

11. Lay Nwe, T., Wei Foo, S., De Silva, L.: Speech Emotion Recognition Using Hidden Markov Models. *Speech Communication* 41(4), 603–623 (2003)
12. Boula de Mareuil, P., Célérier, P., Toen, J.: Generation of Emotions by a Morphing Technique in English, French and Spanish. In: *Proc. Speech Prosody*, pp. 187–190 (2002)
13. Navas, E., Castelruiz, A., Luengo, I., Sánchez, J., Hernández, I.: Designing and Recording an Audiovisual Database of Emotional Speech in Basque. In: *Proc. of the LREC* pp. 1387–1390 (2004)
14. Navas, E., Hernández, I., Luengo, I.: An objective and subjective study of the role of semantics in building corpora for TTS. *IEEE transactions on Speech and Audio Processing* 14, 1117–1127 (2006)
15. Saratxaga, I., Navas, E., Hernández, I., Luengo, I.: Designing and Recording an Emotional Speech Database for Corpus Based Synthesis. In: *Basque. Proc. of the LREC*, pp. 2127–2129 (2006)
16. Boersma, P., van Heuven, V.: Speak and unSpeak with PRAAT. *Glott International* 5(9–10), 341–347 (2001)
17. Boersma, P., Weenink, D.: Praat: doing phonetics by computer (Version 4.3.16) [Computer program] (2005), <http://www.praat.org/>
18. Hozjan, V., Kacic, Z.: Context-Independent Multilingual Emotion Recognition from Speech Signals. *International Journal of Speech Technology* 6(3), 11–320 (2003)
19. Iida, A., Campbell, N., Higuchi, F., Yasumura, M.: A Corpus-based Speech Synthesis System with Emotion. *Speech Communication* 40(1,2), 161–187 (2003)
20. Burkhardt, F., Sendmeier, W.F.: Verification of Acoustical Correlates of Emotional Speech using Formant-Synthesis. In: *Proc. ISCA Workshop on Speech and Emotion*, pp. 151–156 (2000)
21. Nogueiras, A., Moreno, A., Bonafonte, A., Mariño, J.B.: Speech emotion recognition using hidden Markov models. In: *Proc. EUROSPEECH* pp. 2679–2682 (2001)
22. Iwai, A., Yano, Y., Okuma, S.: Complex emotion recognition system for a specific user using SOM based on prosodic features. In: *Proc. INTERSPEECH*, pp. 1341–1344 (2004)
23. Nicholson, J., Takabashi, K., Nakatsu, R.: Emotion Recognition in Speech Using Neural Networks. *Neural Computing & Applications* 9(4), 290–296 (2009)
24. Ververidis, D., Kotropoulos, C.: Automatic Speech Classification to five emotional states based on gender information. In: *Proc. of 12th. EUSIPCO*, pp. 341–344 (2004)
25. Vogt, T., André, E.: Improving automatic emotion recognition from speech via gender differentiation. *Proc. Language Resources and Evaluation Conference (LREC 2006)*, pp. 1123–1126 (2006)
26. Theune, M., Meijs, K., Heylen, D., Ordeman, R.: Generating expressive speech for storytelling applications. *IEEE Transactions on Audio, Speech and Language Processing* 14(4), 1137–1144 (2006)

Prosodic and Gestural Expression of Interactional Agreement

Eric Keller¹ and Wolfgang Tschacher²

¹ IMM-Lettres, University of Lausanne, Switzerland

² Psychology, University of Bern, Switzerland

Abstract. Conversational interactions are cooperatively constructed activities in which participants negotiate their entrances, turns and alignments with other speakers, oftentimes with an underlying long-term objective of obtaining some agreement. Obtaining a final and morally binding accord in a conversational interaction is of importance in a great variety of contexts, particularly in psychotherapeutic interactions, in contractual negotiations or in educational contexts. Various prosodic and gestural elements in a conversational interaction can be interpreted as signals of a speaker's agreement and they are probably of importance in the emergence of an accord in a conversational exchange. In this paper, we survey the social and psychological context of agreement seeking, as well as the existing literature on the visual and prosodic measurement of agreement in conversational settings.

1 Introduction

In our increasingly complex and interconnected societal structure, obtaining agreements is often a considerable challenge. In comparison to the societal structures predominating in the industrialized countries until World War II, current structures are far more complex, geographically more diffused, and culturally more diverse. As a result, individuals responsible for commercial, governmental or educational decisions are often unknown to individuals working at the periphery of an organization, they may speak a different language, and their value systems may be unlike that of their employees or of their customers. Also, many more families in Europe have been constituted from diverse linguistic and cultural backgrounds. Such differences increase the opportunities for disagreements, which naturally leads to an increased potential for lack of understanding, disagreements and disputes in the execution of joint objectives.

At the same time, individuals in our modern societies have more options available to them than in previous generations. It is far easier today to change products, employers or life partners than it was in previous generations. Although many of these options are desirable for personal advancement and for an optimized matching between job offer and employee availability, there is no denying that the diversity of options available to modern individuals frequently complicates obtaining agreements between employers and employees, between vendors and customers as well as between the different members of a family structure.

Research on marriage, family and psychotherapeutic issues is probably the context where issues of agreement, disagreement and its various connoted elements (heated discussion, verbal conflict, verbal abuse, verbal attack, verbal aggression, conflictual conversations, etc.) have received the most systematic attention, particularly during the past few years. Research in this specialized field of interpersonal relationships provides some valuable insights into (a) the various categories of agreement, (b) the theoretical position and importance of agreements and disagreements within interpersonal interactions, and (c) the identification of prosodic and visual indicators of agreement.

2 Agreement in Family Relationships and in Organizational Structures

2.1 Chronic Aggressions and Conflicts

The most systematic past research on agreement/disagreement has been conducted from a dysfunctional perspective. A fairly clear trail of relationships has been established between the effects of *long-lasting disagreements (destructive and unhelpful comments, verbal attacks, etc.)* and various degrees of *family dysfunction*. These disruptive effects become most evident over longer periods of time. For example, in Shortt et al.'s study on the relationships of 158 young couples at risk for separation (male and female average 21 years), both physical and verbal forms of aggression were examined over the course of 6 years as predictors of separation of the couple [56]. As expected, the likelihood of relationship dissolution was significantly increased in couples where physical aggression was present; also, psychological aggression¹ was strongly correlated with physical aggression (r of about .5 in the two measured time periods). The relationship between verbal and physical aggression was similarly situated in a study conducted by DeMaris on 3,508 married and cohabiting couples [1]. In these and a number of other studies with smaller population samples, strong verbal disagreement was shown to be less destructive of family relationships than was physical aggression, but it figured as an important contributing factor in family dysfunction.

The harmfulness of chronic verbal disagreement appears to show up most clearly at an older age, after long-term exposure. In a study of 729 adults at least 50 years of age, currently married and in their first marriage, marital quality (defined as agreements/disagreements, positive and negative spousal behaviours, overall quality of relationship and marital communication) predicted to a significant degree measures of physical health, defined as chronic health problems, more disability and poorer

¹ Definition of psychological aggression in this study: "Observed psychological aggression was the rate per minute of content codes negative verbal, verbal attack, and coerce during the problem-solving task. Negative verbal behavior was blaming or disapproving of the partner (e.g., 'You really blew that one, didn't you.'). Verbal attack included name-calling, threats, and humiliation of the partner (e.g., 'You're such a loser.'). Coerce was a demand for behavior change that implied impending physical or psychological harm (e.g., 'You'll shnt up if you know what is good for you.')."

perceived health [2]. In particular, chronic negative verbal behaviours, such as excessive demands, being too critical or argumentative, being unreliable or continually agitating one's partner was associated with poorer physical health in the respondents. When these negative behaviours were present, they outweighed the positive spousal behaviours with respect to influencing physical health.

These studies place verbal disagreement and aggression in the family core into a chain of events that can be summarized as follows. The original intention behind verbal aggressions is generally an "attack on the self concept of the receiver in order to deliver psychological pain". These attacks are often used to "intimidate, subjugate and control another human being"². In mature adults, verbal aggression ultimately translates into lower degrees of marital satisfaction³, which leads to negative outcomes in physical and mental well-being in those couples that remain married.

Several elements of this causative translate chain were statistically supported by a study by Gavazzi et al. [3], notably the relationship between (a) the repeated use of *depreciative comments* (i.e., frequent events), (b) the establishment of *dissatisfaction* within the relationship (i.e., a stable state in which verbal information has been cumulated and has been translated into appreciations and evaluations of a marital condition), and (c) the *negative physical and mental health effects* resulting from this state (i.e., a measurable outcome).

Temporary and Non-threatening Disagreements

These references thus indicate that the presence of chronic verbal aggression is an important negative factor in the evolution of personal relationships. Does this also mean that the opposite, verbal support, can act as an important positive factor? Unfortunately, relatively little research has been performed on the benefits of verbal support in family relationships. However, research by Patterson and colleagues on particularly successful enterprises performed in the late 1990's has shown the remarkable impact of successful conversational patterns in organizational structures, and has deepened our understanding of the complex patterns on agreements and disagreements [4].

In the large-scale and long-term research underlying the Patterson et al. account, it emerged that one of the major differences between particularly successful enterprises and other enterprises of similar size orientation was related to the presence of successful internal communication channels (e.g., well-functioning meetings, productive email exchanges and satisfactory face-to-face conversational exchanges, etc.). A fairly extensive network of supportive semantic, emotional and verbal elements was identified in the analyses of such exchanges. It was shown, for example, that above all, an atmosphere of trust and confidence must prevail to enable the discussion of delicate issues (Patterson et al.: "safety in dialog"), that emotions must be bound and be translated into productive verbal statements, that attentive listening and valid interpretation of interlocutors' statements is required to build confidence in the conversational process, and that final conversational accords must ultimately be translated into valid actions. Only when all these elements are present, can a given

² Infante & Wigley [53] and Jacobson & Gottman [54] respectively, cited in Gavazzi et al. [3].

³ The relationship between dissatisfaction and separation could not be demonstrated for the younger adults (21 years old [1]).

communication channel make successful contributions to employee satisfaction and enterprise dynamics⁴.

2.2 Support and Agreement

It emerges from these and related studies that an important difference must be made between the underlying presence or absence of *support*⁵ and the overt presence or absence of *agreement*. Successful communication resulting in generally supported accords involves *both agreements and disagreements formed in the context of support*. In fact, one function of successful social groups is to favour "agreeing to disagree". Nelson & Aboud [5], for example, showed that friends of third- and fourth grade children furnished both more explanations and more criticism of their partners than did non-friends. Also among friends, disagreements provoked more beneficial change on the given experimental tasks than did agreements. Further, the disagreeing friend-pairs presented more mature solutions than did non-friends. Warranted disagreement between well-meaning friends is thus part of a healthy social development pattern which involves an evolution of both agreements and disagreements. The important component of this evolution is that both agreements and disagreements can occur within the framework of a supportive social environment.

Also, some disagreements are imposed by external circumstances. For instance, it has been argued that premature consensus on certain medical treatments has led to standardized treatments that were later shown to be harmful or of doubtful benefit (bloodletting, electroconvulsive therapy, etc.). Also, such premature consensus has led to the marginalization of proponents of opposite views [6]. Warranted disagreement must thus be possible at various levels of societal grouping (families, enterprises, countries, etc.) to favour the evolution of meaningful social solutions, and it should reinforce values that are ethically or morally responsible. As an interesting example of the importance of disagreement, Erath & Bierman [7] showed that children living in violence-favouring families where there was little disagreement between parents showed significantly *more* tendency toward the use of violence outside the family than those where there was disagreement between parents. Apparently, the silent approval of violence by parents reinforced the willingness to use violence outside of the family.

We conclude from this section that successful communication involves patterns of seeking agreements where they are warranted and of permitting, encouraging and supporting disagreements where they are necessary and appropriate. The systematic promotion of supportive communication channels where both agreements and disagreements are welcome can lead to the well-being and a heightened productivity of participants. On the other hand, chronic lack of support in a non-supportive

⁴ Somewhat similar conclusions were formulated by a team working at the Harvard Negotiation Project [55].

⁵ "Support" is taken here in the wider sense of active as well as passive support. In a family situation, support involves the notion of "potential aid in need", while in an enterprise situation, support can be merely "acceptance of common purpose" or "acceptance of divergence". This concept is termed "safety" by Patterson et al. [04]. The crucial elements are that participants in the enterprise must share "mutual purpose" and/or "mutual respect".

environment and the inflexible use of aggressive verbal material have considerable negative long-term effects on psychological and physical welfare of family and organizational participants.

A systematic theoretical structure emerges from these studies that can be summarized in terms of a *tri-phase model* (Table 1). The key dimension is *±support*, and the three phases are *events*, *states* and *outcomes*. Productive and sympathetic *communicative interactions* (events) permit to create an environment of *positive or negative rapport* (a state); if the results of conversational exchanges are seen to translate into actual *positive effects* (outcomes), positive rapport is maintained and improved, and increased participant satisfaction results. Failing communicative interactions can be seen as the inverse of this model. Chronically destructive or unhelpful statements or comments showing lack of understanding lead to stable states of negative rapport and distrust; over a certain period of time, such a state can lead to reduced productivity as well as to psychological and physical ill health.

Table 1. An Overall Interactional Model

	Events	States	Outcomes
+ Support	- constructive or supportive comments - statements showing understanding - constructive or supportive responses	positive rapport	Increased satisfaction and productivity
- Support	- destructive and unhelpful comments - statements showing lack of understanding - aggressive responses - silence, withdrawal	negative rapport	Impaired psychological and physical health

2.3 Agreement in Dyadic Interaction: State of Research in Social and Clinical Psychology

We conclude from the previous sections that conversations contribute in major fashion to an atmosphere of presence or absence of support through complex interplays between manifestations of agreement and disagreement, with the support perceived in such exchanges contributing in crucial manner to important social outcomes. The building blocks of such exchanges are the instances of *agreement* and *disagreement*. Instances of agreement/disagreement provide the basis for support to be perceived or for a pre-existing confidence to be undermined or destroyed.

Although the overall logic governing conversational behaviour is probably reasonably simple, the positive identification and significance of the various socially significant components in a conversation is rarely easy. On lexical and semantic grounds, it is often difficult to make differences between supportive and unsupportive comments (see e.g., neutral and sarcastic comments), and prosodic and visual indicators of presence or absence of support may pre-exist between conversational

partners, but be barely noticeable in a given conversation. Also, participants generally pay little attention to such indicators, since they are primarily bent on transmitting their own intentions in the conversation. Even when participants have been trained to become more sensitive to subtle indicators of approval and disapproval, they may still show resistance or ignorance about the use of such information in the midst of a conversation.

It is therefore of interest to explore external technical means of measuring interaction parameters empirically and as automatically as possible. If such indicators can indeed be measured reliably and linked to perceived indications of the presence and absence of support, they could be used in post-hoc sessions to clarify and support the training of beneficial conversational behaviours, particularly in psychotherapeutic and managerial training contexts. The purpose must thus be to elucidate both conscious and non-conscious elements in a conversation, particularly in dyadic exchanges.

Indeed, there is converging evidence in social psychology that the course of a dyadic interaction is shaped in various ways by *non-conscious* influences. Such influences may arise in all channels of communication. Established empirical findings addressed a wide range of variables, e.g. body configurations and postures [8], hand gestures [9, 10], head movements [11] and emotion-related prosodic features [12]. It is generally acknowledged that a substantial portion of behaviour occurs without conscious awareness (e.g. [13]) but nevertheless may have high impact on the course of a conversation. One special domain of nonverbal research in dyadic interaction deals with the correspondence of nonverbal features between two or more people.

This phenomenon of "synchronization" is found both in living systems and even inanimate nature. Flocking birds or a school of fish provide examples of behavioural manifestations of synchrony in the animal kingdom. In animals, these phenomena have been linked with the stability of perception-behaviour links [14]. Analogous mechanisms appear to play a major role in human interactions as well; they may be considered the behavioural underpinnings of higher-level cognitive appraisals of social situations and processes.

Synchrony in Nonverbal Measures

"Behavioural synchronization is a form of coordinative interaction which is thought to be present in almost all aspects of our social lives, helping us to negotiate our daily face-to-face interaction." [15]. Considering human interaction globally, it is evident that some form of coordination or mutual influence plays a crucial role. Cappella [16] summarized that "Coordination is arguably the essential characteristic of every interpersonal interaction. ... Interpersonal communication requires the coordination of behaviour."

Prior to summarizing findings from social and clinical psychology, we wish to clarify the terminology, which presents high heterogeneity. Bernieri & Rosenthal [17] group most of the manifestations of synchrony in the human domain under the term interpersonal coordination, loosely defined as "...the degree to which the behaviours in an interaction are non-random, patterned, or synchronized in both timing and form." Studies that emphasize temporal aspects such as simultaneous movement, rhythm, or meshing of nonverbal behaviour mainly regard quantitative characteristics. Because of this reliance on kinetic qualities we may classify this type of synchrony as "*movement synchrony*". It deals with "... the precise timing and coordination of

movements between individuals ... while the nature or similarity of movements is irrelevant." [18]. Irrespective of which movements are involved, global quantitative variables such as speed, duration, or complexity of movement become synchronized between the two interacting individuals.

In contrast to movement synchrony, the focus may also be put on static or qualitative features of an interaction; postures, mannerisms, and facial displays may thus be categorized as "*behaviour matching*". Corresponding terms in social psychology are mirroring, mimicry, congruence, or the chameleon effect [13]. In many real world applications, however, these two categories — movement synchrony and behaviour matching — are not disjunctive; commonly a mixture of both categories is observed. If interactants share the same posture (i.e. static synchrony, behaviour matching) and subsequently change their bodily configuration in a temporally coupled manner (dynamic synchrony, movement synchrony), we view synchrony both on quantitative as well as qualitative levels.

Emotional phenomena such as empathy, emotional propagation and emotional contagion have been investigated in a majority of the studies that dealt with synchrony in human interaction. The link between emotional closeness and synchrony has thus received considerable attention. In this vein, Darwin (1872/1965) used the term "sympathy" to refer to imitation based on reflex or habit [19]. Allport [20] stated that "...our understanding of other people is derived from our capacity to imitate, usually in imperceptible ways, the behaviour of the person we are trying to understand ..."; and "empathy becomes simply 'kinaesthetic inference'." The connection of synchrony and empathy has stimulated numerous research efforts. Condon [21] who coined the term interactional synchrony stated that "Synchrony and other forms of behavioural sharing express degrees of closeness or distance between interactants." The notion of nonverbal behaviour correlating with rapport (i.e. a favourable therapeutic relationship) is shared by most psychotherapists [22, 23, 24]. The work of Schefflen [25, 26, 27] suggested ways to conceptualize empathy, rapport and the quality of the therapeutic bond. Yet naturalistic studies have been scarce since many published contributions were descriptive or cited merely anecdotal evidence (e.g. Charny [28]). Empirical research of synchrony in psychotherapy found preliminary evidence in favour of the hypothesis that interactional synchrony is associated with a positive therapeutic relationship and higher agreement between clients and therapists [29, 30, 31].

Taken together, the spectrum of findings in psychology covers various domains of human behaviour and experience. An individual's experience with the phenomenon of synchrony can be traced back to early infancy: Mother-infant studies on imitative behaviour (for a review, see [32]) have shown that even neonates imitate basic facial gestures (see also [33]). Recent neurocognitive research has made a connection between interactional synchrony (including empathy and related psychological phenomena) and a certain system of cortical neurons, the "mirror neurons" [34, 35, 36, 37]. Accordingly, Ramachandran [38] predicted "... that mirror neurons might do for psychology what DNA did for biology: they will create a unifying framework and help explain a host of mental abilities that have hitherto remained mysterious and inaccessible to experiments."

3 Prosodic and Visual Measures of Conversational Interaction Parameters

Past experiments have explored primarily measures of *prosody* and of *visual measures* of head movement in dyadic conversations. The two measures will be discussed separately.

3.1 Prosodic Measures

Prosody is the use of pitch (intonation), amplitude and timing in voice and speech; it is used primarily to signal a person's identity, emotion or attitude, and it can be used secondarily to support other conversational functions, such as agreement/disagreement. The parameters of greatest relevance to support and agreement are those that have been examined with respect to emotion and attitude. Pittam [39] has summarized the most relevant prosodic parameters relating to emotion in the following list:

Parameter	Phenomena
F0 (fundamental frequency, pitch)	variability, perturbation, contour details
Intensity	mean, range, variability
Formants	means, bandwidth
Temporal	speech rate, pausing
Fluency	slurring of articulation
Spectral	noise, proportion of high frequency energy to low LTS contour and frequency range, short-term spectral envelope measures

Within this list, parameters relating to pitch are probably easiest to analyze and have been used in recent studies on agreement. For example, Roth & Tobin showed a number of pitch patterns emerging in recent naturalistic recordings of well- or ill-integrated teachers in New York Inner City schools [39]. Fairly evident patterns of pitch *disagreements*, *alignments* and *integration* were shown in this study. The pitch patterns of well-integrated teachers formed continuous patterns with those of their students, while those of ill-integrated teachers showed discontinuous and independent patterns.

Quantitative evaluations of agreement/disagreement in meetings were performed in [40] and [41]. In the first study, 9854 "spurts" (periods of speech without pauses > 500 ms) from seven meetings were labelled as positive, negative, backchannel⁶ and "other". One fifth of the spurts was hand labelled, and the rest was labelled with the use of a decision tree classifier using lexical categories (e.g., number of words in a spurt, type of expression, and frequency indicators) as well as prosodic categories (e.g., pauses, fundamental frequency, and duration). Adjacency information (which can be seen to be of importance in Figures 1-4) was not used. Both lexical and prosodic indicators provided encouraging learning rates with 78% and 66% accuracy

⁶ Phatic or "back-channeling" comments are used to signal events relating to conversational organization, such as signals or invitations to take a turn [39].

respectively. This result was improved upon by the approach used in the subsequent study where an adjacency search was implemented. The result was improved and a score of 86.9% accuracy was attained using an adjacency analysis of a maximum entropy classification for speakers and a reduced number of only three expression types instead of four (positive/negative/backchannel+other).

3.2 Visual Measures

Previous studies of synchrony in psychology predominantly relied on observer ratings. Contemporary multimedia technology, however, makes computerized quantification of movement increasingly accessible. Computer-based systems eliminate several of the problems commonly encountered when assessing nonverbal behaviour by means of observer ratings, especially the high costs (behavioural observation is time-consuming) and the low objectivity of rating procedures.



Fig. 6. Frame-to-frame head movements for both participants of a psycho-therapeutic dyad are converted into an amplitude graph. Data from Ramseyer & Tschacher, 2006.

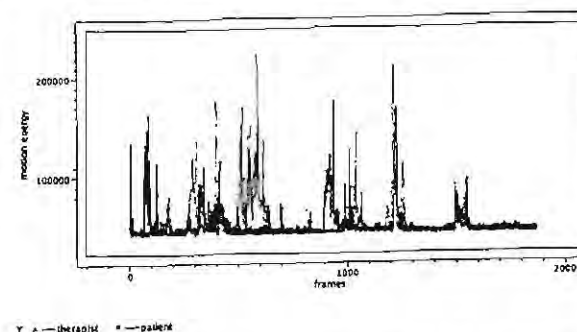


Fig. 7. The delays between head movements are scored. In the left part of the graph, the therapist's movements (grey) follow those of the patient (black), and in the right part of the graph, there is a passage where both participants show coincidental head movements ("synchrony"). Data from [44].

One such automated approach is Motion Energy Analysis (MEA). MEA of a recorded movie is based upon an image differencing algorithm [42, 43]. Each individual image (frame) of a movie has a constant number of pixels that generate a distribution of grey-scale values ranging from 0 (black) to 255 (white). With a fixed camera shot and nothing moving, each pixel retains its grey-scale value from one frame to the next. As soon as any item in a frame moves, the grey-scale distribution changes and can be quantified by differencing subsequent frames. The degree of movement from one frame to the next (the motion energy) equals this difference.

MEA is a simple method to continuously quantify movement in a video stream. Some caveats need to be considered however. First, the camera shot has to remain perfectly steady throughout the sequence; second, lighting conditions must be kept stable; third, the method solely quantifies movement energy, yet is blind to the direction or location of movement. To monitor motion energy of two persons in an interaction setting, two regions of interest (ROI) are defined. Within each ROI, differencing of grey-scale values is performed and recorded separately. If location information is essential, more than two ROIs may be defined (e.g. the faces, hands and arms of interactants). We thus generate two or more continuous time series that encode the amount of movement in these regions. Synchrony is consequently defined as the statistical correlation between the time series. Grammer's research group at the University of Vienna implemented the MEA method in several empirical studies, e.g. of courtship communication [45, 46], physical attractiveness [45] and interpersonal attraction [47].

Recent and current research projects in the second author's department have addressed the process of dyadic relationship formation in psychotherapy. We found synchronization of interactants at the level of questionnaire data (i.e. self-evaluations in post-session reports) [48, 49], in the domain of physiological parameters [50, 51] as well as in nonverbal social behaviour [44].

These findings concern the nonverbal channel. In a pilot study of naturalistic dyadic psychotherapy, we analyzed randomly selected therapy sessions taken from a very large sample of therapies conducted at the psychotherapy research centre of the University of Bern [52]. This database consisted of over 22,000 recorded therapy sessions, each 50 minutes in duration. From this set of sessions, a random sample of 100 sessions displaying different therapy dyads was drawn. An interim analysis at this moment is based on 50 dyads. In their therapeutic interactions, nonverbal synchrony was measured using the MEA approach.

A significant level of synchrony between patient and therapist time series was detected, in comparison to surrogate data. Statistical analyses showed that synchrony computed within the initial 15 minutes of interaction significantly predicted patients' post session evaluations of therapeutic bond quality. In other words, *movement synchrony was linked with therapeutic support and rapport*. Furthermore, the alleged association between the degree of synchrony during therapy and outcome at the end of therapy was corroborated, especially with the outcome measures 'patients' subjective well-being' and 'patients' competence expectancy'. In sum, higher degrees of nonverbal synchrony correlated with better therapeutic relationships as well as with better outcomes at the end of therapy.

4 Conclusion

In this brief review we have considered the process and context of agreement building, as well as the empirical measurement of agreement in the participants. It was seen that the pursuit of agreement building is embedded in and builds upon the pre-existing social framework. If this framework is supportive, or if it is at least characterized by an acceptance of common purpose, conversational transactions have much greater chance of reaching the intended goal of a morally binding agreement. If the framework is non-supportive or conflictual, reaching an agreement may be difficult to impossible.

Prosodic, gestural and postural information may provide a differentiated and independent measure of the process of agreement building which can be of considerable use in clinical and training contexts. Some pilot work has shown excellent correspondence between prosodic indicators of agreement in conversational settings, and between head movement and fundamental accord in psychotherapy settings. The data must be replicated over new studies, and a wider and more precisely circumscribed set of indicators must be defined for this research objective. Also it must be established if these empirical indicators concern overt agreement or relate more to the less evident development of support mechanisms within the conversational relationship.

Acknowledgements

The assistance of the European COST 2102 project as well as the assistance of the Swiss SBF Programme in support of COST is gratefully acknowledged.

References

- [1] DeMaris, A.: Till discord do us part: the role of physical and verbal conflict in union disruption. *Journal of Marriage and Family* 62, 683-692 (2000)
- [2] Bookwala, J.: The role of marital quality in physical health during the mature years. *Journal of Aging and Health* 17, 85-104 (2005)
- [3] Gavazzi, S.M., McKenry, P.C., Jacobson, J.A., Julian, T.W., Lohman, B.: Modeling the effects of expressed emotion, psychiatric symptomology, and marital quality levels on male and female verbal aggression. *Journal of Marriage and Family* 62, 669-682 (2000)
- [4] Patterson, K., Grenny, J., McMillan, R., Switzler, A., Covey, S.R.: *Crucial conversations: Tools for talking when stakes are high*. McGraw Hill, New York (2002)
- [5] Nelson, J., Aboud, F.E.: The resolution of social conflict between friends. *Child Development* 56, 1009-1017 (1985)
- [6] Haaga, D.A., Ahrens, A.H.: A disagreement about the benefits of agreement seeking. *Psychological Inquiry* 3, 244-247 (1992)
- [7] Erath, S.A., Bierman, K.L.: Aggressive marital conflict, maternal harsh punishment, and child aggressive-disruptive behavior: Evidence for direct and mediated relations. *Journal of Family Psychology* 20, 217-226 (2006)
- [8] Bernieri, F.J.: Coordinated movement in human interaction: Synchrony, posture similarity, and rapport. *Dissertation Abstracts International* 49, 4601 (1989)

- [9] Bavelas, J.B., Black, A., Chovil, N., Lemery, C.R., Mullett, J.: Form and function in motor mimicry: Topographic evidence that the primary function is communicative. *Human Communication Research* 14, 275-299 (1988)
- [10] Núñez, R., Sweetser, E.: With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science* 30, 401-450 (2006)
- [11] Altorfer, A., Jossen, S., Wuermle, O.: Eine Methode zur zeitgenauen Aufnahme und Analyse des Bewegungsverhaltens. *Zeitschrift für Psychologie* 205, 83-117 (1997)
- [12] Neumann, R., Strack, F.: Mood contagion: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology* 79, 211-223 (2000)
- [13] Chartrand, T.L., Bargh, J.A.: The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology* 76, 893-910 (1999)
- [14] Dijksterhuis, A., Bargh, J.A.: The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology* 33, 1-40 (2001)
- [15] Kendon, A., Harris, R.M., Key, M.R.: The organization of behavior in face to face interaction. The Hague: Mouton and Co. (1975)
- [16] Cappella, J.N.: Coding mutual adaptation in dyadic nonverbal interaction. In: Manusov, V. (ed.) *The sourcebook of nonverbal measures*, Lawrence Erlbaum, Mahwah, NJ (2005)
- [17] Bernieri, F.J., Rosenthal, R.: Interpersonal coordination: Behavior matching and interactional synchrony. In: Feldman, R.S., Rime, B. (eds.) *Fundamentals of nonverbal behavior*, pp. 401-432. Cambridge University Press, New York (1991)
- [18] Grammer, K., Honda, R., Schmitt, A., Jütte, A.: Fuzziness of nonverbal courtship communication unblurred by motion energy detection. *Journal of Personality and Social Psychology* 77, 487-508 (1999)
- [19] Darwin, C.: The expression of the emotions in man and animals. University of Chicago Press, Chicago (1965)
- [20] Allport, G.W.: *Personality: A psychological interpretation*. Holt, Rinehart, & Winston, New York (1937)
- [21] Condon, W.S.: The relation of interactional synchrony to cognitive and emotional processes. In: Key, M.R. (ed.) *The relationship of verbal and nonverbal communication*, pp. 49-75. Mouton Publishers, The Hague (1980)
- [22] Kriz, J.: *Grundkonzepte der Psychotherapie*. München: Urban & Schwarzenberg (2001)
- [23] Philippot, P., Feldman, R.S., Coats, E.J.: *The social context of nonverbal behavior*. Cambridge University Press, New York (1999)
- [24] Philippot, P., Feldman, R.S., Coats, E.J.: The role of nonverbal behavior in clinical settings. In: Philippot, P., Feldman, R.S., Coats, E.J. (eds.) *Nonverbal behavior in clinical settings*, Oxford University Press, New York (2003)
- [25] Schefflen, A.E.: The Significance of Posture in Communication Systems. *Psychiatry* 27, 316-331 (1964)
- [26] Schefflen, A.E.: Quasi-courtship behavior in psychotherapy. *Psychiatry* 28, 245-257 (1965)
- [27] Schefflen, A.E.: Systems and psychosomatics. An introduction to psychosomatic manifestations of rapport in psychotherapy. *Psychosomatic Medicine* 28, 297-304 (1966)
- [28] Charny, E.J.: Psychosomatic manifestations of rapport in psychotherapy. *Psychosomatic Medicine* 28, 305-315 (1966)
- [29] McDowall, J.J.: Interactional synchrony: A reappraisal. *Journal of Personality and Social Psychology* 36, 963-975 (1978)

- [30] Maurer, R.E., Tindall, J.H.: Effect of postural congruence on client's perception of counselor empathy. *Journal of Counseling Psychology* 30, 158-163 (1983)
- [31] Willis, C.J.: The measurement of mutual nonverbal coordination in the psychotherapeutic process: An exploratory study of the development of an index for clinical use. *Dissertation Abstracts International* 50, 144 (1989)
- [32] Meltzoff, A.N., Prinz, W.: *The imitative mind: Development, evolution, and brain bases*. Cambridge University Press, New York (2002)
- [33] O'Toole, R., Dubin, R.: Baby feeding and body sway: An experiment in George Herbert Mead's. *Taking the role of the other*. *Journal of Personality and Social Psychology* 10, 59-65 (1968)
- [34] Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L.: Premotor cortex and the recognition of motor actions. *Brain Research Cognitive Brain Research* 3, 131-141 (1996)
- [35] Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G.: Action recognition in the premotor cortex. *Brain* 119, 593-609 (1996)
- [36] Gallese, V., Keysers, C., Rizzolatti, G.: A unifying view of the basis of social cognition. *Trends in Cognitive Science* 8, 396-403 (2004)
- [37] Boker, S.M., Rotondo, J.L.: Symmetry building and symmetry breaking in synchronized movement. In: Stamenov, M., Gallese, V. (eds.) *Mirror Neurons and the Evolution of Brain and Language*, John Benjamins Publishing, Amsterdam (2003)
- [38] Ramachandran, V.S.: Mirror Neurons and imitation learning as the driving force behind the great leap forward in human evolution. *EDGE* 69, (2000), available online: http://www.edge.org/3rd_culture/index.html
- [39] Pittam, J.: *Voice in Social Interaction: An Interdisciplinary Approach*. SAGE, Thousand Oaks, CA (1994)
- [40] Roth, W.-M., Tobin, K.: Solidarity and conflict: Prosody as interactional resource in intra- and intercultural communication involving power differences, www.educ.uvic.ca/faculty/mroth/PREPRINTS/Solidarity109.pdf (submitted)
- [41] Hillard, D., Ostendorf, M., Shriberg, E.: Detection of agreement vs. disagreement in meetings: training with unlabeled data. *Proceedings of HLT/NAACL* 2, 34-36 (2003)
- [42] Galley, M., McKeown, K., Hirschberg, J., Shriberg, E.: Identifying agreement and disagreement in conversational speech: use of Bayesian networks to model pragmatic dependencies. In: *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, Article No. 669, Barcelona (2004)
- [43] Sonka, M., Hlavac, V., Boyle, R.: *Image processing, analysis, and machine vision*. Chapman & Hall Computing, New York (1993)
- [44] Bobick, A.F., Davis, J.W.: The recognition of human movement using temporal templates. *IEEE Transactions on pattern analysis and machine intelligence* 23, 257-267 (2001)
- [45] Ramseyer, F., Tschacher, W.: Synchrony - A core concept for a constructivist approach to psychotherapy. *Constructivism in the Human Sciences* 11, 150-171 (2006)
- [46] Grammer, K., Keki, V., Striebel, B., Atzmüller, M., Fink, B.: Bodies in motion: A window to the soul. In: Voland, E., Grammer, K. (eds.) *Evolutionary aesthetics*, pp. 295-324. Springer, Heidelberg (2003)
- [47] Grammer, K., Kruck, K.B., Magnusson, M.S.: The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. *Journal of Nonverbal Behavior* 22, 3-29 (1998)
- [48] Bechinie, M.: BHS - The "broken heart syndrome". Ethological aspects of lovesickness. Unpublished thesis, University of Vienna (1998)

- [49] Tschacher, W., Ramseyer, F., Grawe, K.: Der Ordnungseffekt im Psychotherapieprozess: Replikation einer systemtheoretischen Vorhersage und Zusammenhang mit dem Therapieerfolg. *Zeitschrift für Klinische Psychologie und Psychotherapie* 36, 18–25 (2007)
- [50] Tschacher, W., Grawe, K.: Selbstorganisation in Therapieprozessen Die Hypothese und empirische Prüfung der "Reduktion von Freiheitsgraden" bei der Entstehung von Therapiesystemen. *Zeitschrift für Klinische Psychologie* 25, 55–60 (1996)
- [51] Tschacher, W.: *Prozessgestalten*. Göttingen: Hogrefe (1997)
- [52] Tschacher, W., Scheier, C., Grawe, K.: Order and pattern formation in psychotherapy. *Nonlinear Dynamics, Psychology, and Life Sciences* 2, 195–215 (1998)
- [53] Grawe, K.: *Psychological Therapy*. Seattle: Hogrefe (2004)
- [54] Infante, D.A., Wigley, C.J.: Verbal aggressiveness: An interpersonal model and measure. *Communication Monograph* 53, 61–69 (1986)
- [55] Jacobson, N., Gottman, J.: *When men batter women: New insights into ending abusive relationships*. Simon & Schuster, New York (1998)
- [56] Stone, D., Patton, B., Heen, S.: *Difficult conversations*. Penguin, New York (1999)
- [57] Shortt, J.W., Capaldi, D.M., Kim, H.K., Owen, L.D.: Relationship separation for young, at-risk couples: Prediction from dyadic aggression. *Journal of Family Psychology* 20, 624–631 (2006)

Gesture, Prosody and Lexicon in Task-Oriented Dialogues: Multimedia Corpus Recording and Labelling

Ewa Jarmolowicz, Maciej Karpinski, Zofia Malisz, and Michal Szczyszek

Centre for Speech and Language Processing, Adam Mickiewicz University,
Collegium Novum, Al. Niepodległości 4

61-784 Poznań, Poland

ewa@jarmolowicz.art.pl, maciej.karpinski@amu.edu.pl,

zmalisz@ifa.amu.edu.pl, szczysze@interia.pl

Abstract. The aim of the *DiaGest* Project is to study interdependencies between gesture, lexicon, and prosody in Polish dialogues. The material under study comprises three tasks realised by twenty pairs of subjects. Two tasks involve instructional, task-oriented dialogues, while the third is based on a question answering procedure. A system for corpus labelling is currently being designed on the basis of current standards. The corpus will be annotated for gestures, lexical content of utterances, intonation and rhythm. In order to relate various phenomena to the contextualized meaning of dialogue utterances, the material will also be tagged in terms of dialogue acts. Synchronised tags will be placed in respective annotation tiers in ELAN. A number of detailed studies related to the problems of gesture-prosody, gesture-lexicon and prosody-lexicon interactions will be carried out on the basis of the tagged material.

Keywords: dialogue, gesture, lexicon, derivation, prosody.

1 Introduction

While the holistic approach to dialogue analysis, comprising speech and gesture interaction, can hardly be considered "new" or "recent" [7], [29], [30], the attempts to implement it in corpus-based studies of spontaneous communication still face many difficulties and many fundamental questions have not been fully answered yet [35].

Due to substantial technological progress, standard desktop computers offer not only the possibility of detailed, large-scale audio signal analysis, but also video editing and labelling options. On the side of methodology, after the initial period of revolutionary chaos, an increasing tendency to define corpus labelling standards can be noticed. Nevertheless, even with the new software, hardware and methodological tools in hand, we can hardly cope with the abundant data coming from completely spontaneous, uncontrolled dialogue situations.

One way to obtain at least partially controlled and more manageable data is to put, indirectly, some constraints on the speakers by assigning them a specific task. Task-oriented dialogues are a frequent subject of study because they provide linguistic material which is both relatively well-structured and contains a certain dose of