultaneously. Furthermore, it was conducted within a comparatively naturalistic setting employing an integrated “verbal-plus-nonverbal” persuasive presentation. In addition, it documents the application of multi-method-measurements which is potentially able to uncover differences in the overall persuasiveness of the speeches as well as specific differences in the contribution of certain nonverbal stimuli to this overall persuasiveness.

Method
Prior to the two experiments, an initial representative split-ballot survey was conducted by a professional survey research institute to ascertain which of three short speeches exerted the greatest influence on public opinion. In comparison to texts on the topics “European Constitution” and “Driving with headlights on full beam during the daytime” a passage on globalization exerted the strongest effects. The pro-globalization text was selected for use in the experimental part of the study because the data proved it to be effective in changing survey-participants opinions. The content of the text, presented to respondents in written form, placed emphasis on the positive consequences of globalization for Germany.

The first experiment was conducted from November 13 to 21, 2006. Within the framework of a single factorial experimental design, the mode of presentation of the text was varied. A combination of a conventional questionnaire and RTR-measurement devices (Reinemann, Maier, Faas, and Maurer, 2005) was employed to measure the effect of the presentations on the test subjects; 204 students of a large German University participated in the experiment. Using student samples for experimental purposes is a common practice in social research (Daschmann 2004). With regard to the dual processing models (ELM, HSM), however, the student’s ego-involvement, their individual predispositions (need for cognition, pre-existing opinions towards globalization), and their background knowledge concerning globalization may have impacted the experimental outcomes. Therefore, the potential implications of this specific stimulus-audience-constellation will have to be discussed closely (see paragraph 5 “Discussion and Conclusion”).

In all, the experiment consisted of three experimental and three control settings: The first group of test subjects watched a video in which
a professional presenter delivered the speech on globalization without vocal emphasis and without any emphasis via gestures; reactions being ascertained via a questionnaire and RTR-devices. The second group of test subjects watched a video in which the same presenter delivered the same speech, only this time with vocal emphasis, but without any emphasis via gestures (reactions measured as for group 1). The third group of test subjects watched the same presenter delivering the same speech on video, only this time with both vocal emphasis and emphasis via gestures (reactions measured as for group 1). Alongside three questions on demographical characteristics, the questionnaire consisted of three questions on globalization as well as 15 items for assessing the content of the speech or, respectively, the presentation. Measurements in all cases were made using five-interval rating scales. The items were designed to ascertain three dimensions: the perceived performance of the speaker (e.g. vividness, liveliness), the perceived characteristics of the argumentation (e.g. factual accuracy, thoroughness) and the perceived characteristics of the speaker (e.g. credibility, competence). Respondents were given a five-interval scale on which to record their impressions, whereby a higher value (nearer to 5) indicated that a certain characteristic was perceived to be more evident while a lower value (nearer to 1) indicated that the characteristic in question was perceived to be less evident. The scales were marked with pairs of opposite characteristics, e.g. “self-assured” vs. “insecure”. The scales’ suitability was confirmed via a reliability test. For the RTR-measurements, subjects were instructed to turn the dial to the left on a seven step scale (towards positions 1–3) if they found something in the subsequent presentation not persuasive and to the right (towards positions 5–7) if they found something persuasive. The participants were instructed to rate the overall impression of the speech’s persuasiveness continuously whenever they found a certain passage worth to be evaluated. No predetermined measuring points were defined because it was intended to measure the subjective impression of each single individual as it received and processed the stimulus material. The dial was preset to position 4 (mid-point) representing “neutral” and subjects were instructed to reset that position at the end of the experiment.

The second experiment was conducted on the 19 and 20 June 2007 with the dual aims of confirming the validity of the first experiment and isolating the effects of body language. For this purpose, test subjects were played only the audio tracks (without picture) from the video presentations employed in the first experiment. The experiment consisted of two experimental settings in total: The first group heard the audio track of the video recording employed in the first experimental setting of the first experiment (the speech on globalization without vocal emphasis); their reactions were ascertained via a questionnaire and the
RTR-devices. The second group of test subjects heard the audio track of the video recording from the second experimental setting in the first experiment (speech on globalization with vocal emphasis, for example variable tone, pauses, tempo; measurements as for group 1). The intention of the second experiment was to test whether the questionnaire and RTR-findings varied from those ascertained in the first experiment when only an audio track was presented. To this end, all aspects of the experimental setting other than the stimulus and the measurement tools employed were identical to those in the first experiment.

Findings

Experiment 1: Findings of the questionnaire measurement

The analyses revealed that there were considerable statistically significant differences in audience evaluations of the different versions of the presentation. Table 1 provides an overview of findings. The data show that the audience rated the speech with vocal emphasis and body language as the most *vivid* \( (M = 3.27; \ F = 5.907; \ p < .001; \ \eta^2 = .135) \), most *powerful* \( (M = 3.21; \ F = 9.769; \ p < .001; \ \eta^2 = .205) \), *liveliest* \( (M = 3.24; \ F = 21.495; \ p < .001; \ \eta^2 = .361) \) and most *self-assured* \( (M = 4.0; \ F = 13.252; \ p < .001; \ \eta^2 = .259) \). The difference was particularly significant in comparison with the speech presented monotonously with no vocal emphasis or body language, which was perceived as not very *powerful* \( (M = 2.03) \), rarely *lively* \( (M = 1.5) \) and less *self-assured* \( (M = 2.88) \). These variations point to at least one aspect of audience evaluations being positively influenced by nonverbal support and that is the perceived *performance of the presenter*.

Furthermore, the audience adjudged the two versions of the speech with relatively less body language, i.e. those not supported by gestures or facial expressions, to be more *factually accurate*, more *rational*, more *authentic* and more *thorough*; although only the variations with regard to thoroughness were statistically significant \( (M = 2.23; \ F = 2.616; \ p < .05; \ \eta^2 = .064) \). However, even though the findings for related characteristics like *factual accuracy* and *authenticity* were not statistically significant, a uniform pattern can be discerned: the versions of the speech with reserved body language were rated better in these aspects than the version employing a lot of gestures and facial expressions (version C).

Thus it can be concluded that persuasive presentations do not benefit in *all respects* from the use of gestures and facial expressions. It is clearly evident, therefore, that, although the vividness and liveliness of a presentation depend on audiovisual aids such as spoken emphasis, facial expressions or gestures, it does not necessarily follow that such support
Table 1. Analysis of variance of the questionnaire measurement regarding performance, content, and personality (first experiment)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>A (no vocal emphasis)</th>
<th>B (vocal emphasis)</th>
<th>C (vocal emphasis and gestures)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Opinion about</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Globalization ($\eta^2_p = 0.023$)</td>
<td>$2.84$ (SD = 0.99)</td>
<td>$2.87$ (SD = 0.96)</td>
<td>$3.15$ (SD = 0.87)</td>
</tr>
<tr>
<td>Comprehensible ($\eta^2_p = 0.096$)</td>
<td><strong>3.59</strong>$^a$ (SD = 0.95)</td>
<td><strong>4.10</strong>$^{ab}$ (SD = 0.83)</td>
<td><strong>4.21</strong>$^b$ (SD = 0.74)</td>
</tr>
<tr>
<td>Convincing ($\eta^2_p = 0.037$)</td>
<td>$2.16$ (SD = 1.05)</td>
<td>$2.65$ (SD = 1.11)</td>
<td>$2.48$ (SD = 1.00)</td>
</tr>
</tbody>
</table>

| **Performance** |                      |                   |                                 |
| Self assured$^1$ ($\eta^2_p = 0.209$) | **2.88**$^a$ (SD = 1.07) | **3.58**$^b$ (SD = 0.92) | **4.00**$^b$ (SD = 0.75) |
| Vivid ($\eta^2_p = 0.095$) | $2.41$ (SD = 1.34) | $2.55$ (SD = 1.23) | $3.27$ (SD = 1.01) |
| Lively ($\eta^2_p = 0.337$) | $1.50$ (SD = 0.92) | $1.90$ (SD = 1.11) | $3.24$ (SD = 1.17) |
| Powerful ($\eta^2_p = 0.176$) | $2.03$ (SD = 1.09) | $2.61$ (SD = 1.05) | $3.21$ (SD = 1.05) |
| Professional ($\eta^2_p = 0.049$) | $2.25$ (SD = 1.08) | $2.74$ (SD = 1.06) | $2.73$ (SD = 0.91) |

| **Content** |                      |                   |                                 |
| Authentic ($\eta^2_p = 0.016$) | $2.72$ (SD = 1.14) | $2.84$ (SD = 1.10) | $2.52$ (SD = 0.94) |
| Thorough$^1$ ($\eta^2_p = 0.088$) | $2.47$ (SD = 1.19) | $2.48$ (SD = 1.06) | $1.82$ (SD = 0.77) |
| Rational ($\eta^2_p = 0.026$) | $3.16$ (SD = 0.88) | $3.13$ (SD = 0.81) | $2.85$ (SD = 0.94) |
| Factually accurate | $3.25$ (SD = 1.24) | $3.52$ (SD = 1.09) | $2.97$ (SD = 1.05) |
| Interesting ($\eta^2_p = 0.106$) | $2.38$ (SD = 1.10) | $2.74$ (SD = 1.06) | $3.21$ (SD = 0.89) |

| **Personality** |                      |                   |                                 |
| Honest ($\eta^2_p = 0.006$) | $2.84$ (SD = 1.05) | $2.97$ (SD = 1.02) | $2.79$ (SD = 0.89) |
| Likeable ($\eta^2_p = 0.128$) | **2.03**$^a$ (SD = 0.93) | **2.52**$^{ab}$ (SD = 1.06) | **2.97**$^b$ (SD = 1.08) |
| Credible ($\eta^2_p = 0.022$) | $2.63$ (SD = 1.01) | $2.84$ (SD = 0.90) | $2.48$ (SD = 1.03) |
| Trustworthy ($\eta^2_p = 0.047$) | $2.34$ (SD = 0.83) | $2.84$ (SD = 0.97) | $2.45$ (SD = 1.06) |
| Competent ($\eta^2_p = 0.031$) | $2.56$ (SD = 0.98) | $2.97$ (SD = 0.88) | $2.79$ (SD = 0.93) |

$^a$The values in each column differ significantly ($p < .05$), if they do not share the same superscript.

$^1$Where inhomogenous variances occurred, the Post-Hoc-Test Tamhane T2 was used, otherwise Bonferroni was chosen.

The effects of verbal and nonverbal elements will lead to the content of the speech itself being evaluated more favorably. The use of body language can even impact negatively on certain aspects of how a presentation is perceived. In fact, there was only one area in which the speech presented with body language fared better than the other two versions in this experiment: the audience found version C more interesting ($M = 3.21; F = 5.992; p < .001; \eta^2 = .136$) than the other versions, in particular in comparison to version A, which was rated as significantly less interesting ($M = 2.38$). It is to be assumed that there is a correlation between this characteristic and others such as vividness and liveliness.

In contrast to the audience ratings ascertained for performance and content, evaluations of the speaker failed to reveal any uniform trends and, moreover, yielded hardly any statistically significant findings. The presentation with vocal emphasis and body language, for example (ver-
sion C), was considered more “likeable” overall ($M = 2.97; F = 5.301; p < .001; \eta^2 = .122$) than the speech without vocal emphasis, facial expressions or gestures ($M = 2.03$) – this can be interpreted as an assessment of the speaker. Other key persuasion variables typically defined as speaker characteristics (for example credibility, trustworthiness or competence; cf. Hovland et al., 1953; Perloff, 2003), appear, in the investigation at hand, not to be positively affected, or at least not significantly, by the types of vocal emphasis or body language employed here. The effects profiles are too inconsistent to enable any reasonably certain conclusions to be drawn.

**Experiment 1: Findings of the RTR-measurement**

In addition to the questionnaire findings described above, the RTR-measurements provide a detailed picture of relative audience perceptions of the different presentation types over the course of the presentation. Test subjects were instructed to judge how persuasive they found the speeches presented and record their impressions via a seven-interval scale marked on the dial of the RTR-devices. Each version of the speech lasted approximately 90 seconds, with dial positions on the measurement devices being recorded 81 times during the course of each presentation. Figure 1 shows firstly that subjects reacted to the presentation with vocal emphasis and body language (version C) most positively, evaluating this version as the most persuasive overall.

Secondly, we can see that subjects reacted most negatively to the speech without vocal emphasis or body language (version A): it was considered the least persuasive. The speech with vocal emphasis but without body language (version B) is not significantly different. Apart from slight differences in the length of the presentations, which explain why the curves are not exactly parallel$^{14}$, figure 1 also shows thirdly that, if we consider the general structure and path of the curves, reactions to the speeches were fundamentally consistent for each presentation: the pattern of peaks and troughs is practically identical across all three curves. After roughly one quarter of the speech, the subjects’ reactions are neutral with a slight positive tendency. This is followed in the second part of the speech by a downturn in the curves, signalizing that subjects found the arguments presented at this point, concerning the positive consequences of globalization in developing countries, less convincing than they had the introduction. Towards the end of the middle third, there is a brief rise across all three groups. This positive reaction corresponds to the point where the speaker describes globalization as an opportunity for Germany. A couple of seconds later, the curve dips again, only for it to rise once more near the end when the speaker refers to a
The effects of verbal and nonverbal elements

Figure 1. First experiment — RTR-curves for the speech without emphasis (A), with emphasis (B) and with emphasis and gestures (C).

quote from the former German president, Roman Herzog\(^\text{15}\). It is only between the 65\textsuperscript{th} and the 73\textsuperscript{rd} seconds that we see a rise in the perceived persuasive power of the speech with gestures which is not evident in the other two curves — even taking into consideration the slight time lag. This single variation does not take anything away, however, from the general similarity of the curves.

The curves follow the same path during the course of the speeches, i.e. the three versions produce identical “effect patterns”. From this, we can conclude that the text and/or the speaker’s arguments dictate the structure of reactions, and thus the overall effect of the speech, in great detail. The key finding of the first experiment is, therefore, that nonverbal elements are not the main driving force behind the persuasive power of a speech. On the contrary, audience perceptions of the persuasive power of a speech depend much more directly on what the speaker said than on how he said it. At the same time, the use of emphatic speech in combination with body language is by no means without effect. Far from it, the findings clearly show that emphasis and body language increase the persuasive power of a presentation: the average RTR-curve for the corresponding version rises above those for the other versions. The presentation makes a more positive overall impression. Thus it is clear that the content of a speech predetermines the general pattern of effects it
generates and that presentation techniques can merely enhance (or also detract from) the overall effect. Thus, presenters who speak emphatically and utilize body language can, in so doing, ensure that their weaker arguments come across as less weak and that their stronger arguments are further strengthened. Persuasion models such as the ELM and the HSM assume that both persuasion pathways (central and peripheral route; systematic and heuristic elaboration) are part of a continuum of elaboration behind and that there is a tradeoff between elaboration valence and peripheral cues: as elaboration declines, the impact of peripheral cues increases and vice versa; O’Keefe, 2008: 1477). Therefore, utilizing body language as means of persuasion may be generally promising in persuasive settings: Nonverbal elements, which are regarded as important cues within low-involvement reception, may in addition have a positive effect in high-involvement situations, too.

**Experiment 2: Findings of the questionnaire measurement**

First of all, it is clear that the differences in reactions to the audio only versions with vocal emphasis and without vocal emphasis (both without body language) were more significant than had been the case between the corresponding audiovisual versions from the first experiment. The audience found the version with vocal emphasis *more vivid* (0–5 rating scale: $M = 3.06$ vs. $3.53$; $T = -2.002$; $p < .05$; $\eta^2 = .064$), *more likeable* ($M = 2.29$ vs. $3.00$; $T = -2.636$; $p < .05$; $\eta^2 = .106$), *more self-assured* ($M = 3.81$ vs. $4.40$; $T = -2.965$; $p < .05$; $\eta^2 = .130$), *more interesting* ($M = 2.87$ vs. $3.53$; $T = -2.587$; $p < .05$; $\eta^2 = .102$), *more powerful* ($M = 3.03$ vs. $3.70$; $T = -2.144$; $p < .05$; $\eta^2 = .072$) and *livelier* ($M = 2.39$ vs. $3.67$; $T = -4.461$; $p < .05$; $\eta^2 = .252$). In contrast, the version with vocal emphasis was also perceived to be less *factually accurate* ($M = 3.35$ vs. $2.77$; $T = 2.269$; $p < .05$; $\eta^2 = .080$) than the version without vocal emphasis. The effect of vocal emphasis is thus very similar to that of gestures in the first experiment. Enhancing what is said via nonverbal stimuli can lead to improvements in how the audience rate the speaker’s performance and personality, but can, at the same time, also have negative effects on how the content of the speech itself is perceived. Both experiments ascertained that patterns in audience perceptions and presentation effects depend on how various persuasive stimuli interact. For the purposes of comparison between the audiovisual and the audio-only presentations, it made sense here to consider the two experimental settings from the second experiment as additional extensions within the single factorial framework of the first experiment. The analyses were therefore based on, *on the one hand*, a comparison of the questionnaire findings for the first “video” group (video of the speech without vocal
emphasize or body language) and the questionnaire findings for the first “audio” group (audio track without vocal emphasis) and, on the other hand, on a comparison between the findings for the second “video” group (video of speech with vocal emphasis without the use of body language) and the findings for the second “audio” group (audio track of the speech with vocal emphasis). In accordance with the rationale of experiments in the field of the social sciences, any statistically significant differences in responses between the groups compared here can be attributed with sufficient probability to differences in the stimulus, i.e. in the mode of presentation.

The comparison between the video version without vocal emphasis and the corresponding audio version reveals one finding that was not present in the first experiment at all. Namely, the analysis shows a clear and statistically significant difference in the test subjects’ opinions toward globalization across the two groups: the group that heard the audio version saw more positive consequences of globalization for Germany ($M = 3.39$), while the group that watched the video presentation saw more negative consequences ($M = 2.84; T = -2.219; p < .05; \eta^2 = .075$).

Furthermore, the test subjects also found the audio version of the speech easier to comprehend ($M = 3.59 \text{ vs. } 4.03; T = -2.0.40; p < .05; \eta^2 = .064$) and more convincing ($M = 2.16 \text{ vs. } 2.77; T = -2.615; p < .05; \eta^2 = .1$) than the video version without vocal emphasis. A comparison of the items for the perceived performance and personality of the speaker revealed fewer differences than above. Also, all the significant differences that were found were, in any case, consistent with the previous analyses: in comparison to the video version without emphasis, the audience found the audio version without emphasis more vivid ($M = 2.41 \text{ vs. } 3.06; T = -2.332; p < .05; \eta^2 = .081$), more authentic ($M = 2.72 \text{ vs. } 3.29; T = -2.047; p < .05; \eta^2 = .065$), more self-assured ($M = 2.88 \text{ vs. } 3.81; T = -3.779; p < .001; \eta^2 = .19$), more powerful ($M = 2.03 \text{ vs. } 3.03; T = -3.347; p < .01; \eta^2 = .155$), more professional ($M = 2.25 \text{ vs. } 2.81; T = -2.142; p < .05; \eta^2 = .07$) and livelier ($M = 1.50 \text{ vs. } 2.39; T = -3.510; p < .01; \eta^2 = .168$).

With just one exception, the comparison of the two versions of the presentation with vocal emphasis – one with and one without visual presentation of the speaker – paint a similar picture. In comparison to the video version, the audio version was considered more vivid ($M = 2.55 \text{ vs. } 3.53; T = -3.467; p < .01; \eta^2 = .168$), more self-assured ($M = 3.58 \text{ vs. } 4.40; T = -3.948; p < .001; \eta^2 = .209$), more interesting ($M = 2.74 \text{ vs. } 3.53; T = -3.133; p < .01; \eta^2 = .143$), more powerful and livelier ($M = 2.61 \text{ vs. } 3.70; T = -3.952; p < .001; \eta^2 = .201$ and, respectively, $M = 1.90 \text{ vs. } 3.67; T = -6.092; p < .001; \eta^2 = .386$). The sole exception being the item: factually accurate. The audience perceived the speech as
significantly more factually accurate ($M = 3.52$) when they could see the speaker than when they could only hear his words ($M = 2.77$; $T = 2.704$; $p < .01$; $\eta^2 = .11$).

Overall, the comparison of the two experiments shows that audience perceptions of the speaker and the speech are not necessarily positively affected by the visual presence of the speaker. In the context of this experiment, vocal emphasis proved more effective when the audience was not engaged in processing visual impressions while listening, suggesting that the body language employed by the speaker during the speeches presented here was not always entirely convincing. The two outliers amongst the findings, specifically the improvement in the opinion of the audience of the audio version of the speech without vocal emphasis toward globalization and the worsening of the perceived factual accuracy of the audio version presented with vocal emphasis (compared to the two corresponding video versions) may be attributable to interplay between emphasis and mode of presentation. This is certainly an issue that requires clarification via further analyses and, where necessary, further experiments.

**Experiment 2: Findings of the RTR-measurement**

When comparing the audiovisual and audio only versions, the RTR-findings again offer the opportunity for a detailed analysis of differences in the perceived persuasive power of the two speeches over time. Figure 2 shows three of the key findings for the investigation at hand: A comparison of the two curves confirms the questionnaire finding that vocal emphasis by no means always guarantees an improvement in the perceived persuasive power of a presentation. In fact, in the second part of the speech, the audience found the presentation with vocal emphasis considerably less convincing than the version without vocal emphasis. Furthermore, the version with vocal emphasis takes over the role played by the version with both vocal emphasis and body language in the first experiment (compare figure 1): during the first part of the speech, the perceived persuasiveness curve for this version is clearly higher than that of the other two versions, even exceeding the average value of 4, only to fall away dramatically and to sink below the other two curves at some points. This would suggest that forms of visual and auditory emphasis are not cumulative: but rather that the perception and effects profiles generated by a speech differ depending on the presence or absence of individual presentation techniques such as vocal emphasis, facial expressions or gestures respectively. In other words, it is entirely possible that a presentation without any visual support will be perceived as more persuasive than one with such support. There is thus no universal hierarchy
The effects of verbal and nonverbal elements

Figure 2. Second experiment — RTR-curves for audio-only speech without emphasis (A) and with emphasis (B).

of effects that prescribes the superiority of one mode of presentation over another in all conceivable communicative contexts. Otherwise, the version with vocal and visual emphasis from the first experiment would have to have been clearly more persuasive than the version with only spoken emphasis presented in the second experiment. It was not. The extent to which measurement effects and the scales employed have influenced findings cannot, however, be determined without recourse to specialist methodology studies.

In addition, the second experiment confirms one of the key findings from the first experiment, namely that the mode of presentation is not solely responsible for determining the effect of the content. There is not merely a close similarity between the two the perceived persuasiveness curves for the audio versions despite differences in the deployment of nonverbal persuasion techniques; these two curves also bear a striking resemblance to the curves from the first experiment: a sharp rise in perceived persuasive power in the first third, followed by a clear drop off towards the middle and a slight recovery in the final third of the speech. According to popular assumptions mentioned in the introduction about general dominance of the form of presentation relative to the content of a speech in dictating reactions, any similarities between the findings for the five independently observed experimental groups should be far less
prominent than those ascertained here. Were there to be any truth at all in the so-called communication pyramid (whereby a mere 7 percent of the effect of a speech is attributable to its content), there would have to be very distinct differences in the five graphs. In particular, the effects patterns for the first (audio visual) and second (audio only) experiments would have to be fundamentally different. The confirmation of these findings in the second experiment provides further evidence for dismissing the assumptions mentioned.

Summary

In summary, the following picture emerges: both experiments point to the conclusion that the speaker’s performance, as well as certain qualities of the speech presented, can profit from vocal emphasis and/or support from facial expressions and gestures. Audiences find speeches with vocal emphasis and/or with both vocal emphasis and the support of facial expressions and gestures, for example, more vivid, more powerful, more self-assured, and livelier. Regardless of whether video versions or audio only versions (or both) are compared, the effects pattern ascertained was identical.

Nonverbal support did not, however, necessarily lead to recipients finding the speech itself more factually accurate, more authentic, more thorough or more rational. On the contrary, certain nonverbal persuasive tools, especially facial expressions and gestures, led to the speech being rated less favorably, at least when it came to impressions of the content or the speech itself. Here too, the pattern was practically identical regardless of which versions were compared (audio, visual or audio and visual).

On the basis of the RTR-persuasiveness findings, those speeches presented with the aid of vocal emphasis and supported by body language were generally more favorably received overall. However, despite higher average levels for the versions employing persuasive techniques, the general response pattern was the same across all versions. Thus, the content was more or less responsible for determining whether certain parts of the speech were perceived to be relatively more or less persuasive than others. It follows that it was the content and not nonverbal elements that was responsible for determining the general effects pattern, the acceptance or rejection of arguments presented and the persuasiveness of individual passages — nonverbal techniques, however, enhanced existing effects.

Furthermore, the audio-only versions were rated more favorably than the audiovisual versions in several respects, providing further evidence that there is no general primacy of the audiovisual presentations over
other formats. Even though such a finding is very closely related to the configuration of the experimental stimulus in the study at hand, it is nevertheless quite clear that the persuasive power of the content of a speech can benefit significantly from the audience not being distracted by visual impressions or, indeed, engaged in simultaneously taking in and processing visual, audio and content based elements. This provides further support for a point that has already gained general acceptance among researchers, for instance in the context of the development of dual processing theories (ELM, HSM): different forms of presentation engage different information processing and evaluation pathways in recipients and messages are therefore perceived and exert their effects very differently depending on composition, response context and audience characteristics. In the present study, one plausible explanation would be, for example that a text on a complex phenomenon such as globalization is more persuasive when read out or presented in the form of an audio recording.

It follows from the findings presented here that there is in fact no such thing as a universal hierarchy of effectiveness and that presentations armed with nonverbal techniques are not automatically more effective than others. On top of this, were the assumptions that form the basis of the so-called communication pyramid true, the content itself would have to have exerted a much less significant influence over the general pattern of effects of the different versions than was found to be the case here. Moreover, the versions of the speech presented with the assistance of nonverbal elements would not only have to have been slightly more persuasive overall but the RTR-findings would probably have to have produced significantly different effects patterns. In particular, the comparative analysis of the two experiments (audio visual vs. audio only) would have to have revealed more considerable differences.

Discussion and conclusion

Against the background of the dual processing theories (ELM, HSM) it must be pointed out that the topic of the speech and the audience selected for the experiments represent a specific constellation that makes it impossible to draw all too general assumptions from the findings presented here. First of all, it may be argued that the composition of the student sample itself has an effect on the experiments’ results. It is characteristic for students that they are higher educated, trained in intellectual information reception and evaluation and probably more critical towards speeches. Furthermore, it can be assumed that they are highly involved in the topic of the investigated speech and share specific pre-existing (potentially negative) opinions about globalization (more than
60 percent of the students located themselves on the left of the political spectrum, which is generally more globalization-skeptic) and that they have more prior knowledge of the topic than the average citizen. Moreover, the absence of potentially distracting factors within the laboratory setting may have additionally stimulated a more systematic and intellectual stimulus evaluation. Therefore, it is likely that these students are highly motivated to process the stimulus-speech through the central route of persuasion.

The dual processing theories, however, state that non-verbal cues are especially impactful when the audience is heading for the peripheral processing of the persuasive communication, which is typical for audiences that are very little involved with the topic or audiences that lack the (intellectual) capacity to process the content of the speech. This may be the reason for the finding, that there was a higher impact of non-verbal cues on more emotional (peripheral route) aspects of the speech rather than on more rational (central route) aspects: Our analyses show, for example, that recipients found the presentation with vocal emphasis and/or with both vocal emphasis and the support of facial expressions and gestures more vivid, more powerful, more self-assured, and livelier. Nonverbal support did not, however, necessarily lead to recipients finding the speech itself more factually accurate, more authentic, more thorough or more rational. These findings may be explained by the presumably high involvement-levels within the student samples.

In addition, it may be argued that a potentially higher level of ego-involvement among students may be the reason for the overall finding that the speech’s content was the decisive for the persuasion outcomes whereas the nonverbal elements were rather secondary. This criticism, however, is not very plausible, because the reception of political communication content per se is rather the domain of high-involved audiences: Who would voluntarily watch a politician’s speech in the German Bundestag on TV if there is not a touch of interest and involvement in the topic? Can we imagine a person that deliberately decides to visit a panel discussion on the topic of globalization and its effects on Germany when this topic is regarded as entirely irrelevant and not at all interesting by this person? In Germany, reading high-quality newspapers and political magazines, for example, is still the domain of the intellectual elite (perhaps 10 percent of the population; Kepplinger, 2010b) and it is very likely that this will not change. It is therefore plausible to assume, that the findings of the experiment can be generalized at least for typical audiences of political communication content.

This does not mean, however, that there are not many open questions left which have to be addressed in further studies: Many different variables may have impacted on the findings: the content and topic of a
The effects of verbal and nonverbal elements

speech; image and characteristics of the speaker (e.g. his attractiveness); the audience’s mood, predispositions, prior knowledge and other characteristics of the public; characteristics of the setting such as time, location, historical context etc.). Within the framework of the dual processing-theories each of these variables is probably likely to exert effects on the persuasion outcomes. Future studies on the effects of speeches could use the methodological approach of this study (in particular the combination of questionnaire and RTR-measurement) for testing the effect of a large number of stimulus variations (e.g. different attractive speakers, a more naturalistic setting outside the laboratory, different topics for the speeches and so on). Here, producing a potentially more convincing stimulus presentation would be an important project: the finding, that vocal emphasis proved more effective when the audience was not engaged in processing visual impressions while listening (audio-only presentation) and that the body language employed by the speaker during the speeches presented here was not always entirely convincing, suggests that choreographing the nonverbal performance of the speaker may be one of the most important challenges for further research.

In sum, the findings of this study confirm that persuasive and persuasive communication respectively can only productively be observed and analyzed in the correct setting and context. Thus any attempts to come up with simple, universal recipes for successful and effective presentations are doomed to failure — there are no simple, universal explanations for responses to or the effects of persuasive communication. Furthermore, it has also been confirmed that it is not the composition of auditory and visual stimuli alone that determines whether or not an audience finds a speech convincing. The findings also show that attributing a rudimentary 5, 7 or 10 percent share of the overall persuasive effect of a speech to its content, in other words to the arguments and entreaties of the speaker, is a gross simplification. That is not to suggest, however, that the content of persuasive communication is solely responsible for dictating its effects; such a claim would not only be equally unrealistic but would also be in direct contradiction with the primary objection raised in the previous paragraph. Nevertheless, it is highly likely that the content of a speech has a defining influence on at least the structure of its effect, i.e. the peaks and troughs in persuasive power. Moreover, it is very likely that auditory and visual elements, as long as they are not completely inappropriate, contradictory or lacking credibility, can enhance existing effects. These elements can positively or negatively affect the speech and the way how arguments are perceived: e.g. they can lessen the effect of weaker arguments and add strength to stronger ones (or vice versa). This reading of the data is supported by the dual processing theories which state that there may be a tradeoff between both
modes of stimulus reception and evaluation. But in the majority of cases the nonverbal elements will probably not fundamentally change the overall pattern of effects and perceptions prescribed by the content.

Finally, it has also been confirmed that the use of individual persuasive tools such as vocal emphasis, pauses, gestures, facial expressions etc. does not have a positive influence on the effect of persuasive communication per se, to address the many normative assessments in practical literature. Thus a speaker is not necessarily better (specifically: “more effective” or “more persuasive”) because he raises an eyebrow, has a smile on his face, lowers his voice, etc. In many cases, such nonverbal stimuli can even have a negative impact, a fact not only supported by simple common sense but also clear from the findings ascertained. The universal validity of models such as the “persuasion pyramid” is refuted by the findings presented here and by the general premises of the dual processing theories.

The findings of both experiments indicate that nonverbal elements of persuasive communication can, under certain circumstances, have a beneficial effect on the persuasiveness of individual arguments as well as on the persuasive effects of presentations as a whole. They also show, however, that this is not always the case and, further, that the content of a speech plays a much more significant role than is sometimes assumed in dictating the structure of persuasive effects. What related research on nonverbal communication or persuasion research has found in prior studies has been confirmed by the presented experiments: The raising of an eyebrow, pointing of a finger or timing of a pause does not have any major or lasting persuasive effect, at least not on a highly involved audience (which can be reasonably assumed to be the case, at least in a political context).

Bionotes

Nikolaus Jackob is Assistant Professor at the Department of Communication, Johannes Gutenberg-University Mainz, Germany.

Thomas Roessing is Assistant Professor at the Department of Communication, Johannes Gutenberg-University Mainz, Germany.

Thomas Petersen is Senior Researcher at The Allensbach Institute for Public Opinion Research, Germany.
Notes


2. For example, Ekman und Friesen (1967; 1969) as well as Frey (1999) reached the same conclusions some 2000 years on, i.e. that the face is a rich source of information and can have a major influence on the effect of communication (see also Hassin and Trope 2000).

3. Subjects were recruited by their fellow students on a quota basis, split according to sex and subject studied. The youngest participant was 19, the oldest 32 ($M = 22.94; SD = 2.49$). 53.8 percent of the participants were female, 46.2 male. The largest share of participants studied philosophy or philology respectively, followed by social sciences and law. 62.5 percent of the participants placed themselves politically on the left, 17.3 declared themselves as belonging to the right wing of the political spectrum (scale 0–9; $M = 4.09; SD = 1.49$).

4. The smallest group consisted of 27 test subjects, the largest group 41. The groups displayed no differences during the analyses that could be put down to anything other than coincidence and thus can be seen as comparable in the sense of the logic of the experiment. The first control group was presented the text in written form as in the representative survey conducted prior to the experiments. Another control group also saw the presentation with vocal emphasis as in the second experimental setting; however no RTR measurements were taken. The third control group did not receive the text in any form, test subjects were merely asked about their opinion to globalization using an adapted questionnaire.

5. The speech was presented by a professional TV-journalist. The presenter was male, tall, aged between 30 and 40. He presented the speech in the classical pose of an orator behind a standing desk in front of a blue screen – with the upper part of his body clearly visible. He was shot frontally (medium close-up), the camera perspective was fixed during the session.

6. In persuasive communication it is rather unlikely that a speaker presents a speech completely without facial expressions, gestures, and/or vocal emphasis – a stimulus speech completely free of these nonverbal elements would be unrealistic and would potentially affect external validity. The versions of the stimulus-speech differed in the degree of nonverbal support: In the first video the presenter spoke monotonously, he contained himself, laid his hands on the standing desk, and had a neutral expression in his face. In the third video the presenter spoke accentuated and fluently, used a large variety of hand gestures (e.g. pointing with a finger, shrugging the shoulders), and supported his argument with facial expressions (e.g. knitted brows, smiling mouth). In the second video body language was similar to the first one – but the presenter spoke accentuated and fluently like in the third version.

7. The performance and personality scales scored Cronbachs Alpha values of .751 and .776, at once confirming their reliability. The scale for assessing the content of the speech attained a slightly lower score of .650. This is mainly due to the item “interesting”, which is strongly correlated with items in both of the other scales. However, scale-construction was never a goal of this study, therefore the reliability coefficients are interpreted just as a corroboration of the premises concerning the categorization of items.

8. Given the focus on the overall persuasiveness of the different presentations, the following analyses do neither concentrate on detailed comparisons of video-sub-
sections and/or single verbal or non-verbal entities (e.g. certain emphases, facial expressions, hand gestures, or persuasive figures) nor on a comprehensive interpretation of interrelations between presentation and content. For comparing the overall persuasiveness of the speech-versions a detailed persuasive examination to the split second is not required.

9. Evidence for the reliability and validity of this RTR-measurement procedure provide Reinemann et al. (2005).

10. \( n = 61; 31 \) test subjects in the group that heard the speech without vocal emphasis and \( 30 \) in the group that heard the version with vocal emphasis.

11. Thanks to the identical matching procedures, the corresponding groups from the first experiment could be used for comparison. The second experiment was conducted in the same room as the first one with identical measuring procedures. The groups from the first and second experiments displayed no significant differences in terms of age structure, gender, political orientation or course of study and are thus comparable for the purposes of this experiment.

12. A single factorial ANOVA with Bonferroni Post-Hoc-Tests as well as Tamhane T2 where variances were disparate.

13. \( Df = 4,152 \) for all F-Tests.

14. All three speech versions took as long as about 90 seconds — they differed only in few seconds from each other. It has been argued in persuasion research that the speed of a speech and the speaking rate have an impact on the persuasion process. In face of the minimal differences, however, it cannot be expected that these differences have a confounding influence on the RTR-findings presented above.

15. The correlation between the curves for the speech with and the speech without vocal emphasis (versions A and B, both without body language) is \( r = .63 \). The correlation between the curves for the speech without vocal emphasis and the speech with both vocal emphasis and gestures (versions A and C) is \( r = .49 \). And, the correlation between the pair of speeches with vocal emphasis (B and C) is \( r = .56 \). When calculated with a time lag based on certain points within the content of the speech so that the groups are synchronized according to content and the curves are thus more accurate parallel representations of the presentations, the correlations rise to \( r = .73, r = .61 \) and \( r = .78 \) respectively.

16. \( Df = 59 \) for all T-Tests.

References


