IntechOpen

Types of Nonverbal Communication

Edited by Xiaoming Jiang

Types of Nonverbal Communication

Edited by Xiaoming Jiang

Published in London, United Kingdom



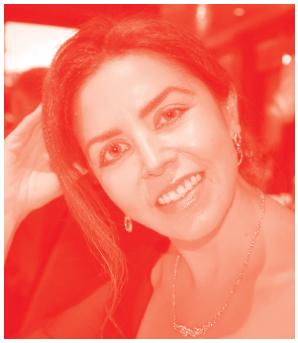




IntechOpen





Supporting open minds since 2005

Types of Nonverbal Communication
http://dx.doi.org/10.5772/intechopen.83004
Edited by Xiaoming Jiang

Contributors
Mahesh Goyani, I-Jui Lee, Manuela Valentini, Ario Federici, Maria Chiara Mancini, Izidor Mlakar, Matej Rojc, Darinka Verdonik, Simona Majhenič, Bashir Ibrahim, Usman Ambu Muhammad, Xiaoming Jiang

First published in London, United Kingdom, 2021 by IntechOpen
IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, 5 Princes Gate Court, London, SW7 2QJ, United Kingdom
Printed in Croatia

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Additional hard and PDF copies can be obtained from orders@intechopen.com

Types of Nonverbal Communication
Edited by Xiaoming Jiang
p. cm.
Print ISBN 978-1-83962-705-7
Online ISBN 978-1-83962-709-5
eBook (PDF) ISBN 978-1-83962-710-1

Meet the editor

Dr. Xiaoming Jiang is a professor from the Institute of Linguistics, Shanghai International Studies University, China. He obtained a BS in Psychology from East China Normal University and a Ph.D. in Cognitive Neuroscience from Peking University. He worked as a research fellow at the School of Communication Sciences and Disorders, McGill University, Montreal, Quebec, Canada, and as a senior speech scientist in nuance communication. His research utilizes experimental methodologies to uncover social and interpersonal aspects of human communicative processes including speaker confidence and trustworthiness. While serving as the editor for journals such as *Frontiers in Psychology* and *Frontiers in Communication*, he published nearly forty peer-reviewed articles as lead author in high-impact journals such as *NeuroImage*, *Journal of Experimental Psychology: Human Perception and Performance*, and *Speech Communication*.

Contents

Preface

The use of nonverbal cues in social activities is essential for human daily activities. Successful nonverbal communication relies on the acquisition of rules of using cues from body movement, eye contact, facial expression, tone of voice, and more. These nonverbal cues, with high relevance to evolutionary and socio-adaptive implications, are demonstrated to strengthen, complement, and conflict with our verbal messages, and exert immediate impact on a perceiver's inferences, evaluations, and decisions based on these cues. Although human nonverbal communication has been studied for decades, it is time to ask to what extent we have adapted our communication to a new age of artificial intelligence.

This book examines nonverbal behavior based on research efforts with state-of-the-art methodological approaches. It discusses how nonverbal communications are classified, how nonverbal cues adapt and function as social behavior, how nonverbal cues are used in different social domains and practices, how individuals in different cultures and groups express and understand nonverbal behaviors of their own and of others, how nonverbal communication is distinct in special populations, and how the use of nonverbal cues can be measured in different applied settings.

With contributions from scientists from disciplinary backgrounds including psychology, artificial intelligence and computer science, communication sciences and disorders, linguistics and literature, and philosophy and sociology, this book characterizes research on fundamental issues related to the processing, learning, categorizing, and consequences of nonverbal communications, as well as showcases solutions to applied questions on developing and adapting methods to measure nonverbal behaviors. Three unique questions are addressed by six independent chapters contributed by researchers from multiple scientific fields.

The first two chapters contribute to the classification and training of nonverbal communication with advanced technologies. Chapter 2 by Dr. Mahesh Goyani presents a novel computer vision technique to recognize human facial expression, which is efficient in dealing with both local and global appearance-based features. The author demonstrates the performance of such an approach to be robust in various real-world scenarios like recognizing expressions from low resolution, with small training samples, and in the presence of noise. Chapter 3 by Dr. Lee I-Jui analyzes the mechanisms and benefits of applying augmented and virtual reality techniques to the training of nonverbal social communicative skills in autistic children, highlighting the role of novel technologies that create an immersive setting for nonverbal development in a special population.

The next two chapters give a selective overview of the factors underlying the learning and evaluating of nonverbal communications in educational settings and in digital worlds. Chapter 4 by Dr. Manuela Valentini et al. discusses how preschool and school-aged children coordinate their body language during the acquisition of communicative capacities, putting forward the role of cultural factors in children's grasping of certain conventions to communicate one's non-verbal behaviors.

Chapter 5 by Dr. Xiaoming Jiang provides an empirical study on how human-like animal stickers are perceived when used as nonverbal communicative tools on social media (e.g., Chinese WeChat). Importantly, the author shows how the user's gender, interpersonal sensitivity, and attitudes towards the ethical use of animals affects various perceptual attributes of animal stickers.

The final chapters characterize the latest advancements that uncover the psychological nature (in particular the communicative intents of the users of nonverbal cues) underlying nonverbal communication in conversations and other applied settings. Chapter 6 by Dr. Izidor-Mlakar et al. analyzes the association between dialogue acts and nonverbal communicative intents, especially during turn management. The authors demonstrate how to use ELAN to perform multimodal analysis of audiovisual stimuli of dynamic conversations. Chapter 7 by Dr. Ibrahim Bashir et al. presents a range of functions of silence in the conversation and how these functions of silent speech may change as a function of sociocultural norms.

This book is a unique and timely contribution to the field of human nonverbal communication. It is designed for a large audience for a variety of purposes, including students and professors in academic institutions for teaching and research activities as well as active researchers in industries related to human communication for applying relevant knowledge to the development of communication products, for both healthy individuals and special populations. We would like to thank the staff at IntechOpen, especially Ms. Dolores Kuzelj for her whole-hearted assistance over the entire process of preparing and publishing the book, and Ms. Ana Simcic for contacting me about this great opportunity to serve as the book's editor.

Xiaoming Jiang
Shanghai International Studies University,
Institute of Linguistics,
Shanghai, China

Section 1

Introduction

Chapter 1

Introductory Chapter: On the Road towards the Social-Adaptive Implication of Nonverbal Communication

Xiaoming Jiang

1. The adaptive use of nonverbal cues under various communicative contexts

The nonverbal communication is by nature socially adaptive and has high relevance to the real-world application. Latest research efforts have witnessed a boost in the empirical studies of nonverbal communication which have yielded significant discoveries. Minimally verbal and nonverbal cues can be effective in predicting individuals with autism [1], cerebral palsy [2], dementia [3], and traumatic brain injury patients [4]. The interpersonal resonation of nonverbal behaviors has a unique role in the early detection of psychiatric status. The synchronized nonverbal acts can predict some level of severity in social anxiety disorder [5]. The synchrony of movement and facial expressions is a diagnostic features of depression [6].

The communication of nonverbal signals can be essential in understanding patient-doctor interaction, and more recently in particular, in nursing houses between nurses and older adults [7], and about chronic disease management consultations between patients with cancer and their oncologists [8]. The perception of nonverbal communication skills in the emergency department is associated with the department's service quality and patient satisfaction [9, 10]. The doctor-patient synchrony can mediate the strength of the social group effects on the social group effects on the intergroup trust [11].

In terms of the application in social media and political communication, the video creators' nonverbal communication can impact viewer's intention to subscribe online media [12]. Sex differences existed in nonverbal cues marking political debates [13]. Communicating nonverbal cues in public speaking can be trained for the spokesperson to face the media in times of crisis [14]. The perceived debate style from nonverbal cues of a leadership role has an immediate impact on the perceiver's response on social media (e.g. Twitter [15]). As for the consequences in the real-world decision making, nonverbal presence can even predict the evaluations of one's hirability [16] in the recruiting situation, and can predict whether the volleyball game which the target player participated in is successful or not [17]. Moreover, the deception-related cues can bias legal decisions in the justice system [18], and have demonstrated to be useful in predicting forensic cases such as bribery [19].

Nonverbal cues are typically multimodal and often accompany linguistic messages with many forms to achieve an efficient communication. Nonverbal behaviors can reveal the speaker's characteristics of dominance, trust, composure [20], eloquence [21], persuasion [22], and guilt [23] and can facilitate one's understanding

of a variety of social emotion, and can even be an index of one's cultural competency [24]. Individual differences in speaker's voice pitch signals are consistently manipulated in speech and in various vocalizations such as screams, roars and pain cries [25]. A speaker's gaze patterns predict group interactions and his/her charisma [26]. Emoticon can sometimes be unintentional cues of the writer's emotion in email communication [27]. The perception of a speaker's emotion can rely on subtle nonverbal behaviors and its interaction with the contextual cues. Voice quality and breathiness conveyed in the moan of pleasure play roles in determining the meaning of human nonverbal vocalization in the perceiver [28]. The perception of vocal emotional cues is not context-invariant and appears to be modulated by sound context [29].

Nonverbal communication has also increasingly received attention in studies on education and human development. Scales are developed to quantify the developmental trajectory of nonverbal communication during childhood [30]. The context-aware augmentative and alternative communication system has been developed to assist school children with intellectual disability [31]. The social-communicative gestures at baseline can predict verbal and nonverbal development for children with autism [32].

A crucial question is whether the virtual technology, popularly growing during COVID-19, enhances or hinders interpersonal communication [33]. The prevalence of mobile technology has affected the children's ability to read nonverbal emotional cue of others [34]. The understanding of nonverbal communication cues of human personality traits can be essential in human-robot interaction [35]. Algorithms have been developed based on nonverbal behavioral features. Automatic recognition of nonverbal mimicry has been achieved in medical video consultations [36]. Virtual negotiation can be affected by nonverbal cues of speakers which is altered by automatic algorithm [37]. The facial expressions can be efficiently modeled and generated by FACSHuman, a software program [38].

2. The present book: Featuring the nonverbal communication in the real-world application

These novel discoveries, summarized based on a cross-journal search and post-hoc comparisons of research papers, call for a proposal to systematically organize and plan a new book to address the nonverbal communication, its socio-adaptive nature and its various types.

In this book, we aim to respond to this call with several attempts. Our first endeavor is to reveal the novel approach to the classification and the training of nonverbal communication with advanced technologies.

Our second goal is to give a selective overview on the factors underlying the learning and the evaluation of nonverbal communications in educational settings and in digital worlds.

Our last effort is to characterize the latest advancement that uncovers the psychological nature underlying nonverbal communication in conversations and other applied settings.

Despite the topics relevant to other specific types of nonverbal communication (e.g. in drama and performing arts, etc.), the interesting works featured in this book, together with other contributions to the burgeoning field of nonverbal communication, have already demonstrated a promising new line of research to embed the study on nonverbal communication within an evolutionary and socio-adaptive perspective.

Author details

Xiaoming Jiang
Institute of Linguistics, Shanghai International Studies University, China

*Address all correspondence to: xiaoming.jiang@shisu.edu.cn

References

[1] Koegel, Lynn Kern; Bryan, Katherine M.; Su, Pumpki L.; Vaidya, Mohini; Camarata, Stephen. (2020). Definitions of Nonverbal and Minimally Verbal in Research for Autism: A Systematic Review of the Literature. JOURNAL OF AUTISM AND DEVELOPMENTAL DISORDERS. 50, 2957-2972.

[2] Soriano, Jennifer U.; Hustad, Katherine C. (2021). Speech-Language Profile Groups in School Aged Children with Cerebral Palsy: Nonverbal Cognition, Receptive Language, Speech Intelligibility, and Motor Function. DEVELOPMENTAL NEUROREHABILITATION. 24, 118-129.

[3] Brown, Ellen L. x; Agronin, Marc E.; Stein, Jennifer R. (2020). Interventions to Enhance Empathy and Person-Centered Care for Individuals With Dementia. RESEARCH IN GERONTOLOGICAL NURSING. 13, 158.

[4] Raukola-Lindblom, Marjaana; Vuorinen, Elina; Vartiainen, Riitta (2020). Recovering social participation - experience with a relational group intervention for traumatic brain injury patients. NEUROREHABILITATION. 47, 55-64.

[5] Asher, Maya; Kauffmann, Amitay; Aderka, Idan M. (2020). Out of Sync: Nonverbal Synchrony in Social Anxiety Disorder. CLINICAL PSYCHOLOGICAL SCIENCE. 8, 280-294.

[6] Altmann, Uwe; Bruemmel, Maria; Meier, Julija; Strauss, Bernhard. (2021). Movement Synchrony and Facial Synchrony as Diagnostic Features of Depression A Pilot Study. JOURNAL OF NERVOUS AND MENTAL DISEASE. 209, 128-136.

[7] Wanko Keutchafo, Esther L.; Kerr, Jane; Jarvis, Mary Ann. (2020). Evidence of nonverbal communication between nurses and older adults: a scoping review. BMC NURSING, 19, 53.

[8] Hamel, Lauren M.; Moulder, Robert; Harper, Felicity W. K.; Penner, Louis A.; Albrecht, Terrance L.; Eggly, Susan. (2021). Examining the dynamic nature of nonverbal communication between Black patients with cancer and their oncologists. CANCER, 127, 1080-1090.

[9] Orcajada Munoz, Irene; Amo Setien, Francisco Jose; Diaz Agea, Jose Luis; Hernandez Ruiperez, Tomas; Adanez Martinez, Maria de Gracia; Leal Costa, Cesar. (2021). The communication skills and quality perceived in an emergency department: The patient's perspective. INTERNATIONAL JOURNAL OF NURSING PRACTICE. 26.

[10] Stefanini, Alessandro; Aloini, Davide; Gloor, Peter; Pochiero, Federica (2021). Patient satisfaction in emergency department: Unveiling complex interactions by wearable sensors. JOURNAL OF BUSINESS RESEARCH. 129, 600-611.

[11] Goldstein, Pavel; Losin, Elizabeth A. Reynolds; Anderson, Steven R.; Schelkun, Victoria R.; Wager, Tor D. (2020). Clinician-Patient Movement Synchrony Mediates Social Group Effects on Interpersonal Trust and Perceived Pain. JOURNAL OF PAIN. 21, 1160-1174.

[12] Park, Jinsu; Lee, Yuri. (2021). Luxury haul video creators' nonverbal communication and viewer intention to subscribe on YouTube. SOCIAL BEHAVIOR AND PERSONALITY. 49, e10119.

[13] Bramlett, Josh C. (2021). Nonverbal Communication in Political Debates. JOURNAL OF BROADCASTING & ELECTRONIC MEDIA. 65, 182-184.

[14] De Waele, Aurelie; Claeys, An-Sofie; Opgenhaffen, Michael. (2020). Preparing to face the media in times of crisis: Training spokespersons' verbal and nonverbal cues. PUBLIC RELATIONS REVIEW. 46, 101871.

[15] Bucy, Erik P.; Foley, Jordan M.; Lukito, Josephine; Doroshenko, Larissa; Shah, Dhavan, V; Pevehouse, Jon C. W.; Wells, Chris. (2020). Performing populism: Trump's transgressive debate style and the dynamics of Twitter response. NEW MEDIA & SOCIETY. 22, 634-658.

[16] Tu, Min-Hsuan; Gilbert, Elisabeth K.; Bono, Joyce E. (2021). Is beauty more than skin deep? Attractiveness, power, and nonverbal presence in evaluations of hirability. PERSONNEL PSYCHOLOGY.

[17] Durdubas, Deniz; Martin, Luc J.; Koruc, Ziya (2021). An examination of nonverbal behaviours in successful and unsuccessful professional volleyball teams. INTERNATIONAL JOURNAL OF SPORT AND EXERCISE PSYCHOLOGY. 19, 120-133.

[18] Denault, Vincent. (2020). Misconceptions About Nonverbal Cues to Deception: A Covert Threat to the Justice System? FRONTIERS IN PSYCHOLOGY, 11, 573460.

[19] Baranov, Anatoliy N. (2020). SPEECH AND GESTURE IN FORENSIC LINGUISTICS IN BRIBERY CASES: TOWARDS SEMANTICS AND PRAGMATICS OF "CLOSED" DISCOURSES. Science Journal of Volgograd State University. Linguistics. Science Journal of Volgograd State University. Linguistics. 19, 64-76.

[20] Burgoon, Judee K.; Wang, Xinran; Chen, Xunyu; Pentland, Steven J.; Dunbar, Norah E. (2021). Nonverbal Behaviors Speak Relational Messages of Dominance, Trust, and Composure. FRONTIERS IN PSYCHOLOGY. 12, 624177.

[21] Rockwell, Stephen J. (2021). Eloquence Embodied: Nonverbal Communication among French and Indigenous Peoples in the Americas. JOURNAL OF AMERICAN HISTORY. 107, 985-986.

[22] Vaughan-Johnston, Thomas I.; Guyer, Joshua J.; Fabrigar, Leandre R.; Shen, Charlie (2021). The Role of Vocal Affect in Persuasion: The CIVA Model. JOURNAL OF NONVERBAL BEHAVIOR.

[23] Julle-Daniere, Eglantine; Whitehouse, Jamie; Mielke, Alexander; Vrij, Aldert; Gustafsson, Erik; Micheletta, Jerome; Waller, Bridget M. (2020). Are there non-verbal signals of guilt? PLOS ONE. 15, e0231756.

[24] Wong, Sarah; Plowman, Christina; Puri, Davina; Nwibe, Ife . (2021). Why we have to move beyond the idea of cultural competency. MEDICAL EDUCATION ONLINE 26, 1841398.

[25] Pisanski, Katarzyna; Raine, Jordan; Reby, David. (2020). Individual differences in human voice pitch are preserved from speech to screams, roars and pain cries. ROYAL SOCIETY OPEN SCIENCE. 7, 191642.

[26] Maran, Thomas; Moder, Sebastian; Furtner, Marco; Ravet-Brown, Theo; Liegl, Simon (2020). From self-report to behavior: Mapping charisma onto naturalistic gaze patterns. PERSONALITY AND INDIVIDUAL DIFFERENCES. 152, 109562.

[27] Blunden, Hayley; Brodsky, Andrew. (2021). Beyond the Emoticon: Are There Unintentional Cues of Emotion in Email? PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN. 47, 565-579.

[28] Anikin, Andrey (2020). A Moan of Pleasure Should Be Breathy: The Effect of Voice Quality on the Meaning of Human Nonverbal Vocalizations. PHONETICA. 77, 327-349.

[29] Liuni, Marco; Ponsot, Emmanuel; Bryant, Gregory A.; Aucouturier, J. J. (2020). Sound context modulates perceived vocal emotion. BEHAVIOURAL PROCESSES. 172, 104042.

[30] Oryadi-Zanjani, Mohammad Majid. (2020). Development of the Childhood Nonverbal Communication Scale. JOURNAL OF AUTISM AND DEVELOPMENTAL DISORDERS. 50, 1238-1248.

[31] Chan, Rosanna Yuen-Yan; Sato-Shimokawara, Eri; Bai, Xue; Yukiharu, Motohashi; Kuo, Sze-Wing; Chung, Anson. (2020). A Context-Aware Augmentative and Alternative Communication System for School Children With Intellectual Disabilities. IEEE SYSTEMS JOURNAL. 14, 208-219.

[32] Laister, Dominik; Stammler, Magdalena; Vivanti, Giacomo; Holzinger, Daniel. (2021). Social-communicative gestures at baseline predict verbal and nonverbal gains for children with autism receiving the Early Start Denver Model. AUTISM, 1362361321999900.

[33] Ruben, Mollie A.; Stosic, Morgan D.; Correale, Jessica; Blanch-Hartigan, Danielle. (2020). Is Technology Enhancing or Hindering Interpersonal Communication? A Framework and Preliminary Results to Examine the Relationship Between Technology Use and Nonverbal Decoding Skill. FRONTIERS IN PSYCHOLOGY, 611670.

[34] Uhls, Yalda T.; Broome, Julia; Levi, Shani; Szczepanski-Beavers, Jade; Greenfield, Patricia. (2020). Mobile Technologies and Their Relationship to Children's Ability to Read Nonverbal Emotional Cues: A Cross-Temporal Comparison. CYBERPSYCHOLOGY BEHAVIOR AND SOCIAL NETWORKING. 23, 465-470.

[35] Shen, Zhihao; Elibol, Armagan; Chong, Nak Young. (2020). Understanding nonverbal communication cues of human personality traits in human-robot interaction. IEEE-CAA JOURNAL OF AUTOMATICA SINICA. 7, 1465-1477.

[36] Wu, Kaihang; Liu, Chunfeng; Calvo, Rafael A. (2020). Automatic Nonverbal Mimicry Detection and Analysis in Medical Video Consultations. INTERNATIONAL JOURNAL OF HUMAN-COMPUTER INTERACTION. 36, 1379-1392.

[37] Baten, Raiyan Abdul; Hoque, Ehsan. (2021). Technology-Driven Alteration of Nonverbal Cues and its Effects on Negotiation. NEGOTIATION JOURNAL. 37, 35-47.

[38] Gilbert, Michael; Demarchi, Samuel; Urdapilleta, Isabel. (2021). FACSHuman, a software program for creating experimental material by modeling 3D facial expressions. BEHAVIOR RESEARCH METHODS.

Section 2

Nonverbal Communication in Technology

Chapter 2

Nonverbal Communication through Facial Expression in Diverse Conditions

Mahesh Goyani

Abstract

In this chapter, we investigated computer vision technique for facial expression recognition, which increase both - the recognition rate and computational efficiency. Local and global appearance-based features are combined in order to incorporate precise local texture and global shapes. We proposed Multi-Level Haar (MLH) feature based system, which is simple and fast in computation. The driving factors behind using the Haar were its two interesting properties - signal compression and energy preservation. To depict the importance of facial geometry, we first segmented the facial components like eyebrows, eye, and mouth, and then applied feature extraction on these facial components only. Experiments are conducted on three well known publicly available expression datasets CK, JAFFE, TFEID and in-house WESFED dataset. The performance is measured against various template matching and machine learning classifiers. We achieved highest recognition rate for proposed operator with Discriminant Analysis Classifier. We studied the performance of proposed approach in several scenarios like expression recognition from low resolution, recognition from small training sample space, recognition in the presence of noise and so forth.

Keywords: facial expression recognition, multi-level Haar, local mean binary pattern, local Haar mean binary pattern, classification

1. Introduction

1.1 Background

Communication is not possible without some channel. Communication between human is modeled by two ways: multiplicity and modality. Multiplicity defines more than one way for the communication and modality defines the way human senses are used to perceive signals from outer world [1]. Speech and vocal information are communicated through the auditory channel, whereas facial expression is communicated via the visual channel. Organs such as nose, ear, skin provides the different modalities for the communication. Multi-channel communication is highly robust; failure of one channel can be compensated by another channel.

Facial expression provides an important behavioral measure for studies of emotion, cognitive processes, and social interaction [2]. For a human being, recognition of face and facial expression is a trivial task. We discriminate the faces with almost

no effort in a fraction of a second. But it is equally challenging to teach a machine to perform the same task.

Expressions are not mere changes in muscle position, rather a complex psychophysiological process. The psychological process of thoughts emerging in mind is followed by a physiological process in which the thoughts render as expressions on the face by means of muscle deformation. The muscle movement lasts for a brief period of about 250 ms to 5 sec. Hence recognizing expressions from the spontaneous image is harder compared to posed still images [3].

Recognition of pure expression is difficult to wide range of expressions, as well as a same expression might have different intensities. Schmidt and Cohn [4] noted 18 unique classes of the expression smile. Intensity of expression can vary from gentle to peak.

1.2 Expression representation

According to the psychological and neurophysiological studies, there are six basic emotions. Each basic emotion is associated with one unique facial expression. Facial expressions can be represented using: discrete category model/Judgmental Coding Scheme (prototype expressions) or Facial Action Coding System (FACS) model.

1.2.1 Judgmental coding scheme

As the name suggests, the model classifies the expressions on subjective judgment. Prototypic expressions are a subjective measure of the texture like wrinkles, bulges, furs on the face, which is useful for judging the expression. Ekman and Friesen categorized expressions into six classes: Happiness (HA), Sadness (SA), Surprise (SU), Anger (AN), Fear (FE), and Disgust (DI) [5], which are portrayed in **Figure 1**. These six expressions serve as a ground truth label, and instead of distinguishing the comprehensive facial features, most FER systems attempt to recognize a small set of these prototypic expressions.

The other way of describing expression is using geometry of the face. Judgmental coding scheme based algorithms use appearance features for expression recognition. Descriptive coding scheme described in next section uses geometry of the face for expression recognition, which is more robust compared to judgmental coding scheme. However, extraction of exact Action Unit (AU) is challenging.

1.2.2 Descriptive coding scheme

Later, in 1978, Ekman and Friesen developed the Facial Action Coding System (FACS) [6], which describes and encodes the facial expressions based on the movements of the facial muscles. These codes are called action units (AUs). FACS

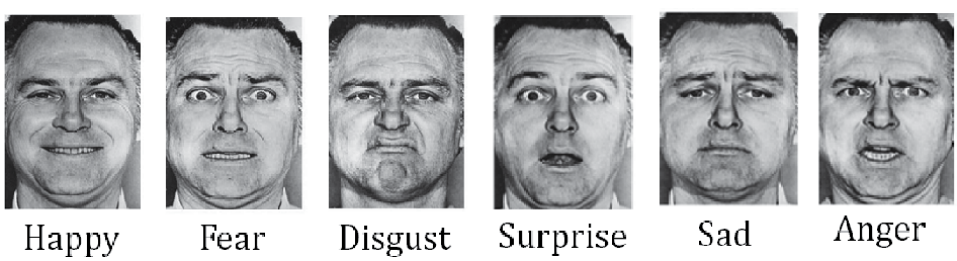


Figure 1.
Basic six expressions postulated by Ekman and Friesen.

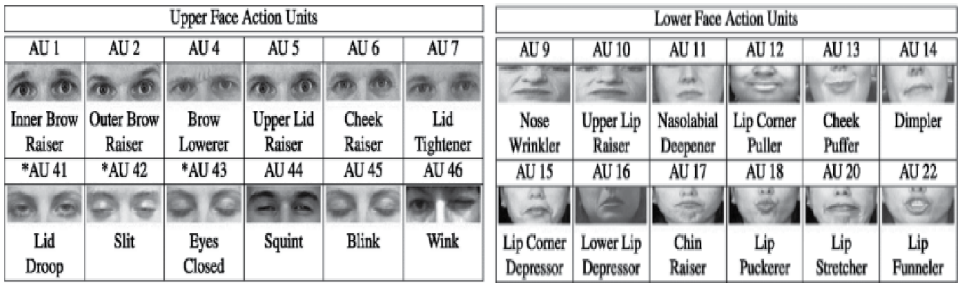


Figure 2.
Few of the upper and lower face action units.

identifies the facial muscles that cause changes in particular facial expression, thus enabling facial expression analysis. FACS consists of 44 action units describing the facial behaviors. FACS and the six prototypic expressions form the foundation for facial expression analysis and recognition research. **Figure 2** shows few of the upper and lower face action units.

Action Units can be additive or non-additive. If the appearance of AU is independent, then AU is said to be additive. When expression changes, different AU's get activated. In some expression, AU's are mixed and hence change the appearance of some AU during muscle deformation. AUs are said to be non-additive if they modify each other's appearance [7]. Each expression can be represented as a combination of one or more additive or non-additive AUs. For example, 'fear' can be represented as a combination of AUs 1, 2 4, 5, 7 and 26. Ekman and Friesen reported more than 7000 such combinations of the AUs [8]. In order to get the expression estimation, the FACS code needs to be converted into the Emotional Facial Action System. Even a good trained coder takes one to three hours of time to label one minute video on a frame by frame basis [9].

1.3 Applications

FER plays a vital role in many applications, such as human-computer interaction, indexing and retrieving images based on expressions, emotion analysis, image understanding, synthetic face animation, etc. A comprehensive study on recent advancements in affect recognition and its applications to HCI can be found in a survey by Zeng et al. [10].

Online Multiplayer Games (MOG) are increasingly becoming popular. Many FER based MOGs have been studied and proposed [11]. Applications of FER are not just limited to the physiological domain; rather it has touched many aspects of engineering, medical, social communication, entertainment, and automation. Application area of FER covers a broad spectrum, including grading of physical pain, smile detection [12–14], driver fatigue detection [15], patient pain assessment [16], video indexing, robotics and virtual reality [3], depression detection [10] etc.

Bartlett et al. [17] have successfully used their face expression recognition system to develop an animated character that mirrors the manifestations of the user. They have also managed to deploy the recognition system on Sony's Aibo Robot and ATR's RoboVie [17]. Anderson and McOwen developed an interesting application called the 'EmotiChat' [18]. It provides set of emojis for the easier and quick communication. The FERS is connected to this chat application, and it automatically inserts emoticons based on the user's facial expressions. Recently, Microsoft developed a fascinating Emotion API [19], which detects the face from an image and finds the weight of each expression.

1.4 Scopes and challenges

The main issue in the design of ideal automated expression analyzer is the degree of automation. All the stages - face detection, facial representation, and expression classification – should be fully automated. However, incorporation of these operations in the system depends on an application where the analyzer is to be used. Real-time performance is not expected if the analyzer is to be utilized for the study of behavioral science. Whereas, running time of the systems is an important issue for advanced user interfaces, in which delay of few seconds makes the system non-effective or non-usable [20].

Expression recognition in the low-resolution environment is almost unaddressed. Real-time videos like conference recordings, surveillance videos etc. are normally available in low resolution. Precise recognition of expression in such environment is challenging task. In 2004, Tian [21] used geometric and appearance based features to perform expression recognition in low-resolution images. Bartlett et al. [8] evaluated the performance of Gabor features and achieved noticeable accuracy. Later, in 2009 Shan et al. [22] investigated Gabor and LBP features for FER in a similar environment. Jabid et al. [23] evaluated the performance of Local Directional Pattern (LDP) features for low-resolution images.

To provide the standardized platform, Facial Expression Recognition Analysis (FERA) challenge events are being held by Social Signal Processing Network (SSPNET) in conjunction with Face and Gesture Recognition Group. Two such editions of FERA were held in 2011 at Santa Barbara [24], California and in 2015 at Ljubljana, Slovenia [25]. FERA 2017 is to be held in Washington, USA in March 2017. FERA brings the researchers across the globe under a common roof to understand and solve the issues of FER.

Based on the study of previous work, we list out following challenges in the field of facial expression recognition:

- Approaches are evaluated for person dependent databases only.
- Generalizing approach for spontaneous expressions is still an open area.
- The Very little contribution is made for the occluded facial expression recognition.
- Work cited in the different literature is addressing only one or two databases.
- Facial expression recognition under noisy environment is rarely addressed.
- The system is expected to run effectively on profile views.
- Expression recognition under low-resolution environment is still an almost unaddressed issue.

2. Multi level Haar wavelet based system

Texture and geometry convey complementary yet important information for FER. Studies [10] have shown that facial expression information is equally conveyed by geometric fiducial points and texture features. It has been observed that expression might have similar texture features bud different geometric features and

vice versa [26]. Experiments have shown that combination of both the type of features could prove better for implementation of FERS [27].

The proposed method detects the facial components, which in turn, effectively reduces the computation and improves the accuracy. A prototype human face is shown in **Figure 3**. Rectangle with dotted border indicates the region of interest which can further be used for feature extraction. The Larger region itself may contain smaller regions of interest within it.

2.1 Pre-processing

Preprocessing of the face and locating Region of Interest (ROI) is a crucial step for robust feature extraction. Segmentation of local facial components leads to significant reduction in computation cost of both - the feature extraction and classification. Following Tian [21], Shan et al. [22] and Baughrara et al. [28], we used fixed eye distance-based approach to normalize the face. Shan et al. [22] fix the distance between eyeballs to 52 pixels and face is cropped and normalized to 150×110 pixels using prior knowledge of facial geometry. Most of the literature have preferred manual or semi-automated approach for eye registration. However, our approach is completely automatic. We used iterative approach for eye registration. At first, eye pair is detected using cascade object classifier proposed by Viola-Jones (Refer **Figure 4**). Eye segment is thresholded and complemented using global threshold estimation.

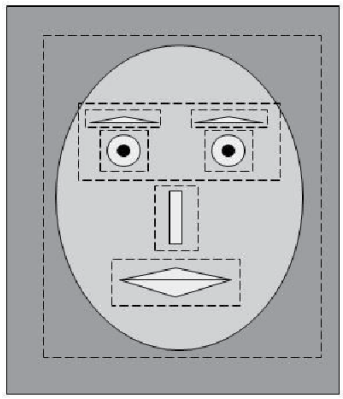


Figure 3.
Face prototype and few regions of interest.

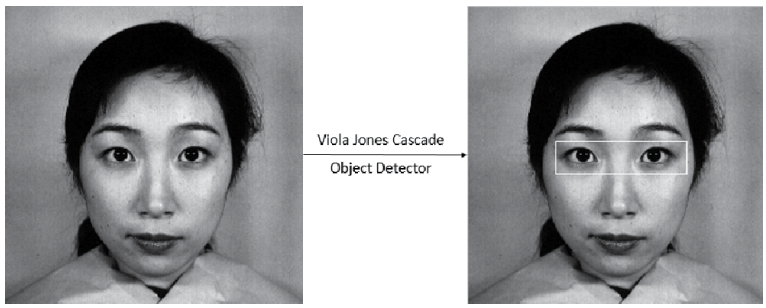


Figure 4.
Eye detection using Viola-Jones cascade object detector.

The binary image contains some unwanted small regions which satisfy the global threshold. From the prior knowledge, areas with less than 65 pixels are removed, so that binary image contains only the eyeball region. The thresholded eye region may not be connected due to the difference in skin tones of the subjects. Morphological erosion operation is applied with 3 × 3 structuring element having all 1's to connect areas around the eyeball. Let A represents the binary image of eye strip and B is the structuring element. In integer grid space E, erosion of the binary image A is defined as,

$$A \oplus B = \{z \in E \mid B_z \subseteq A\} \tag{1}$$

Where, B_z is the translation of B by the vector z, i.e.

$$B_z = \{b + z \mid b \in B\}, \forall x \in E \tag{2}$$

Centroid of both eyes is computed after applying erosion. Let (l_x, l_y) and (r_x, r_y) represents the spatial coordinates of the centroid of left and right eyes respectively. Even images are acquired in a controlled environment; head of certain subjects are not in exact upright frontal position. Such faces introduce alignment error, so we performed eyeball registration by measuring the angle of the line joining the eyeballs. If the face is perfectly vertically positioned, then the slope of the line joining eyeballs would be θ. Otherwise, it would be non-zero, and the face is aligned by performing negative rotation of the angle around the z-axis. Let Δx and Δy represent the difference of x and y coordinates of eyeballs. Thus, $\Delta x = r_x - l_x$ and $\Delta y = r_y - l_y$. Angle is estimated by taking tan^{-1} of the slope of the line,

$$\theta = tan^{-1}\left(\frac{\Delta y}{\Delta x}\right) \tag{3}$$

If the angle is greater than the prescribed threshold, then the image is rotated by negative rotation angle, and the process is reiterated from the eye pair detection phase. **Figure 5** demonstrates the angle estimation for slant face.

Once the angle threshold is adjusted within the range, the image is rescaled such that distance between eyeballs maintained at 52 pixels. Scaling factor is computed by normalizing the required eye distance by actual eye distance.

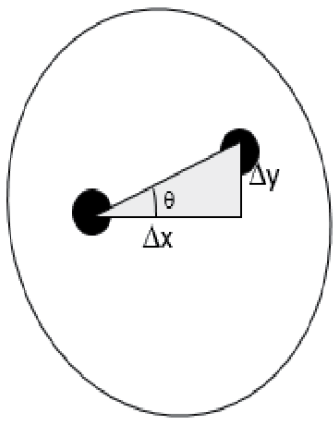


Figure 5.
Angle estimation from eyeball for eye registration.

$$SaclingFactor = \frac{52}{r_x - l_x} \tag{4}$$

Using advanced knowledge of facial geometry, we crop the facial components based on eyeball position and distance between them. Dimensions used for our experiments are portrayed in **Figure 6**.

This process registers the eye of all the images used in the dataset. The registration process significantly improves the performance. The spatial features would be more correlated now. We evaluate the performance of upper and lower facial

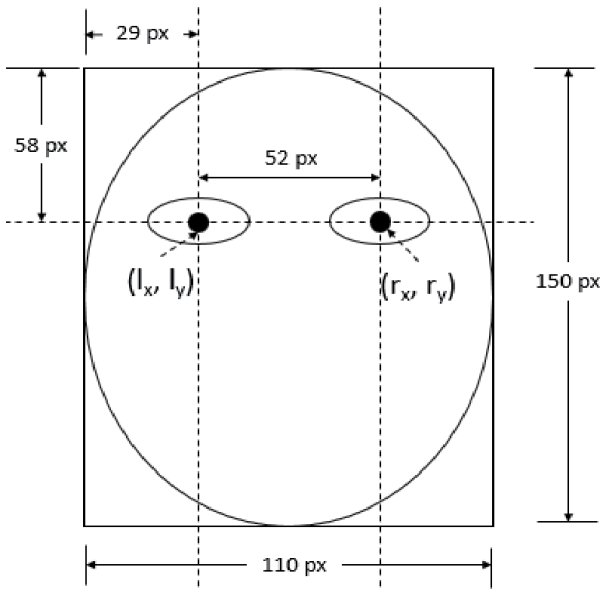


Figure 6.
Facial dimensions estimated from eyeball distance.

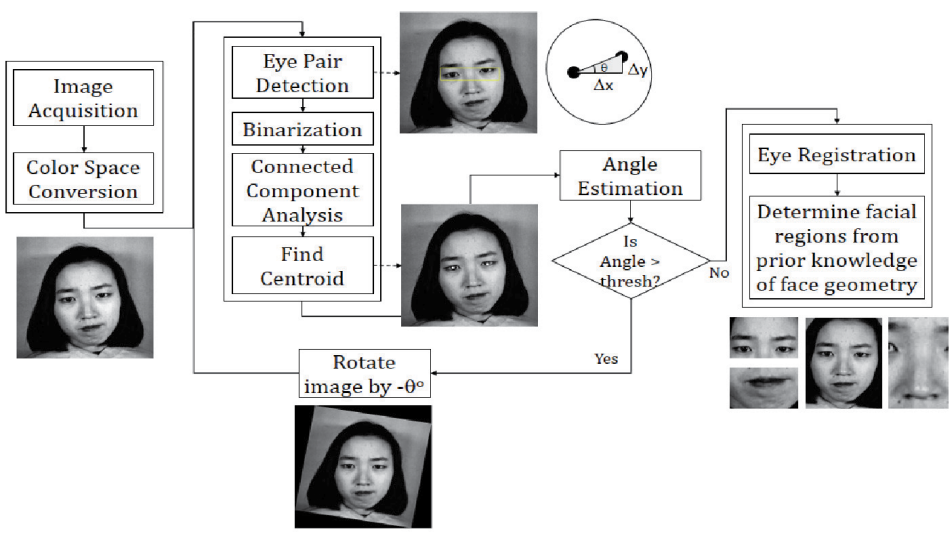


Figure 7.
Preprocessing flow for face and facial component extraction.

Component	Face	Eye	Mouth	Nose
Resolution	150 × 110	60 × 90	40 × 60	70 × 40

Table 1.
Dimensions of cropped facial components.

regions for expression recognition, and hence we also cropped top and bottom face regions. The entire process is explained in **Figure 7**.

Extracted geometric components are normalized and send to feature extraction module. The normalized size of the individual component is listed in **Table 1**.

2.2 Feature extraction

Haar functions were introduced by mathematician Alfred Haar [29]. A Haar wavelet is the simplest type of wavelet. It decomposes the image into one low-frequency band and number of high-frequency bands, known as coarse signal and detail signals respectively. Results are analogs to the output of low pass and high pass filters. Corse signal is an approximation of luminance and chrominance distribution of the original signal. In discrete form, Haar wavelets are related to a mathematical operation called the Haar transform. The Haar transform serves as a prototype for rest of all wavelet transforms. It provides a natural mathematical structure for describing the patterns [30].

The digital image is a discrete signal, which is a function of time with values occurring at discrete positions or time intervals. A discrete signal of length N is represented as $f = (f_1, f_2, \ldots, f_N)$. The values $f_1, f_2, \ldots, f_N$ are the approximation of analog signal g, measured at the time intervals $t = t_1, t_2, \ldots, t_N$. Components of signal f are obtained as,

$$f_1 = g(t_1), \quad f_2 = g(t_2), \quad \ldots, \quad f_N = g(t_N) \tag{5}$$

Haar wavelet decomposes the signal into two sub signals called *running average* or *trend* and *running difference* or *fluctuation*. The first trend sub signal, $a1 = (a_1, a_2, \ldots, a_{N/2})$, for the signal f, is computed by taking a running average of a pair of components of f. Mathematically, we can compute a_i as,

$$a_m = \frac{f_{2m-1} + f_{2m}}{\sqrt{2}} \tag{6}$$

for m = 1, 2, ..., N/2. Multiplication of average by $\sqrt{2}$ is needed in order to ensure that the Haar transform preserves the energy of a signal.

The other sub signal is called the first fluctuation, which is denoted by $d1 = (d_1, d_2, \ldots, d_{N/2})$, and it is computed by taking a running difference of a pair of values of f. In general,

$$d_m = \frac{f_{2m-1} - f_{2m}}{\sqrt{2}} \tag{7}$$

for m = 1, 2, ..., N/2. The Haar transform is performed in several stages, or levels. The first level is the mapping H_1 defined by,

$$f \overset{H_1}{\rightarrow} (a^1 \mid d^1) \tag{8}$$

The mapping H_1 in Eq. (8) has an inverse. Its inverse maps the transformed signal $(a^1|d^1)$ back to the signal f, via the following formula:

$$f = \left(\frac{a_1 + d_1}{\sqrt{2}}, \frac{a_1 - d_1}{\sqrt{2}}, \ldots, \frac{a_{N/2} + d_{N/2}}{\sqrt{2}}\right) \tag{9}$$

For Multi-Level Haar feature extraction, the same decomposition is repeatedly applied to latest trend signal in each iteration.

2.3 Experimental setup

We conduct the experiments on three widely used comprehensive datasets, Cohn-Kanade (CK) [31], Japanese Female Facial Expression (JAFFE) [32] and Taiwanese Facial Expression Image Database (TFEID) [33]. Existing datasets rarely address the issues of spontaneous expressions. Most of the time, images are acquired under a static environment with a fixed illumination source. On the other hand, real-life scenarios are very different. We addressed all possible issues by considering images of different ethnicity, age, pose, illumination, and occlusion in WESFED. WESFED dataset was created by collecting the images from google. Random images have been processed, faces were detected using Viola Jones Cascade face detector. Each face was manually labeled by 10 persons. SVM based model was also created to classify the expression of cropped face. Majority voting based scheme was employed to label the face. WESFED dataset contains subjects from various country, different age groups, different head positions, varying illumination condition and so on. Details of a number of images used for the experiment from all datasets are listed in **Table 2**.

We considered basic seven expressions anger (AN), disgust (DI), fear (FE), happy (HA), sad (SA), surprise (SU) and neutral (NE), for our experiment. Subjects from all four datasets with all seven expressions are depicted in **Figure 8**.

2.4 Result analysis

2.4.1 Optimal parameter selection

The performance of the algorithm is bound to many parameters like a number of features, the number of images used to train the model, regions size used to compute the features, etc. Derivation of optimal combination of parameters follows here:

A. Estimation of Number of Eigenvectors

More often than not, kernel methods generate a large feature vector. It has few obvious disadvantages. Training the classifier with such a large input vector is time-

	AN	DI	FE	HA	SU	SA	NE	Total
CK	110	120	100	280	130	220	320	1280
JAFFE	30	29	32	31	31	30	30	213
TFEID	34	40	40	40	39	36	39	268
WESFED	130	60	66	204	133	145	182	920

Table 2.
Number of images used for experiment from datasets.

consuming and often leads to poor generalization. Dimensions of the feature vector has been reduced by Linear Discriminant Analysis (LDA). To handle the issue of singular matrix problem, we have employed Principal Component Analysis (PCA). We applied PCA on original feature vector, and LDA is applied on PCA subspace. For C class problem, LDA produces (C – 1) features.

To find the optimal number of features, we varied the number of eigenvectors from 20 to 200 in step of 20. **Table 3** shows the performance of discussed approach on JAFFE against various classifiers with 2-fold cross-validation strategy. Performance is reported for two template matching strategy – Chi-Square (CS) and Cosine distance (CO), and two machine learning classifiers –Least Squares Support

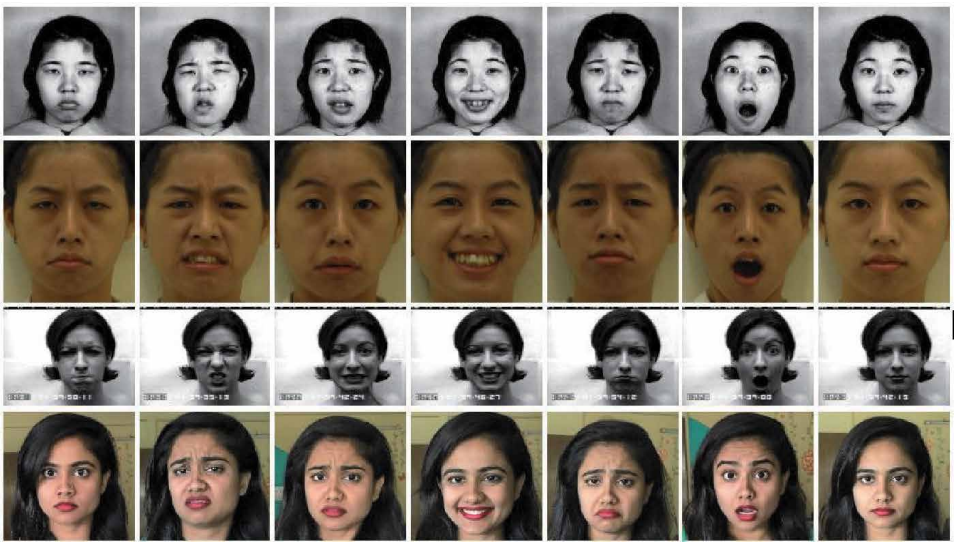


Figure 8.
Snapshots of various expressions from JAFFE (first row), TFEID (second row), CK (third row) and WESFED (fourth row) datasets.

	Template Matching		Machine Learning		AP[a]	API[b]
#	CS	CO	LS-SVM	DA		
20	76.43	76.24	71.57	73.48	74.43	0.00
40	91.10	90.33	90.90	91.86	91.05	16.62
60	94.62	93.95	93.19	94.90	94.17	3.12
80	96.05	95.86	95.38	95.76	95.76	1.60
100	96.62	96.52	96.52	96.62	96.57	0.81
120	96.62	96.62	96.43	96.62	96.57	0.00
140	97.00	97.00	97.00	97.00	97.00	0.43
160	97.00	97.00	97.00	97.00	97.00	0.00
180	97.00	97.00	97.00	97.00	97.00	0.00
200	97.00	97.00	97.50	98.00	97.38	0.38

[a] *Average Performance.*
[b] *Average Performance Improvement.*

Table 3.
Effect of eigenvectors on performance (%), dataset: JAFFE.

Vector Machine (LS-SVM) with RBF kernel and Discriminant Analysis (DA) classifier. Performance of all four classifiers is averaged to find Average Performance (AP) and Average Performance Improvement (API) is analyzed to select the optimal number of directions for PCA projection.

To choose the optimal number of eigenvectors, we averaged the performance of all four classifiers. To balance the accuracy-computation trade-off, we choose 140 eigenvectors for the further analysis.

JAFFE dataset contains only female subjects. To add the gender-specific variation, we also performed the same experiment on TFEID dataset, which includes 50% male and an equal number of female. Although TFEID contains male and female both gender, JAFFE and TFEID do not have ethnicity diversion. All subjects in both datasets belong to the same ethnicity. To test the robustness of algorithm against various diversities, we also conduct the experiment on comprehensive CK dataset. In CK, 65% of the subjects are female, and 35% are male. 15% of subjects belong to African-American background and 3% subjects belong to Asian or the Latino-American background. Images in CK contains large variations in illumination. We also test the accuracy of the system for our in-house dataset WESFED. We conducted all experiments on all three datasets with common parameters and results are shown in **Figure 9**.

B. Estimation of Region Size.

In the prototypic facial expression, textures such as wrinkles, bulges, furs play a crucial role. To extract the local texture features, we divided face image into M × N regions. To find the optimal number of regions, we divide images into 1 × 1, 3 × 3,

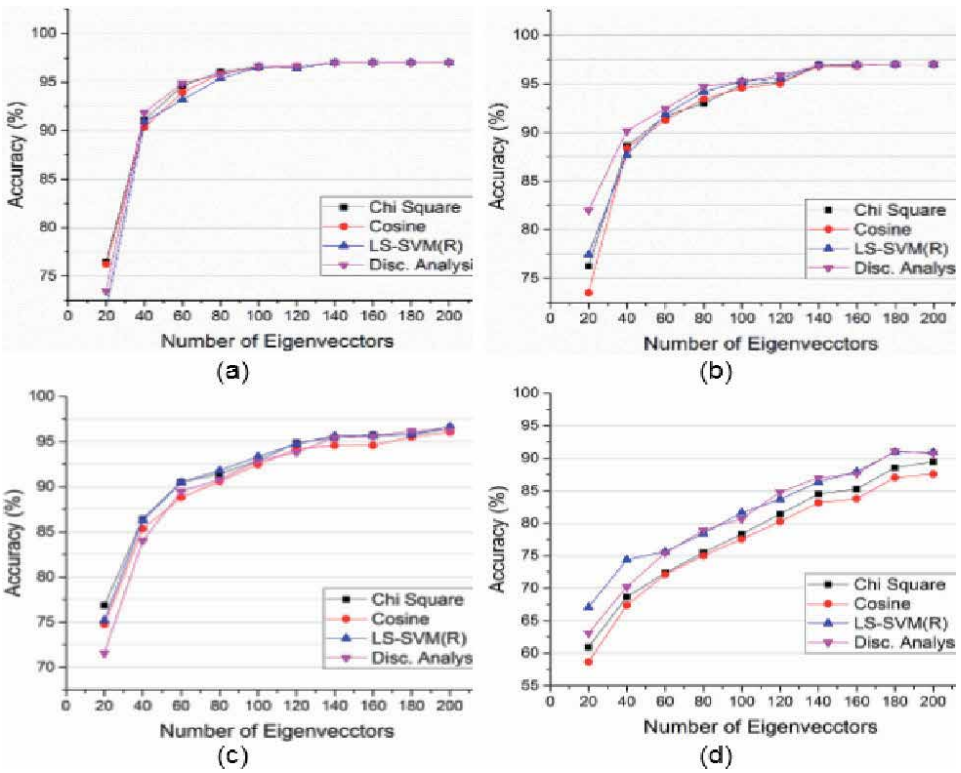


Figure 9.
Plot of a number of eigenvectors vs. accuracy (%) for (a). JAFFE, (b). TFEID, (c). CK and (d). WESFED datasets.

5 × 5, 7 × 6 and 9 × 8 blocks. Bartlett et al. [8], Tian [21], Shan et al. [22], Jabid et al. [23] have also conducted experiments in these neighborhoods. Larger regions size fails to capture the texture of small size. Small regions effectively capture local and spatial relationship. However, after a certain point, the smaller regions introduce unnecessary computation and feature vector becomes too large to train the classifier efficiently. We used 100 eigenvectors in the experiment. Performance behavior on JAFFE and TFEID dataset for a different number of blocks is stated in **Figure 10**.

With 1 × 1 region, we can only derive the holistic features, which also suffers from the feature localization. From the **Figure 10**, we observed that algorithm performs well for 7 × 6 regions on both datasets. From above results, we chose a number of regions to be 7 × 6, as it gives a proper balance between accuracy and computational time. With more regions, dimensions of feature vector grow tremendously and PCA also takes more time to compute covariance matrix, eigenvector, and eigenvalues of huge feature matrix.

C. Selecting Classifier.

Certain classifiers are good at classifying specific features only. We evaluated the performance of MLH feature descriptor against various template matching and machine learning based classifiers. We tested out a system for L2 norm, Chi-Square, Cosine, Correlation and k-NN based template matching classifiers. We also measured the performance using various machine learning classifiers like Artificial

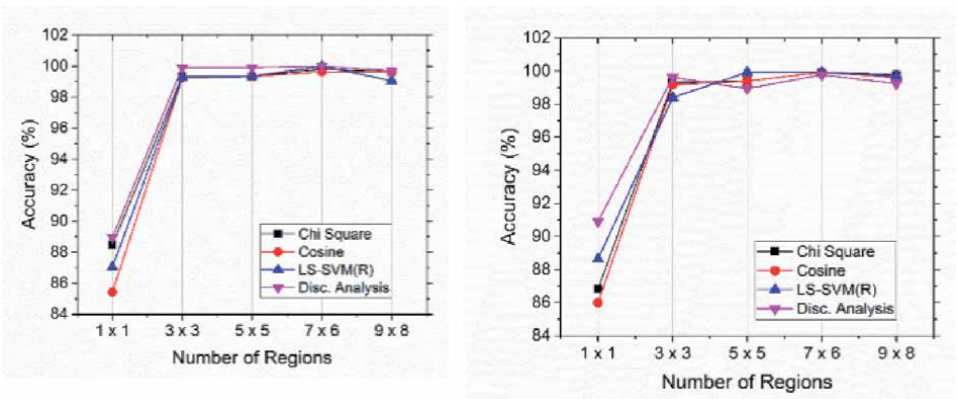


Figure 10.
Effect of a number of regions on performance on JAFFE (left) and TFEID (right) dataset. Response for 7 × 6 regions is better compared other tested regions.

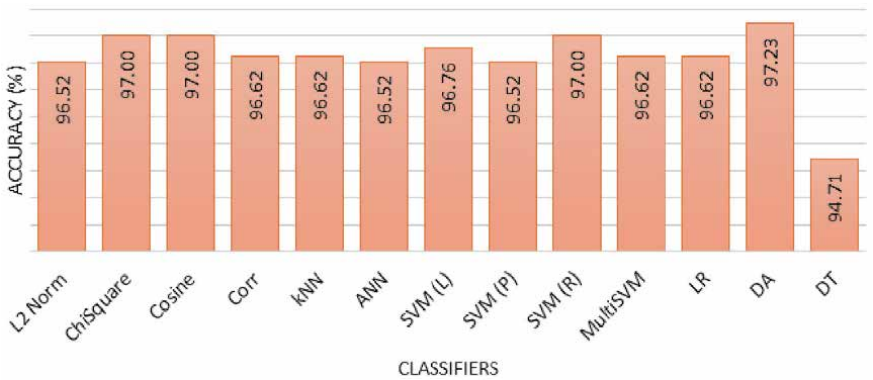


Figure 11.
Performance comparison of different classifiers, dataset: JAFFE.

Neural Network, Least Square Support Vector Machine (with linear, polynomial and RBF kernel), Multi-SVM (extension of binary SVM to multi-class SVM), Logistic Regression, Discriminant Analysis and Decision Tree. Results of all classifiers are compared in **Figure 11**.

Chi-square and Cosine measure gives the best classification results among all template matchers. Discriminant Analysis classifier achieves the highest accuracy among used machine learning classifiers for chosen parameters. A particular instance of execution is shown here; in general, the performance of LS-SVM is very close to that of Discriminant Analysis. For further analysis, we used two template matching (Chi-Square and Cosine) and two machine learning (Discriminant Analysis and LS-SVM with RBF kernel) classifiers.

D. Performance Analysis in Small Training Sample Space.

An important aspect of learning methods is that they should generalize well on unknown data. The success of any classifier depends on how quickly adapts to new and unseen patterns. K-fold cross-validation is the most commonly used validation technique. A lot of work done in the past reports the use of 10-fold validation, wherein 90% samples are used for training and the rest are used for testing. Reduction in a number of training samples has shown to negatively impact the performance. Discrimination capability of the proposed methods has been evaluated with six different cross-validation methods, varying the training samples from 90%, 80%, 70%, 50%, 30% and 10%. Even with 10% training samples, it exhibits far better accuracy compared to many state of the art methods. **Figure 12** exhibits the behavior of the system for various validation strategies. Varying number of sample

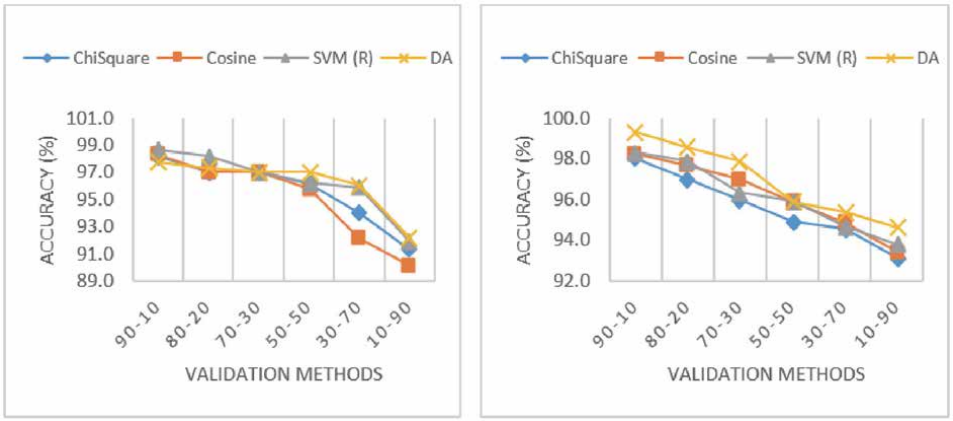


Figure 12.
Performance of various classifiers against different validation methods for JAFFE (left) and TFEID (right) dataset.

Parameter	Chosen value
Number of eigenvectors	140
Number of regions	7 × 6
Validation method	2-Fold (50% training – 50% testing)
Template matching classifier	Chi-square, cosine measure
Machine learning classifier	LS-SVM, discriminant analysis

Table 4.
Optimal parameters chosen for result analysis.

size is used to see the generalization of algorithm. If algorithm can give better results even with small number of training samples implies proposed algorithm is able to effectively capture the discriminating features of the image.

Based on the experiments, we choose parameters for the further analysis as shown in **Table 4**.

2.4.2 Expression recognition using facial components

It is observed that beauty is a factor which affects the reminiscence of the face. Faces with higher beauty factor are remembered for a long time. Similarly, certain facial regions have more influence on recognition rate. We evaluated the importance of upper and lower facial regions in expression recognition. Eye, eyebrow and forehead lines show different geometrical movement during certain expressions. The texture on facial component surface carries essential discrimination information. In anger state, eyebrows pulled down, upper and lower lids pulled up, and lips may be tightened. In the fear state, eyebrows and upper eyelids are pulled up, and mouth is stretched. During disgust state, eyebrows are pulled down; nose gets wrinkled and upper lip is pulled up. Similar changes can be observed in other expressions too. We performed expression recognition using MLH features extracted from the only eye, only mouth, eye + mouth, and face. Results are stated in **Figure 13** for JAFFE and TFEID datasets.

Results show that performance of FER system with features extracted from upper face regions is slightly better than features extracted from mouth region. However, a fusion of both the features outperforms results of individual components. Although nose remains in almost same shape and position, for few expressions like disgust and anger, its appearance changes. While the full face is used for feature extraction, these changes are also incorporated and highest recognition rate is achieved.

2.4.3 Expression recognition from noisy images

Images acquired in real-time are often noisy. A robust system should be able to handle the noise. Salt and pepper, Gaussian and speckle noise are the common noise introduced in the image. We conducted the experiment by manually adding noise in the images. Noise is added in half of the randomly selected images. The performance of the system in a noisy environment is evaluated with various noise parameters like mean and variance. The amount of various noise is controlled by the

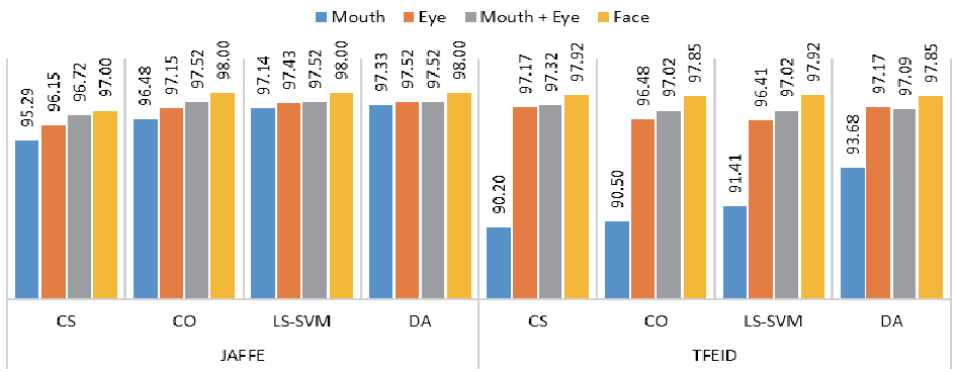


Figure 13.
Performance evaluation of various facial components on JAFFE and TFEID dataset.

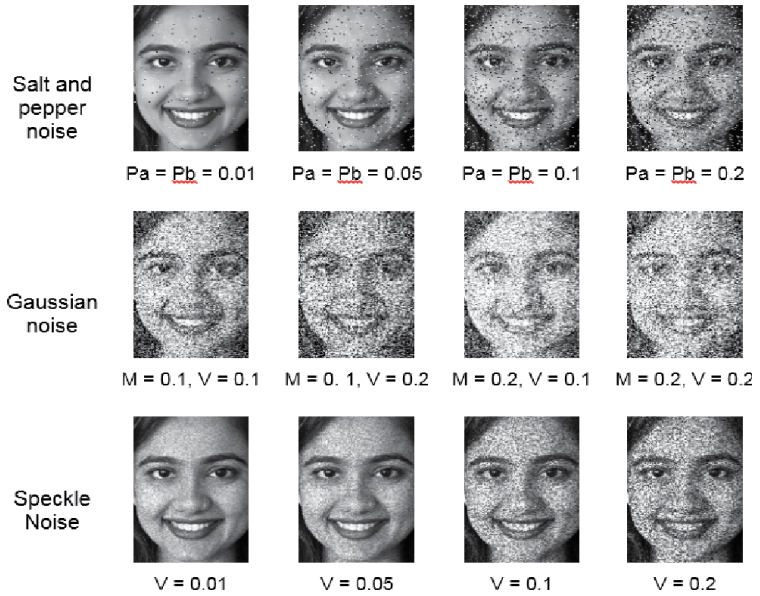


Figure 14.
Performance evaluation of various facial components on JAFFE and TFEID dataset.

	Template matching		Machine learning	
Variance	**Chi-square**	**Cosine**	**LS-SVM (RBF)**	**Discriminant analysis**
Pa = Pb = 0.01	93.43	93.43	93.52	93.81
Pa = Pb = 0.05	93.52	93.43	93.24	93.52
Pa = Pb = 0.1	94.00	94.00	93.81	94.00
Pa = Pb = 0.2	93.43	93.14	91.05	92.38

Table 5.
Results on JAFFE in noisy environment, noise: Salt & Pepper.

probability of salt (Pa), the likelihood of pepper (Pb), variance (V) and mean (m). Effect of different types of noises with varying probability is shown in **Figure 14**.

Wavelets have shown good applications to noise removal. The selection of wavelet depends on the energy conservation in approximation subband. Haar possesses the nice property of signal compaction and energy preservation and hence they can prove an ideal choice for noise reduction. Salt and pepper noise has very high impact on the illumination of affected pixels. Robustness to noise is inherent in Haar. Performance of proposed method in presence of salt and pepper noise is shown in **Table 5**.

Gaussian noise is controlled by two parameters, mean and variance. As can be seen from the **Figure 14**, Gaussian noise corrupts images visually higher than other two noises. And hence it has a more diverse effect on accuracy and performance degrades compared to the presence of salt and pepper noise. Intensity disturbance created by speckle noise is less compared to Gaussian and hence the effect on performance is also less compared to Gaussian.

Due to wrinkles and aging, skin of aged people have more texture than younger ones. Suck skin texture may introduce noise effect in feature vector due to high variability in skin texture. Perhaps, the noise reduction works better with younger people.

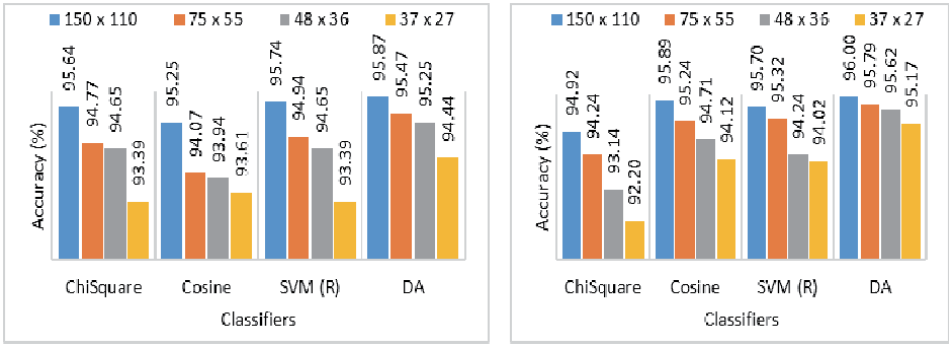


Figure 15.
Recognition rate of MLH in low resolution on JAFFE (left) and TFEID (right) datasets.

Resolution	150 × 110	75 × 55	48 × 36	37 × 27
Chi-square	95.6	94.8	94.7	93.4
Cosine	95.3	94.1	93.9	93.6
LSSVM (R)	95.7	94.9	94.7	93.4
DA	95.9	95.5	95.3	94.4
Average	95.6	94.8	94.6	93.7

Table 6.
Average recognition rate over all four classifiers.

2.4.4 Expression recognition in low resolution

Resolution can have significant effect on quality of the image. High resolution images may not be available always. Applications such as surveillance applications, home monitoring, smart meeting produces low resolution videos, which makes facial expression recognition difficult [21]. Very little work has been done on low resolution images. In our experiment, we studied the performance of MLH operator in four different resolutions: 150 × 110, 75 × 55, 48 × 36 and 37 × 27. Low-resolution images are derived by down-sampling the original images. Results on JAFFE and TFEID are portrayed in **Figure 15**.

For JAFFE dataset, the average recognition rate of all four classifiers for 150 × 110 resolution is 95.6%, which is 1.9% higher than the recognition rate in case of 37 × 27 resolution, which has an average recognition rate of 93.7%. Performance degradation with lower resolution is stated in **Table 6**. Results confirm that the performance decreases with lower resolution.

It is apparent that recognition of expression becomes difficult from low-resolution images. Even for a human it gets difficult. **Table 6** shows that the performance degradation for 75 × 55 is 0.8% but it is as high as 1.9% for 75 × 55 resolution.

3. Conclusions

This chapter presents preprocessing technique for face registration. Head pose angle is estimated and the head is rotated if needed to make it up-right frontal pose. Eyeballs are aligned in order to register the face. In this chapter, we have also

proposed Multi-Level Haar (MLH) based facial expression recognition system. The proposed method extracts level-1 and level-2 approximation coefficients of various facial components and feature vector are derived by concatenating these coefficients. Dimensions of the obtained feature vector are reduced by projecting it in PCA subspace followed by LDA subspace. Performance of the algorithm is evaluated in different scenarios like low resolution, noisy environment, low training sample space etc. Due to nice properties of Haar, and proper alignment of features due to preprocessing method, proposed method is able to achieve high recognition rate in diverse scenarios.

The work could be extended to various real world applications. For example, expression based song selection application can help users to create playlist of songs based on their mood and expressions. Face recognition based bio-metrics can be made more robust and secure by incorporating expression along with face. In class room, engagement level of the students can also be analyzed based on their facial expression, which could help teachers to understand the mood of students and change the teaching style. Facial expression based surveillance systems in shopping mall can be helpful to understand the customer's feedback from their expressions. Recommendation system based on facial expression can auto suggest the products to the customers. Facial expression plays very crucial role in nonverbal communication in society at different levels.

Author details

Mahesh Goyani
Gujarat Technological University, Gujarat, India

*Address all correspondence to: mgoyani@gmail.com

References

[1] Pantic M, Rothkrantz LJM. Automatic analysis of facial expressions-the state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000:1424-1445. DOI: 10.1109/34.895976

[2] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski: Classifying facial actions. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1999;974–989. DOI: http://dx.doi.org/10.1109/34.799905

[3] B. Fasel and J. Luettin: Automatic facial expression analysis: A survey. Pattern Recognition. 2003;259–275. DOI: http://dx.doi.org/10.1016/S0031-3203(02)00052-3

[4] Schmidt KL, Cohn JF. Dynamics of facial expression: Normative characteristics and individual differences. In: Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '01); 22–25 August 2001; Tokyo. Japan: IEEE; 2001. pp. 547-550

[5] P. Ekman and W. V Friesen: Constants across cultures in the face and emotion. Journal of Personality and Social Psychology. 1971;124–129. DOI: http://dx.doi.org/10.1037/h0030377

[6] Ekman P, Friesen WV. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press; 1978

[7] Ekman P, Friesen WV. Facial Action Coding System. A Human Face: Salt Lake City; 1977

[8] Bartlett MS, Littlewort G, Frank M, Lainscsek C, Fasel I, Movellan J. Recognizing facial expression: Machine learning and application to spontaneous behavior. In: Proceedings of IEEE computer society conference on computer vision and pattern recognition (CVPR '05); 20-25 July 2005; San Diego, CA, USA. IEEE. 2005:568-573

[9] Bartlett MS, Littlewort G, Frank M, Lainscsek C, Fasel I, Movellan J. Fully automatic facial action recognition in spontaneous behavior. In: Proceedings of IEEE 7th International Conference on Automatic Face and Gesture Recognition (FGR '06); 10–12 April 2006; Southampton. UK: IEEE; 2006. pp. 223-230

[10] Zeng Z, Pantic M, Roisman GI, Huang TS. A survey of affect recognition methods: Audio, visual. and spontaneous expressions: IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009:39-58

[11] Zhan C, Li W, Ogunbona P, Safaei F. Facial expression recognition for multiplayer online games. In: Proceedings of ACM International Conference Proceeding Series (Australasian Conference on Interactive Entertainment); 04–06 December 2006; Perth. Australia: ACM; 2006. pp. 52-58

[12] Freire D, Santana MC, Déniz-Suárez O. Smile detection using local binary pattern and support vector machine. In: Proceedings of 4th International Conference on Computer Vision Theory and Applications; 5–8 February. Vol. 2009. Portugal: Lisboa; 2009. pp. 398-401

[13] J. Whitehill, G. Littlewort, I. Fasel, M. Bartlett, and J. Movellan: Toward practical smile detection: IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009;2106–2111. DOI: http://dx.doi.org/10.1109/TPAMI.2009.42

[14] Shan C. Smile detection by boosting pixel differences. IEEE transactions on image processing. 2012:431-436

[15] Q. Ji, Z. Zhu, and P. Lan: Real-time nonintrusive monitoring and prediction of driver fatigue: IEEE Transaction on Vehicular Technology. 2004;1052–1068. DOI: http://dx.doi.org/10.1109/TVT.2004.830974

[16] B. Gholami, W. M. Haddad, and A. R. Tannenbaum: Relevance vector machine learning for neonate pain intensity assessment using digital imaging: IEEE Transactions on Biomedical Engineering. 2010;1457–1466. DOI: http://dx.doi.org/10.1109/TBME.2009.2039214

[17] Bartlett MS, Littlewort G, Fasel I, Movellan J. Real time face detection and facial expression recognition: Development and applications to human computer interaction. In: Proceedings of Computer Vision and Pattern Recognition Workshop; 16–22 June. Vol. 2003. USA: Madison, Wisconsin; 2003. pp. 53-53

[18] K. Anderson and P. W. McOwan: A real-time automated system for the recognition of human facial expressions: IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics. 2006; 96–105. DOI: http://dx.doi.org/10.1109/TSMCB.2005.854502

[19] Azure Cognitive Services [Internet]. 2016. Available from: https://azure.microsoft.com/en-in/services/cognitive-services/face/ [Accessed: 2020-06-05]

[20] M. Pantic and L. J. M. Rothkrantz: Automatic analysis of facial expressions: The state of the art: IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000;1424–1445. DOI: http://dx.doi.org/10.1109/TSMCB.2005.854502

[21] Tian YTY. Evaluation of face resolution for expression analysis. In: Proceedings of Computer Vision and Pattern Recognition Workshop; 27 June-02 July. Vol. 2004. Washington, D.C.: USA; 2004. pp. 1-7

[22] C. Shan, S. Gong, and P. W. Mcowan: Facial expression recognition based on local binary patterns - a comprehensive study: Image and Vision Computing. 2009;803–816. DOI: http://dx.doi.org/10.1016/j.imavis.2008.08.005

[23] T. Jabid, M. H. Kabir, and O. Chae: Robust facial expression recognition based on local directional pattern: ETRI Journal. 2010;784–794. DOI: http://dx.doi.org/10.4218/etrij.10.1510.0132

[24] Valstar M, Jiang B, Mehu M, Pantic M, Scherer K. The first facial expression recognition and analysis challenge. In: Proceedings of IEEE International Conference on Face & Gesture Recognition; 27 June-02 July. Vol. 2004. Washington, D.C.: USA; 2004. pp. 921-926

[25] Valstar MF, Almaev T, Girard JM, McKeown G, Mehu M, Yin L, et al. FERA 2015 - second facial expression recognition and analysis challenge. In: Proceedings of 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition; 04–08 May 2015; Ljubljana, Slovenia. 2015. pp. 1-8

[26] M. Song, D. Tao, Z. Liu, X. Li, and M. Zhou: Image ratio features for facial expression recognition application: IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics. 2010; 779–788. DOI: http://dx.doi.org/10.1109/TSMCB.2009.2029076

[27] Pantic M, Patras I. Dynamics of facial expression: Recognition of facial actions and their temporal segments from face profile image sequences. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics. 2016: 433-449

[28] H. Boughrara, M. Chtourou, C. Ben Amar, and L. Chen: Facial expression recognition based on a mlp neural network using constructive training

algorithm: Multimedia Tools and Applications. 2016;709–731. DOI: http://dx.doi.org/10.1007/s11042-014-2322-6

[29] Stankovir RS, Falkowski BJ. The Haar wavelet transform: Its status and achievements. Computers and Electrical Engineering. 2003:25-44

[30] Papageorgiou C, Poggio T. A trainable system for object detection. International Journal of Computer Vision. 2000:15-33

[31] T. Kanade, J. F. Cohn, and Y. Tian: Comprehensive database for facial expression analysis: Proceedings of 4th IEEE International Conference on Automatic Face and Gesture Recognition; 28–30 March 2000; Grenoble, France; 2000. p. 46–53. DOI: http://dx.doi.org/10.1109/AFGR.2000.840611

[32] M. J. Lyons: Automatic classification of single facial images: IEEE Transactions on Pattern Analysis and Machine Intelligence. 1999;1357–1362. DOI: http://dx.doi.org/10.1109/34.817413

[33] Chen L-F, Yen Y-S. Taiwanese Facial Expression Image Database: Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan. Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University. Taiwan: Taipei; 2007

Chapter 3

How to Use the Advantages of AR and VR Technique to Integrate Special Visual Training Strategies in Non-Verbal Communication Skills Training for Children with Autism

I-Jui Lee

Abstract

Social communication is the basis of human interaction, as people transmit messages, their emotions, and some symbolic social content through oral or non-verbal communication. However, it is difficult for most children with autism to master, or understand, these non-verbal social communication skills. In recent years, Augmented Reality (AR) and Virtual Reality (VR) technology (A/VR technology) have been widely used in the social training of children with autism. This technology creates many learning states that are difficult to achieve in the traditional social training of autism, and they are mainly achieved through the sensory advantages of A/VR technology in the visual presentation and simulation of surreal state situations. This technology can provide more attractive interactive games and play strategies to help autistic children learn social reciprocity skills. Therefore, this paper discusses how A/VR technology can be used to construct a surreal visual training strategy in the non-verbal social training of autistic children and, by using different teaching strategies and visual training applications, to help them to improve their mastery of non-verbal social communication. Based on the findings, future researchers are offered some suggestions and references for the application of A/VR technology in the social training of autism.

Keywords: children with autism, augmented reality and virtual reality technology, non-verbal social communication, pretend-play and symbolic play, self-reference and imitative learning, role perspective exchange of the first person and the third person, social situation simulation, social reciprocity skills

1. Introduction

1.1 Congenital social communication obstacles in children with autism

The common obstacles for the social skills of children with autism include the following [1]: they have difficulty in their social reciprocity with others, they have

difficulty in maintaining mutual social relations, they are unable to share their feelings with others, they are unable to think like, and understand the emotions, of others, from their perspective [2, 3], and they have a lack of empathy [4, 5]. Children with autism find it difficult to establish effective social relations due to their social skill barriers [6–9], which lead to them to feel uneasy and to retreat when interacting with a crowd [7], and indirectly lead to social isolation [5].
A person's social skills refers to his/her ability to successfully and independently participate in social interaction, to establish and maintain social relations with others [10, 11], and to meet the social reciprocity conditions and communication skills required in different situations [12]. Having social skills is one of the important abilities that people need, in order to get along with others in society. It includes the ability to observe and listen to others, to make basic eye contact with people [13], to give social body feedback [14], to have complex social interactions and to manage their emotions [15]. Social skills are often constructed on the basis of some non-verbal social communication behaviors [16], such as facial expressions, head or hand movements [2], posture and eye expressions, etc. However, children with autism generally lack these important social communication skills [17, 18].

1.2 Social communication skills for children with autism

The social communication skills that autistic children need to master are mainly divided into various levels: (1) social initiation, (2) social reciprocity, and (3) non-verbal social communication [19, 20]. Social initiation refers to the ability to actively interact with others and to initiate social content. Social reciprocity refers to the ability to have good social communication and interaction with others, to be able to have appropriate behavioral responses, and to engage in a dialog. A person's non-verbal social communication ability refers to his/her ability to understand the deeper meaning, or symbolic metaphors, of other people's body language and expressions in non-verbal communication [21]. Among them, the ability to reciprocate socially summarizes most of a person's comprehensive ability and behavior performance in social skills, and it is also considered as the most important comprehensive ability in social contact [22, 23]. For example, people can judge the environmental background and social situation of the other party, while understanding the other party's emotion and discourse and making an appropriate social response [24]; it includes the ability to understand other people's emotions [25], their deep feelings, and their implications [26]. Therefore, social reciprocity is often regarded as the focus of learning social skills for autistic children.

1.3 Non-verbal social orientation covered by social reciprocity ability

The DSM-V [27] of the fifth edition of the diagnostic manual of the American Psychiatric Association (APA) points out that one of the most important symptoms of autistic patients is the behavioral disorder of social reciprocity. The key to this behavior is whether they can master some more subtle non-verbal social communication clues, including facial expressions, eye contact, social body movements, social ceremony actions, or emotional gestures and eye expressions [17, 28].
These social features cover a large number of slight visual characterizations and non-verbal communication cues, which provide the basic elements of strong social

reciprocity for ordinary people to interact with others, to share their feelings, to think and understand their behavior, from their perspective, as well as their intentions, emotional feelings and empathy [3]. Moreover, this ability requires considerable visual information for grasping and making social judgments [17, 18]. However, this ability is very difficult for autistic children, as they often have difficulty in handling non-verbal social communication proficiently, and they cannot carry out effective social interaction and communication with others. Similarly, it is a challenge for them to actively make eye contact with others and to make appropriate social responses [29]. Such symptoms will cause autism patients to become estranged from people and will make it difficult for them to effectively integrate into society [30].

1.4 Social training methods and teaching strategies are widely used in autism nowadays

There are many social training methods [31] and teaching strategies to help autistic children learn to judge more "social situations", to understand different "non-verbal social cues", and to make an appropriate social reciprocal performance; these include behavioral interventions [32], pivotal response training (PRT) [33], scripting [34], story-based interventions, etc. [35].

Among them, the most common training strategy is the use of Social Stories™ [36], or the social intervention training method, based on situational stories [35]. The Social Stories™ method can help children with autism to better understand and follow social rules and routines. In addition, it can promote a better self-awareness, help them to gain insight into the perspectives of others, and to understand how their behavior impacts others. This model appropriates social interaction by describing a situation by using the relevant social cues, it gives the perspectives of others, and it suggests an appropriate response. Generally-speaking, social stories or scripts are created to test for autism and they are used to assess the effects of multiple-baseline designs, across the different participants of Social Stories™, for intervention or training [36]. By using the social story method, combined with role play, autistic children can be guided to handle a situation and understand the relationship between social objects and other social communication states [37]. This includes how to respond with appropriate social reciprocal behavior, according to different social objects, and to grasp social skills [38]. For example, according to other people's non-verbal social cues (such as their posture and eye expressions, facial expressions, or all kinds of social body movements), they will try to figure out the social intentions of the other person, or to carry out a role exchange and a role disguise game [39], which is the same as a psychological intervention training strategy [21]. This kind of teaching method can help autistic children to master the situation, as well as the interpersonal interactions in social stories, and it guides them to observe, and familiarize themselves with, the social behavior and skills that an adult society should have [36]. Through targeted social behavior intervention training [32], the social reciprocity and judgment of autistic children in a social situation can be gradually improved, and their ability to communicate socially with others can be enhanced [17, 40].

1.5 Current limitations of social training for autism

The existing social intervention training, such as role-playing, using the social story method, or interpersonal skills on empathy, requires considerable imagination

and pretend-game skills [41] to guide autistic children to figure out and understand "social relations" and "social behavior cognition". Among them, empathy is the ability to understand and comprehend the emotions and behavior of another person. This skill, which can be trained, facilitates interpersonal relationships.

However, children with autism are reluctant to participate in the situation guidance of intervention training, when they cannot see the scene, or anything to help them understand the situation or game [42]. For example, starting from general social training, game intervention activities, such as role-playing or perspective-taking, are training methods that are used for autistic children to perceive the feelings of other people and to generate empathy. However, due to their inherent social defects and weak imagination, they cannot be effectively implemented, which leads to the exclusion of autistic children from teaching and training, making it even more difficult to attain effective training results [43].

In addition, the ability of role play or transposing their thinking is a very difficult skill for autistic children, and they cannot understand their social relationship with others, from the perspective of the others, by using empathy or their imagination [43]. Moreover, they are often unable to identify the non-verbal social cues in different complex social situations. It is difficult to effectively teach autistic children to understand special and complex social situations, or to develop further empathy. It is a complicated and abstract social structure for them, which makes it challenging for them to acquire such concepts and social skills [17, 18].

1.6 Key factors of good social intervention training strategy

For children with autism, the key point of a good social intervention strategy is to provide powerful visual media in this training game, to attract their attention and maintain their interest [17]. Compared with oral expression, or the written communication performance, image information is one of the best ways of learning. Temple Grandin, who is one of the most well-known and accomplished adults with autism [44], once described in his work that all the words and dialogs presented before the eyes of autistic patients are like pictures, which represents the unique visual learning approach and learning ability of autistic patients.

However, past research has pointed out that the traditional way of matching a picture with a text is not attractive to autistic children, and the effect is not significant [21]. The reason for this phenomenon is that traditional teaching strategies lack interaction and a clear teaching framework, and it is difficult for them to present or demonstrate an appropriate interactive situation. For example, traditional storybook combines social storytelling methods for teaching; however, storybooks can only provide fragments of situational pictures. (Situational pictures, like the normal photographic capture of a real scene, which we call a "montage" or a "stop motion video", include some people's interactions and social reciprocity behavior in a specific place. This material is used to help the therapist to teach the children and to indicate their roles and their interaction with it). However, it is difficult to show continuous details from situational pictures, and traditional storybooks lack an interactive mechanism, and many details can only be imagined, which causes autistic children to get bored or to lose interest.

2. A/VR technology's great breakthrough in a visual sensory experience

With the intervention of Augmented Reality (AR) and Virtual Reality (VR) technology, these training problems have been improved, to varying degrees. Due to an autistic child's innate visual and learning advantages, A/VR technology

provides users with a great breakthrough by providing a visual sensory experience [45]. This includes the most basic interactive game operation for complex multi-person perspective exchanges, self-reference imitations, and a variety of situation simulations [42, 43, 46].

2.1 Overview of AR and VR technology and their benefits and differences

AR technology is an interactive display platform that superimposes virtual objects in a real environment. Using the computer, it generates 3D image display, sound, text and animation effects to enhance the user's visual sensory experience. Therefore, it is good for pretend-play and symbolic play games to use AR technology, because they can augment the virtual 3D materials by overlaying them on the real objects to make it look like another thing. For example, by using AR technology, we can add the 3D model or animation overlay to real matchboxes to symbolize cars (**Figure 1**). By using the AR app, the matchboxes will look like the shape of a car on the screen, which can help autistic children to pretend and imagine that the matchboxes are cars. This concept can be easily understood from Bai's (2014) research article [47].

Another VR technology provides a more immersive and realistic environment in which people can experience the environment. VR provides visual images on the media to help children with autism to have an in-depth experience, to master their imagination, and to experience visual sensory stimulation. Thus, VR technology has another benefit in that it lets the children with autism become immersed in another world and to empathize with another person's perspectives, feelings and thoughts. Therefore, the goals of AR and VR technology can be very different in the way that it presents the virtual environment, by either enhancing or changing the perceptual world.

The biggest difference between AR and VR is that AR integrates with real objects (**Table 1**), while VR is completely immersed in a virtual environment. AR presents an interaction with the objects in the actual field, such as the social skills training

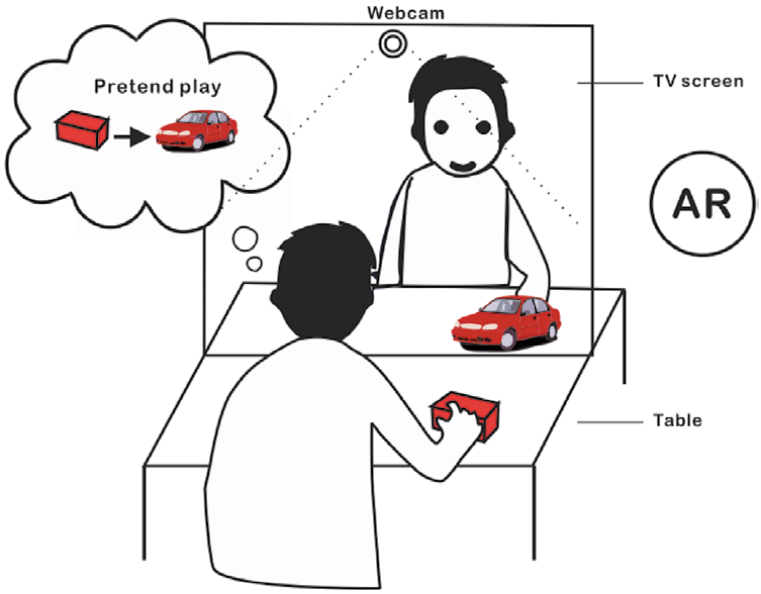


Figure 1.
AR technology is used to help children with autism to elicit pretend-play and imagination skills, such as imagining a square matchbox as a car. The original picture was produced and re-drawn from Bai [47].

		AR technology	VR technology (including VR CAVE)
1	Technical Description	AR technology can add additional visual information on the surface of specific objects, space environments, or image information, and its application in teaching has been effective	Enhancing the perceptual state of an immersive perception environment.
2	Technology core	AR technology can help autistic children with their expressions through self-reference and watch their social status with others through pretend and symbolic games, or from the perspectives of different persons	VR can provide different situations for autistic children to speculate about their current immersion feelings and to master environmental information through the immersion environment.
3	Execution strategy	Additional virtual information is superimposed on physical space objects, e.g. the 3D virtual car model is overlaid on the box, making it look like a real car. It is usually combined with board games or pretends to operate.	A head-mounted display is usually used to exchange perspectives in immersive situations, to strengthen and experience empathy
4	Learning mechanics	Pretend-play and symbolic play (imagination training), self-reference and imitative learning	The role perspective exchange of the first person and the third person (taking a perspective and empathy construction), and social situation simulation
5	Social Skills Enhancement	Strengthens the mastery of (1) pretend play and symbolic play (imagination training), (2) self-reference and imitative learning.	Strengthens the behavioral cognitive training of (3) role perspective exchanges of the first person and the third person (taking a perspective and empathy construction), and (4) social situation simulations

Table 1.
The differences of AR and VR technology in social skills training for autism.

that is used in pretend games or iconic games, while VR strengthens the experience of simulation and the environment of situational feelings, which means that they have different functions and characteristics in social training for autism. Therefore, we will show some examples in the following chapter, and explain how AR and VR technology are applied in the social skills training of autism.

2.2 Advantages of A/VR technology in the social intervention training of autism

A/VR technology for the social intervention training of autistic children has developed to another different sensory level and visual cognitive stage and, from the general game interactions, different teaching interaction strategies have gradually developed. One by one, they correspond with the different social aspects of autistic patients, such as teaching them how to pretend to be others and to symbolize things through AR technology games [47]. In addition, empathy, imagination, and non-verbal social communication skills are the abilities that ordinary children should have when they are involved in social interaction with adults. These abilities will also be indirectly reflected in the performance of their social skills. Therefore, A/VR technology has considerable advantages in the social training strategies of autistic children, not only because it can reproduce and simulate different situations, social roles, and other perspectives, but also because it is advantageous for the different psychological levels of vision. For example, VR can provide different

situations for autistic children to speculate about their current immersion feelings [43, 48], and the 3D modeling technique in AR technology also allows them to practice social reciprocity with specific characters through pretend-playing and role-playing [43]. In addition, AR technology can help autistic children to express themselves and watch their social status with others by playing pretend and symbolic games and seeing things from the perspective of different people [49]. A/VR technology has greatly broken through the traditional training framework, but at the same time, it has extended the training on the basis of the existing teaching strategies and theories, giving it a certain theoretical basis for development, as well as a learning framework [50].

2.3 AR technology solves the problem of a weak imagination in children with autism

Because AR technology can add additional visual information onto the surface of specific objects, space environments, or image information, its application in teaching has been effective [51]. For autistic children, AR technology can help to overcome their weak imaginations, as it is capable of presenting visual information beyond words or pictures [52]. Moreover, the superimposed information can make autistic children perform the operation and teaching application of disguise and symbolic games. Below are some special teaching and training cases to illustrate and analyze the follow-up research.

2.4 AR technology can provide autistic children with a training situation by using the content of social games in disguise and symbolic play

In the process of learning how to interact and play with others, impersonation and symbolic games play an important role for normal children [53]. The children can try to figure out different situations and things and even have a tacit understanding between each other, by pretending to be in a game [54, 55]. For example, in the game of pretending to have a specific role or occupation, such as a doctor, nurse, or salesman, normal children can use the available items at hand as accessories, in the process of their diagnosis and treatment, or they can use the items on display in stores, such as using bananas as telephones [56], using matchboxes to symbolize cars (**Figure 1**), using sticks to symbolize a king's scepter, or using a bottle as a microphone. In general, normal children can achieve the operation of each other's games by defining different objects, formulating game rules and interactive methods, and even developing them into more complex game mechanisms. For example, children are often "playing house "or "playing grown-up" games (playing house, is a traditional children's game. It is a form of make-believe, where players take on the roles of a nuclear family). They are game forms that help children to understand each other's gestures and social interactions and to grasp metaphorical symbols. In the process of playing these games, they need to use a lot of imagination, to follow the rules of the game, and to have the ability to understand the concept of the symbolic form and the ability to guess the intention [37, 52].

These games not only retain specific symbols and rules, but they also integrate many different ways of interaction. In addition to the game, some more subtle interactive information helps children to learn and familiarize themselves with more complex social skills in the process of the game. These techniques can convey different information through oral or non-verbal communication and they can also involve social and emotional connections. People will use some conventional objects or gestures to replace the content of complex dialogs. For example, certain body movements, including nodding, shaking hands, hugging, high-fives, and

shoulder-clapping, are used to express the participants' concern and interaction. Moreover, different body movements will naturally convey the distance and intimacy of their social relations. However, there is a certain degree of complexity that is required to implement such a social approach with autistic children, by using traditional teaching strategies. Because autistic children find it difficult to understand complex and abstract social communication codes, the symbolic concepts and social cognition content need to be translated [57, 58].

Therefore, this is an opportunity to use AR technology to help autistic children to understand and participate in pretend-games and try to use their imagine to decode the social signals, because pretend games usually require an imagination to figure out some state that does not exist on actual objects, for example, imagining a square matchbox as a car (**Figure 1**), or imagining some actions and processes in the game, such as using a toy car on the table to imagine a real car running on the road. In this way, AR technology can today superimpose the 3D model and animation of the car and overlap it on the square matchboxes, to let the matchboxes look like real cars. In this way, it can help autistic children to master the skills of pretend-play, by using visual methods.

3. How to enhance social interaction skills in children with autism by visualization

3.1 How to help children with autism translate complex and abstract non-verbal social communication codes

Non-verbal social communication ability covers a wide range of social cognitive levels and sensory ability integration, including judging the other party's emotional expressions, their body gestures, empathy, and the ability to pretend to play games and imagine things [59]. These non-verbal social cues convey the way people communicate and their ability to convey emotions [60]. For autistic children, such content needs to be visualized, or even structured, to achieve a better performance of social details and social connections. How to translate abstract and complex non-verbal social cues is a research topic that makes researchers curious, because non-verbal social communication often represents some symbolic social communication symbols and behavior performance, which are conveyed to others through body movements or facial illustrations that involve the relationship between physical and social communication symbols [21]. For example, when a man hands a red rose to a young woman, the series of hand and body movements and their facial expressions convey each other's behavior, intentions, social relationship, and even their emotional feelings [28] (**Figure 2**). The use of situational pictures can facilitate and develop a child's ability to adapt to social relationships and to understand the overall interactive situation described in this series of story plots, by repeated learning and guessing. However, these abstract non-verbal communication expressions appear to be difficult for autistic children to understand [61], as they are unable to deconstruct the information content and social communication codes conveyed by such non-verbal communication. They need to make an extra effort and give extra attention, in order to detect these different social context fragments and non-verbal social cues.

3.2 How to help autistic children deconstruct visual media and focus on specific non-verbal social communication codes

As described in a previous study, attention is one of the key factors in social cognition and in the deconstruction of social communication codes [62]. The study

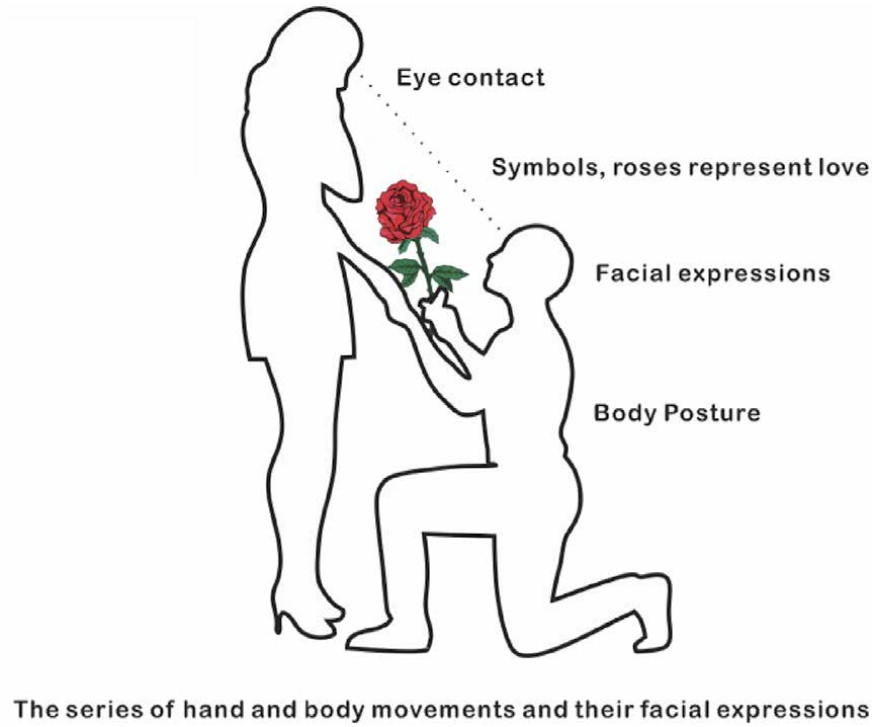


Figure 2.
When a man hands a red rose to a young woman, the series of hand and body movements and their facial expressions convey each other's behavior, intentions, social relationship, and even their emotional feelings [28].

points out those autistic children often find it difficult to detect and perceive these key non-verbal social cues. However, traditional social training, whether by using video modeling [4, 53] or specific social story methods [36], did not have effective outcomes or help autistic children to deconstruct and understand the social cues in situational pictures. In such training, therapists or special education teachers guide autistic children to observe specific situational pictures or play the videos repeatedly, through which they gradually become familiar with the social details in the story. After multiple practices, autistic children can gradually grasp such social information and apply the information in their daily lives. This kind of teaching strategy is a very common approach in social training; however, it is considered to be a rather passive method and does not attract the interest of an autistic child, even with films or picture storybooks.

3.3 How to help children with autism to master the key to visual guidance and the contextual information behind it

On the basis of our previous study [21] extracted non-verbal social cues from a series of dynamic life films and compiled them into a social storybook. Through the fixed visual structure of the storybook and the sequential page-by-page learning framework, autistic children can clearly grasp the visual pictures and clues of the occurrence of the social situations. A fixed and structured visual information deconstruction method, which researchers call Stop Motion Video (SMV), can guarantee a simultaneous and stable visual interpretation, just like traditional social storytelling [18].

Although the key frames are structural, they are quite fragmentary [28]. Such social storybooks lose most of the social context clues and story content. At the same time, static images do not easily attract the attention of autistic children and

encourage them to read. Therefore, by combining them with AR technology, researchers can use the key social context clues in social storybooks as visual primers, and they can also use the explicit learning framework and visual structure of social storybooks, to help autistic children deconstruct and master the specific non-verbal language in a series of complex social films. For example, the social interaction state, a handshake, a hug, or special eye contact, and the expressions of male and female protagonists in the film, are employed to guide autistic children to master the key events of the complete story by using a visual primer that is constructed on the fixed screen. The situational information that is guided by the back of the protagonist can then be played in a dynamic video and superimposed on the social storybook, for autistic children to refer to and read repeatedly, thus making up for the lack of information in social storybooks. At the same time, focusing on non-verbal social cues by using AR technology, successfully increases the attention and motivation of autistic children. The film that is triggered by AR technology promotes the opportunity and vision for autistic children to understand the more complex stimulation of social cues [21].

3.4 How to help autistic children master the relationship between non-verbal social communication messages and symbols

Non-verbal social communication covers a wide range of levels, including the structural relationship between people and symbols. Different body movements that correspond to the extraction of different symbols, also endow the situational story with a specific social narrative framework and clues [63]. For example, a man with a bunch of flowers kneeling on one knee before a woman may represent a marriage proposal, while taking out a diamond ring represents a wedding ceremony. These social cues and symbols will be continuously expanded upon in the process of a child's growth, as a part of their acquired learning. Generally, children obtain more social experiences when interacting with adults, which help them to apply the non-verbal social communication information and signs, and to establish a connection between such information and signs [28].

As indicated by the playing-house game, children become familiar with the strategy of the disguise and the symbolic game, which leads them to deliberate on the situation and to feel during the process. However, these social information and symbols are full of metaphorical and social content for autistic children, who find it difficult to decipher the invisible and abstract social communication codes. Therefore, additional visual aids can help them. For example, the abstract concepts, such as honor, friendship, peer recognition, praise, or respect for others, could help them to understand some social implications and emotional components, and thus to eventually develop the ability to handle the mechanism of social interaction [37]. Generally, children can figure out these skills through symbolic and pretend-game strategies, such as giving a badge to represent honor, bowing to each other as a symbol of respect, offering a handshake as a symbol of friendship, giving high-fives for peer recognition, or giving a thumbs-up as a symbol of praise. Such content often appears in children's game stories. However, the question is whether such ways can trigger and motivate autistic children to devote more attention, and their imagination, to specific non-verbal social cues and to successfully link them to a symbolic ceremony.

3.5 What are the benefits of AR on social skills training on children with autism?

Nowadays, AR can play a very important role [37]. The relevant literature points out that it offers a significant breakthrough in the disguise and symbolic game

strategies, because the additional visual information given by AR can enhance an autistic child's attention and mastery of specific social cues [21]. At the same time, it also gives some abstract symbolic content that is different from any visual sensory interpretation, such as giving a visual perception of abstract symbolic content at the moment when the perceptual acousto-optic feedback of an event is triggered [37]; for example, when courage is acquired, AR systems provide animated responses for children with autism.

3.6 AR can help autistic children to comprehend the abstract and metaphorical social relationships

We can understand the benefits of AR technology from different perspectives. For example, in metaphorical social relationships, it is difficult to explain to autistic children about the relationship between you, me, and him. We have always called them pronouns, which are a common part of speech, but they are the concepts that make children with autism most likely to encounter confusion. Autistic children often reverse their status of use, which may be one of the reasons why it is difficult for them to change their role status or feel empathy for others.

Moreover, it is not easy to explain the relationship between social intimacy and intimacy. Usually, when one teaches ordinary children about the differences between the concept of family members and their relatives (such as uncles and aunts) and even neighbors and friends, we can easily convey a distant relationship by using an oral description or some other explanations. However, autistic children find it difficult to grasp such a concept, because of their relationship with society; so they become obstacles in the transformation of the concept of communication. An oral description lacks a clear visual framework and image structure to help them understand. Therefore, it is necessary for autistic children to deconstruct different social concepts and situations by using visualization and images. One can use a Concept Map (CM) plus AR technology to deconstruct different social concepts and situations, by linking different role objects (**Figure 3**). Autistic children can then understand such social connections through visual aids. In addition to the social content that can be taught, AR provides auxiliary content, so that the visual interface on static images can be extended and generate more visual stimuli, which can subsequently attract autistic children to invest in more attention and observation [21, 49].

Through AR technology, we can give different visual interpretations to abstract social concepts. In the past, autistic children could only understand social content through role-playing and social stories. However, with the intervention of AR technology, some social content can be presented and interpreted in different ways, and can extend the sensory level given by static images or inanimate objects, such as teddy bears, dolls, or a doll in human form (**Figure 4**). The therapist uses these entity dolls to explain the situational dialog and emotional feelings between different characters to the autistic children, by using disguise and symbolic game strategies. However, in the process of a pretend-game, the dialog content, body movements, and the palm of the eyes are included. This becomes a very difficult task for autistic children, because their imagination is inherently weak, and they do not easily associate and interact with each other. The rigid thinking mode makes it difficult for them to understand each other's communication, as well as the real intention behind the words and the social meaning that they hope to convey.

3.7 AR technology can be used to give new life to role-playing games

Through the superposition of 3D animation and situational sound, AR technology can easily solve the problem of a weak imagination in autism. A therapist can

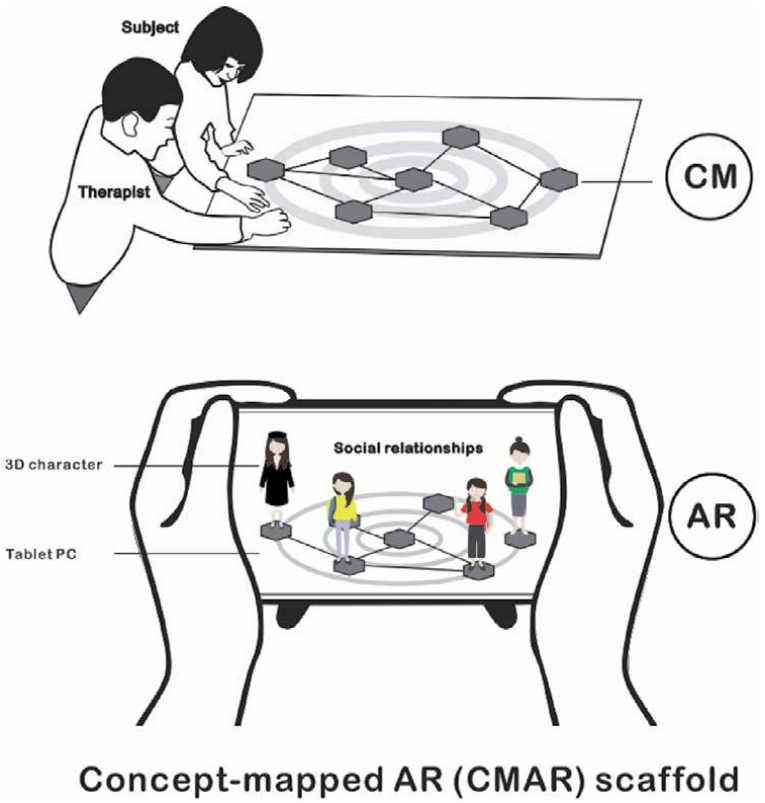


Figure 3.
A concept map plus AR technology provides additional visual aids to help autistic children to construct abstract concepts and enhance their learning motivation and attention [17].

easily make a lifeless doll appear to be alive (**Figure 4**). Through the man–machine interface and context design, the social interaction content can be successfully deduced [64]. In addition to increasing the motivation and attraction of autistic children in learning, the whole process also involves the sensory connection of entity interaction, which could help them to acquire this social concept. When autistic children play with the entity doll, they can link the social cognitive relationship between the senses and the body, which reinforces their understanding of a social relationship. In addition, AR technology gives a deeper feeling to the physical interaction and sensory manipulation required by the disguise game. From the perspective of the third person, the social interaction relationship between the disguised role and AR technology can become a more complete social training mode and make the social training of autistic children, from a visual sense, gradually push to the level of a psychological feeling. Next, the researchers will explain how to further use AR technology to transform the transposition thinking ability of autistic children and to trigger their inner ability to show empathy.

3.8 AR technology can help autistic children to transform and achieve transpositional thinking and empathy training

Empathy is seen as the ability to perceive another person's emotions, intentions, and feelings [65]. In the growth process of ordinary children, it is an innate ability and can indirectly become a skill for social interaction between people [66]. If they have this ability, people will understand another person's feelings through

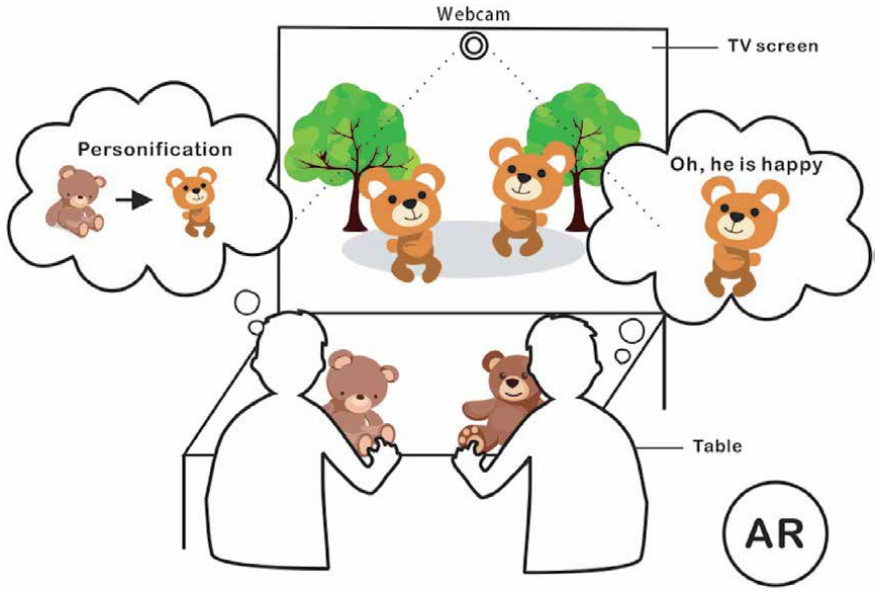


Figure 4.
Personification is when something non-human is given a human characteristic/personality. However, personification is sometimes given a broader definition. Therefore, in this section, we focus on enhancing the autistic person's empathy towards others and trying to think, feel, and understand the emotions of another person, from their perspective. The AR system will turn an inanimate object into a real creature.

transpositional thinking and then be able to predict and adopt his/her behavior representation and social communication mode [67]. However, this ability is difficult to achieve in autistic children and it is closely related to non-verbal communication cues. In general, normal adults can understand the feelings and intentions of others through their facial expressions, body movements, voice intonation, and even some of their gestures and their eyes, which are all non-verbal communication clues that could transmit abundant social information. People can use them to perceive the emotions and intentions of others and to make further judgments. Through the accumulation of experience and learning, people's judgments thus become more accurate and effective.

3.9 Shortcomings of traditional transposition thinking training

In the teaching strategy of the transposition thinking ability training for autistic children in the past, facial pictures [68], or asking autistic children to look in the mirror, was often used for repeated judgment training [69]. However, this method has not been effective, especially when autistic cases enter a real-life field, whereby such training is difficult to implement and the effect is not obvious. The researchers pointed out that it is necessary for autistic children to be proficient in this skill by using more flexible and authentic sensory stimuli. This also shows that it is difficult for autistic children to use simple images and integrate their creative imagination into a more profound on-the-spot experience, and such a perceptual environment can only be achieved in real life. However, some autistic children have a fear of the real environment, and the complexity of the information in the environment makes those with poor communication more afraid and likely to retreat, which means it cannot really achieve the purpose of training [70]. Therefore, transpositional thinking skills and empathy (also known as defects in the Theory of Mind (ToM) - the ability to interpret another person's interests, intentions, and emotions) has

always been the focus of researchers in the training of autism [71]. However, there are many difficulties in the training and implementation of a person's transpositional thinking ability. The first difficulty to be faced is how to let autistic children figure out how to take on another role? It is very difficult for them to show their abilities, because they cannot understand why therapists ask them to show the emotions and feelings of another character, or even to pretend to be another character, because they have a weak imagination and rigid logical thinking [72]. This is a rather difficult task for autistic children to understand. These communication methods are not in line with their real identity, they are abstract, and cannot be interpreted through a visual mode. For autistic children, who understand things through vision, it is only when they can truly see, or feel, the social status of different roles that they can solve the current training problems [43].

3.10 AR technology can help autistic children to realize a social interaction experience from multiple perspectives and senses

With the application of AR technology, a person's transpositional thinking ability has been greatly improved. AR can simulate the facial expression of patients through self-reference (**Figure 5**) [46], Unlike the traditional training strategy, where autistic patients are asked to look into a mirror to guess, and learn about, emotional expressions, which is especially futile for them [69], AR technology can give correct facial expressions and superimpose them on the faces of autistic children [46]. The training methods listed above (**Figure 5**), whether they are looking at their own facial expressions in a mirror or imitating their own social actions, are collectively referred to as self-referential imitation, especially the visual reference of social cues (such as facial expressions, body movements, voice intonation, etc.). Whether viewed from their own first-person perspective, or from the perspective of others (a third-person perspective), it is a very special visual sensory experience for autistic children [46].

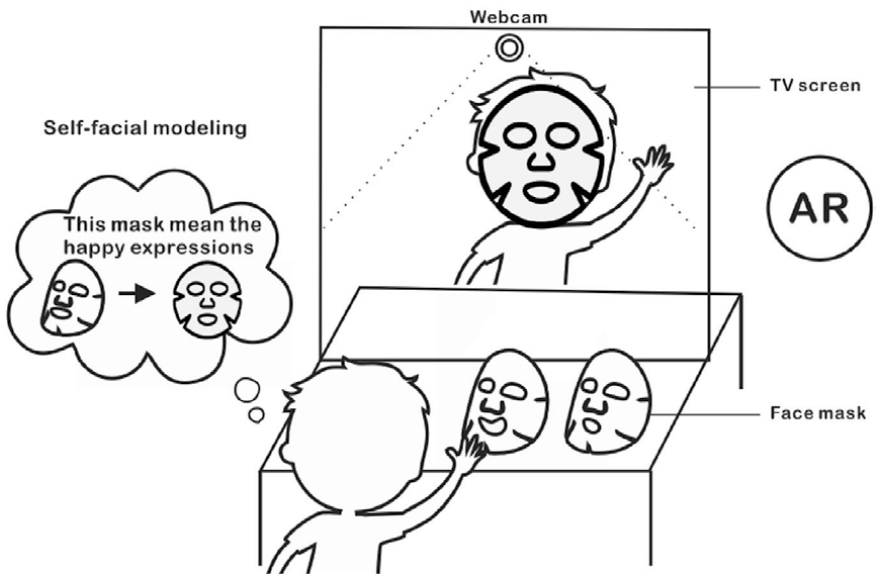


Figure 5.
AR technology can achieve a self-face reference training strategy to solve the dilemma of autism in social skills training [46].

3.11 The difference in the implementation of AR and VR for helping children with autism to empathize and improve their transpositional thinking

AR can help children with autism to implement empathy skills, by using board games. The intervention method of AR technology is usually to superimpose the content of different media on an object. For example, AR can superimpose different 3D virtual facial expressions on the autistic patient's face to help them think about the different kinds of emotions that are present in themselves. We call this as self-facial modeling (**Figure 5**), which was previously impossible to achieve, but it can be done by using AR technology. In addition, AR can provide some virtual 3D decorations for children to play with, or it can augment some virtual 3D animations on their body by self-reference, which are shown on a screen, to help a child pretend he is a king or a specific character (**Figure 6**). In this form, AR will become a tool to help autistic children to imagine and pretend, and they will gain the ability to empathize from it. This strategy lets autistic children play the game of empathy through the mechanism of pretend-play. Most of the AR methods are used to manipulate or passively perceive the comparison between real objects and the virtual content, in order to obtain the feeling of empathy.

From another perspective, VR provides a completely immersive picture experience, but it has different effects, compared to AR. The current VR technology can change a person's judgment of the senses, which allows one to fully integrate another person's perspectives, and this method can make one see the world through another person's eyes. In addition, through the exchange of different perspectives, the individual can even see the state of interaction between himself and another person, through the eyes of others, or he can experience and observe the type of social interaction from a different personal perspective. This approach greatly changes and breaks through the VR of the past, as a mechanism for on-site

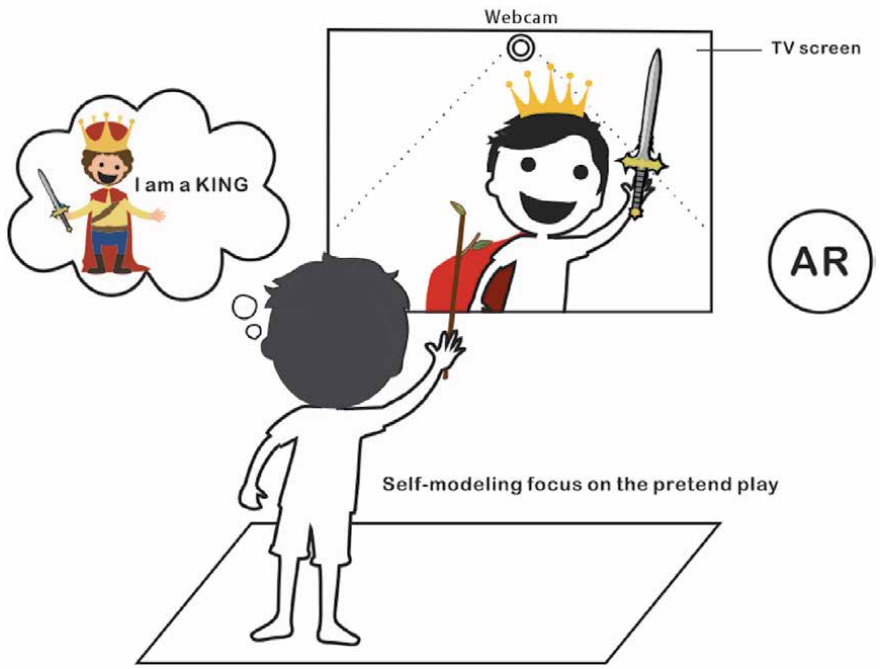


Figure 6.
Using AR technology to help children with autism to elicit pretend-play and symbolic play (a crown means a king, or courage). An autistic child can also play a self-role game; its meaning is different from a table game, because participants can join this scene face-to-face and play with others who are also self-imitating.

		AR technology	VR technology (include CAVE VR)
1	Viewing angle (perspective view)	From the third-person perspective, a person must try to figure out the feelings of the posing object and watch the superimposed animation to understand the interactive content.	The game experience method, from the first-person perspective, helps autistic children to engage directly in face-to-face social reciprocity activities with virtual characters, and to master skills.
2	Special visual situation to achieve empathy and transpositional thinking skills	Visually, it can be transformed into a state of empathy in a fake game, by comparing the relationship between the virtual and the real state. For example, through the teddy bear doll in the hand and the realistic animation of AR, the doll becomes a real bear and one must to try to figure out his feelings.(**Figure 4**.)	Visually, it can switch to the perspective of different characters, such as viewing things from the perspective of friends, or viewing yourself from the perspective of others, and feeling the feelings of others.
3	Teaching and implementing the mechanism	Self-imitation, self-modeling, pretend-play game, symbolic game operation,	A multi-person perspective exchange, direct face-to-face interaction

Table 2.
Differences in the implementation of empathetic and transpositional thinking skills between AR and VR.

experience, but with the performance of empathy, and it helps users to think from the perspective of others and to master the skills and abilities of empathy. In addition, the empathy experience provided by VR is felt from the first-person perspective and is achieved through the mechanism of exchanging perspectives (**Table 2**).

4. Some specific situations in VR presents

Similar concepts can also be presented in the creation of science and technology art. For example, the work is called the inter dis-communication machine created by the Hachiya [73]. The Dis-Communication Machine is a communication system that is aimed at transmitting and receiving sensual experiences used by two people wearing head-mounted displays. The 'machine' projects one wearer's sight and sound perception of the environment into the other one's display, thus confusing the borders between the identities of 'you' and 'me'. Through the exchange with the world, from the perspective of the different role objects, the participants can understand and feel the sensory world visually [73]. In the past, this visual sensory effect was achieved through the exchange and transmission link of entity lens images. Today, VR technology can achieve this surreal sensory experience. The case subjects can view others (first person) from their own perspective and interact with virtual characters, or view themselves from the perspective of others, or even watch their interaction with others (third person) from the perspective of another person (**Table 3**). These switching perspectives successfully achieve the visual thinking mode of transpositional thinking in the virtual world. For autistic children, the empathy training strategy can promote the physical sensory experience synchronously. Whether immersed in a specific virtual environment, or watching themselves interact with virtual characters, from God's perspective, they have achieved different visual sensory experiences, thus forming another training application of self-reference. This training framework has also been preliminarily verified by

experiments, which proves that autistic children can still achieve the common perception of psychological and physiological perceptions, with the help of technology. Autistic children can effectively perceive and express empathy, which encourages them to use the mechanism of social reciprocity, and enhances their social skills [43].

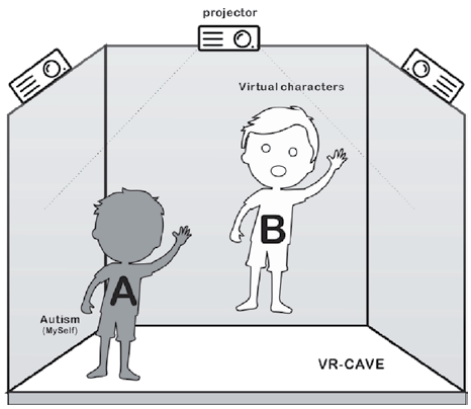
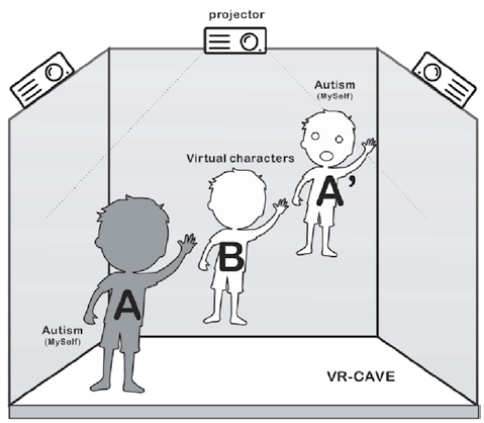
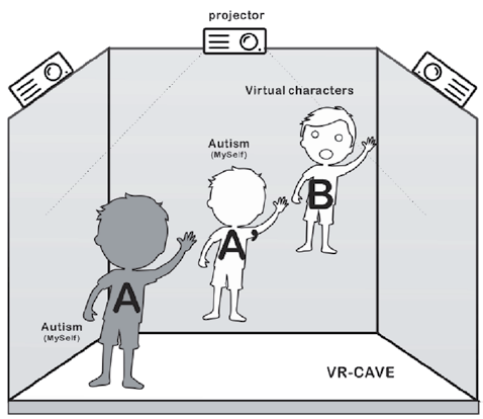
	Personal perspective	Visual picture	Experimental scenes
1	First person	Generally faces the perspective of others and interacts with virtual characters	
2	Third person (side view) (put yourself in it)	From the perspective of the third person, we can see the social interaction between ourselves and the virtual role, and we can distinguish them by looking from the back of the virtual role (see the back of the virtual character, and yourself in front). [43]	
		Or the perspective is looking at the past from behind yourself (see your background and the front of the virtual character) [43]	

Table 3.
The self-reference application of a synchronous visual sensory world achieved by CAVE-like immersive 3D technology (a-autistic participant, A"-(participant's virtual avatar captured by front or back camera), B-virtual character).

4.1 Unique visual and sensory advantages of VR technology

Researchers believe that VR technology has some unique visual advantages in the training of autistic children, which can form surreal visual and sensory benefits [43], such as a role perspective exchange of the first person and the third person (taking a perspective and empathy construction), and social situation simulation (**Table 3**). Compared to the existing teaching strategies, VR technology has considerable differences and has made major breakthroughs [45]. As a compensation mechanism of visual sense and additional training advantages given to the visual characteristics of autistic children, VR technology can encourage them to face more social objects and to observe the behavior and reactions of others. Moreover, these external stimuli can increase the possibility and opportunity of real social interaction between autistic children and their peers. Therefore, expanding these nonverbal communication training units, as well as visual assistance and additional sensory compensation, can help autistic children to reinterpret and solve their problems, even though they have a non-verbal social disorder and a weak imagination. The unique visual and sensory advantages of these VR technologies have repeatedly reminded researchers that they should make good use of VR technology, so that it can be applied and brought into play in these unique special education fields. In the next section, we will further explain and analyze the visual and sensory advantages of these VR technologies.

4.2 VR technology can help autistic children to widely connect to a real social environment and different social situation stimulations

As described above, the past training in social reciprocity behavior would make autistic children come into wide contact with real social environments and different social situations, which can promote their skills in social reciprocity with others. VR technology has sensory advantages in situation simulation and social training, especially in the early stages of its application. It is often used to solve the problems that autistic children are afraid to face in the real world, and to help them to simulate and experience different situations, such as going to school, taking busses, shopping, or to solve different levels of social fear problems [70]. However, most of these training styles are still in the stage of scenario simulation, which is also the most common and best operating mechanism of VR technology. Situation simulation solves most of the social training problems that cannot be repeated in real situations, such as helping autistic children learn how to complete a certain task under different situations (such as customers checking-out in front of a cash register, or ordering meals for customers by using an ordering machine). These training mechanisms are mostly aimed at solving the problems that autistic children need to face and overcome in real society, and to encourage them to be able to enter the workplace, in order to integrate better into the society [30].

Most researchers recognize that VR technology has several major advantages. For example, the situation simulation training developed by VR technology can be used for repeated situation training and task operations and can help the children to overcome some situations, which may be difficult to do in the real world. Another example is that it allows autistic children to experience social interactions by using different roles. However, it is undeniable that social interaction in the face of a real life environment cannot be replaced, and so VR technology is generally recognized as an early-training framework that can help autistic children to integrate into a real-life environment. When an autistic patient is unable to successfully face a real social field, he is considered to be in need of training through role-playing and placing him in social situations. He can then carry out situation experience and

social simulation training under the game mechanism of specific social situation guidance and role exchange. However, the inborn social defects of autistic people often make them resist such game mechanisms. The main reason is that they are not interested in the interactive content, or that it is difficult to make a social conjecture of game interaction. These problems are related to their innate weak imagination, their inability to think transpositionally, or their own rigid behavior problems. Therefore, it is necessary for VR technology to move from the framework of situation simulation and to gradually cross over to the psychological and sensory level, because then it will obtain better results. The following section will provide some suggestions for AR and VR, respectively.

5. Discussion

5.1 AR technology can be used to guide social visual cues and provide additional extended information

Whether in non-verbal social communication or in different visual image information, AR technology is considered to be a scanner. Through a hand-held tablet, PC or other imaging devices, any image information can be converted into a scanning icon that is suitable for AR application (**Figure 7**). Moreover, through lens sensing in front of a hand-held tablet or PC, the AR app can overlay dynamic information images (such as 3D animations, additional visual information assistance, videos and audio media, etc.) behind the static images. Researchers call this a framework approach for visual cue guidance, which means that it constructs an autistic child's sensory cognition of the visual guidance through a fixed visual framework (they are usually static images, such as social storybooks, which are commonly used in autism training). With a fixed visual cue framework and sequential story guiding the content, children with autism can quickly grasp the main structure of social stories through visual images and they can then extend to more complex and diverse social content, or context, details. This division of labor can greatly reduce the cognitive load of autistic children in their understanding of information, through the hierarchical information of AR technology, and it can also

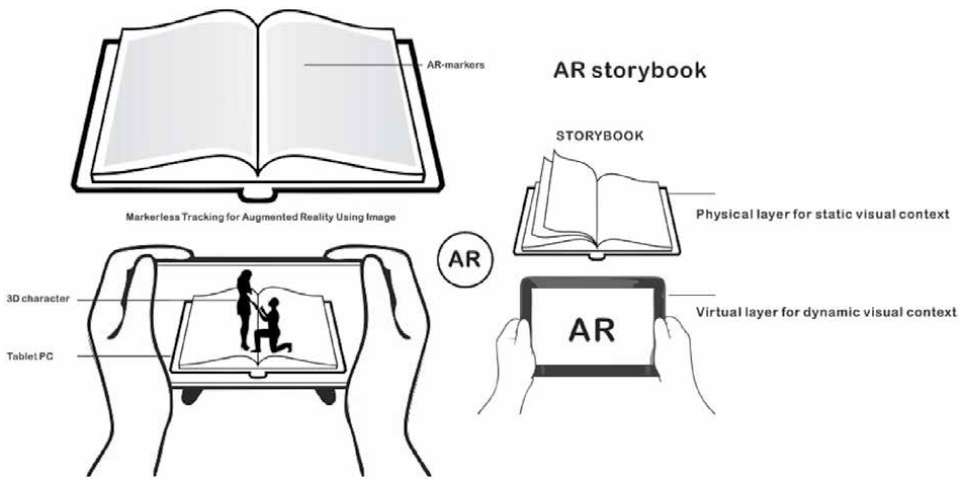


Figure 7.
AR realizes hierarchical visual indexing. Firstly, the fixed visual structure is used to help autistic children master the social context framework. Secondly, AR technology is used to link further videos and audio information [21].

attract the attention of autistic children with the help of the audio-visual effects of AR technology (**Figure 7**). The information provided by AR includes different layers of visual information, and complex social information is deconstructed by using a visual architecture. Moreover, with the support of AR technology and visual cues, the extended social information content is supported and improved [21, 28].

5.2 AR technology can help autistic children to master complex social content through a structured visual architecture and an entity interaction strategy

Visual information can be deconstructed, and then presented by linking CM (**Figure 8**) or sequence learning [38, 74]. Since people's reception of sensory stimuli depends on the interaction of their vision, hearing, and body senses, autistic children are endowed with a touchable "concept of objects", which can help them to clarify the symbolic meaning that is conveyed "behind the things". Moreover, such a form can also be used to teach autistic children to understand the state of social stories. For example, the storybook itself is a symbol that can help to construct structured concepts. By reading the situational stories in the storybook, we can help autistic children to construct the visual state of situational concepts and the sequence of events. Just like operating tools with people, we can construct a smooth operation mechanism between the mind and body, through repeated practice. Although the mechanisms of tactile, visual, and physical perceptions in sensory acceptance are different, autistic children can use their common sensory memory to

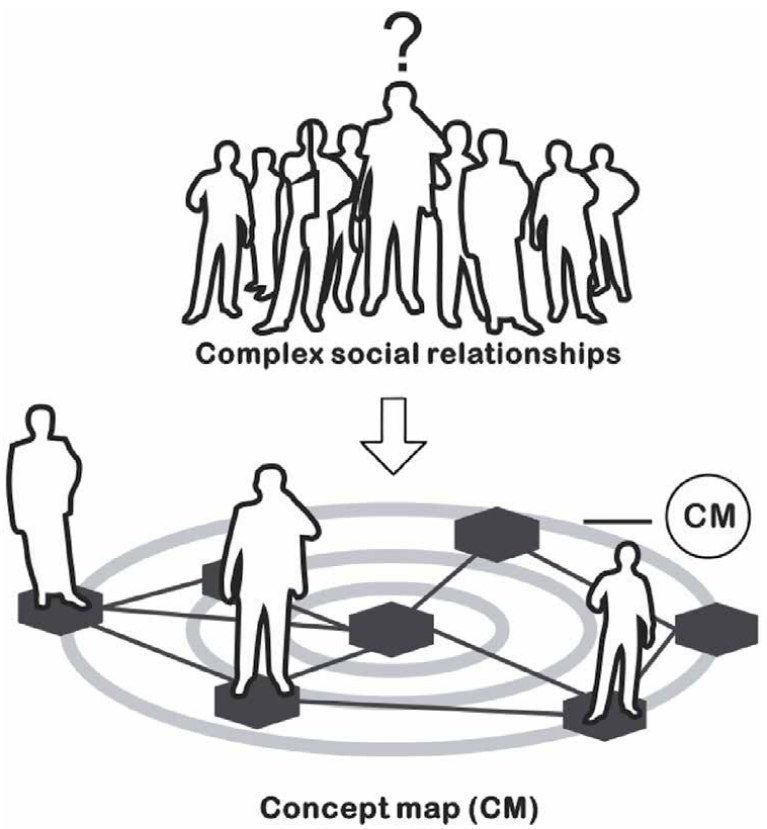


Figure 8.
Concept mapping can help autistic children to form visual and abstract social relations and help them to understand and master the content of a social situation [17].

grasp complex social content and sensory information, so that they could learn to deconstruct continuous social actions and interaction concepts.

Through the interaction strategy of a structured visual concept and entity, a visual framework, or visual primer, can be established to help autistic children understand and deal with abstract social relations, such as the distance between relatives in social relations, the social connection between different roles, or role playing and the speculation of transpositional thinking (**Figure 8**) [49]. In addition, the 3D character animations provided by AR also have the advantage of multiple viewing angles. Autistic children can watch the social relations of different character objects by different role identities and then speculate on the game. At the same time, through the game's dolls in the board games, such a learning framework can become a key training strategy for understanding social games, or disguising games, with the aid of the concept mapping and visual content that is superimposed by AR technology [50]. By combining the game operating mechanism with the board game concept, players can easily pretend to take on the role of God (or the third-person perspective). In this way, we can at the same time solve the problem of the weak imagination of autistic children. As AR animation directly makes up for the visual sensory information, which is difficult to show in the operation content, it provides effective social animation content to help autistic children understand the social status among different roles (such as hugging family, waving at friends, schoolmates clapping encouragement), supplemented by fixed visual guidance, which constitutes a more stable and dynamic social interaction structure [17].

5.3 VR technology can provide an alternative mechanism of different multiple viewpoints, an immersive virtual environment, and an immersive theater game environment

VR technology is a great breakthrough in the field of autism social training. It can change the visual training strategy under different sensory conditions and through multiple viewpoints ('multiple viewpoints' means that there is a perspective shift of being able to look at something through another person's eyes. One of the biggest influences that multiple viewpoints have is that they enable us to see how others view the world, which enables one to experience another person's sense of sight face-to-face; of course, we just let the autistic person wear the head mount display, in order to display the other person's viewpoint via a camera), and an immersive theater game environment. Moreover, it contributes to the scale, from being a table game size to a sensory experience with the whole immersion environment, and even to the development of a semi-immersive Mixed Reality (MR) training mechanism. Such content has a different training purpose and different game strategies on the basis of training. Through visual experience and fixed visual guidance, researchers can enlarge the scale of board games to the framework of immersive theater games. Using the same visual concept strategy, we can develop a social interaction mode, with the first-person perspective and God's third-person perspective, and then, through to the immersive and semi-immersive interaction experience, we can construct the participants' on-the-spot sensory experience of the social interaction objects. These visual senses can be used to guide and induce empathy through the visual sensory mechanism provided by VR technology, and they are applied in the following situations, as described previously: (1) first-person and third-person role perspective exchange (empathy and empathy construction), and (2) social situation simulation. The operation of this mechanism requires more situational guidance and sensory conditions, combined with different game and entity interaction strategies, to achieve the training effect, and such a training form needs to be designed and interpreted through a situational script. Through such a

mechanism, autistic children can gradually acquire social skills and non-verbal communication and it will equip them to live their own lives.

6. Research limitations

In real daily life, children with autism have to face many obstacles, with regard to social reciprocity, most of which are reflected in the mastery of some non-verbal social behavior. These social actions and reciprocity behaviors are not impossible to achieve in real life, but because of the innate social barriers of autism, they cannot master social skills in a more natural or appropriate way. In such a state, it is difficult to use traditional methods, such as VM or storybooks, to train them. However, AR or VR, as a visual method and teaching strategy, helps them to "see" the situation, to feel, and to have a status in these states. It even helps children with autism to understand the feelings, the states, or social reciprocity of others by multi-person perspective exchange, self-reference imitation, and a variety of situation simulations. These social actions and reciprocal behaviors usually have to be felt and communicated through empathy and inner mechanisms, and are not tangible skills (because the skill of empathy is not like riding a bicycle, as long as you can master it with proficiency, it needs the ability to function effectively in a dynamic environment in which multiple, and substantively different, perspectives must be maintained). Because empathy is an inner state and a social cognition mental skill, A/VR technology provides an attempt to help autistic children to master this ability, but not all feelings can be expressed. Children with autism still need to rely on repeated training, or to learn this skill with the growth of life experience, so this article is only a stepping-stone to help researchers to think of other ways and other methods of thinking about social training issues, to get rid of traditional memorization method, and to become a more flexible training strategy (**Table 4**).

	Difficult to achieve non-verbal social skills training in the real world	**Main technical assistance**	**Non-verbal social skills to be strengthened (training purpose)**	**Teaching strategy description**
1	Self-facial expression reference and limitations	AR	Master emotions and facial expressions by observing the face and facial expressions, such as head movements (refer to the **Figure 5**)	It is very important for children with autism to observe the correct expression state that they should make through AR technology, and, especially when they cannot master the correct emoji, AR can solve this problem
2	Give life to items in the pretend-game	AR	Master the state of social and physical interactions through pretend-play games. (non-verbal social communication skills, such as eye contact, social body movements, social ritual behaviors) (refer to the **Figure 4**)	In the real world, some inanimate objects cannot truly interact, but with the assistance of AR technology or VR technology, these inanimate objects can become real characters for interaction (for example, an inanimate teddy bear becomes an alive role)
3	See the relationship between each other through	AR	Improve the grasp of social relationships with the aid of visual information (refer to the **Figure 3**)	There are many intangible social relationships and statuses in the real world. In the AR interface, the social distance between each other can be

	Difficult to achieve non-verbal social skills training in the real world	Main technical assistance	Non-verbal social skills to be strengthened (training purpose)	Teaching strategy description
	virtual information			presented visually. For example, through a combination of the AR and CM strategies, the relationship between each other can be represented by images.
4	Observe the detailed state of different body movements from all angles	AR/VR	Use visual information aids to enhance social interaction and behavior control (facial expressions, social actions) (refer to the **Figure 7**)	In the real world, there are many details relating to body language. Only in the virtual environment can you repeatedly figure out and observe from different perspectives.
5	Watching and practicing social body movements repeatedly	AR/VR	Use visual resources to enhance the mastery of social reciprocal behaviors (social actions, including nodding, shaking hands, hugging, high fives, and shoulder clapping)	It is impossible to practice social action repeatedly in the real world, but the virtual mechanism can.
6	Socialize as someone else	VR	Grasp the state of empathy through different personal perspectives (refer to the **Table 2**)	In the real world, it is difficult to see things as another person, such as thinking about things in different roles, and autism can only be taught through visual images (thinking from the perspective of others). VR technology can achieve this function.
7	Instantly see your social interaction status with others	VR	Grasp the state of empathy through different personal perspectives (refer to the **Table 2**)	In the real world, only through video recordings and video referencing can you see the state of interaction between yourself and others at the same time, but it can be achieved in real time under the framework of VR.

Table 4.
Non-verbal social skills training that is difficult to perform in the real world.

7. Conclusion

The situational script design and social story context are often two of the most critical and characteristic training materials in the training process. The situational script involves many important training items, including the setting of social reciprocity behavior, the social context content, non-verbal communication body action, the conversation content, the symbolic metaphor story, the role interaction mechanism, and so on. Through the interpretation and visual deduction of A/VR technology, we can achieve the training framework on a visual level and then promote the transformation of social behavior in learning. This depends on a different scenario script design and social story writing. In the execution, subjects are often obtained from the participants' daily life and applied to script writing. Therefore, A/VR technology has become an interactive platform and game carrier for social training and it also gives these social materials and scripts another visual interpretation. Through the interposition of A/VR technology, autistic children can pay attention to the information details of different non-verbal social cues and

transform their communication feelings in the visual images, so that the visual information can be deconstructed and presented sensitively. This kind of training strategy breaks away from the traditional story-book and pure film training mode and expands the understanding level of sense and perception. It has been released in the training of autistic children and has become another new training framework.

The visual training strategy of A/VR technology has gained its initial achievement in the non-verbal social training of children with autism, and the application of this technology has solved many limitations in traditional teaching. However, in the application and practice of this new technology, researchers must adopt a more rigorous experimental design and take a cautious attitude. Especially for autistic children with mental development defects, the use of this technology must consider more clinical and experimental evidence and it needs to be evaluated by more therapists and physicians. Therefore, we look forward to the breakthrough and development of A/VR technology in the field of special education. At the same time, we are also careful to apply the various visual effects of A/VR technology on social interaction stimulation (e.g., body language, social situations, social reciprocity, and relevant non-verbal communication in social interactions). We hope that it will be used better and applied in the future and that its advantages in working with these special children with autism will be displayed.

Acknowledgements

This work is supported by the Ministry of Science and Technology of Taiwan (107-2218-E-027 -013 -MY2/109-2221-E-027-069 -MY2). The authors thank the referees very much for their valuable comments and suggestions on this paper.

Author details

I-Jui Lee
Ergonomic Design and Technology Lab, Department of Industrial Design, National Taipei University of Technology, Taipei, Taiwan

*Address all correspondence to: ericlee@ntut.edu.tw

References

[1] Fodstad JC, Matson JL, Hess J, Neal D. Social and communication behaviours in infants and toddlers with autism and pervasive developmental disorder-not otherwise specified. Developmental Neurorehabilitation. 2009;***12***(3):152-157

[2] Bölte S, Hallmayer, J. (Eds.). *Autism Spectrum Conditions: FAQs on Autism, Asperger Syndrome, and Atypical Autism Answered by International Experts*. Hogrefe Publishing; 2011

[3] Baron-Cohen S, Tager-Flusberg H, Lombardo, M. (Eds.). *Understanding Other Minds: Perspectives from Developmental Social Neuroscience*. Oxford university press; 2013

[4] Bellini S, Akullian J. A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. Exceptional Children. 2007; ***73***(3):264-287

[5] Hillier A, Fish T, Cloppert P, Beversdorf DQ. Outcomes of a social and vocational skills support group for adolescents and young adults on the autism spectrum. Focus on Autism and Other Developmental Disabilities. 2007; ***22***(2):107-115

[6] Bauminger N. The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: Intervention outcomes. Journal of Autism and Developmental Disorders. 2002;***32***(4): 283-298

[7] Bellini S. Social skill deficits and anxiety in high-functioning adolescents with autism spectrum disorders. Focus on autism and other developmental disabilities. 2004;***19***(2):78-86

[8] Church C, Alisanski S, Amanullah S. The social, behavioral, and academic experiences of children with Asperger syndrome. Focus on autism and other developmental disabilities. 2000;***15***(1): 12-20

[9] Whitehouse AJ, Durkin K, Jaquet E, Ziatas K. Friendship, loneliness and depression in adolescents with Asperger's syndrome. Journal of Adolescence. 2009;***32***(2):309-322

[10] McFall RM. A review and reformulation of the concept of social skills. Behavioral assessment. 1982

[11] Segrin C. Social skills deficits associated with depression. Clinical Psychology Review. 2000;***20***(3): 379-403

[12] Ahlers KP, Gabrielsen TP, Lewis D, Brady AM, Litchford A. Supporting individuals with autism spectrum disorder in understanding and coping with complex social emotional issues. School Psychology International. 2017; ***38***(6):586-607

[13] Senju A, Johnson MH. Atypical eye contact in autism: Models, mechanisms and development. Neuroscience & Biobehavioral Reviews. 2009;***33***(8): 1204-1214

[14] Centelles L, Assaiante C, Etchegoyhen K, Bouvard M, Schmitz C. From action to interaction: Exploring the contribution of body motion cues to social understanding in typical development and in autism spectrum disorders. Journal of Autism and Developmental Disorders. 2013;***43***(5):1140-1150

[15] Giovagnoli G, Postorino V, Fatta LM, Sanges V, De Peppo L, Vassena L, et al. Behavioral and emotional profile and parental stress in preschool children with autism spectrum disorder. Research in Developmental Disabilities. 2015;***45***: 411-421

[16] Jarrold C. A review of research into pretend play in autism. Autism. 2003; **7**(4):379-390

[17] Lee IJ, Chen CH, Wang CP, Chung CH. Augmented reality plus concept map technique to teach children with ASD to use social cues when meeting and greeting. The Asia-Pacific Education Researcher. 2018;***27***(3): 227-243

[18] Lee IJ. Augmented reality coloring book: An interactive strategy for teaching children with autism to focus on specific nonverbal social cues to promote their social skills. Interaction Studies. 2019;***20***(2):256-274

[19] Bellini, S., & Merrill, A. (2017). Social challenges of children and youth with autism spectrum disorders. *Autism spectrum disorders. Foundations, character-isti □, and 呵. 告 ctive strategies*, 201-222.

[20] Oh-Young C, Filler J. A meta-analysis of the effects of placement on academic and social skill outcome measures of students with disabilities. Research in Developmental Disabilities. 2015;***47***:80-92

[21] Chen CH, Lee IJ, Lin LY. Augmented reality-based video-modeling storybook of nonverbal facial cues for children with autism spectrum disorder to improve their perceptions and judgments of facial expressions and emotions. Computers in Human Behavior. 2016;***55***:477-485

[22] Bellini S. *Building Social Relationships: A Systematic Approach to Teaching Social Interaction Skills to Children and Adolescents with Autism Spectrum Disorders and Other Social Difficulties*. AApC publishing; 2008

[23] Moyes R. *Incorporating Social Goals in the Classroom: A Guide for Teachers and Parents of Children with High-Functioning Autism and Asperger Syndrome*. Jessica Kingsley Publishers; 2001

[24] Gresham FM, Elliott SN. *Social skills rating system: Manual*. American Guidance Service; 1990

[25] Ryan C, Charragáin CN. Teaching emotion recognition skills to children with autism. Journal of Autism and Developmental Disorders. 2010;***40***(12): 1505-1511

[26] Smith A. Cognitive empathy and emotional empathy in human behavior and evolution. The Psychological Record. 2006;***56***(1):3-21

[27] American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

[28] Lee IJ, Chen CH, Lin LY. Applied Cliplets-based half-dynamic videos as intervention learning materials to attract the attention of adolescents with autism spectrum disorder to improve their perceptions and judgments of the facial expressions and emotions of others. Springerplus. 2016;**5**(1):1-12

[29] Sucksmith E, Allison C, Baron-Cohen S, Chakrabarti B, Hoekstra RA. Empathy and emotion recognition in people with autism, first-degree relatives, and controls. Neuropsychologia. 2013;***51***(1):98-105

[30] Smith-Demers A, McCrimmon AW, Shaw KS. Vocational training for adolescents and adults with autism Spectrum disorder: Effectiveness of the practical assessment exploration system (PAES). Emerging Perspectives: Interdisciplinary Graduate Research in Education and Psychology. 2017;***1***(1):1-19

[31] Wang, P., & Spillane, A. (2009). Evidence-based social skills interventions for children with autism: A meta-analysis. *Education and Training in Developmental Disabilities*, 318-342.

[32] Boyd BA, McDonough SG, Bodfish JW. Evidence-based behavioral interventions for repetitive behaviors in autism. Journal of Autism and Developmental Disorders. 2012;***42***(6): 1236-1248

[33] Stahmer AC. Teaching symbolic play skills to children with autism using pivotal response training. Journal of Autism and Developmental Disorders. 1995;***25***(2):123-141

[34] Krantz PJ, McClannahan LE. Social interaction skills for children with autism: A script-fading procedure for beginning readers. Journal of Applied Behavior Analysis. 1998;***31***(2):191-202

[35] Sansosti FJ, Powell-Smith KA, Kincaid D. A research synthesis of social story interventions for children with autism spectrum disorders. Focus on autism and other developmental disabilities. 2004;***19***(4):194-204

[36] Gray CA, Garand JD. Social stories: Improving responses of students with autism with accurate social information. Focus on autistic behavior. 1993;***8***(1):1-10

[37] Wang, W.Z., & Lee, I. J. (2020, July). Social Intervention Strategy of AR Combined with Theater-Based Games to Improve the Performance of Autistic Children in Symbolic Play and Social Skills. In *International Conference on Human-Computer Interaction*. Springer, Cham.

[38] Hsu, H. T., & Lee, I. J. (2020, July). Using augmented reality technology with serial learning framework to develop a serial social story situation board game system for children with autism to improve social situation understanding and social reciprocity skills. In *International Conference on Human-Computer Interaction* (pp. 3-18). Springer, Cham.

[39] Krasny L, Williams BJ, Provencal S, Ozonoff SJ. Social skills interventions for the autism spectrum: Essential ingredients and a model curriculum. Child and Adolescent Psychiatric Clinics of North America. 2003;***12***(1):107-122

[40] Hwang B, Hughes C. The effects of social interactive training on early social communicative skills of children with autism. Journal of Autism and Developmental Disorders. 2000;***30***(4): 331-343

[41] Mills B. Autism and the imagination. Autism and representation. 2008:117-132

[42] Lee, I. J. (2020). Kinect-for-windows with augmented reality in an interactive roleplay system for children with an autism spectrum disorder. *Interactive Learning Environments*, 1-17.

[43] Tsai, W. T., Lee, I. J., & Chen, C. H. Inclusion of third-person perspective in CAVE-like immersive 3D virtual reality role-playing games for social reciprocity training of children with an autism spectrum disorder.

[44] Grandin, T. (2008). *Thinking in Pictures, Expanded Edition: My Life with Autism*. Vintage.

[45] Didehbani N, Allen T, Kandalaft M, Krawczyk D, Chapman S. Virtual reality social cognition training for children with high functioning autism. Computers in Human Behavior. 2016; ***62***:703-711

[46] Chen CH, Lee IJ, Lin LY. Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders. Research in Developmental Disabilities. 2015;***36***:396-403

[47] Bai Z, Blackwell AF, Coulouris G. Using augmented reality to elicit pretend play for children with autism. IEEE Transactions on Visualization and Computer Graphics. 2014;***21***(5):598-610

[48] Ip HH, Wong SW, Chan DF, Byrne J, Li C, Yuan VS, et al. Enhance emotional and social adaptation skills for children with autism spectrum disorder: A virtual reality enabled approach. Computers & Education. 2018;***117***:1-15

[49] Huang, Y. C., & Lee, I. J. (2019, July). A study on the development of a mixed reality system applied to the practice of socially interactive behaviors of children with autism Spectrum disorder. In *International Conference on Human-Computer Interaction* (pp. 283-296). Springer, Cham.

[50] Huang, Y. C., & Lee, I. J. (2020, July). Using augmented reality and concept mapping to improve ability to master social relationships and social reciprocity for children with autism Spectrum disorder. In *International Conference on Human-Computer Interaction* (pp. 19-37). Springer, Cham.

[51] Lahiri U. Scope of Virtual Reality to Autism Intervention, In *A Computational View of Autism* (pp. 83–130). Cham: Springer; 2020

[52] Dragomir, M., Manches, A., Fletcher-Watson, S., & Pain, H. (2018, October). Facilitating pretend play in autistic children: Results from an augmented reality app evaluation. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 407-409).

[53] Lillard AS. Why do the children (pretend) play? Trends in Cognitive Sciences. 2017;***21***(11):826-834

[54] McCune L. Developing symbolic abilities. Symbolic transformation: The mind in movement through culture and society. 2010:173-192

[55] Lillard AS, Lerner MD, Hopkins EJ, Dore RA, Smith ED, Palmquist CM. The impact of pretend play on children's development: A review of the evidence. Psychological Bulletin. 2013;**139**(1):1

[56] Weisberg DS. Pretend play. Wiley Interdisciplinary Reviews: Cognitive Science. 2015;***6***(3):249-261

[57] Rutherford MD, Rogers SJ. Cognitive underpinnings of pretend play in autism. Journal of Autism and Developmental Disorders. 2003;***33***(3): 289-302

[58] Jarrold C, Boucher J, Smith P. Symbolic play in autism: A review. Journal of Autism and Developmental Disorders. 1993;***23***(2):281-307

[59] Ekman P. Emotional and Conversational Nonverbal Signals, In *Language, knowledge, and representation* (pp. 39–50). Dordrecht: Springer; 2004

[60] Burgoon JK, Guerrero LK, Manusov V. *Nonverbal Communication*. Routledge; 2016

[61] Simmons, S. (2017). What we Can't Understand: The Non-Verbal Communication among Autistic Children.

[62] Van Wermeskerken M, Grimmius B, van Gog T. Attention to the model's face when learning from video modeling examples in adolescents with and without autism spectrum disorder. Journal of Computer Assisted Learning. 2018;***34***(1):32-41

[63] Noens I, van Berckelaer-Onnes I. Making sense in a fragmentary world: Communication in people with autism and learning disability. Autism. 2004;***8*** (2):197-218

[64] Huang, J. Y., Hung, W. H., Hsu, T. Y., Liao, Y. C., & Han, P. H. (2019). Pumping life: Embodied virtual companion for enhancing immersive experience with multisensory feedback. In *SIGGRAPH Asia 2019 XR* (pp. 34-35).

[65] Davis MH. *Empathy: A Social Psychological Approach*. Routledge; 2018

[66] Decety J, Meidenbauer KL, Cowell JM. The development of cognitive empathy and concern in preschool children: A behavioral neuroscience investigation. Developmental Science. 2018;***21***(3): e12570

[67] Knapp, M. L., Hall, J. A., & Horgan, T. G. (2013). *Nonverbal Communication in Human Interaction*. Cengage Learning.

[68] Blum-Dimaya A, Reeve SA, Reeve KF, Hoch H. Teaching children with autism to play a video game using activity schedules and game-embedded simultaneous video modeling. Education and Treatment of Children. 2010;***33***(3):351-370

[69] Schulte-Rüther M, Otte E, Adigüzel K, Firk C, Herpertz-Dahlmann B, Koch I, et al. Intact mirror mechanisms for automatic facial emotions in children and adolescents with autism spectrum disorder. Autism Research. 2017;***10***(2):298-310

[70] Maskey M, Rodgers J, Ingham B, Freeston M, Evans G, Labus M, et al. Using virtual reality environments to augment cognitive behavioral therapy for fears and phobias in autistic adults. Autism in Adulthood. 2019;***1***(2):134-145

[71] Baron-Cohen S, Leslie AM, Frith U. Does the autistic child have a "theory of mind"? Cognition. 1985;***21***(1):37-46

[72] Wolfberg PJ. *Play and Imagination in Children with Autism*. Teachers College Press; 2015

[73] Hachiya K. Inter dis-communication machine. Prix Ars Electronica. 1993;***96***: 138-139

[74] Barton EE, Choi G, Mauldin EG. Teaching sequences of pretend play to children with disabilities. Journal of Early Intervention. 2019;***41***(1):13-29

Section 3

Nonverbal Communication in Media and Education

Chapter 4

The Body Speaks Society, School and Culture

Manuela Valentini, Maria Chiara Mancini and Ario Federici

Abstract

How can we help all children, since birth, become effective communicators and interpreters? Why should nonverbal behaviour be of interest? The aim of this research is to reflect on the importance of every element of the analogical language, related to a target audience of preschool and school children aged between 0 and 8 years that is always little studied. The ability to communicate is an essential skill that has roots in early childhood; preschool children especially prefer the body as means of communication, from birth. Children learn to know the analogical language by observing the one of the parents and by imitating him. It is worth to underline the essential role of school that, beyond the family context, is the privileged environment for the development and learning of communication, both verbal and non-verbal. However, non-verbal languages are determined by cultures, that is, they are not equal for all regardless of cultures, but they change depending on cultures themselves; understand cultural foundations of the communication, in today's multicultural and pluralistic world, is an essential help to handle an appropriate conversation.

Keywords: non-verbal communication, early development, children, parents, teachers, cross-cultural effects

1. Introduction

Communication is a basic need for all living beings since their birth. It is the foundation of society itself, as it implies an interaction between interlocutors and therefore an exchange, that can be social, interpersonal, verbal, nonverbal, analogical, digital and so on. During the evolution process, the majority of the animal species developed the ability of conveying and receiving messages that could be understood by every member of the same species. These messages may be about: reporting of food or warning of danger, sexual desire, prohibition linked to the social hierarchy, the will to play (which may be found in cubs) and so on. Even newborns are able to communicate in a comprehensible way and the adults who take care of them are able to understand and reply to their messages.

All these forms of communication belong to the nonverbal type and are extremely varied and complex. In each species, the nervous system has evolved so that it could decode and produce nonverbal and at times very complex messages. The human brain is "naturally" suitable to communicate in a nonverbal way within our species. Moreover, the human form of nonverbal communication, which until 4–5 million years ago was identical to the one of chimpanzees, has evolved even

more, becoming extremely sophisticated and allowing humans to interact and cooperate in a more and more well-structured way [1].

How preschool and school children communicate today? Do they prefer verbal or nonverbal channel? Do they use their body? What is the task of Kindergarten and Primary school teachers to encourage the nonverbal channel?

Communicating means sending, passing down, handing over, notifying, making see, making feel, describing, meeting, spending, infecting, sharing, connecting, share with others what belongs to us. Communicate means so many things and can mean nothing: it all depends on the meaning we want to give to the word itself and on the way we act to achieve the aim [2].

Communication is linked to the exchange of information between two or more entities capable of produce and receive signals, but, to do it at best, it is necessary to pay attention to have consistency between the explicit message, given by the verbal language, and the implicit message, transmitted by using the nonverbal language. Although it may seem obvious, the two types of message do not go hand in hand, so that communication can leave the interlocutor in a situation of ambiguity and uncertainty towards the other one. In that case, he tends to rely on the message got nonverbally, compared to the verbal one, which is more spontaneous and more difficult to hide [3].

A communicative event, besides the fundamental abilities of "know the language", "know how to do the language" and "know how to do with the language", requires complementary abilities. The complementary ability par excellence is the emotional intelligence, discovered in 1990 by Salovery and Mayer, who defined as "the ability to control ourselves and others feelings and emotions, distinguish between them and use this information to guide his own thoughts and actions".

Therefore, communication is the process of transmission of an idea or emotion from one person to another through facial expressions, gestures, speech or via means of communication like writing, telephone, radio, television, etc. Communication skill can be defined as "one's ability to express his/her emotions, ideas, beliefs and attitudes comprehensibly way and relevantly" [4].

About it, Mehrabian (1967), having regard to the results of his research, has pointed out the importance of nonverbal communication for years. His studies determined that of any given message, only thirty–five percent is verbal; the remaining sixty–five percent is communicated through various nonverbal channels. The nonverbal component of a message includes movements and body position which we unconsciously sense but often overlook. The major part of any message is conveyed through non-verbal channels, but then why do the education system, from the first childhood, takes the most of his time and efforts in ensuring that children master the thirty–five percent of communication, which means words?

The ability to communicate is an essential skill that has roots in early childhood. Children begin communicating from birth, but they need attention of adults, be they parents or caregivers, so they can develop communicative skills to express themselves clearly and confidently. Parents are the young child's first teachers of communication who help him to master non-verbal and verbal communication through listening, watching and responding to the sounds, communicative gestures and language the child uses [5]. The child learns to know the body language by observing the one of the adult and by imitating him, that's why the first step to improve communicative abilities is to recognise non-verbal signs that we are sending in relation to the verbal ones sent.

Therefore, this research try to investigate and consider all aspects of nonverbal communication in today society, focusing on a target audience of preschool and school children.

2. Nonverbal communication

In 1967, Mehrabian claims that, daily, the human being communicates through three different channels (**Figure 1**):

- verbal, composed of words and content, that accounts on the communication for 7%;
- paraverbal, composed of tone of voice and rhythm of word, that accounts for 38%;
- nonverbal, composed of gesture, mimicry, posture, proxemics, that accounts for 55%.

Therefore, verbal channel, even though it is considered the main communication channel, accounts for 10% in living beings communication; more than 90% of what is communicated is not given by words, so it is not verbal, it does not come out of the mouth and it is not received by ears.

This distribution is not related to a particular age group or to a temporary vital moment; in fact, at the moment when, growing, the use of word is acquired, the nonverbal transmission aspects of messages are not lost, they become a particular feature of every person. Many of these attitudes are linked to a characteristic of the person itself, while others are more properly identifiable in each of us, at the moment we are in a specific situation, with a specific mood [3].

Nonverbal communication is a highly efficient and pervasive means of interpersonal exchange [6]; therefore, in a communication, the aspects that have to be considered are not only the verbal ones (digital language), that is, related to the meaning of words and on <what>, but also the nonverbal ones (analogical language), that is, related on <how> [7].

Know the nonverbal language is therefore very important to communicate well, to transmit or interpret others messages, to hear and understand at the same time.

The essential aim of a communication is that it results effective; the communication effectiveness, encouraged by nonverbal communication and his main elements,

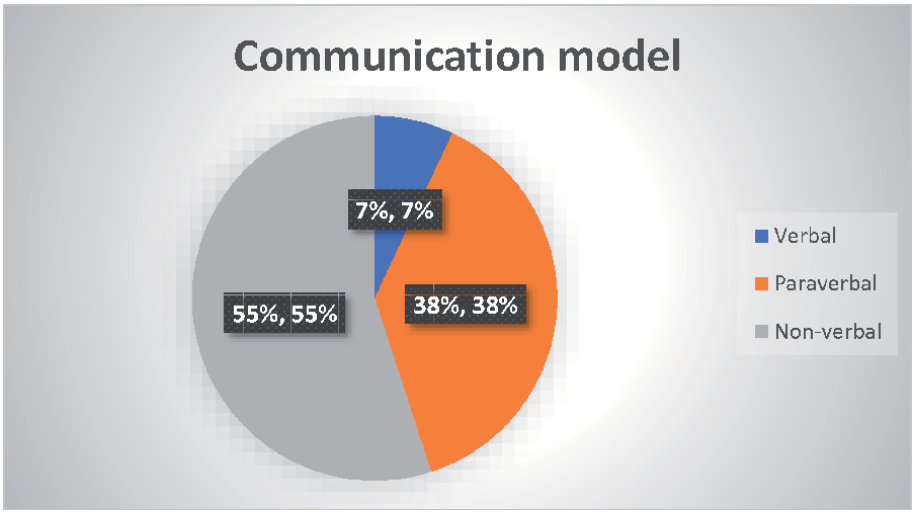


Figure 1.
Source: own elaboration.

comes from the mood of the interlocutor and from the aim that the interlocutor wants to give to the communication. If these components are missing, the risk is to obtain the opposite effect or at least different from intentions.

What makes a message incisive and productive, combined with a clear and effective verbal channel, is the body. This is the first means with which the individual experience the environment, it is the first communication channel between him and the others, between his inner and the outside world. The body expression, intended as the spontaneous nonverbal demonstration of every of our acts, precedes and comes with communicative and artistic techniques [8].

Body language is a universal language code, with a no precise syntax, but with strong cultural influences, in fact, to be understood, it has to be inserted in his reference cultural context. The socio-cultural context is the main learning form of body language, but it is passed on even by daily and family relationships and it remains by lived experiences [1].

Voluntarily or involuntarily the body is the means by which we appear to others. It communicates, through subjective ways, our most deep and true feelings, it reveals our moods and our emotions: that's the reason why it becomes essential the importance of being conscious of what our body can tell about us, so that to send or understand better messages and have a better relationship with ourselves and with others.

Human beings, since birth, are capable of emotion and recognise feelings. The major need of the human being is to have the possibility to reveal his own emotional language, by communicating his own feelings. But too often, words are not enough to express completely the sense and the meaning of feelings; in this way, they come out through the body language, through the face, through attitudes, etc. Expressions and external modifications represent, therefore, internal expressions and mutations, with which they are always in connection.

So, an excited body allows people establish relations with each other through the deepest level of communication [9].

2.1 Components of non-verbal communication

Nonverbal communication is composed of different components, each of which, in turn, includes various elements describing small parts [3]. The main components of non-verbal communication are 4:

- paralanguage;
- kinesics;
- proxemics;
- haptics.

Being able to recognise all these parameters gradually, from the first development, within the communication, means realise the inner state of a person to handle a dialogue and maintain a climate of confidence, by adjusting our intervention.

2.1.1 Paralanguage

Paralinguistic is the study of intonations and inflexions of the voice. Paralinguistic system concerns the prosody, that is, the trend and the dynamic of phonatory flux [10]. This is considered the main element of a communication,

both to characterise the real meaning and his significant. The voice and his characteristics such as duration, intensity, tone, intonation, modulated by the reference culture, can reveal subjective and relational feelings and emotions such as anxiety, fear, anger, pain and joy, especially through the rhythm modulation, the intonation and the speech intensity. In addition, in communication, the trend of the rhythm is modified by pauses, that are distinguished in empty and full pauses. Full pauses are the typical interjections (mmm, ehm…) without a verbal meaning, that are inserted between two sentences; empty pauses, on the other hand, represent silence.

2.1.2 Kinesics

Kinesics involves all communicative acts that are expressed by body movements, by the expression of the face and by gestures, both voluntary and involuntary, mostly linked to emotions. These signals sent by the body most of the time can be interpreted.

Posture

Posture, which means the position of the body or parts of it, is a very significant factor or communication because it reveals different information about the subject: basic attitude, mood and self-image, but interlocutors must be "good readers" to avoid being deceived by appearances. It is dynamic and it changes continuously during the interaction, by signalling and showing internal emotions or particular attitudes to the parties involved. Posture is determined by different factors: the psychic component, since it depends on emotional dynamics inherent in the relationship between the individual and all others; anatomical components, as is the morphological and muscular conformation or pathologies that determine the posture; proprioceptive capacity, on which depends the static and dynamic self-perception that affects the position of the body [1, 3, 11].

Clearly, there is no model posture, but on the contrary, it varies according to all social and context elements, to the situation, to people and to the dialogue type [12]. Some studies have shown even the presence of the "Chameleon effect": the tendency to adopt the partner positions during the social interaction [11].

Expressions

The face expression is the communicative means par excellence and it can be defined as such thanks to mimicry and look, which allow the inscription of state of mind and emotions on the face, as well as interpersonal attitudes. In fact, the face is the most observed part of the body by interlocutors; particularly, look is the channel by which the nonverbal message is sent and helps to perceive nonverbal messages sent by others [13]. This last is characterised by intensity, duration and direction; these elements vary according to the level of intimacy between the parties, to behind emotions and social value in a specific cultural context.

Eye contact, during a conversation, serves to send and collect information, has the function of acquire the partner feedback; in addiction, it promotes the climate of cooperation and is a power signal to obtain consent. Eye contact is the first important step to start every interpersonal relationship, whether it is positive or negative. In fact, reciprocity and frequency of looks represent an indication of directness, transparency and trustworthiness of the interlocutor and are linked to positive emotion, like joy and tenderness, while negative emotions like disgust or embarrassment require a swerve of the eye from the interlocutor. Eye contact plays a central role also in conversation scheme to modulate empathy [14] and to suggest the speaking time between speakers. Therefore, eye contact is essential to collect and receive information, whether parties intend to communicate with the eyes or not.

On the other hand, smile, one of the fundamental signs of human species, plays important functions in social interactions: can act as regulator of social relations, as

promoter of relational affinity by representing an attitude of relational availability or interpersonal closeness, as information tool [15]. It is strongly related to social interaction, not the emotional state. Smiles are divided in "Duchenne[1]" (natural, genuine) and "Non-Duchenne" (false, forced): the first smiles involve both muscles lifting the corners of the mouth and muscles surrounding the eyes and reflecting the positive affect; the second smiles, concern only muscles lifting the corners of the mouth and are considered like masked smiles, often produced in absence of positive affects [16].

In 2010, Bonaiuto enhance that <<through eyebrow and forehead movements, eyes displacement, use of the smile, in fact, the face express emotions and interpersonal attitudes>>.

Gestures

Finally, gestures are movements of hands or the body, coordinated motor actions, that allow a more rapid and efficient communication system which goes beyond words and verbal language [17]. Gestures are immediate, since with one only gesture a lot of information is sent which would need lot of words, and incisive, because requiring a minimum sending time with respect to a verbal message.

Additionally, gestures can be conscious and unconscious. Conscious gestures are movements made voluntarily to make clearer the message to the interlocutor, unconscious gestures are activated regardless our intention and express the current mood (for example cold sweat, red cheeks, etc.).

Gestures can be directly related to the speech *"cospeech gesture"*: accompany words, by making them more effective and by precising the sentences meaning; or can transmit a verbal meaning independently from the speech [18]. All these movements of hands and body are divided in gesticulation, pantomime, expressive gestures, emblems, motor gestures and sign language.

2.1.3 Proxemics

Proxemics is the semiologic discipline that studies individuals behaviour, space and distances inside a communication [11]. The proxemic components of communication is therefore in strong relation with space and its occupation; that's why it is called proxemic space: this space has to be considered as the minimum subjective space that a person draws around her. Even though this distance could apparently seem causal, actually there are very precise rules and it vary according to the type of the existing relationship [11].

In this regard, E. T. Hall[2] defined four interpersonal zones:

- the intimate area (0–45 cm) reserved to some close relatives and the partner,
- the personal area (45–120 cm) used for relatives and friends,
- the social area (1,2–3,5 cm) used for interactions between acquaintances or formal meetings,
- the public area (over 3,5 cm) synonym for public relations of official occasions.

[1] G. B. Duchenne, important French neurologist of XIX century, discovered, through electricity, that smiles resulting from real happiness do not use only mouth muscles, but also those of the eyes. These "genuine" smiles are known as Duchenne smiles in his honour.

[2] Edward T. Hall, American anthropologist, worked on proxemics and in 1966 published his essay "The hidden dimension", in which he collected his researches results. His deductions derive from the observation of animals and people behaviour.

The physical distance helps to understand the existing relationship between interlocutors, even is the radius of these zones varies necessarily according to the crowding, to the age, to sex, and to the importance of interlocutors. Moreover, the physical distance signs to every individual the relationship built with beloved persons, friendly people and the world in general, by making him conscious of the relationships he is living.

2.1.4 Haptics

Haptics focuses on body contact in managing relational and communicative dynamics. Communicative messages expressed through physical contact can be codified or spontaneous; codified communicative forms are those universal and conventional, such as handshake, kisses on cheeks, while spontaneous communicative forms are not conventional and they are subjective. Contact, additionally, can be mutual, like a handshake, or individual, wanted by only one person and that, for one reason or another, is not returned. In fact, the body contact, which influences the nature and the quality of the relationship and which express various interpersonal attitudes, has multiple effects: can, on one hand, facilitate empathy and socialisation, on the other hand, create discomfort, harass, annoy and raise negative reaction. Moreover, some zones of human body allow the contact only between people in close relationship and, equally to distances and proxemics, the contacts show a major or minor intimacy between individuals.

Finally, haptic system does not consider only gestures with which we touch others and we enter into relationship with them, but even self-contact gestures, as is, those referred to ourselves. These behaviours allow to the individual to adapt to external stimulus, because they allow the manipulation of a part of the body (for example the gesture of rub the arms to calm down) [11]. Psychoanalyst Anna Maria Sepe has identified the most frequent self-touch gestures and she group them as follow: touching hair (gesture used as tranquilliser), rubbing the root of the nose (gesture that express rejection), touching the throat (gesture that communicate anxiety and agitation).

2.1.5 Chronemics and vestemics

Last studies concerning communication involve also chronemics and vestemics between the main components of nonverbal communication.

Chronemics is the science that studies the perception, the organisation and the manifestation of the time in human contexts. Every individual has a personal rhythm and does not know that of the other, or he takes for granted that every rhythm is the same as his. This can lead to discomfort and dystonia at the moment in with rhythms don' match; communication, instead, could result synchronic at the moment in which there is a regular and flowing communicative flux [10].

Vestemics is the system referred to the physical structure, to clothing and personal ornaments, therefore we can speak about physical appearance; this last, subjected to constant changes of fashion, involves, as well as clothes, even makeup, hairstyle, accessories, belonging objects, status symbol signals. Consequently, tends to influence what others think about us, as it leads to self-presentation, it helps to provide our image in interpersonal relationships, both private or public.

Ultimately, is useful, while observing a person, do not focus only on one of the elements of nonverbal communication (only hand gestures) or on one dimension (only kinesics behaviour), but consider all dimensions and also the person in his/her uniqueness and context in which the person is [3].

2.2 Non-verbal communication functions

As already said, nonverbal communication is used to send messages of the body, to express emotional states, attitudes, emotions, one's personality, to strengthen verbal expressions, to send feedback and signals in order to achieve better empathy with the interlocutor [1].

It has function:

- Individual-expressive, to satisfy interior affective needs;
- Cognitive, to relate the subject with the external world;
- Psychological-developmental, to build the individual autonomy and awareness relating to the external world;
- Cultural-social, to transmit history and culture of a society.

Nonverbal communication is therefore used for:

- Establish interpersonal relationships: it can be considered as "relationship language"; through nonverbal language there is the possibility to influence others behaviour: according to the way we act in front of the interlocutor, he will act accordingly. If, for example, we act in a way that expresses a sense of superiority, it is at least possible that the who is in front of us will act in a defiant attitude. It has been observed, In the wake of Mehrabian studies, how nonverbal language has major power than verbal communication, in influence others behaviour [3]. Consequently, the most used signals are gestures, voice tone, posture, mimicry and look, that allow to express our feelings with regard to other persons and create relationships between interlocutors. These signals are the first that have been used to communicate and represent the first means of expression in the relationship infant-mother.
- Show the personality: it allows to express the image of ourselves and introduce to others. Face expression, look, gestures and movements make it clear the personality of who is in front of us. Sometimes we try to send controlled o modified nonverbal signals to give a better presentation of us.
- Support verbal language: nonverbal signals, linked to what it is said, influence this last and give more information to the interlocutor. Nonverbal signals complete the verbal communication and have a meta-communicative function, as they supply elements to interpret verbal expression.
- Express emotions: interlocutors are capable of understand and recognise body status and moods through the analysis of face, eyes, gestures, postures and voice tone.

2.3 Cultural influence

Culture has strong and pervasive influences in communication [18]. Verbal and non-verbal languages are determined by cultures, that is, they are not natural and equal for all regardless of cultures but they change depending on cultures themselves [19, 20]. Understand cultural foundations of the communication, in today's multicultural and pluralistic world, is an essential help for human beings

to handle a conversation. The lack of knowledge of cultural aspects, that extend beyond of the purely linguistic knowledge, can lead to the rise of misunderstandings and communicative problems for interlocutors [20], who stop the social coordination and increase the conflict potential [18]. It's easier for people coming from expressive cultures (Mediterranean cultures, Latin American ...) to judge as untrustworthy, inscrutable, smart, misleading or elusive those belonging to reserved culture. At the same time, it's easier for people coming from reserved cultures (Asian, northern European...) to judge as arrogant, noisy, immature or vulgar.

Hall, in 1968, remarks that people of different cultures, as well as speaking different languages, live different sensory worlds, set according to cultural influences. Therefore, it is usual that in multicultural communication a different perceptive way could lead to contrasting meanings and obstacle to communication. In such context, recognise a nonverbal signal, do not always lead to an adequate interpretation of its meaning. Consequently, it's easier to misunderstand, especially when we suppose, perhaps unconsciously, that a particular behaviour has the same meaning in every culture.

Nonverbal communication components vary considerably according to the reference culture: native and stranger speaker communicative competences depend on different cultural background, is therefore clear what are the inconveniences that could arise after have used or interpreted gestures that can be innocent in a language but rude and insulting in another. In fact, gestures, can be more or less conventional, but it is important to know that they show relevant cultural variations, especially for what concerning emblems (gestures with the verbal meaning that id different from word, like the OK signal) and the sign language, used by deaf-mute. The knowledge of gestures and their specific use can be very useful, particularly to know when use them as to not appear ridiculous or rude.

Face expressions and smiles are universal and biologically instinctual, in fact, they have the same meaning in all nations, but culture plays an important role in moderating their use [18]. For what concerning eye contact, it has strong cultural influences: for example, Eastern people tend to look a stranger person for a long time, thing that can be particular according to Western people.

The use of silence, on the other hand, has an ambiguous and ambivalent nature, in fact his meaning varies according to the situational context, the type of relationship of the participant of the communication act, the reference culture [10]; as a matter of fact, Westerners, who are constantly surrounded by a life characterised by din, perceive silence as resignation, powerlessness, and paradoxically connect it to death or disgrace; the Eastern culture instead, is still perceiving and living silence as a force that reaches the bottom of someone's inner reality, digs down into it, and that is able to protect the individual. Just as much as the voice tone that, used in a conversation both high or low, varies according to the discussing topic and the cultural influence, but in general, as shown by Matsumoto and Hwang in 2013, expressive cultures encourage the use of higher vocal tones, while in restricted cultures, the voice tone is weaker and lower [18]. So, different vocal tones can cause embarrassment, discomfort and misunderstandings, especially if interlocutors are from different cultures [20].

Proxemics first of all depends on culture; for example, during a communication situation, distant for an Asian person, is shorter than the one of a European person. Moreover, inside the same culture, proxemics is influenced by the type of the existing relationship: there's a restricted space, same for everyone in the world, inside of which are included only affectively close people. This space is enlarged in case of working relationships and increases more and more as quickly as people in front of us are less known and are perceived as more far away [1].

Body contact and the effects, favourable or not to relations, depend on the subject and his personal inclination, but also cultural differences play a crucial role. There are countries in which contact is frequent and intense (like those Arabian and Latin American) and others in which the contact is even forbidden.

In public relation, vestemic system influences the communication and leads to relations of dominance and persuasion between individuals. Especially in modern society, more and more people give importance to external appearance, not only given by the way of clothing and dressing up, but also given by the purely physical condition, by considering it a benchmark in the relationship choice. Different persons have major or minor confidence even through the way of dressing and the choice of personal objects that often act as status symbol [21].

Finally, it is important that the interaction between interlocutors is adequate from the point of view of chronemics (as is, that respects others' rhythms), so that the communication results effective: each individual has a personal rhythm and does not know the ones of other people, or presumes that they are identical to his; this can lead to awkwardness and dystonia, and it happens when rhythms do not correspond with each other, while when a regular and fluid communication flow is established, communication is synchronous [10]. In modern world, this aspect of communication is very neglected, mostly in Western culture. The time, in fact, in the fast culture if western world is not considered as a resource or an asset to spend in communication, but it is actually lived in a completely frenetic way, in a temporary perspective always future-oriented and economically respecting the "time is money" rule. On the other hand, in Eastern culture, the time has not yet been charged of these negative meaning and it is not lived with such anguish, people live in view in which "time is relation".

In conclusion, being expert in nonverbal communication and knowing the main analogical and intercultural differences, allows to establish effective and beneficial communicative relationships, even ignoring completely o partially interlocutors language.

3. Children and communication

Communication, especially from the nonverbal point of view, assumes a basic position in personal and social development of every individual. With children communication assumes a relevant aspect even in relationship because it is necessary to adjust and moderate every message according to their age. Knowing and understanding how the child's brain matures, when it reaches specific competences, when it is capable of doing specific types of mental operations, can be a very big help for adults, both parents and teachers, which daily interact with children. Working on the way to communicate of the adult with the child at school and at home (but even during free time) can truly make the difference for significance relationships of life. Talk in a respectful, empathic and friendly way to children is an attitude that helps a lot in little daily difficulties and during learning times, that brings real harmony affective moments with children, which ask constantly to the reference adult to help and understand them.

For a healthy development, the child, since birth, needs to live daily communicative situations that are challenging, rewarding and most of all clear. With the beginning of kindergarten, the student, who is living a great development moment both physical, cognitive and especially lexical, gets in touch constantly with peers and adults, with which he spends time and establishes solid and significant relationships, by communicating regularly, giving and receiving information [22]. It is therefore important that kindergarten gives to students effective communicative

exchanges, that are adequate to subjects age, by increasing in them cognitive and relational curiosity. The key role is the one of the teacher: as communication is an important component of early years education, pre-school teachers also need to have adept communication skills [4]. He has to transit to students not only didactic information, but also rules applicable in the society, for example he has to be able to teach them to communicate effectively to solve in the right way their problems or inconveniences, without the use of verbal or physical violence. However, a good communication is based on empathy and listening, both of the teacher and the student.

Focusing on languages development steps, it must be kept in mind that pre-school children, between 3 and 5 years, still need to learn using properly the verbal and nonverbal coherence principle[3]; to do so they have to concentrate and rely on nonverbal signals, despite their salience during the incoherent communication.

2003 research by Eskritt and Lee suggests that children below 5 years old do not have this capacity yet, that can be acquired during primary school; therefore, during the pre-school age, for a lot of children is difficult to understand an interlocutor's mixed message (in which the verbal shows to be different from the non-verbal or even contradictory), finding thems99elves in a position where they have to choose a source on which they can concentrate. Most of the time, the verbal part of the message is chosen, while sometimes people rely on the non-verbal part, but only when these messages are salient, exaggerated. The research has shown also that preschool children have just start recognise the importance of intention of the speaker in communication; with the rising of the mind theory beyond the preschool years, children can develop the capacity of deduce the communicative intention of a speaker when he/she communicates in an incoherent way [23].

In addition, during this period, children, according to Boyatzis and Satyaprasad research [24], have developed major capacities for what concerning decoding, so the interlocutor emotions and gestures understanding, rather than coding, as is the production thereof [24].

So, many researchers argue that preschool years are a critical period for the development of the language; is therefore important to understand all factors that can potentially influence or interact with language development during this period. Assessment and monitoring of language and communication development should be integrated into each country's healthcare settings, alongside easy and timely access to quality early interventions and services, in case the child experiences difficulties in communication and language development [5, 25].

3.1 Non-verbal communication in 0-8 years-old children

Preschool children, which was mentioned earlier, prefer the body as means of communication; this allows them to handle themselves, to be conscious of themselves and support the development of the person in all dimensions. The way to communicate nonverbally can have long term effects on how children listen, act, process information and talk to others; therefore, it is necessary to analyse nonverbal characteristics in order to determine emotions shown by aforementioned [16].

Gestures

Word and gesture, as already demonstrated, are intimately interwoven. This strong relationship between gesture and word, emerges soon in life and gets stronger when children learn the language. The gesture, influenced by cultural

[3] This principle assumes that verbal and nonverbal behaviours transmit normally the same message, but when there is a discrepancy, adults tend to assume as trustworthy the nonverbal message, rather than the verbal one (Rotenberg et al., 1989).

differences and personal social status, has a particularly important role during starting language development. Children, when using gestures when they are not yet able to produce word and combining words and gestures when their repertoires are still limited, increase their communicative potential and communicate over the ability of their talking way [17].

Up to 3 years infants do not combine gestures with words, then they arrive to a turning point in which they start to combine them and synchronise le production of gestures and words. The start of this combination, in correspondence of kindergarten, announces a new linguistic developing phase of children and expands notably the range of their communicative systems. These characteristics of semantic integration and temporal sync, characterise the relationship between gesture and language even in adults. In fact, children from 3 years use gestures like adults [26].

As for adults, even for children, the use of gestures during a speech supports the collection of information from the memory and it lightens, by acting as an instrument, the charge of cognitive work of vocal process, by facilitating the lexical research [25].

In addition, Goldin-Meadow, Cook and Mitchell's 2009 research, showed a strong relation between gesture learning: in fact, the use of gestures helps children learning, therefore, the teacher has to stimulate it. For example, Clark et al., in 1974, showed that indicative gestures of adult speakers have helped understanding adults speech. Tfouni and Klatzky, in 1983, discovered that indicative gestures help children understanding the meaning of deictic words as, for example, "this". In 2008, McGregor reported a series of studies that show that hand gestures used in communication adult-child have been functional for the acquisition of an early vocabulary. Gesturing, by students, increases their implicit knowledge and prepares them to learning [17, 25].

As already said, gestures can be more or less conventional, but every child is different, so it is important to pass time with him to learn his gestures and meanings. It is fundamental, during Primary school, to promote the transition from the dimension of spontaneous expression to that of intentional gestural communication [1].

Expressions

Smile is one of the first mechanisms of social intelligence that is developed by children, to the point that 90% of infants has already smiled during the first two months of life. When one relates to such young subjects, a smile can reduce many barriers during the communication; a smile is also able to create a comfortable situation, that "connects" the subjects, provides and opens the communication channels. This aspect has strong cultural roots; therefore, the essence and intensity of children smiles varies according to the sociality and the culture to which they belong [16]: as a matter of fact, in the same occasions we smile neither in the same way nor with the same frequency. At first, children find it easy to smile and laugh, but their culture can inhibit their spontaneity, especially in girls, insisting on the fact that a woman should be introverted and reserved with advancing age. To the boys, on the other hand, could be taught to hide their feelings, as a sign of masculinity, with the result that, besides the tears, they will inhibit also a smile or a laugh. Ultimately, it seems that a child inherits a raw action, and then refines it with social experience. A person who belongs to an expansive culture could judge insincere a smile of another person who belongs to a more reserved culture, when in reality it's just an attenuated form of expression, conveyed by the cultural contest of belonging. In this case, knowing how to be on the specific wavelength of the culture with which one gets in touch becomes essential.

Look is linked to the internal emotive component of children: the expression of his own emotions, in fact, supply information on the internal status of the child that, at the same time through look direction (that is eye movement),

check the stimulation that he sends or perceives. Therefore, the child look can have a different functional meaning according to social-scientific relevance of stimulus. So, focusing on the combined role of the smile, the look and the emotive expression, researchers examining the early affective development in developing subjects, claim that children smile more when they look people and interact with them [14].

In general, children do not have the same capacity as adults to filter, cover and hide their own emotions, so they tend to show always their feelings. Facial expressions generally arise what a child is thinking or feeling and it is very easy to understand when he likes something or not.

Proxemics

The space is a very important aspect, in fact, since birth are necessary relationships that are built inside an intimate and personal space. But age is an aspect that affects the space use; in fact, a research has shown that from 3 years old to almost 21, personal space increase gradually and constantly. This has been demonstrated both in naturalistic and experimental contexts. However, Proxemics is not influenced only by age, but also by cultural, familiar and gender differences [11]. The teacher, in fact, has to try to establish physical contact with his/her students, but respecting times and modality of each one of them; if not he/she risks to invade the student personal space by creating inconveniences and negative reactions, that affect the relation and the development of the child.

Throughout history, experts have succeeded in formulate various strategies to try to improve non-verbal communication in children, included those with disabilities. For example, children, especially those with ASD, often need to be taught about eye contact or to hold body contact. Eye contact can be stimulated and taught by holding objects, that they draw their attention, in front of own eyes; as time goes by, ASD individual will look up independently, when he/she chase something.

These strategies are important because, first of all, allowing all of them to learn how to communicate with each other without the barrier of verbal language, creates a sense of empathy, patience, understanding and equality.

3.2 The teacher of kindergarten and primary school

The presence of motivated, prepared teachers that are also careful to children groups specificity, is an essential quality factor for the construction of a friendly, inclusive, sure and well organised environment, also capable to stimulate trust in parents and community [27]. So, the teacher has to create conditions in order that students can learn in the best conditions, by identifying his/her role of speaker, harmoniser and stimulus in relationships between students and knowledge. Context is a fundamental element for the child, because, when he realises of being in a suitable space, he starts to communicate in a relaxed manner, not only verbally, but also with his own body; in addition, the child can establish good relations with his friends in a conscious way [15].

Teachers communication is of great importance in children educative process; in fact, according to Pianta studies, have effective communicative capacities is very important for their development and education. It was determined that pre-school children who perceive their relation with their teachers as positive, warm and close had better school adaptation. However, school adaptation of students who perceive their relation with their teachers as negative were poorer compared to others. In similar studies, Brich and Ladd in 1997, and Justice, Cottone, Mashburn, RimmKaufman in 2008, it was revealed that when child-teacher relationship is negative, children exhibit negative attitudes towards school, have poor academic

success, spend less time at school environment, and avoid from being at school and cooperation in the class [4].

It is fundamental that the teacher, starting from kindergarten, is capable of transmit messages effectively, as is, that is an expert in oral expression and in verbal explanation of certain concepts. But it is not enough: it's not only about what it said, but how is said [28]. The positive or negative attitudes and behaviours that the teacher assumes towards the child, and the verbal and non-verbal messages, which he sends, are closely correlated with the confidence, autonomy and emotions of assertiveness, which the child begins to develop in this period [4]. Nonverbal messages, in fact, are a fundamental component of communication in the learning process.

For example, find, in the morning, a smiling teacher, who says hi to students and parents by welcome them with open arms, rather than with crossed arms and pour, modifies the sent message making the difference for positive relationships. Greeting is done in both examples but emotions that arise are totally different [23]. So, it must understand, that it is not only important to refer a message, but even find the right way to transmit it. This can happen only by accompanying and controlling gestures and expressions.

3.2.1 Nonverbal communication components in kindergarten and primary school teachers

The teacher could be conscious of which are nonverbal messages circulating inside the class; this consciousness makes the teacher more watchful to messages sent by the child, it allows him/her to be more effective in the communication of didactic information and he/she is capable to implement the psychologic degree of closeness between student and teacher [28].

A good teacher could try to improve every aspect of his/her nonverbal communication. First of all, to make the communication style effective, it is fundamental that he/she know how to vary voice elements, as tone, rhythm, timbre, volume, cadence. Otherwise, the risk is to be monotonous, demoted, to annoy the children and not be able to attract their attention.

For what concerning postures, it is necessary to be in an upright position but not rigid, with the bust slightly extended forward; in this way children understand that the teacher is "approachable", receptive and friendly [22]. In addition, staying stationary in class and to not transmit the presence, motivation and care to children, on the other hand, walking in the class or suddenly get up, allow to achieve every student and reactivate a declining attention. Finally, it is important to speak to children in the face, avoiding to direct to them with the back turned or looking the floor. This make the student more involved in the communication and contributes to make him feel a real interlocutor.

As already demonstrated, the face of individuals transmits, besides their intention, emotions, thoughts and feelings. In the education sector, face mimicry constitutes one on communication means between the teacher and the group he has in front of him/her, and it becomes one of the most important factor to determine the environment. About it, the teacher, through his facial expression, communicates his/her mood and links to the topic a specific emotion. So, be able to relax face muscles, by smiling and keeping an eye contact, will have the effect of transmit a joy, warmth and emotive closure to children, who, by feeling relaxed and considered, will enforce the learning process. Particularly, smile can result contagious and contribute to the creation on a peaceful environment in the class; eye contact, instead, allow the teacher to have a total control in managing relationships and concentrate the attention and interest of students on him/her, facilitating

communication. Is necessary that the teacher distributes his/her look in a circular manner, to avoid to ignore any part of the classroom, and that stops himself on every student to prove him not only to look him but to watch him. Look the student is therefore important because it increase the probability that children return the look and follow the speech carefully; in addition, through eye contact, that's the possibility to collect precious feedback, as attention level, interest, tiredness degree, understanding [15].

So as gestures, lessons containing gestures improve interest a more specific learning: new reasoning forms, generalisation of new types of problems, conservation of knowledge, etc. In addition, as is known that gestures can promote learning, teachers can consider to improve their students to gesture, that have the potential to activate implicit knowledge and make students particularly receptive to education [17].

Finally, the teacher figure, as well as have and use correctly nonverbal communication elements in class, has to be a good listener. If the teacher is a good listener, he makes sense of not only what children say but also their facial expressions and gestures, and body language. She/he also attracts attention to the messages she/he communicates with his body language [4]. However, the first step to arrive to a good communicative intervention is to use an "active listening" with regard to children, which consists in thinking on the student message by only absorbing it, without emit personal messages. So, active listening allows the teacher to feel important and receive information, advices and interests. In 1991, Gordon claims that active listening improves discussions in class; it allows to support the class in case of problems, both in-school and after-school, as this strategy helps the child to express lived feelings, which are considered and respected. The teacher, thanks to this type of listening, is capable of observe very carefully interests, training needs, strengths of his/her students, that can be used to support the group-class in learning.

4. Conclusions

Since ancient times, human beings have felt the need to group in societies so establish relationships and socialise with others. This process is possible due to dialectical exchanger and in general to communication. Every communication form is possible due to signs, that link to a concrete and intelligible referent (as word), and to symbols. These two systems form the set of expressive forms (symbolic), given by language, science, myth, religion, philosophy, art, etc. It should however be noted that communication is not a simple linguistic act, but it is an action inserted in a dynamic process that a person does through another as the consequence of the information or the message sent and/or received, of transformations and changes, more or less significant, in behaviours, in attitudes, in moods, in the ways to interact with our and someone else's social world [1, 8].

Inside the communication, as already said, it is possible to distinguish two forms, which have different characteristic and appear in totally different way. Verbal communication is deliberate and arbitrary, it serves to argue, describe and narrate, by giving information about expressed topics. On the other hand, nonverbal communication is mostly unconscious, not intentional and can be ambiguous; it gives information about the subject expressing it, but not about the expressed concepts. In addition, it is important to notice that, while verbal communication tends to be characterised by partners that, in turn, speak and listen, sending and receiving nonverbal signs normally occurs simultaneously. Communication comes before and results more effective and relevant through nonverbal channel, with respect to the verbal one [6, 20].

In conclusion, it is possible to underline how communication is the fundamental base of interpersonal relationships and how the most spontaneous way to transmit something about us to another person is through verbal and nonverbal language. This last, most of the time, is not controlled by the transmitter, but allows to send important information about us, our mental status, about the humoral one, about characteristics of our own personality and even about our own story. By sending this type of information it is possible, in some way, to influence the other, by suggesting contents that the recipient can accept or not. One more time this underlines the importance of communication in all the different forms and the need to explore the knowledge of all various aspects of it [3].

Ongoing has already been highlighted the cultural influence in various communication elements; in summary, in intercultural communication it must surely learn to be aware of differences in nonverbal forms and other cultural aspects. Knowledge of nonverbal language of its meaning and verbal expressions, to which it corresponds, in fact, it can help to achieve a better communicative competence and develop a major comprehension and tolerance through interlocutors. More information and multicultural communicative competences we manage to acquire, and it will be easier to adopt strategic behaviours for a successful communication.

Nonverbal communication characterises most of the communicative process that can occur between two or more persons: it is always present and it can reveal lots of particulars related to a person with which there is communication. These components are not often very considered as, in adults, the importance is mostly given to the verbal aspect of the communication; thing that do not occur with children. These last, since childhood, tend to trust mostly the nonverbal attitude, the face expression, smiles, looks, gestures and the interlocutor position in front of them [29], mostly when the interlocutor is incoherent.

Kindergarten welcomes children between 3 and 5 years, to which teaches verbal and nonverbal communication prerequisites. Children of these ages, do not even develop totally the language, so, in order that their communication is incisive, they need the use of the body and its parts to get in touch with peers and adults. It is clear that the knowledge of these arguments taught at kindergarten are not the only element for teacher to keep in mind and maybe neither the most important to the transmission of competences. In fact, for teachers, not only of kindergarten but of all levels, it is fundamental base the work on an effective communication, in which occur, all communication elements, both verbal and nonverbal. Sure, is that everything concerning nonverbal communication development it is easily implemented when competences that have to be transmitted are mastered.

Accordingly, positive teacher-child communication makes it easier for children to adapt to school environment and it positively affects their emotional, social and even cognitive development during this adaptation. Several studies reported that positive teacher-child communication influences children's academic, social success and even their school life in the future [4]. Therefore, for a teacher is strongly important the interpersonal relationship and the communication with his/her students. Interpersonal relationship and communication centrality, in its complexity, are one of the major interesting aspects of the teacher job, since his/her behaviour in various situations often serves to children and people around him/her as a model; so, the more coherent are all body movements, the more believable will be the teacher body language and so his/her own person. The contact with the class, in fact, is established firstly through an effective body language.

This chapter has focused mostly on cultural differences and on children analogic language. At the end of this path inside the communication, it is possible to claim that an optimal development of nonverbal communication elements, sins kindergarten, is fundamental in order that children learn and grow. Investing in language

development, and in particular that of the body from early childhood (ECD), is one of the best investments a country can make [5]. A password to get first and better inside their world, that it is even ours.

Author details

Manuela Valentini[1*], Maria Chiara Mancini[2] and Ario Federici[3]

1 Department of Humanistic Studies, University of Urbino Carlo Bo, Urbino, Italy

2 Graduated in Primary Education Science, University of Urbino Carlo Bo, Italy

3 Department of Biomolecular Sciences, Division of Exercise and Health Sciences, University of Urbino Carlo Bo, Urbino, Italy

*Address all correspondence to: manuela.valentini@uniurb.it

References

[1] Valentini M, Castriconi S. L'altro linguaggio. Perugia: MargiacchiGaleno. 2007 ISBN: 978886494793

[2] Majello C. L'arte di comunicare. Milano: FrancoAngeli; 1993. ISBN: 9788820490775

[3] Trani A. L'influenza del linguaggio non verbale nella comunicazione interpersonale. Un'indagine conoscitiva con adolescenti [Internet]. 2015. Available from: http://rivistapiesse.it/2015/06/04/linfluenza-del-linguaggio-non-verbale-nella-comunicazione-interpersonale-alice-trani/#:~:text=Attraverso%20il%20linguaggio%20non%20verbale%20si%20pu%C3%B2%20influenzare%20il%20comportamento,egli%20si%20comporter%C3%A0%20di%20conseguenza.&text=Ulteriore%20funzione%20della%20CNV%20%C3%A8,emozioni%20(Argyle%2C%201992). [Accessed: 2015-06-04]

[4] Erbay F, Ömeroğlu E, Çağdaş A. Development and Validity-Reliability Study of a Teacher-Child Communication Scale. Educational Sciences: Theory & Practice. 2012; v12 n4 p3165-3172 Aut. ISSN: 1303-0485

[5] World Health Organization Regional Office for Europe. Framework on Early Childhood Development in the WHO European Region. 2020. DOI: WHO/EURO:2020-504-40239-53897

[6] Patterson M, Quadflieg, S. The physical environment and nonverbal communication. APA handbooks in psychology®. APA handbook of nonverbal communication. 2016; 189-220. DOI: https://psycnet.apa.org/doi/10.1037/14669-008

[7] Del Nista PL, Parker J, Tasselli A. Meeting. Firenze: G. D'Anna. 2014a. ISBN: 9788857711621

[8] Federici A, Valentini M, Tonini Cardinali C. Il corpo educante. Roma: Aracne. 2008. ISBN: 9788854819641

[9] Del Nista PL, Parker, J., Tasselli, A. Piùchesportivo. Firenze: G. D'Anna. 2014b. ISBN: 9788857712536

[10] Ferranti C. I sistemi di comunicazione non verbale [Internet]. 2014. Available from: http://docenti.unimc.it/clara1.ferranti/teaching/2014/13354/files/comunicazione-non-verbale

[11] Matsumoto D, Hwang H C, Frank M. The body: postures, gait, proxemics and haptics. APA handbooks in psychology®. APA handbook of nonverbal communication. 2016; 387-400. DOI: https://doi.org/10.1037/14669-015

[12] Bafera, V. Il ruolo della postura e della comunicazione non verbale nella selezione del personale [Internet]. 2014. Available from: http://www.igorvitale.org/2014/04/19/il-ruolo-dellapostura-e-della-comunicazione-non-verbale-nella-selezione-del-personale/ [Accessed: 2014-04-19]

[13] Adams Jr R, Nelson A. Eye Behavior and gaze. APA handbooks in psychology®. APA handbook of nonverbal communication. 2016; 335-362. DOI: https://doi.org/10.1037/14669-013

[14] Carvajal F, Iglesias J. Looking behavior and smiling in down syndrome infants. Journal of Nonverbal Behavior. 2000; 24, 225-236. DOI: https://doi.org/10.1023/A:1006693121491

[15] Sgambelluri R. La comunicazione non verbale in età scolare: aspetti didattico-inclusivi. Formazione & Insegnamento. European Journal of Research on Education and

Teaching. 2017; v15 n2 p187-196. DOI: 107346/—fei-XV-02-17_17

[16] Mui P, Goudbeek M, Swerts M, Hovasapian A. Children's Nonverbal Displays of Winning and Losing: Effects of Social and Cultural Contexts on Smiles. Journal of Nonverbal Behavior. 2016; 41(1), 67-82. DOI: https://doi.org/10.1007/s10919-016-0241-0

[17] Cartmill E A, Goldin-Meadow S. Gesture. APA handbooks in psychology®. APA handbook of nonverbal communication. 2016; 307-333. DOI: https://doi.org/10.1037/14669-012

[18] Matsumoto D, Hwang H C.The cultural bases of nonverbal communication. APA handbooks in psychology®. APA handbook of nonverbal communication 2016; 77-101. DOI: https://doi.org/10.1037/14669-004

[19] Balboni P E. Parole comuni, Culture diverse. Guida alla comunicazione interculturale. Venezia: Marsilio. 1999. ISBN: 8831771280

[20] Dalla Vecchia M. Linguaggi verbali e non verbali. EL.LE Educazione Linguistica. Edizioni Ca'Foscari. 2016; v5 n1 marzo p43-52. DOI: 10.14277/2280-6792/ELLE-5-1-16-2

[21] Toro M B. Parliamo di... vestemica [Internet]. 2015. Available from: https://mariabeatricetoro.com/2015/05/10/parliamo-di-vestemica/ [Accessed: 2015-05-10]

[22] D'Angiò G. La comunicazione efficace con i bambini [Internet]. 2009. Available from: http://www.lascuolapossibile.it/articolo/la-comunicazione-efficace-con-i-bambini/ [Accessed: 2009-01-05]

[23] Eskritt M, Lee K. Do actions speak louder than words? Preschool children's use of the verbal-nonverbal consistency principle during inconsistent communications. Journal of Nonverbal Behavior. 2003; 27, 25-41. DOI: https://doi.org/10.1023/A:1023614123081

[24] Boyatzis C, Satyaprasad C. Children's facial and gestural decoding and encoding: relations between skills and with popularity. Journal of Nonverbal Behavior. 1994; 18(1), 37-55. DOI: https://doi.org/10.1007/BF02169078

[25] Cameron H, Xu X. Representational Gesture, Pointing Gesture, and Memory Recall of Preschool Children. Journal of Nonverbal Behavior. 2011; 35(2), 155-171. DOI: https://doi.org/10.1007/s10919-010-0101-2

[26] Meyer M, Gelman S. Generic Reference is Less Marked Than Specific Reference in Children's Gestures. Journal of Nonverbal Behavior. 2015; 40, 65-79 (2016). DOI: https://doi.org/10.1007/s10919-015-0220-x

[27] Comitato Scientifico Nazionale (del MIUR). Indicazioni nazionali per il curricolo della scuola dell'infanzia e del primo ciclo di istruzione [Internet]. 2012. Available from: https://www.miur.gov.it/documents/20182/0/Indicazioni+nazionali+e+nuovi+scenari/

[28] Fantechi C J. Comunicare a scuola [Internet]. 2018. Available from: https://fisicaperlascuola.wordpress.com/2018/11/05/comunicare/ [Accessed: 2018-11-05]

[29] Stocco J. "Noi riusciamo a capirci!" Il ruolo della comunicazione non verbale fra bambini parlanti idiomi diversi [thesis]. Locarno: Scuola universitaria professionale della Svizzera italiana-Dipartimento formazione e apprendimento; 2011.

Chapter 5

Perceptual Attributes of Human-Like Animal Stickers as Nonverbal Cues Encoding Social Expressions in Virtual Communication

Xiaoming Jiang

Abstract

Communicative expression is a cross-species phenomenon. We investigated the perceptual attributes of social expressions encoded in human-like animal stickers commonly used as nonverbal communicative tools on social media (e.g. WeChat). One hundred and twenty animal stickers which varied in 12 categories of social expressions (serving pragmatic or emotional functions), 5 animal kinds (cats, dogs, ducks, rabbits, pigs) and 2 presented forms (real animal vs. cartoon animal) were presented to social media users, who were asked to rate on the human likeness, the cuteness, the expressiveness and the matchness of each intended expression against the given label. The data shows that the kind of animal that is expected to best encode a certain expression is modulated by its presented forms. The "cuteness" stereotype towards a certain kind of animal is sometimes violated as a function of the presented forms. Moreover, user's gender, interpersonal sensitivity and attitudes towards the ethic use of animals modulated various perceptual attributes. These findings highlight the factors underlying the decoding of social meanings in human-like *animal* stickers as nonverbal cues in virtual communication.

Keywords: Nonverbal communication, Virtual communication, Animal, Anthropomorphism, Stickers, Interpersonal reactivity index, Animal Attitude Scale

1. Introduction

Nonverbal communication is essential in online social interaction. Many interpersonal communications are carried out in virtual scenarios on the internet in the digital age. The understanding and expression of subtle variations in communicative meanings are significant for the users of online social media. WeChat is one of the most commonly used social software in China and has been prevalent among many digital chatting platforms all over the world. One salient benefit for the Wechat user is that people can communicate quickly and conveniently with a large variety of nonverbal expressions, in particular, through virtual symbols such as stickers. Stickers are complex images that have typically a larger size than emoticon and emoji, often with an animated virtual character, and has also considered

equivalent to "biaoqing" (the Chinese words for indicators of social expressions) [1]. One can use stickers to express our thoughts, emotions, intentions and stances on WeChat. Using stickers can complement the lack of human nonverbal cues (e.g. human facial and vocal expressions) in communications under virtual environments such as online social media [2].

The character of the sticker can be a person, but is often a non-human character with anthropomorphic (or human-like) features (which can be animals or other human-created cartoon characters). It is intriguing how communicators evaluate interpersonal meanings delivered by the sticker users and the associated personal characteristics through the anthropomorphic animal stickers.

As a typical way to create a virtual character, anthropomorphism is a process of assigning human identity and human personality to non-human objects [3] and has extended from religious divine targets to animals, naturally occuring objects, and robots [4]. Anthropomorphism makes the target more vivid, affective and human like, which is an effective communication tool [5]. The anthropomorphic image is more attractive because we are familiar with human-like animals or other targets [6]. Therefore, anthropomorphic stickers are well seen in social media and is often embedded in a verbal message or used on its own.

Evidence has already shown that stickers can ease interpersonal interactions in virtual environments, complement textual information exchange, and facilitate the expression of emotions and communicative meanings [2]. Adding anthropomorphic features to stickers can further enhance their ability to represent the speaker's emotions, attitudes and thoughts, and serve as an effective cue for the perceivers to understand the speaker's message, given that humans have the tendency to understand the non-human character with their knowledge of how to communicate to humans [7].

In addition to the inclusion of anthropomorphism, categories of non-human objects may impact the human perception of stickers. In the WeChat sticker store, most stickers are animals or animal cartoons. Psychological studies have revealed the emotional bondings between the human and the animals, and showed that humans displayed empathy and attitudes towards animals [8, 9]. The relationship between human perceivers and stickers with animal characters is therefore worthy of attention. Humans hold different stereotypical opinions towards different kinds of animals which is often culturally-specific [10, 11]. For example, dogs and pigs are typically associated with a positive stereotype of commitment and faithfulness and rabbits are typically associated with a negative stereotype of timidness in the Chinese culture. Dogs were perceived cuter than cats and humans and attracted different frequency and length of lookings than humans [12]. In WeChat, stickers have different representations of animal characters and can be categorized in cartoons or in real forms. It is shown that the contextual reality of a picture in different presented forms affected the perceptual outcome and its neural responses that were related with empathy, with the pictures showing a stronger activity than the cartoon ones [13]. The facial expressions of different presented forms altered the perceiver's neural responses, with the real expression eliciting a response in the relatively late stage of evaluating the emotional value of the stimuli whereas the cartoon expression eliciting a stronger response indexing the face recognition [14]. These findings are not sufficient in informing us about how a cartoon or a real form of animal could modulate the associated evaluative outcomes when they are the targets of the stickers.

Moreover, the expressions encoded by the characters may affect human perception of stickers. Studies on Weibo have documented the capacity of encoding emotional and social expressions in stickers and online tools to perform sentiment analysis such as MoodLens to decode emotions conveyed in the emoji [15]. The perceived cuteness of anthropomorphic animal stickers can be affected by different

expressions which may in turn affecting the human likeness [16]. Increased tendency to decode higher-level cognitive states has been found in animals perceived to be closer to humans (e.g. pets such as cats and dogs) [17]. The universal mechanism underlying the encoding of expression across species suggests an ability for animal stickers to express different social meanings [18].

The present study aims to investigate the factors that affect the human perception of various attributes (i.e. human likeness, expressiveness, cuteness and matchness of the stickers towards the intended expressions) of social expressions of anthropomorphic animal stickers. We assessed the impact of different types of anthropomorphism on human perception of different attributes towards the animal stickers, by presenting stickers with cartoons or real animals as the characters. We are also interested in how different targets of anthropomorphism (cartoon vs. real animals) interact with the categories of social expressions and animal kinds in attribute perception, and whether these perceptions are sensitive to the individual differences. Based on previous literatures on real-world communication, we predict that individual attitudes towards animals [19] and their reactivity towards social contexts [20] may also affect how they evaluated anthropomorphic animal stickers.

2. Methods

2.1 Participants

Twenty-one students who aged 18-35 years old were recruited via WeChat advertisements and the campus forum from Tongji University in China (Gender: 11 females; Age: 21.3 ± 2.12 years). They comprise a wide variety of academic majors and none reported to have suffered from any psychotic or neurological disorders. All participants were given informed consent before the study and each was compensated 30 RMB for their participation. The study was approved by the Ethics Committee of Tongji University.

2.2 Material

One hundred and twenty animal stickers were selected from the "Store of Expression" on the WeChat which provides many sticker packages for people to use in chat. A sticker package normally consists in stickers of different expressions and features of one animal character. Typically, the intended expression of each sticker was labeled a key word in each package. For example, the sticker package called *mitaomao* by *bujuexiaoxiao* has 24 stickers with different expression, and its character is a cat. The expression types were the most commonly used on WeChat and twelve expressions were pre-selected by the authors after they screened 59 animal sticker packages. These social expressions represented typical communicative meanings in interpersonal interaction each of which served a certain pragmatic function (i.e. agreement, commitment, refusal, having fun, greeting, gratitude), or emotional function (i.e. shyness, fear, happiness, sadness, anger, and grievance). These expressions were either labeled by the creator of the sticker and validated by the authors. Each expression was further represented by five animal kinds (cat, dog, pig, duck, rabbit) in two forms (cartoon, real). These five animal kinds were selected as the most common kinds in stickers. All stickers were adjusted into the same size (250 × 250 pixels). All was in static form and contained verbal messages on the stickers (see **Figure 1** for the demonstration of a set of stickers to express agreement). Verbal messages were all Chinese words except for the stickers of a greeting expression.

Figure 1.
Exemplar stickers that express intended agreement by five kinds of animals in both cartoon and real forms.

2.3 Procedure

All participants completed the attribute evaluation questionnaire, followed by two psychological scales, in a sequential order. All tests were completed online via the platform of the Chinese Wenjuanxing (https://www.wjx.cn/) in a quiet computer room in Tongji University. Before they started to rate stickers, they were asked to report the length of using social media per day, the frequency of using WeChat per day, the frequency of using emotive stickers per day and whether they used animal stickers.

2.3.1 Evaluating attributes of animal stickers

Participants evaluated four attributes on each of the 120 animal stickers on 7-point Likert scales. For each sticker, they were firstly given the sticker image

together with the label of its intended expression, and were asked the degree to which the label matched what is shown in the sticker from "1- to the least degree" to "7-to the most degree". This rating was followed by three sequential ratings on how human-like the animal character is on the sticker, how cute the animal character is on the sticker, and how expressive the sticker is. The stickers were randomized and the four evaluation questions were presented in the same order per sticker. At the end of the evaluation, all stickers of cats and dogs in the cartoon forms were presented to participants who were asked to identify which animal category the character belonged to. The identification task aimed to ensure the participants can recognize the correct animal category without ambiguity. Overall, average hit rates were 97.62% ± 6.53% (for dogs) and 96.83% ± 12.77% (for cats), suggesting that the accuracy of disambiguating between dogs and cats displayed in cartoon forms was high.

2.3.2 Evaluating individual differences

At a subsequent session, each participant received two scales which aimed to evaluate their attitudes towards animals and sensitivity towards the interpersonal relationships: the Animal Attitude Scale (AAS-10; [19]) and the Interpersonal Reactivity Index (IRI; [21, 22]). The AAS aims to measure the attitudes towards the ethic use of nonhuman species and consists in 10 items. Participants had to choose an answer on a 5 point scale from strongly agree to strongly disagree. The IRI consists in four subscales, each measuring a distinct component of empathy: (1) Fantasy (FS), which measures the tendency to imagine oneself to be the characters in books, film etc.; (2) Perspective Taking (PT), which assesses the tendency to cognitively take the perspective of another; (3) Empathic Concern (EC) which indicates the feeling of emotional concern for others; (4) Personal Distress (PD), which quantifies negative feelings in response to the distress of others [21]. The IRI has 28 items and the participants had to choose an option whether the statement is consistent with their opinion on a 5-point scale. Six scores were obtained for each participant, including the total scores for the AAS and the IRI and the scores for the sub-scales of IRI. The internal consistencies were calculated for each scale through the cronbach's alpha. The consistency was high for the AAS (=.73) and for the IRI (=.70). The consistency were from medium to high for the subscales of IRI (.59 for FS, .76 for PT, 42 for EC, and .56 for PD).

2.4 Data analysis

2.4.1 Perceptual attributes judgments

The repeated measures ANOVAs were performed on each of the judgments (Degree of Match, Human likeness, Cuteness, Expressiveness). The models treated these measures as dependent variables each for one model, and included Expression, Animal Kind and Form as fixed factors. Follow-up analyses were planned whenever an interaction was significant. Statistic values from pairwise comparisons were corrected for multiple comparisons using the Bonferonni method.

2.4.2 Individual differences: correlations and impacts on perceptual attributes judgments

To assess the relations between individual difference measures, Pearson correlations were conducted on participants' age, the score of AAS, IRI and its sub-scales.

To assess the relations between measures on one's social media usage, different categorical measures were first transformed into the ranking scale *from short to long duration* (4 levels) for the Length of Using Social Media Per Day (Length), *from less to more frequent* (4 levels) for the Frequency of Using WeChat Per Day (Frequency of Wechat Usage) and for the Frequency of Using Emotive Stickers Per Day (Frequency of Sticker Usage). Spearman correlations were applied to these ranking scores.

To assess the impacts of individual differences on social attribute judgments, new ANOVAs based on what was described in 2.4.1 were performed with one individual difference measure included as an additional fixed factor and all other individual difference measures included as controlling factors. Besides measures of AAS and IRI, participant sex was also considered as the fixed factor. Follow-up analysis was planned whenever a significant interaction between these measures and Expression/Form/Kind was shown. Linear regression models were built to assess the effects of scale measures on each expression. The effects of categorical measures were analyzed with ANOVA. The analyses were performed in R 3.6.0 within the R studio 1.2.1335.

3. Result

3.1 Self-reports on social media usage

All reported to be the user of WeChat. On the length of social media use, 13 reported to use WeChat for 1-3 hours per day, 4 reported 3-6 hours per day, 3 reported 6 hours or more per day and 1 reported less than 1 hour per day. On the frequency of social media use, 6 reported to use WeChat every ten minutes, 10 reported every half an hour, 3 reported every hour and 2 reported every two to three hours.

All reported to have used emotive stickers, among whom 18 reported to have used animal stickers. On the frequency of using stickers, 4 reported to use stickers highly frequently, 9 reported often, and 8 reported occasionally.

3.2 Social attribute ratings

3.2.1 Match-ness

The ANOVAs revealed significant interactions (Expression x Animal: $F(44, 880) = 2.95$, $p < .0001$; Expression x Form: $F(11, 220) = 3.58$, $p < .0001$; Animal x Form: $F(4, 80) = 6.58$, $p < .0001$; Expression x Animal x Form: $F(44, 880) = 2.15$, $p < .0001$).

The stickers of real animals showed a higher matchness in dogs to express refusal[1] relative to cartoon animals. The stickers of cartoon animals showed higher matchness than real ones in dogs for fear, commitment, gratitude, and shyness[2]; in cats for happiness, fear, commitment and greeting[3]; in ducks for happiness, agreement, commitment, grievance, gratitude, greeting, shyness, fun and anger[4]; in

[1] refusal: $t = 2.21$, $p = .03$

[2] fear: $t = 3.47$, $p = .001$; commitment: $t = 3.77$, $p = .0005$; gratitude: $t = 3.22$, $p = .003$; shyness: $t = 2.12$, $p = .04$

[3] happiness: $t = 5.89$, $p < .0001$; fear: $t = 2.27$, $p = .03$; commitment: $t = 4.85$, $p < .0001$; greeting: $t = 2.83$, $p = .007$

[4] happiness: $t = 3.28$, $p = .002$; agreement: $t = 2.52$, $p = .02$; commitment: $t = 2.36$, $p = .02$; grievance: $t = 2.97$, $p = .005$; gratitude: $t = 2.18$, $p = .04$; greeting: $t = 2.90$, $p = .006$; shyness: $t = 2.80$, $p = .008$; fun: $t = 2.84$, $p = .007$; anger: $t = 2.77$, $p = .008$

rabbits for happiness, fear, agreement, commitment, grievance, gratitude, greeting, shyness, sadness, fun, and anger[5], and in pigs for commitment and anger[6].

For stickers with real animals, the matchness of the intended expression was highest for dogs than other animals to express refusal[7], happiness[8] and agreement[9]. The matchness was higher for cats than other animals to express grievance[10], fun[11], and anger[12]. The matchness was higher for ducks than rabbits when expressing fear[13]. Moreover, the matchness was highest for anger and grievance and was lowest for happiness in cats[14]. The matchness was highest for refusal and lowest for gratitude in dogs[15]. The matchness was higher for sadness than greeting and grievance in ducks[16].

For stickers with cartoon animals, the matchness was lower for rabbits than other animals to express refusal[17], higher for rabbits than pigs to express grievance[18], and lower for dogs than other animals to express sadness[19]. The matchness was higher for commitment than sadness in dogs[20], and was lowest to express refusal for rabbits[21].

3.2.2 Human likeness

The interactions of Expression x Form ($F(11, 220) = 1.97$, $p = .03$) and Animal x Form ($F(4, 80) = 4.15$, $p = .002$) were significant. Stickers with cartoon animals were judged as more human-like than those of cartoon animals for happiness, fear,

---

[5] happiness: $t = 5.40$, $p < .0001$; fear: $t = 2.56$, $p = .01$; agreement: $t = 4.49$, $p < .0001$); commitment: $t = 3.6$, $p = .001$); grievance: $t = 3.89$, $p = .0004$; gratitude: $t = 2.45$; $p = .02$; greeting: $t = 2.47$, $p = .02$; shyness: $t = 3.60$, $p = .0009$; sadness: $t = 2.57$, $p = .01$; fun: $t = 3.78$, $p = .0005$; anger: $t = 2.33$, $p = .02$

[6] commitment: $t = 2.69$, $p = .01$; anger: $t = 4.47$, $p < .0001$

[7] dog>cat: $t = 3.91$, $p = .002$; dog>duck: $t = 3.24$, $p = .01$; dog>pig: $t = 3.24$, $p = .01$; dog>rabbit: $t = 3.66$, $p = .004$.

[8] dog>cat: $t = 4.18$, $p = .001$; pig>cat: $t = 4.09$, $p = .001$; dog>duck: $t = 2.93$, $p = .03$; dog>rabbit: $t = 3.64$, $p = .004$; pig>duck: $t = 2.84$, $p = .04$; pig>rabbit: $t = 3.55$, $p = .01$.

[9] dog>rabbit: $t = 4.02$, $p = .001$.

[10] cat>duck: $t = 4.30$, $p = .001$; cat>pig: $t = 2.89$, $p = .04$; dog>duck: $t = 3.27$, $p = .013$

[11] cat>rabbit: $t = 2.94$, $p = .03$.

[12] cat>pig: $t = 4.67$, $p < .0001$; cat>rabbit: $t = 4.12$, $p = .001$; dog>pig: $t = 2.93$, $p = .03$

[13] duck>rabbit: $t = 3.21$, $p = .02$

[14] anger>agreement: $t = 3.86$, $p = .01$; grievance>agreement: $t = 3.57$, $p = .02$; anger>commitment: $t = 4.42$, $p = .01$; anger>fear: $t = 3.67$, $p = .02$; anger>happiness: $t = 5.74$, $p = .01$; anger>refusal: $t = 3.95$, $p = .01$; grievance>commitment: $t = 4.14$, $p = .01$; shyness>commitment: $t = 3.57$, $p = .02$; grievance>fear: $t = 3.39$, $p = .04$; fun>happiness: $t = 4.51$, $p = .01$; gratitude>happiness: $t = 3.48$, $p = .03$; greeting>happiness: $t = 3.39$, $p = .04$; grievance>happiness: $t = 5.45$, $p = .01$; grievance>refusal: $t = 3.67$, $p = .02$; sadness>happiness: $t = 4.42$, $p = .01$; shyness>happiness: $t = 4.90$, $p = .01$.

[15] refusal>commitment: $t = 3.56$, $p = .02$; refusal>fear: $t = 4.05$, $p = .01$; fun>gratitude: $t = 3.50$, $p = .03$; grievance>gratitude: $t = 3.68$, $p = .01$; happiness>gratitude: $t = 3.68$, $p = .01$; refusal>gratitude: $t = 5.43$, $p = .01$; sadness>gratitude: $t = 3.31$, $p = .05$; refusal>shyness: $t = 3.59$, $p = .02$.

[16] sadness>greeting: $t = 3.43$, $p = .03$; sadness>grievance: $t = 3.34$, $p = .05$

[17] dog > rabbit: $t = 3.44$, $p = .007$; duck > rabbit: $t = 3.90$, $p = .002$

[18] rabbit > pig: $t = 2.90$, $p = .04$

[19] duck>dog: $t = 2.92$, $p = .034$; rabbit > dog: $t = 2.92$, $p = 03$

[20] commitment>sadness: $t = 3.58$, $p = .02$

[21] agreement>refusal: $t = 3.90$, $p = .01$; anger>refusal: $t = 3.59$, $p = .02$; commitment>refusal: $t = 5.38$, $p = .01$; grievance>fear: $t = 3.69$, $p = .01$; fun>refusal: $t = 4.64$, $p = .01$; gratitude>refusal: $t = 4.43$, $p = .01$; grievance>greeting: $t = 3.38$, $p = .04$; grievance>refusal: $t = 6.43$, $p = .01$; happiness>refusal: $t = 5.59$, $p = .01$; sadness>refusal: $t = 5.38$, $p = .01$; shyness>refusal: $t = 5.70$, $p = .01$.

and commitment[22]. The increased human-like judgment of cartoon animals was also revealed for dogs, ducks and rabbits[23].

3.2.3 Cuteness

The ANOVAs revealed a significant effect of Animal ($F(4, 80) = 30.25$, $p < .001$), with stickers of dog being judged of higher cuteness than those of pig ($t = 2.90$, $p = .03$). The interactions were significant (Expression x Animal: $F(44, 880) = 2.56$, $p < .0001$; Expression x Animal x Form: $F(44, 880) = 1.41$, $p = .04$).

Stickers of cartoon animals were judged as cuter than those of real animals in cats[24], ducks[25] and rabbits[26].

For stickers of real animals, ducks were judged as cuter than other animals to express happiness[27], and commitment[28]. Cats were judged as cuter than other animals to express gratitude[29], greeting[30], and shyness[31]. Fear and refusal were judged as less cute than expressions of gratitude and shyness in cat[32].

For stickers of cartoon animals, cats were judged as cuter than other animals to express happiness[33], and sadness[34]. Pigs and ducks were judged as less cuter than other animals to express agreement[35], commitment[36], grievance[37], fun[38]. Refusal was judged as less cute than gratitude and happiness in cats[39]. Sadness was rated less cuter than other expressions in dog[40]. Commitment was judged as cuter than anger and refusal for rabbit[41].

---

[22] happiness: $t = 2.62$, $p = .009$; fear: $t = 2.32$, $p = .02$; commitment: $t = 2.39$, $p = .02$

[23] dog: $t = 2.62$, $p = .009$; duck: $t = 3.18$, $p = .002$; rabbit: $t = 3.47$, $p = .0006$

[24] happiness: $t = 2.59$, $p = .01$; commitment: $t = 2.29$, $p = .03$

[25] happiness: $t = 2.09$, $p = .04$; shyness: $t = 2.06$, $p = .05$

[26] agreement: $t = 3.30$, $p = .002$; commitment: $t = 5.84$, $p < .0001$; grievance: $t = 3.07$, $p = .004$; gratitude: $t = 2.39$, $p = .02$; fun: $t = 2.16$, $p = .04$

[27] cat>duck: $t = 2.91$, $p = .03$; dog>duck: $t = 3.18$, $p = .02$; pig>duck: $t = 4.24$, $p = .001$; rabbit>duck: $t = 3.27$, $p = .01$

[28] dog>duck: $t = 3.12$, $p = .02$

[29] cat>dog: $t = 4.37$, $p = .0003$; cat>duck: $t = 4.37$, $p = .0003$; cat>pig: $t = 2.86$, $p = .04$; cat>rabbit: $t = 3.12$, $p = .02$

[30] cat>duck: $t = 3.43$, $p = .008$; cat>pig: $t = 3.26$, $p = .01$

[31] cat>dog: $t = 3.69$, $p = .003$; cat>duck: $t = 3.85$, $p = .002$; cat>pig: $t = 4.36$, $p = .0003$

[32] gratitude>fear: $t = 3.58$, $p = .02$; shyness>fear: $t = 3.76$, $p = .01$; gratitude>refusal: $t = 4.68$, $p = .01$; greeting>refusal: $t = 3.39$, $p = .04$; shyness>refusal: $t = 4.86$, $p = .01$.

[33] cat>duck: $t = 2.90$, $p = .04$; cat>pig: $t = 2.90$, $p = .04$, gratitude: cat>dog: $t = 2.96$, $p = .02$; cat>duck: $t = 4.54$, $p = .001$; cat>pig: $t = 2.96$, $p = .03$; rabbit>duck: $t = 3.16$, $p = .02$

[34] cat>dog: $t = 3.70$, $p = .003$; cat>duck: $t = 2.96$, $p = .03$

[35] dog>pig: $t = 3.31$, $p = .01$; rabbit>pig: $t = 3.01$, $p = .03$

[36] cat>duck: $t = 3.91$, $p = .002$; dog>duck: $t = 3.62$, $p = .004$; pig>duck: $t = 2.84$, $p = .04$; rabbit>duck: $t = 5.47$, $p = .001$

[37] cat>duck: $t = 2.91$, $p = .04$; dog>duck: $t = 3.64$, $p = .004$; dog>pig: $t = 3.00$, $p = .03$; rabbit>duck: $t = 4.46$, $p = .0002$; rabbit>pig: $t = 3.82$, $p = .002$

[38] cat>duck: $t = 2.94$, $p = .03$; dog>duck: $t = 2.94$, $p = .03$

[39] gratitude>refusal: $t = 3.42$, $p = .03$; happiness>refusal: $t = 3.32$, $p = .05$

[40] agreement>sadness: $t = 4.69$, $p = .01$; commitment>sadness: $t = 3.37$, $p = .04$; fun>sadness: $t = 3.78$, $p = .01$; grievance>sadness: $t = 3.88$, $p = .01$

[41] commitment>anger: $t = 3.43$, $p = .03$; commitment>refusal: $t = 3.63$, $p = .02$

3.2.4 Expressiveness

The ANOVAs revealed significant interactions (Expression and Animal: $F(44, 880) = 1.89$, $p = .0004$; Animal x Form: $F(4, 80) = 7.83$, $p < .0001$; Expression x Animal x Form: $F(44, 880) = 1.53$, $p = .01$).

Stickers of real animals were judged as more expressive than those of cartoon animals in cats[42] and dogs[43]. Stickers of cartoon animals were judged as more expressive than those of real animals for rabbit[44], duck[45] and pig[46].

For stickers of real animals, dogs were judged as more expressive than other animals to express refusal[47] and happiness[48]. Cats were judged as more expressive than ducks to express grievance[49] and anger[50]. Moreover, anger and fun were judged as more expressive than other expressions in cats[51]. Refusal was judged more expressive than gratitude in dogs[52]. Shyness was judged more expressive than agreement in rabbits[53].

For stickers of cartoon animals, cats were judged as less expressive than other animals for happiness[54], dogs were judged as less expressive for sadness[55], and ducks were judged as less for greeting[56]. Grievance was judged as less expressive than other expressions in pigs[57].

3.3 Individual differences in social attributes ratings

3.3.1 Associations between individual difference measures

The Pearson correlation revealed significant positive associations between AAS and IRI total scores ($r = 0.47$, $p = 0.03$), between the scores of IRI and the subscales[58] and between the scores for FS and PD ($r = 0.63$, $p = 0.002$). Neither significant effects between Length, Frequency of WeChat Usage and Frequency of Sticker Usage, nor significant effects between these usage-related measures and ASS or IRI scores were shown ($ps > .1$).

---

[42] grievance: $t = 2.06$, $p = .05$

[43] refusal: $t = 3.36$, $p = .002$

[44] happiness: $t = 2.44$, $p = .02$; shyness: $t = 3.93$, $p = .0003$; sadness: $t = 2.34$, $p = .02$

[45] agreement: $t = 2.79$, $p = .008$; commitment: $t = 2.03$, $p = .05$; grievance: $t = 2.25$, $p = .03$; greeting: $t = 3.56$, $p = .001$); fun: $t = 2.55$, $p = .01$

[46] anger: $t = 2.37$, $p = .02$

[47] dog>cat: $t = 3.53$, $p = .006$; dog>pig: $t = 2.97$, $p = .03$; dog>rabbit: $t = 3.05$, $p = .02$

[48] dog>cat: $t = 2.84$, $p = .04$; dog>rabbit: $t = 3.19$, $p = .02$

[49] cat>duck: $t = 3.63$, $p = .004$

[50] cat>pig: $t = 3.35$, $p = .01$; cat>rabbit: $t = 3.44$, $p = .007$

[51] anger>gratitude: $t = 3.74$, $p = .01$; anger>happiness: $t = 3.74$, $p = .01$; anger>refusal: $t = 3.56$, $p = .02$; fun>gratitude: $t = 3.65$, $p = .02$; fun>happiness: $t = 3.65$, $p = .02$; fun>refusal: $t = 3.48$, $p = .03$

[52] refusal > gratitude: $t = 3.94$, $p = .01$

[53] shyness > agreement: $t = 3.43$, $p = .03$

[54] dog>cat: $t = 3.26$, $p = .01$; duck>cat: $t = 3.17$, $p = .02$

[55] duck>dog: $t = 3.07$, $p = .02$; pig>dog: $t = 2.98$, $p = .03$

[56] duck>dog: $t = 3.36$, $p = .01$; duck>rabbit: $t = 3.18$, $p = .02$

[57] fear>grievance: $t = 3.38$, $p = .04$; sadness>grievance: $t = 3.56$, $p = .02$

[58] IRI-FS: $r = 0.73$, $p = 0.0001$; IRI-PD: $r = 0.71$, $p = 0.0003$; IRI-EC: $r = 0.59$, $p = 0.005$; IRI-PT: $r = 0.43$, $p = 0.05$

3.3.2 Matchness

The interactions of Sex x Animal ($F(4, 80) = 3.01$, $p = .02$) and Sex x Form ($F(1, 20) = 69.31$, $p < .01$) were significant. The matchness score was higher for women than men and such gender difference was more pronounced in pigs than ducks[59], and was more pronounced in in stickers of real relative to cartoon animals[60].

The interaction of Length x Form ($F(3, 60) = 12.29$, $p < .01$) was significant. The matchness score was higher for individuals who used social media 3-6 hours than those who used 1-3 hours in stickers of cartoon animals ($t = 2.70$, $p = .03$).

The interactions Frequency of WeChat Use x Animal ($F(20, 400) = 2.78$, $p = .0009$) and Frequency of WeChat Use x Form ($F(5, 100) = 23.98$, $p < .01$) were significant. For dogs, the matchness was higher for individuals who used WeChat every 30 minutes than those who used that every 10 minutes ($t = 3.14$, $p = .009$). For cats and rabbits, the matchness was higher for those who used WeChat every 30 minutes than those who used that every 10 minutes and those who used WeChat every 1 hour[61]. For pigs, the score was higher for individuals who used WeChat every 30 minutes than those who used WeChat every 1 hour ($t = 2.79$, $p = .03$).

For stickers of real animals, the matchness was higher for those who used that every 1 hour than those who used WeChat every 30 minutes and those who used WeChat every 2-3 hours[62]. For stickers of cartoon animals, the matchness was higher for individuals who used WeChat every 30 minutes than those who used WeChat every 1 hour ($t = 2.61$, $p = .04$).

The interactions of Frequency of Stickers x Animal ($F(12, 240) = 1.97$, $p = .05$) and Frequency of Stickers x Form ($F(3, 60) = 26.04$, $p < .01$) were significant. For cats, the matchness score was higher for individuals who used stickers very frequently than those using stickers often ($t = 2.65$, $p = .02$). For pigs, the matchness was higher for those who used stickers very frequently than those who used stickers often and those who used stickers occasionally[63]. For stickers of cartoon animals, the matchness was higher for those who used stickers frequently than those using stickers often ($t = 3.08$, $p = .006$).

The interaction AAS x Animal was significant ($F(52, 1040) = 3.80$, $p = .004$). Individuals with higher AAS produced lower matchness for dogs and rabbits[64]. The interaction IRI x Expression was significant ($F(132, 2640) = 2.27$, $p = .009$). Individuals with higher IRI produced lower matchness for the expression of agreement, greeting and refusal[65]. The two-way interaction EC x Expression was significant ($F(99, 1980) = 2.36$, $p = .007$). Individuals with higher EC produced higher matchness for the expression of anger, agreement, commitment, gratitude, greeting, fun, fear, shyness and refusal[66].

---

[59] ducks: $t = 2.35$, $p = .02$; pigs $t = 2.99$, $p = .003$

[60] real: $t = 3.36$, $p = .0008$; cartoon animals: $t = 2.52$, $p = .01$

[61] every 30 minutes > every 10 minutes: cats: $t = 2.76$, $p = .03$; dogs: $t = 2.83$, $p = .02$; every 30 minutes > every 1 hour: cats: $t = 2.97$, $p = .02$; dogs: $t = 3.06$, $p = .01$

[62] every 1 hour > every 2-3 hours: $t = 2.81$, $p = .02$; every 1 hour > every 30 minutes: $t = 3.11$, $p = .01$

[63] very frequently > often: $t = 3.06$, $p = .007$; very frequently > occasionally: $t = 3.00$, $p = .008$

[64] dogs: $b = -0.16$, $t = -2.32$, $p = .02$; rabbits: $b = -0.28$, $t = -3.74$, $p = .0002$

[65] agreement: $b = -0.16$, $t = -2.91$, $p = .004$; greeting: $b = -0.15$, $t = -2.73$, $p = .007$; refusal: $b = -0.17$, $t = -2.82$, $p = .005$

[66] anger: $b = 0.30$, $t = 2.46$, $p = .01$, agreement: $b = 0.35$, $t = 2.83$, $p = .005$, commitment: $b = 0.41$, $t = 3.19$, $p = .002$; gratitude: $b = 0.38$, $t = 3.25$, $p = .001$, greeting: $b = 0.38$, $t = 3.11$, $p = .002$, fun $b = 0.30$, $t = 2.54$, $p = .01$; fear: $b = 0.28$, $t = 2.16$, $p = .03$; shyness: $b = 0.26$, $t = 2.26$, $p = .03$; refusal $b = 0.52$, $t = 3.85$, $p = .0002$

3.3.3 Human likeness

The interaction between Sex x Animal was significant ($F(4, 80) = 7.66$, $p < .01$). Men judged stickers more human like than women and such effect was more pronounced in cats than dogs[67].

The interactions Length x Animal ($F(12, 240) = 2.08$, $p = .002$) and Length x Form ($F(3, 60) = 5.80$, $p = .0006$) were significant. For cats, ducks and pigs[68], those who used social media for 3-6 hours produced a higher rating than those who used that for 1-3 hours. For stickers with real animals, those who used social media for 1 hour per day rated higher than those who used that for 1-3 hours ($t = 3.00$, $p = .01$). For stickers with cartoon animals, those who used social media for 3-6 hours rated higher than individuals who used it for 1-3 hours and those who used it for 6 hours[69].

The interactions of Frequency of Stickers Use x Animal ($F(12, 240) = 2.23$, $p = .02$) and Frequency of Stickers Use x Form ($F(3, 60) = 20.58$, $p < .01$) were significant. For pigs, those who used stickers highly frequently rated higher than those who used stickers often ($t = 2.55$, $p = .03$). For stickers of cartoon animals, those who used stickers often rated higher than those who used stickers highly frequently and those who used stickers occasionally[70].

The interaction IRI x Form was significant ($F(12, 240) = 4.05$, $p = .04$). For stickers of real animals, those who displayed higher IRI score produced lower human-likeness scores ($b = -0.11$, $t = -2.68$, $p = .007$). The interactions Form x FS ($F(10, 200) = 12.69$, $p = .0004$) and Animal x FS ($F(40, 800) = 3.26$, $p = .01$) were significant. For stickers of cartoon animals, those with higher FS score judged less human-like ($b = -0.16$, $t = -2.31$, $p = .02$). Those showing higher FS score judged less human-like on rabbits ($b = -0.19$, $t = -2.34$, $p = .02$). The interaction PT x Expression was significant ($F(154, 3080) = 1.83$, $p = .04$). Those with higher PT score judged the agreement expression to be less human like ($b = -0.30$, $t = -2.74$, $p = .007$).

3.3.4 Cuteness

The interaction of Sex x Form was significant ($F(1, 20) = 101.47$, $p < .01$). Women judged stickers of real animals to be cuter than men ($F(1, 20) = 78.95$, $p < .01$; $t = 2.09$, $p = .04$).

The interactions of Length x Animal ($F(12, 240) = 3.99$, $p < .01$) and Length x Form ($F(3, 60) = 8.04$, $p < .01$) were significant. For cats, individuals who used social media 3-6 hours rated cuter than those who used that 1-3 hours ($t = 2.64$, $p = .04$). For ducks, the cuteness was lower for those who used that for 1-3 hours than those who used it for 1 hour and those for 6 hours[71]. For pigs, the cuteness was higher for those who used social media for 6 hours than those using that for 1-3 hours ($t = 3.47$, $p = .003$). For stickers of real animals, the cuteness was higher for those who used social media for 1 hour than those for 1-3 hours ($t = 2.76$, $p = .03$). For stickers with cartoon animals, the cuteness was higher for those who used social media 6 hours than those who used that for 1-3 hours ($t = 3.08$, $p = .01$).

The interactions Frequency of Stickers Use x Animal ($F(12, 240) = 2.61$, $p = .008$) and Frequency of Stickers Use x Form ($F(3, 60) = 39.44$, $p < .01$) were

[67] cats: $t = 2.62$, $p = .009$; dogs: $t = 1.97$, $p = .05$

[68] cats: $t = 4.04$, $p < .001$; ducks: $t = 2.96$, $p = .02$; pigs: $t = 3.05$, $p = .01$

[69] 1-3 hours: $t = 6.11$, $p < .001$; 6 hours: $t = 2.70$, $p = .03$

[70] highly frequently: $t = 3.08$, $p = .006$; occasionally: $t = 2.95$, $p = .009$

[71] 1 hour per day: $t = 3.01$, $p = .01$; 6 hours: $t = 2.72$, $p = .03$

significant. For cats, the cuteness was higher for individuals who used stickers very frequently than those who used stickers often ($t = 2.55$, $p = .03$). For stickers of cartoon animals, the cuteness was higher for those who used stickers very frequently than those using them often ($t = 2.43$, $p = .04$).

The interaction of AAS x Animal was significant ($F(52, 1040) = 3.17$, $p = .01$). Those who showed a higher AAS judged dogs and rabbits[72] less cuter.

The interaction of IRI x Animal was significant ($F(48, 960) = 5.10$, $p = .0004$). Those with higher IRI judged ducks ($b = -0.16$, $t = -2.91$, $p = .004$) and pigs ($b = -0.13$, $t = -2.80$, $p = .005$) less cute than those with lower IRI. The interaction Animal x PD ($F(36,720) = 15.86$, $p < .01$) was significant. Those who showed higher PD produced a higher rating of cats, dogs, rabbits and pigs[73] than those with lower PD.

3.3.5 Expressiveness

The interaction Sex x Form was significant ($F(1, 20) = 34.20$, $p < .01$). Females produced higher expressiveness score as compared with males and such difference was more pronounced for stickers of cartoon animals than for those of real animals[74].

The interaction of Length x Animal ($F(12, 240) = 1.87$, $p = .03$) was significant. The expressiveness of cats was higher for those who used social media 6 hours than those who used 1 hour, 1-3 hours and 3-6 hours[75]. The expressiveness of dogs, ducks and rabbits[76] was higher for those who used social media 6 hours per day than those who used 1-3 hours. The expressiveness of pigs was higher for those who used social media 6 hours than those who used 1 hour and 1-3 hours, and was higher for those who used social media 3-6 hours than those who used 1-3 hours[77].

The interaction AAS x Animal was significant ($F(52, 1040) = 3.75$, $p = .005$). Those who demonstrated higher AAS revealed lower expressiveness for dogs, ducks and rabbits[78].

The interactions FS x Expression ($F(110,2\,200) = 1.90$, $p = .04$), FS x Animal ($F(40, 800) = 3.19$, $p = .01$) and FS x Form ($F(10, 200) = 6.24$, $p = .01$) were significant. Individuals with higher FS produced higher expressiveness ratings for anger, agreement, commitment, greeting, grievance, fun, shyness and refusal[79]. Those with higher FS produced higher expressiveness for cats, ducks and pigs[80]. Those with higher FS produced higher expressiveness and such effects were stronger for stickers of cartoon than for those of real animals[81]. The interaction of PD x Form ($F(9, 180) = 5.03$, $p = .03$) was significant. Individuals with higher PD produced lower expressiveness rating when cartoon and real animals were presented in stickers[82].

---

[72] dogs: $b = -0.17$, $t = -2.28$, $p = .02$; rabbits: $b = -0.15$, $t = -2.15$, $p = .03$

[73] cats: $b = 0.22$, $t = 2.22$, $p = .03$); dogs $b = 0.41$, $t = 4.00$, $p < .01$; rabbits: $b = .21$, $t = 2.10$, $p = .04$; pigs: $b = 0.26$, $t = 2.39$, $p = .02$

[74] cartoon animals: $t = 2.65$, $p = .008$; real animals: $t = 2.35$, $p = .02$

[75] 1 hour: $t = 3.51$, $p = .003$, 1-3 hours: $t = 5.40$, $p < .001$; 3-6 hours: $t = 3.45$, $p = .003$

[76] dogs: $t = 2.70$, $p = .03$; ducks: $t = 3.66$, $p = .001$; rabbits: $t = 3.71$, $p = .001$

[77] 1 hour: $t = 3.54$, $p = .002$; 1-3 hours: $t = 4.79$, $p < .001$; 1-3 hours: $t = 3.01$, $p = .01$

[78] dogs: $b = -0.29$, $t = -4.06$, $p < .01$; ducks: $b = -0.19$, $t = -2.61$, $p = .009$; rabbits: $b = -0.39$, $t = -5.39$, $p < .01$

[79] anger: $b = 0.24$, $t = 2.44$, $p = .02$; agreement: $b = 0.33$, $t = 3.42$, $p = .0008$; commitment: $b = 0.23$, $t = 2.15$, $p = .03$; greeting: $b = 0.26$, $t = 2.59$, $p = .01$; grievance: $b = 0.21$, $t = 2.22$, $p = .03$; fun: $b = 0.32$, $t = 3.60$, $p = .0004$; shyness: $b = 0.35$, $t = 3.60$, $p = .0004$; refusal: $b = 0.36$, $t = 3.58$, $p = .0005$

[80] cats: $b = 0.30$, $t = 3.85$, $p = .0001$; ducks: $b = 0.38$, $t = 4.29$, $p < .01$; pigs: $b = 0.28$, $t = 3.42$, $p = .0007$

[81] cartoon: $b = 0.25$, $t = 3.33$, $p = .0009$; real animals: $b = 0.32$, $t = 4.09$, $p < .01$

[82] cartoon: $b = 0.27$, $t = 3.04$, $p = .002$; real animals: $b = 0.49$, $t = 5.36$, $p < .01$

4. Discussion

The study mainly investigated the role of human-like animal stickers in encoding social expression. Four perceptual attributes (the matchness between the intended and the perceived expression, the human likeness, the cuteness and the expressiveness) were demonstrated to be modulated by the discrete expression types, by which animal served as the virtual target of the sticker, and by whether the animal was a real or a cartoon character.

4.1 Perceptual attributes of animal stickers

Animal kinds interplayed with forms of presentation in affecting the perceptual attributes of animal stickers. Rabbits, ducks and pigs are generally judged more expressive, more human like and cuter than others in cartoon forms. Cats and dogs are perceived more expressive in real forms. These exploratory findings based on a group of social media users suggest the expected expression stereotypically associated with certain animal kinds maybe affected by whether the animal is perceived as a real or a virtual character. The matchness rating reflects the degree a given label fits the intended expression and maybe associated with the most expected communicative expression encoded by an animal. The expressiveness rating reflects the perceived amounts of cues that are associated with the expression, and may be associated with the expected easiness of encoding certain expression by an animal. As is shown in the matchness and expressiveness, the cats are more expected to convey anger and grievance; the dogs are expected to convey refusal; and ducks are expected to convey sadness when they are presented in real animal forms. However, when they are presented in cartoon forms, dogs are expected to convey commitment and rabbits are expected to convey refusal. Besides, the expected "cuteness stereotype" is sometime violated as a function of forms of presentation. Rabbits are considered cuter when expressing anger and refusal in the cartoon forms. Cats are considered cuter in real animal forms when expressing shyness and in cartoon forms when expressing gratitude. Although not directly tested in the present study, it is possible the amount of anthropomorphic features (e.g. the perceived similarity to human based on physical likeness, familiarity, cultural stereotype as human like) may explain different expectations towards different kinds of animals presented in different forms [23]. Pending further research, these data draws a first sketch on how animals encode social expressions that serve different communicative functions in stickers.

4.2 Individual characteristics and evaluation of animal stickers

The second aim of the study was to explore the individual differences in the judging the perceptual attributes of the animal sticker. Consistent with previous studies showing a female advantage in recognizing social signals and inferring meanings from these signals [24–26], our data showed females perceived the intended expression of the animal stickers to match to a greater extent with the labels and cuter relative to males for real animals, but perceived stickers to be more expressive relative to males for cartoon animals. One exception is in human likeliness which demonstrated a male advantage. Despite a higher frequency of using Stickers to communicate (Female: 18% - occasionally; 55% - often; 27% - highly frequently; Male: 60% - occasionally; 30% - often; 10% - highly frequently), the female did not consider certain animals (cats and dogs) to be more human like. It is assumed that a certain motivation may underlie the use of less human-like animal stickers as a communicative strategy; nevertheless, this assumption needs to be

evaluated in future studies, given that this study did not include a set of person stickers to serve as a baseline.

Interestingly, showing a higher level of concern towards animals was associated with a reduced perception of matchness, cuteness, and expressiveness but not human likeness in animal stickers. The stickers showing such individual differences were mainly featured by dogs, rabbits and ducks. Increased cuteness perception was seen in animals bearing more baby features [12]. More inference of cognitive states is demanded on animals sharing more human characteristics [17]. However, neither these cognitive processes seem to well explain the current pattern given an opposite pattern would have be shown. However, this finding maybe associated with the emotional or empathic response triggered by those who showed higher sensitivity towards the ethic use of animals, given an increased intention of protection may lower the users' emotional responses (and their ability to recognize expressions from these stickers) but increase their empathic response [27].

Higher interpersonal sensitivity was associated with a reduced perception of matchness of certain types of intended expressions with the actual labels regardless of real or cartoon animals. The IRI was shown to modulate one's sensitivity towards nonverbal cues such as face [28], voice [29, 30] and body awareness [31], and the mechanisms underlying constructing pragmatic representations beyond literal expressions [32, 33]. These affected expressions marked certain pragmatic functions in human communication (e.g. agreement, refusal, greeting) and may reflect a mismatch between the expected and the actual nonverbal cues in the animal stickers, for example, an expression of greeting between two interlocutors is not commonly seen in animals. The unexpected use of pragmatic expressions can be generally observed in real animals which typically do not display human characteristics, and lead to a lower perceptual rating in human-likeness by those demonstrating higher sensitivity. On a similar note, the unexpected use of social expressions by certain animals less familiar to humans (e.g. pigs, ducks, rabbits) is often given a lower cuteness rating by those showing higher sensitivity.

Fantasy, the tendency to transpose oneself imaginatively to the feelings and actions of fictitious characters in books and movies was shown to be associated with the expressiveness of the nonverbals cues in the sticker. Increased fantasizing ability was shown to modulate the behavioral acceptability of an underspecified sentence which can make sense after one engages a pragmatic inference (e.g. Sentence: *even such a person*$_{[\text{underspecified}]}$ *can afford an expensive house.* Inference: That person is poor). The extent to which an inference is engaged can be systematically involved in the medial prefrontal cortex [33], a region critical to decoding nonverbal meaning in social communication [34]. Consistent with previous literature, our data shows the social expressions that mark pragmatic functions were modulated by fantasy; moreover, the more imaginative to the cartoon animals, the more expressive the stickers were perceived [34].

Empathic concern has been defined to assess one's other-oriented feelings of sympathy and concern for unfortunate others. Evidence suggests the sensitivity of such tendency with the perceptual accuracy of emotional cues in nonverbal communication, and we demonstrated the modulation of EC on the expression of anger, fear and shyness. Furthermore, the matchness of the expected pragmatic function to the actual label was also modulated by such individual difference. Emotional concern has been considered an essential part when interactants encode speech acts (here agreement, commitment, greeting and refusal; and also see [35]) and complex social emotions (Apology: [36]; Gratitude; [37]; Guilt: [38]). The decoding of certain nonverbal cues of emotional consequences in animal stickers may also require higher EC.

4.3 Experience in using social media and evaluation of animal stickers

The heavier use of social media per day and the more frequent use of emotive stickers generally enhanced the ratings of the perceptual attributes of animal stickers, such as human-likeness, cuteness and expressiveness, although different kinds of animals appeared to benefit from different amounts of social media use.

Importantly, a clearer dissociative pattern can be seen between stickers of real and cartoon animals. For the cartoon animals, the longer time the social media is used (and the more frequently emotive stickers is used), to a larger extent was the intended expression judged to match the actual label, more human like and cuter was the sticker. However, an opposite pattern was shown in the expression of real animals, with the heavier use of social media producing lower ratings. The added experience of using social media does not exert a unified impact on the perceptual attributes of the social expression in animal stickers. On one hand, such experience may enhance the general acceptability of imaginary characters (cartoon animals) to encode human expressions which results in positive evaluations towards these animals. On the other hand, the perceived unexpected use of human expression by a real animal gets more salient and may cognitively result in a conflict which requires monitoring and resolution (for example, to reason why a real-animal expresses a human-like expression; to think of a conversational context when such use can be accommodated). Further studies could be developed to see whether the experience of using social media is associated with the motivation of using an animal sticker [39].

5. Conclusion

We reported the perceptual attributes of a set of *animal* stickers varying in social expressions (of pragmatic and emotional communicative functions), animal kinds (of different levels of familiarity to humans) and presented forms (real vs. virtual). We also correlated the individual characteristics (interpersonal sensitivity and attitudes towards animals) and the experience in social media use with the attributes. Our data shows the expression that is expected to be best encoded by certain animals is modulated by its presented forms. The cuteness stereotype towards an animal is sometimes violated as a function of presented forms. Moreover, gender, dispositional empathy towards humans and concerns towards animals modulated the perceptual attributes of nonverbal cues in the social expression of animal stickers. These findings highlight the role of anthropomorphism for animal stickers to encode social meanings in nonverbal cues; and put forward a novel avenue of research on the effectiveness and the mechanisms of human-like *animal* stickers and the related forms in virtual communication.

Acknowledgements

Special thanks are given to Li Xin for her great work on data collection, analysis and an earlier version of this manuscript. This study was supported by the grants from Natural Science Foundation of China (31971037), Shanghai Planning Office of Philosophy and Social Sciences (2018BYY019), and the "Shuguang Program" supported by Shanghai Education Development Foundation and Shanghai Municipal Education Committee (20SG31).

Author details

Xiaoming Jiang
Institute of Linguistics, Shanghai International Studies University, Shanghai, China

*Address all correspondence to: xiaoming.jiang@shisu.edu.cn

References

[1] de Seta, G. (2018). Biaoqing: The circulation of emoticons, emoji, stickers, and custom images on Chinese digital media platforms. First Monday, 23(9).

[2] Tang Y., Hew K.F. (2018) Emoticon, Emoji, and Sticker Use in Computer-Mediated Communications: Understanding Its Communicative Function, Impact, User Behavior, and Motive. In: Deng L., Ma W., Fong C. (eds) New Media for Educational Change. Educational Communications and Technology Yearbook. Springer, Singapore.

[3] Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. Psychological Review, 114(4), 864-886.

[4] Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social cognition unbound: Insights into anthropomorphism and dehumanization. Current Directions in Psychological Science, 19, 58-62.

[5] Karlsson, F. Critical Anthropomorphism and Animal Ethics. J Agric Environ Ethics 25, 707-720 (2012).

[6] Nowak, K. L. (2017). "The Influence of Anthropomorphism and Agency on Social Judgment in Virtual Environments." Journal of Computer-Mediated Communication **9**(2).

[7] Xu, L., Yu, F. Wu, J., Han, T., Zhao, L. (2017). Anthropomorphisim: From "it" to him. (Nirenhua: Cong ta dao ta). Progress in Psychological Science, 25(11), 1942-1954. (in Chinese)

[8] Paul, E. S. (2000). "Empathy with Animals and with Humans: Are They Linked?" Anthrozoös **13**(4): 194-202.

[9] Signal, T. D. and N. Taylor (2007). "Attitude to Animals and Empathy: Comparing Animal Protection and General Community Samples." Anthrozoös **20**(2): 125-130.

[10] Hsieh, S. C. (2006). A corpus-based study on animal expressions in Mandarin Chinese and German. Journal of Pragmatics 38, 2206-2222.

[11] Wright, J., Smith, A., Daniel, K., Adkins, K. (2007). Dog breed stereotype and exposure to negative behavior: Effects on perceptions of adoptability. Journal of Applied Animal Welfare Science, 10, 255-265.

[12] Borgi M, Cogliati-Dezza I, Brelsford V, Meints K and Cirulli F (2014) Baby schema in human and animal faces induces cuteness perception and gaze allocation in children. *Front. Psychol.* **5**:411.

[13] Gu, X., Han, S. (2007). Attention and reality constraints on the neural processes of empathy of pain. NeuroImage, 36, 256-267.

[14] Zhao, J., Meng, Q., An, L., Wang, Y. (2019). An event-related potential comparison of facial expression processing between cartoon and real faces. PLoS ONE, 14, e0198868.

[15] Zhao, J., L. Dong, J. Wu and K. Xu (2012). MoodLens: an emoticon-based sentiment analysis system for chinese tweets. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. Beijing, China, ACM**:** 1528-1531.

[16] Meese, J. (2014). "It belongs to the Internet": Animal images, attribution norms and the politics of amateur media production. M/C Journal.

[17] Eddy, T., Gallup Jr., G., Povinelli, D. (1993). Attribution of cognitive states to animals: Anthropomorphism in

comparative perspective. Journal of Social Issues, 49, 87-101.

[18] Darwin, C. (1965). The expression of the emotions in man and animals. Chicago: University of Chicago Press. (Original work published 1872).

[19] Herzog, H., S. Grayson and D. McCord (2015). "Brief Measures of the Animal Attitude Scale." Anthrozoös **28**(1): 145-152.

[20] Pulos, S., J. Elison and R. Lennon (2004). "The hierarchical structure of the Interpersonal Reactivity Index." Social Behavior and Personality: an international journal **32**: 355-359.

[21] Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. JSAS Catalogue of Selected Documents in Psychology, 10, No. 85.

[22] Davis, M. H. (1994). Empathy: A social Psychological Approach. Wisconsin: Brown & Benchmark.

[23] Hamm, M., Mitchell, R. (1997). The interpretation of animal psychology: Anthropomorphism or behavior reading? Behavior, 134, 173-204.

[24] Jiang, X. & Pell, D. M. (2015). On how the brain decodes speaker's confidence. Cortex, 66, 9-34.

[25] Hampson, E., van Anders, S., Mullin, L. (2006). A female advantage in the recognition of emotional facial expressions: test of an evolutionary hypothesis. Evolution and Human Behavior, 27, 401-416.

[26] Schirmer, A., Lui, M., Maess, B., Escoffier, N., Chan, M., Penney, T. (2006). Task and sex modulate the brain responses to emotional incongruity in Asian listeners. Emotion, 6, 406-417.

[27] Gunnthorsdottir, A. (2015). Physical attractiveness of an animal species as a decision factor for its preservation. Anthrozoos, 14, 204-215.

[28] Choi, D., Nishimura, T., Motoi, M., Egashira, Y., Matsumoto, R., Watanuki, S. (2014). Effect of empathy trait on attention to various facial expressions: Evidence from N170 and late positive potentials (LPP). Journal of Physiological Anthropology, 33, Article Number 18.

[29] Jiang, X. & Pell, D. M. (2016). Feeling of another knowing: how "mixed messages" in speech are reconciled. Journal of Experimental Psychology: Human Perception and Performance, 42, 1412-1428.

[30] Jiang, X., Sanford, R., Pell, D. M. (2017). Neural systems for evaluating speaker (un)believability. Human Brain Mapping. 38, 3732-3749.

[31] Morganti, F., Rezzonico, R., Cheng, S., Price, C. (2020). Italian version of the scale of body connection: Validation and correlations with the interpersonal reactivity index. 51, Complementary Therapies in Medicine, 51, 102400.

[32] Jiang, X., Zhou, X. (2015). Who is respectful? Effects of social context and individual empathic ability on ambiguity resolution during utterance comprehension. Frontiers in Psychology, 6, Article 1588.

[33] Li, S., Jiang, X., Yu, H., Zhou, X. (2014). Cognitive empathy modulates the processing of pragmatic constraints during sentence comprehension. Social, Cognitive and Affective Neuroscience, 9, 1166-1174.

[34] Jiang, X. (2018). Prefrontal cortex: Role in communicating language during social interaction. Prefrontal Cortex, Dr. Ana Starcevic (Ed.), InTech.

[35] Rothermich, K., Giorio, C., Falkins, S., Leonard, L., Roberts, A. (2021). Nonliteral language processing across

the lifespan. Acta Psychologica, 212, 103213.

[36] Beyens U, Yu H, Han T, Zhang L and Zhou X (2015) The strength of a remorseful heart: psychological and neural basis of how apology emolliates reactive aggression and promotes forgiveness. *Front. Psychol.* 6:1611. doi: 10.3389/fpsyg.2015.01611

[37] Yu, H., Gao, X., Zhou, Y., Zhou, X. (2018). Decomposing gratitude: Representation and integration of cognitive antecedents of gratitude in the brain. Journal of Neuroscience, 23, 4886-4898.

[38] Yu, H., Duan, Y., Zhou, X. (2017). Guilt in the eyes: Eye movement and physiological evidence for guilt-induced social avoidance. Journal of Experimental Social Psychology, 71, 128-137.

[39] Liu S and Sun R (2020) To Express or to End? Personality Traits Are Associated With the Reasons and Patterns for Using Emojis and Stickers. *Front. Psychol.* 11:1076. doi: 10.3389/fpsyg.2020.01076

Section 4

Nonverbal Communication in Psychology

Chapter 6

Can Turn-Taking Highlight the Nature of Non-Verbal Behavior: A Case Study

Izidor Mlakar, Matej Rojc, Darinka Verdonik and Simona Majhenič

Abstract

The present research explores non-verbal behavior that accompanies the management of turns in naturally occurring conversations. To analyze turn management, we implemented the ISO 24617-2 multidimensional dialog act annotation scheme. The classification of the communicative intent of non-verbal behavior was performed with the annotation scheme for spontaneous authentic communication called the EVA annotation scheme. Both dialog acts and non-verbal communicative intent were observed according to their underlying nature and information exchange channel. Both concepts were divided into foreground and background expressions. We hypothesize that turn management dialog acts, being a background expression, co-occur with communication regulators, a class of non-verbal communicative intent, which are also of background nature. Our case analysis confirms this hypothesis. Furthermore, it reveals that another group of non-verbal communicative intent, the deictics, also often accompany turn management dialog acts. As deictics can be both foreground and background expressions, the premise that background non-verbal communicative intent is interlinked with background dialog acts is upheld. And when deictics were perceived as part of the foreground they co-occurred with foreground dialog acts. Therefore, dialog acts and non-verbal communicative intent share the same underlying nature, which implies a duality of the two concepts.

Keywords: non-verbal behavior, non-verbal communicative intent, multimodal analysis, background expressions, regulators, deictics, turn-taking, dialog acts, ISO 24617-2

1. Introduction

Turn-taking is an indispensable part of spontaneous and authentic human communication. Despite its significance, it is not always as obvious and straightforward as one might want it to be. Rather, it is sometimes conveyed by elusive and subtle cues. These cues can be of verbal or non-verbal nature, but, in successful communication, all of them can be picked up by the human observer. To facilitate effective natural communication between machines and humans, significant effort must be put towards understanding and recognizing the inter-dynamics and intent of non-verbal communication, of which turn-taking is also a part.

The theory of dialog acts offers one possible way to gain insight into the functionality of verbal and non-verbal expressions of communication. Dialog act (hereinafter DA) theory has its origins in speech act theory [1, 2]. But despite its name, DA theory is not merely a theoretical concept. As Bunt [3] emphasizes, its goal is to provide a computational model of language in actual use. According to Searle [2], a DA represents the meaning of an utterance at the level of illocutionary force, and hence, it constitutes the basic unit of linguistic communication.

There are numerous DA annotation schemes, some of which are more purpose-specific, such as the Verbmobil scheme, which is based on business appointment-scheduling dialogs [4], the TRAINS scheme, annotating dialogs about train freight management [5], or the Coconut annotation scheme, with dialogs about buying dinning or living room furniture [6], while the ISO 24617-2, the DIT++, the DAMSL and the Switchboard annotation schemes, for example, cover various topics and apply to a wider range of material. The Switchboard scheme was created for a corpus of various authentic, spontaneous telephone calls in the United States and defined 42 types of DAs [7]. The DAMSL scheme, moreover, filled the need for applying multiple tags to a single segment [8] and was the first multidimensional scheme [3]. The concept of dimensions is best described by the ISO 24617-2 annotation scheme, whereby it is defined as a "class of DAs with the same type of semantic content" ([9]: 2). In comparison to multidimensional schemes, one-dimensional schemes use several tags, which are, however, mutually exclusive. Multidimensional schemes are, therefore, more appropriate for the annotation of naturally occurring dialogs. Another example of a multidimensional scheme is the DIT++ annotation scheme, which is partly based on the DAMSL scheme. It distinguishes between general-purpose and dimension-specific functions, which together form a set of ten dimensions – the Task/Activity dimension, the Auto-Feedback, the Allo-Feedback, the Turn Management, the Time Management, the Contact Management, the Own Communication Management, the Partner Communication Management, the Discourse Structuring Management, and the Social Obligations Management dimensions [3]. Furthermore, the DIT++ is not limited to verbal communication only; it also considers non-verbal communication, such as head gestures and prosody. The ISO 24617-2 annotation scheme is partially based on the DIT++ taxonomy. As Bunt [10] elaborates, it was created as a consolidation of selected taxonomies with the aim of avoiding confusion among the several existing annotation schemes and their inconsistent terminology [9]. Moreover, in addition to its multidimensionality, the ISO scheme strives to be a domain-independent scheme. Regarding dimensions, it contains functionally the same dimensions as the DIT++ with the exemption of the Contact Management dimension, which is not included in the ISO 24617-2. Among these nine dimensions, the scheme specifies 57 different functions. Six of these functions pertain to the dimension of Turn Management, namely, the functions of accepting, taking, grabbing, assigning, releasing, and keeping a turn. The functions are relatively self-explanatory as long as we remember that the function is always carried out by the sender, i.e. the "dialogue participant who produces a dialog act" ([9]: 4). The functions of turn management are all dimension-specific, which means that they cannot be assigned to any other dimension. The scheme also acknowledges the need for subtle characteristics of utterances such as conditionality, modality, (un)certainty, stance, and sentiment, which Petukhova and Bunt [11] raised in their analysis of existing annotation schemes. As a solution, the ISO 24617-2 proposes function qualifiers that can be applied to a DA function. Following its predecessor, the DIT++, the ISO 24617-2 also considers non-verbal behavior in terms of DA annotation. Afterall, in its definition of DAs, the ISO 24617-2 does not discriminate between verbal and non-verbal behavior, since it defines DAs as "a semantic unit of communicative behaviour".

Hence, the ISO 24617-2 is well-suited for the annotation of multimodal material and was implemented in research of non-verbal behavior. Yoshino et at. [12] utilized the scheme to annotate information navigation and attentive listening dialogs to improve natural conversation modeling for caretakers that communicate with the elderly. Navaretta and Paggio [13] explore non-verbal behavior occurring when providing feedback among persons who just met, i. e. in highly spontaneous settings. The one-hour recordings, annotated with the tool Anvil, specifically analyze what kind of head movement or facial expressions accompany a certain subtype of the feedback dimension. Their classification of non-verbal behavior is based on the MUMIN scheme. Petukhova and Bunt [14] utilize almost an hour-long recording from the corpus AMI, which consists of project meetings. They analyze DAs according to the DIT++ and the ISO 24617-2 schemes together with co-occurring non-verbal behavior, which is classified according to the CoGest scheme. In their previous work, Petukhova and Bunt [15] annotate recordings from the AMI corpus according to the DIT scheme. Both the annotation of DAs and the annotation of non-verbal behavior is carried out with the DIT scheme, since, as they emphasize, non-verbal behavior helps us understand the true function of a DA. The pragmatical annotation of the multimodal corpus HuComTech [16], however, is not based on the ISO scheme, yet its main annotation units are very similar. They are referred to as communicative acts, which denote the function or purpose of an utterance (e.g., agreement, turn management, information). The annotation of non-verbal behavior, including facial expressions, eyebrow movement, head movement, touch motions, posture, or emotions, was performed manually and partially automatically with the tool Qannot.

Although there seems to be strong evidence to support the multimodal and multi-signal nature of the human-human interaction, for decades, spoken language understanding has first and foremost focused on speech a priori [17]. The classification of non-verbal behavior by Mlakar et al. [18], draws upon McNeill's [19] common growth point theory, according to which speech and gestures both stem from a common growth point of a concept and mutually influence one another, Pierce's [20] semiotics, that provides analysis of non-linguistic signs and symbols as the meaning of non-verbal behavior, Ekman and Friesen's [21] categories and coding of non-verbal behavior, and Birdwhistell's [22] insights into the importance of kinesics. Moreover, the classification by [18] utilizes the communication management theory [23, 24] and, therefore, also encompasses discourse functions to some extent. Mlakar et al. [18] refer to 'gestures' as behavior generated by moving body parts (i.e., head, hands/arms, face, and posture) performing a communicative purpose, i.e., containing a discourse function, as a non-verbal communication intent (hereinafter NCI). These non-verbal expressions represent the basis of cognitive capabilities and understanding [25]. Namely, although not bound by grammar, non-verbal expressions co-align with language structures and compensate for the less articulated verbal expression models, thus providing a certain degree of clarity of discourse [26]. The non-verbal behavior retains the semantics and at the same time helps in providing suggestive influences and serves for interactive purposes, even such as content expression of one's mental state, attitude, and social functions. The classification proposed by Mlakar et al. [18] positions the role/intent of non-verbal concepts into five main NCI classes of regulators or adapters, deictics or pointers, illustrators, symbols or emblems, and batons.

Cooperrider's [27] classification of gestures, on the other hand, concerns itself with the question of whether the gesture "communicates a critical part of a message" ([27]: 179) or not. He divides gestures into foreground and background gestures. Foreground gestures are those gestures of which we are aware when we perform them, such as a thumb up, whereas background gestures occur

unconsciously, automatically, such as nodding during a telephone call. Therefore, foreground gestures are also in the foreground of the interaction. Among their characteristics, he lists co-occurrence with demonstratives, absence of speech, and a significant effort in their production, i.e., gestures that are bigger and more precise. Contrary to them are background gestures. They are both smaller in size and precision and occur while the sender is speaking. Despite this clear division, Cooperrider [27] emphasizes that the line between foreground and background gestures is anything but straightforward, as some gestures can break the foreground-background barrier. He demonstrates this with pointing gestures, which are generally in the foreground, but when pointing to oneself, they occur in the background. Furthermore, even symbolic gestures can take the background if performed automatically and if they are void of their communicative message. On the other hand, beats occur only as background gestures. One can, therefore, roughly consider illustrators, symbols, and partially deictics as NCI occurring in the foreground, while regulators, beats, and partially deictics can be considered as NCI occurring in the background, while still bearing in mind that the dividing line can always be crossed.

Hence, Cooperrider [27] differentiates between gestures with a semantic or propositional content, i.e., a message that provides some kind of information, and those that are void of it. The same distinction can be made for DAs. There are DAs that primarily convey information that is indispensable for communication, such as the task dimension, and those DAs that primarily do not contain propositional content (hence, they contain metadiscursive content) yet are vital for successful natural communication, such as the turn and management dimensions. Nevertheless, we must apply the same caveat as the one in the background-foreground distinction for gestures, as some DAs can occur either in the foreground or the background. For example, the dimension of managing social obligations can generally be considered part of the foreground, such as the concept of greeting someone upon the first encounter. Still, if a social convention is performed routinely, unconsciously, and is deprived of its semantic content, such as thanking someone for the floor, such a DA can be considered as occurring in the background. The nine DA dimensions can, therefore, roughly be divided into those occurring in the foreground, such as the task and the social obligation management dimension, and those occurring in the background, such as the feedback dimensions, the time and the turn management dimensions, the discourse structuring dimension, and the own- and the partner communication management dimensions.

For successful communication, the message must be as clear as possible. An utterance with a mismatching underlying nature is potentially confusing. For example, to take a turn, which is a typical background DA, one sometimes begins one's utterance with "look". The NCI accompanying "look" is usually a subtle hand gesture (e.g., a referential deictic), completely void of meaning and therefore a background gesture. Whereas when one uses "look" in the propositional sense, one uses a pointing gesture; both the DA and the NCI are, in this case, of foreground nature. To use a pointing (foreground) gesture with the mentioned turn-taking (background) DA in the "look" example would therefore be confusing, steering the collocutor to search for an object in sight, which does not exist. Therefore, to ensure cohesion and for the communication to be more effective, it seems plausible that a non-propositional episodes should require a background DA as well as a background NCI.

In light of this foreground-background link between DAs and the NCI of gestures, we set out to explore whether the theory of DAs can help predict the nature of the NCI of the corresponding unit. Specifically, we hypothesize that turn management DAs correlate with background gestures. Therefore, we propose the following hypothesis:

Turn management DAs, as background expressions, will tend to co-occur with NCI of background nature. In particular, turn management DAs will co-occur primarily with communication regulators.

2. Data and methodology

In order to perform research into authentic non-verbal behavior during turn-taking, we utilized a 57-minute long video recording from the Corpus EVA [18]. Our annotation scheme, adapted from Mlakar et al. [28], outlined in **Figure 1**, was applied in the dataset to perform conversational analysis. For this research, dialog acts were added as a linguistic branch.

The main objective of the scheme is to identify inferred meanings of co-verbal expressions as a function of linguistic, paralinguistic, and social signals (e.g., where and when to gesture) on a symbolic level, and to identify the physical nature (e.g., articulation of body language) and use of the available "imaginary forms" (e.g., how to gesture, how to vocalize), i.e., the level of the interpretation of non-verbal forms. The first layer, in **Figure 1**, the symbolic interpretation, is the focus of this research. It is used to analyze the interpretation of the interplay between various conversational signals, that is, verbal and non-verbal (i.e., DAs, gestures, syntax, discourse markers) at a symbolic level. The second layer, the interpretation of form, is concerned with how information is expressed beyond language, through prosody and embodied expressions, as an abstract concept of a non-verbal conversational expression with a specific communicative intent, i. e. how it is physically realized. For example, the 'form' of a gesture or 'accentuation' of speech. Its primary goal is to provide a detailed description, the closest possible to the physical reality and the entity that will realize it (e.g., an embodied conversational agent). As already mentioned, in this chapter, however, we focus on the first layer. The layer which aims to find patterns and tendencies in how people communicate through joint use of language, prosody, gaze, gesture, facial expressions, and other articulation of the

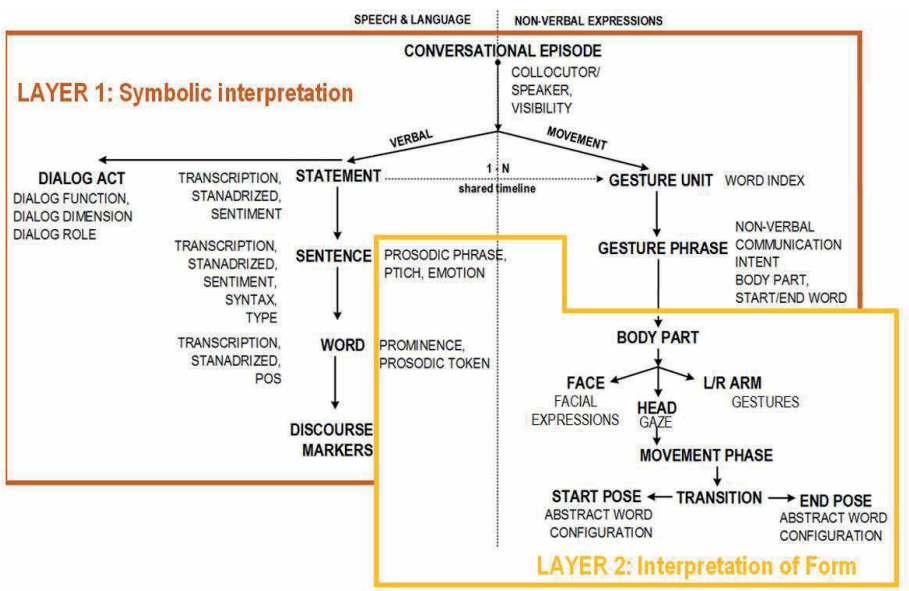


Figure 1.
The topology of annotation in the EVA Corpus: The levels of annotation describing verbal and non-verbal contexts of conversational episodes.

body, specifically focused on turn-taking and analysis of DAs and NCIs overlapping in conversational expressions (episodes).

2.1 The EVA Corpus

The EVA Corpus consists of 228 minutes in total, and includes four video and audio recordings, each 57 minutes long, with corresponding orthographic transcriptions. The discourse in all four recordings is a part of the entertaining evening TV talk show *'A si ti tut not padu'*, broadcast by the Slovene commercial TV in 2010. In this research, we utilize one of the videos.

In total, five different collocutors are engaged in each episode in multiparty discourse. The conversational setting is relaxed and unrestricted. The hosts are skilled interlocutors who engage in witty, humorous, and sarcastic dialog with the guest. Therefore, the discourse is highly spontaneous, authentic, and, in this case, since all the participants know each other privately, also relaxed and full of emotional responses. Overall, the video contains 1,516 utterances, with an average of 303 utterances per speaker. The episode contains 1,999 sentences, with an average of 399.8 per participant. The average sentence duration is 2.8 seconds, whereby the longest is 18.1 seconds, and the shortest is 0.19 seconds. Overall, there are 10,471 words in the episode, and on average, a speaker uttered 2,094 of them, with a mean value of 7.9 words per sentence. While the total length of the recording is just under one hour, the total duration of all utterances without overlapping is 1 hour 33 minutes and 26.3 seconds, which suggests a substantial amount of overlapping speech. Consequently, the dialog is characterized by a vivid and rapid exchange of speaker roles, which makes it ideal for the study of non-verbal behavior that accompanies turn-taking.

2.2 DA annotation

The entertainment show was segmented and transcribed with the transcription tool Transcriber 1.5.1 and annotated in the annotation tool ELAN. The annotation of DAs was performed with the web-based annotation tool Webanno. For the classification of DAs, we applied the ISO 24617-2 scheme, however, it was partially consolidated in accordance with our research's aim. In the dimension of information-providing functions, we specified the function Correction as it does not clarify whether the sender corrects themselves of the interlocutor. Therefore, we added the function CorrectionPartner, which denotes the action of the sender who is correcting the interlocutor. Among the functions Inform or Agreement, we also filled the need for argumentative acts and added the function Argument. For occasions where the sender quotes someone, the function ReportedSpeech was added. Among the directive functions, the Instruct function did not suffice for acts where the sender provides support to the interlocutor or when the sender warns the interlocutor. Therefore, the functions Encouragement and Warning were added. With regard to feedback-specific functions, we merged the AutoPositive and AutoNegative functions into the OwnComprehensionFeedback function. Similarly, we merged the alloPositive, and the AlloNegative functions into the PartnerComprehensionFeedback function. The dimension of discourse structuring provided the function of opening but lacked the closing action, which we added. As regards the dimension that manages social obligations, we merged the InitGreeing and the ReturnGreeting functions into Greeting. The dimension, however, lacked the function of providing and accepting praise or flattery, which is why the functions Praise and AcceptPraise were included. The annotation of sentiment included

the qualifiers Disappointment, Disgust, Emphasis, Hurt, Negative, Positive, Satisfaction, and Surprise.

In line with Cooperrider's [27] foreground-background distinction, we divided DAs according to whether they are conveying a vital part of the message without which the encounter would be void of propositional content or not. Since task-oriented DAs include the functions of information-seeking and -providing, as well as commissive and directive functions, they are part of the foreground. Similarly, the social obligations management DAs perform functions such as greetings, introductions, apologies, thanking, and valedictions. They contain propositional content and can, therefore, be considered part of the foreground. On the other hand, the feedback DAs, turn management DAs, time management DAs, discourse structuring DAs, and own- and partner communication management DAs perform background functions as their main purpose is not to convey information but to steer the dialog or to provide active listenership. For example, when correcting oneself after misspeaking, the act of correction is not in the foreground; it is the underlying information-related DA. Similarly, when helping the interlocutor to find the correct ending to a word, the act of completion is in the background, while the interlocutor's primary utterance that is being completed by the partner is in the foreground. As emphasized in the Introduction, some functions can cross this distinction. Let us consider an example with the function of completion. When people try to demonstrate their connection by finishing each other's sentences, the partner's act of completing the interlocutor's primary utterance is in the foreground, since both interlocutor's purpose of communication was to demonstrate their connection by completing each other's sentences. Nevertheless, for the majority of cases, the proposed distinction of DAs can be applied as proposed.

In terms of the background-foreground distribution of observed DA episodes, we can conclude that the material is well balanced. It consists of 1,897 instances where the primary role of the DA was recognized as of foreground nature, and 2,020 instances were the primary role of the DA was of background nature.

2.3 NCI annotation

The annotation of non-verbal expressions focusing on gestures, mimics, was carried out in Elan. The annotation of each phenomenon highlighted in **Figure 1** (e.g., gesture unit, phrase, NCI) was conducted individually, but by two or three annotators at a time. In terms of annotation disagreement, diverging values were elaborated and argued until consensus was reached. Moreover, before the annotation process began, all annotators were familiarized with the nature of the signal to be annotated and notified with the possible values from which they could choose.

In terms of NCI annotation, we used the following classification:

- **Illustrators (I)** define body movement (embodiment) that illustrates what a speaker is saying. Regarded as foreground behavior, they accompany or reinforce verbal cues and are accompanied by an actual word referent in the speech. Illustrators are further classified into outlines, ideographs, and dimensional illustrators. The outlines (I_O) subclass encompasses embodiments that reproduce a concrete aspect of the accompanying verbal content (explicit referents in speech). The ideographic/metaphoric illustrators (I_i) subclass refers to a concretization of the abstract through a specific shape. The spatial/dimensional (I_d) subclass refers to the spatial movements outlining or depicting dimensional relations. They are used to 'paint' characteristics of entities and actions to further highlight their physical properties.

- **Regulators/adaptors (R)** define embodiments that are primarily used to model the flow of information exchange. Adaptors are regarded as part of background behavior and can be produced even without speech. They exist without a specific speech reference and do not link with a specific speech structure. The regulators are further classified into self-adaptors (R_S), the communication regulators (R_C) subclass, the affect-regulators (R_A) subclass, the manipulators (R_M) subclass, and the social function and obligation regulators (R_O) subclass. Self-adaptors relate to how a speaker continuously manages the planning and the execution of the speaker's own communication. The communication-regulators refer to managing interactions with other interlocutors through systems of turn-taking, feedback, and sequencing, e.g., interactive communication management (ICM). The affect-regulators are either self- or person-addressed and are used to further emphasize or express attitude or emotion regarding a topic, object, or person. Manipulators convey relief or release of emotional tension or outline states of the body or mind, such as anxiety, uncertainty, or nervousness. Finally, social function and obligation regulation primarily deals with embodied behavior used in social settings, such as greetings, goodbyes, introductions.

- **Deictics (D)** include entities that can actually be present in the real environment of the gesturer (e.g., indicating objects, persons, or places) or are ideally present in the discourse content or abstract (e.g., pointing upwards or pointing backward to indicate the past). If deictic expressions are actual word referents with a semantic interlink, they are regarded as part of the foreground. If the semantic link does not exist or is weak, deictic expressions will also be recognized as part of the background. We further distinguish between pointers (D_P), indexes/referential pointers (D_R), and enumerators (D_E).

- **Symbols/emblems (S)** tend to establish a strong semantic link with verbal counterparts. They are regarded as foreground and include all symbolic gestures and symbolic grammars. Their specific meaning is often cultural-specific, as the same emblem can have different meanings in different cultures. Nevertheless, there are cross-cultural hand emblems, which are easily recognizable because, despite their arbitrary link with the speech they refer to, they have a direct verbal translation, which usually consists of one or two words or a whole sentence (often a traditional expression shared in a specific culture).

- **Batons (B)** are those staccato strikes that create emphasis and grab attention, such as a short and single baton that marks an important point in a conversation. Whereas repeated batons can "hammer" a critical concept. Batons are equivalent to beats, however, beats may appear as a more random movement (e.g., outlining rhythm). Batons, on the other hand, may also set the rhythm and signal importance but, more importantly, they also outline the structure of verbal counterparts, e.g., tag a set of words that should be processed together (e.g., to produce a summary of the meaning of an utterance).

In terms of the background-foreground distribution of observed NCIs, we can observe that the material contains predominantly non-verbal behavior "functioning" in the background. Overall, we have observed roughly 1,684 non-verbal expressions, out of which 1,274 belonged to regulators (75.65 percent) and 136 (8.08 percent) to illustrators and symbols. The rest, 275 (16.33 percent), belonged to deictic expressions. The majority of NCI is, therefore, of background nature.

A rough classification of NCIs and DAs according to their underlying nature, which can be of background and/or foreground nature is represented in **Table 1**. It must be emphasized that this classification is purely provisional, as the foreground-background barrier is vague and can, depending on the wider context, be crossed by both NCIs and DAs.

2.4 Annotation agreement

In total, five annotators, two with a linguistic background, and three with a technical background in machine interaction were involved in this phase of annotations. Annotations were performed in separate sessions, each session describing a specific signal. The annotation was performed in pairs, i.e., two or three annotators annotated the same signal. After the annotation, consensus was reached by observing and commenting on the values where the was no or little annotation agreement among multiple annotators (including those not involved in the annotation of the signal). The final corpus was generated after all disagreements were resolved. Procedures for checking inconsistencies were finally applied by an expert annotator.

Before starting with each session, the annotators were given an introductory presentation defining the nature of the signal they were observing and the exact meaning of the finite set of values they could use. An experiment measuring agreement was also performed. It included an introductory annotation session in which the preliminary inconsistencies were resolved. Overall, given the complexity of the task and the fact that the values in **Table 2** also cover cases with a possible duality of meaning, the level of agreement is acceptable and comparable to other multimodal corpus annotation tasks [29].

For the less complex signals, influenced primarily by a single modality (e.g., pitch, gesture unit, gesture phrase, body-part/modality, sentence type), the annotators' agreement measured in terms of Cohen's kappa [30] was high, namely, between 0.75 and 0.9 on the Kappa score. The signals such as, Part-of-Speech, Syntax, Word Segmentation, were annotated (semi)automatically and the two expert annotators (linguists) overviewed the process and corrected the tags manually. The agreement was measured over the agreement on the corrections made. Pitch was annotated completely automatically, no agreement was measured. The only exceptions between less complex, unimodal signals, were Gesture phrase (0.53) and Prosodic phrases (0.71). The disagreements were expected since in some cases it is quite ambiguous to identify where a certain phrase ends and the next stars. Moreover, in a lot of cases, a retraction phase of a gesture can be recognized as stroke phase of the next gesture phrase.

As summarized in Table 3, for the more complex signals that involve multiple modalities for their comprehension (including speech, gestures, and text) the disagreements in interpretation were expectedly higher.

	Background nature	**Foreground nature**
NCIs	Regulators, Batons, Deictics	Illustrators, Symbols, Deictics
DAs	Turn management, Social obligations management, Time management, Discourse structuring, Feedback, Communication management	Task, Social obligations management

Table 1.
A coarse-grained classification of the underlying nature of NCI classes and DA dimensions.

Signal	Kappa score
Word Segmentation (semi-automatic)	0.95
Part-of-Speech (semi-automatic)	0.81
Pitch (automatic)	/
Syntax (semi-automatic)	0.79
Sentence type	0.97
Gesture unit	0.82
Gesture phrase	0.53
Modality	0.88
Prosodic phrases	0.71
Sentiment	0.67
Dialog function	0.64
Dialog dimension	0.71
Intent (semiotic class)	0.48
Emotion label	0.51
Gesture unit	0.75
Movement phase	0.66

Table 2.
Results of the preliminary inter-coder agreement experiment.

3. Case study

Example 1: DAs as part of background and foreground conversational expressions.

Guest: Ampak se izkaže, da ta zdravnik ne zna nič drugega delat kot vedno iste in samo iste (A) obraze in so vsi poklonirani (B) – no to je (1) to. Fajn, ne (2)?
Co-host: (C) Samo v bistvu, a veš, v bistvu sej če pri nas gledaš sj so tud pol vsi glih.

Guest: But it turns out that this doctor can create only one and the same (A) face and nothing else and that they are all cloned (B) – well this is it (1)[1]*. Great, huh (2)?*
Co-host: (C) But actually, you know, actually if you took a look at where we are then they are also all the same.

This segment represents a case of sudden turn release by the main guest. Previously, the participants were discussing the effects of aging, during which several sarcastic comments were uttered. The show's host afterwards tries to transition to the next topic, which is the play the guest was directing, called *The Ugly One*. However, the guest is offended by the co-host's snide remark, where he compares the name of the play and the guest, suggesting that the guest might also be an ugly one. Nevertheless, after being asked to tell the audience about the play, he briefly outlined the plot, which deals with cosmetic surgery in connection with the feeling of self-worth and success. He is still mid-sentence and speaking with a rising intonation (see **Figure 2**: B, *in so vsi poklonirani* "and they are all cloned") when he suddenly takes a deep breath and decides to stop summarizing the play with the words *no to je to* "well this is it". Additionally, he emphasizes that he no longer wishes to talk about the topic, as he adds *fajn, ne*? "great, huh?". With it, he simultaneously elicits feedback, which is yet another way to assign his turn to someone else.

[1] The literal translation of the utterance is »this is this«.

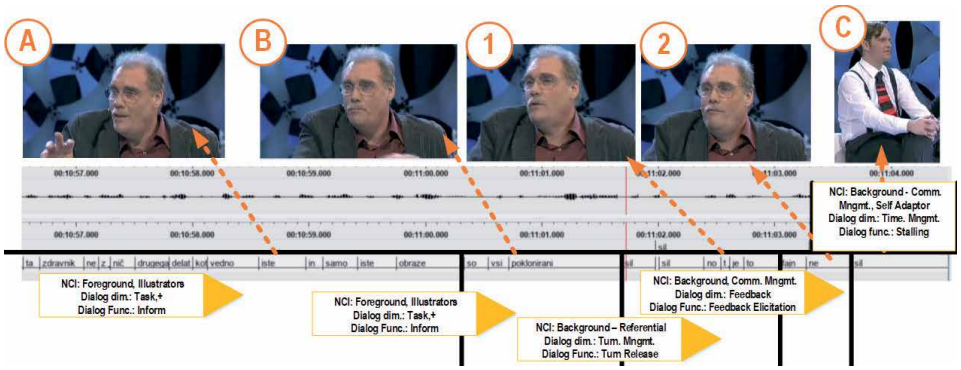


Figure 2.
Multimodal analysis of the conversational expressions: use of DAs in background and foreground expressions.

The utterance "well this is it" cannot be characterized as any other DA than turn management with the function of turn release as it serves no other communicative purpose. The phrase itself is tautological, deprived of any propositional content. The analysis of the accompanying body behavior also corroborates this fact. While outlining the plot of the play, he uses foreground NCI (A, B), namely illustrators, represented by two very prominent hand gestures. As he decides that he no longer wishes to explain the gist of the play, his NCI also changes. The body behavior is no longer prominent but very quick and even difficult to notice. The guest swiftly turns his head slightly to the right and back again (see **Figure 2**: 1) as if he was trying to point to the abstract "this" in "this is it" while still keeping eye contact with the host and co-host. The head movement was classified as a deictic NCI, specifically, a referent, since the guest is referring to the abstract "this". He then adds the utterance *fajn, ne?* "great, huh?" which primarily acts as feedback elicitation, but secondarily also serves the function of turn management. The accompanying body behavior is, again, subtle, just a slight shrug of the right shoulder (see **Figure 2**: 2). It was classified as a communication regulator. The co-host perceives his turn release request and takes the turn by commenting on the essence of the play. However, since the release was unexpected, his response is yet to be formulated. This is highlighted through the use of metadiscourse ("But actually, you know, actually") acting primarily as stalling within the time management dimension. However, of course, stalling functions also as a turn-take maneuver.

Example 2: Ambiguity of DAs in conversational expressions.

Host: eee eee eee(A) no eee (B) dejta ubesedet to midva se mava rada (1)

Host: uh uh uh (A) well uh (B) come on, define this we like one another (1)

The example above is a case of strong turn assigning. As a surprise for the main guest, his stepdaughter was invited to the show. The show's host is trying to determine the correct nomenclature for the relation non-biological father/adoptive daughter, which are specific and probably less frequently heard words in Slovene. However, he is very clumsy when formulating his question, and neither of the guests understands him, but rather fill their answers with humor. The show's host is dissatisfied and tries to change the evolution of the conversation. However, he needs time to formulate proper utterances and thus uses fillers (see **Figure 3**: **A** and **B**). After the first filler (A), which acts as stalling, the content is not completely formulated, which is why he uses the second filler (B). At the same time, however, the guests become impatient. The second filler, therefore, functions not only as a stalling element but primarily

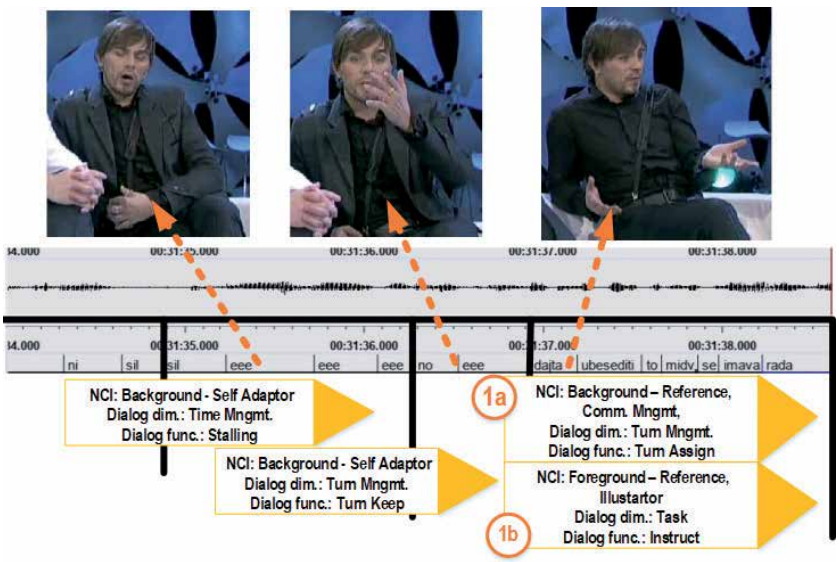


Figure 3.
Multimodal analysis of the conversational expressions: The duplicity of DAs when interpreted as background or as foreground conversational expressions.

as turn keep device. Once he formulated his idea, he begins with the imperative formulation *dajta* (for this purpose best translated as) "come on". This utterance is accompanied by the host's extended and raised left arm, both (temporarily) open hands, slightly raised shoulders, and a protruding head movement (see **Figure 3**: 1). This NCI was classified as a referential deictic (1a), as the host's hands and head are extending towards the guests. At the same time NCI can also be perceived as visualizing the word *dajta,* thus being recognized as an illustrator. From the context of DAs, the utterance can be interpreted as having the underlying function of turn-taking or, due to the imperative formulation, the instruct function within the task dimension. As highlighted in the example, the use of DAs determines the perceived NCI.

Example 3: DAs in turn management within a multiparty conversation.

Guest: grmičevje je zlo nerodno objemat
Co-host: zakaj?
Guest: ful pič ... ful ful te (A)
Host: ful ful pič ful pič
Co-host: ful me
Guest: drevo je fajn men
Host: eee (1)
Co-host: kosmulja (2)
Co-host: nadaljuj (3).
Host: ja (4)

Guest: it's very tricky to hug shrubs
Co-host: why?
Guest: totally pricks ... you get totally totally (A)
Host: totally totally pricks totally pricks
Co-host: I get totally
Guest: trees I like
Host: uh (1)
Co-host: gooseberry (2)

Co-host: continue (3).
Host: yes (4)

This segment illustrates NCI in tree different turn management functions. Prior to the several turn-taking acts, the show's host is mocking the guest for his alleged morning ritual where he hula-hoops in his garden. The co-host humorously adds that he hugs surrounding trees and shrubs in the garden. First, the guest smiles at this mental image, but then his facial expression changes to serious, and he cautions that it is very hard to hug shrubbery. As the co-host asks why this is so, he turns the answer into a comical depiction of how he gets stung by thorns whereby he uses the colloquial Slovene word *ful*, which means *very* or *a lot*. The co-host is fascinated by this word choice of the guest, a theater actor, who just minutes before teased him for not enunciating correctly. He, therefore, mocks his (almost plosive) pronunciation of the word *ful* and the guest joins in in the mocking. The show's host, however, tries to join the conversation (see **Figure 4**:1), but his co-host still continues the mocking by saying "gooseberry" (see **Figure 4**: A) in a very comical manner, trigging light laughter from the guest but befuddlement from the host. The host turns to the co-host, hoping for clarification, the co-host stares back at him and finally tells him to continue with the show. After briefly gathering his thoughts, the host nods, says "yes" and changes the topic.

This excerpt, therefore, contains turn-taking, turn-assigning, and turn-accepting. Following a series of task dimension DAs, the show's host tries to take the turn by uttering the filler *eee* "uh". He fails, as his co-host drowns him out with "gooseberry". There is no NCI accompanying the host's utterance, as he barely moves (see **Figure 4**: 1). We, therefore, classified the NCI as undetermined. An indicator for his turn-take attempt is his gaze, which remains directed towards the guests (see **Figure 4**: 1) throughout the

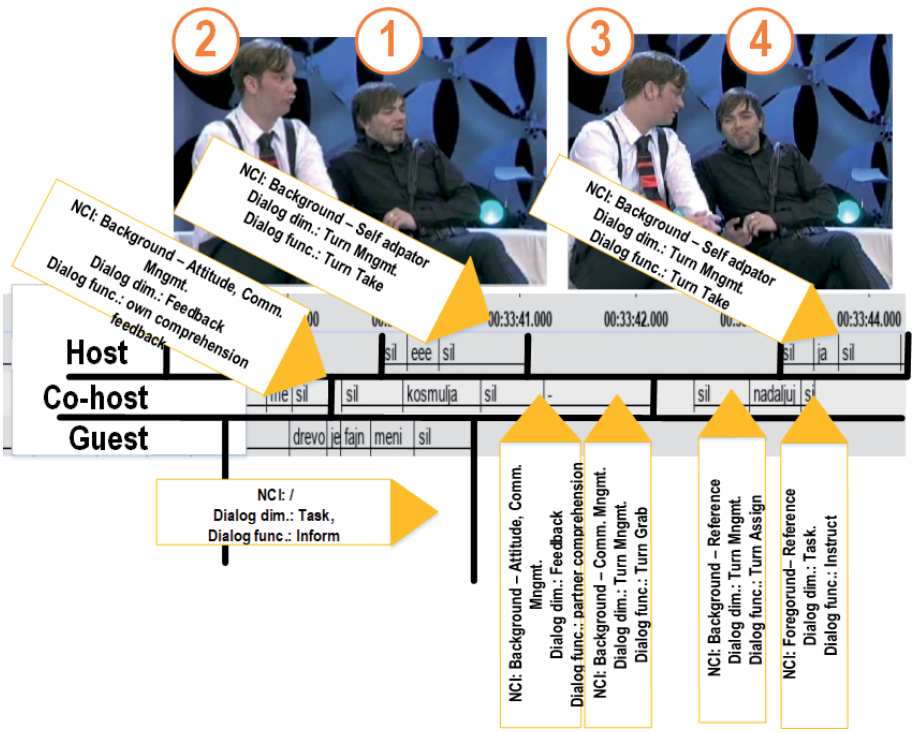


Figure 4.
Multimodal analysis of the conversational expressions: Use of turn-take and turn-grab to mediate the conversation.

co-host's interruption. As he turns to the co-host, he remains speechless and waits for him to elaborate. The co-host subsequent NCI, on the other hand, is a clear referential deictic (3) accompanying the DA of assigning the turn. His gaze towards the host was not enough to prompt a response, so he adds a firm head nod (see **Figure 4**: 3) towards him and verbalizes his intent of assigning the turn to him with "continue". This firm head nod is why the DA was not secondarily classified as a turn release, but as instructing within the task dimension. The host almost simultaneously responds to this NCI with a slight nod himself (see **Figure 4**: 4) and thereupon the verbal confirmation "yes". The nod was identified as a communication regulator NCI. The verbal confirmation functions as a turn accept DA. Both the DA and the NCI are of background nature.

4. Discussion

An effective analysis of the non-verbal behavior that accompanies turn-taking requires material that is authentic and rather informal than formal. The episode from the entertainment show *As ti tut not padu?* offers such a resource. As elaborated in the section *The EVA Corpus*, the notion that the material is highly spontaneous is corroborated by the video's statistics. The sheer amount of overlapping speech indicates a high frequency of turn management acts. The notion that the material is spontaneous and performed in a relaxed manner is further supported in Cooperrider's [27] foreground-background distinction, as the DAs are well-balanced according to their nature. Overall, there are 1,897 foreground DAs and 2,020 background DAs. Therefore, more than half of the DAs are of background nature. This suggests that the material is not task-oriented, but instead serves the purpose of an entertainment show. Moreover, the most frequently observed NCI in the video were NCIs classified as regulators. The group of regulators represents 3/4 of all recognized NCIs in the entertainment episode. Regulators are followed by the group of deictic NCIs representing almost roughly 16 percent of the recognized NCIs. The remaining groups of illustrators, batons, symbols, and undetermined NCI each account for less than a tenth of the recognized NCIs. Again, the dominant NCI groups are of background nature, even if observing only regulators. These findings further support the notion that the material is highly spontaneous, relaxed, and entertaining. Therefore, it was a suitable choice for the analysis of natural turn-taking behavior and its accompanying non-verbal behavior.

Even though it seems relatively elementary to annotate turn management DAs, at times, the process proved to be complex. The acts of turn management are intertwined with stalling and instruction DAs. Examples 1 and 2 show how the stalling function (within the time management dimension) can also act as a turn-taking mechanism. In example 1, on the one hand, the co-host wants to take the floor since the guest suddenly released his turn. On the other hand, he does not know how to start. Hence, it is difficult to determine the primary DA, especially since the remainder of his response can be considered an information-providing DA. Example 2 illustrates how fillers, which generally pertain to the stalling function within the time dimension, act as turn-taking devices. They signal to the interlocutor that one wishes to speak but still requires additional time to properly verbalize one's thoughts. To further complicate the annotation of turn-taking management, in a conversation, each utterance by another person can secondarily be considered a turn-taking DA. Even the act of posing a question, which would primarily be annotated within the information-seeking dimension, can secondarily be annotated with a turn-assign function (since person A, who is asking person B the question, wishes person B to respond). Nevertheless, we did not annotate such secondary cases of turn management as alternative DAs. Only DAs where turn management is key for an utterance were assigned the turn dimension. Consequently, the share of turn

management DAs could be significantly greater, and the ratio between foreground and background expressions notably tilted towards the background spectrum. Again, this only highlights the nature of the material, which is by no means primarily task-oriented.

Since regulators and (partially) deictics are background expressions, we expected them to co-occur with turn management DAs. As noted, deictics are the elusive group of NCIs which can occur both in the foreground and the background. In Example 1, the explicit turn management DA is accompanied by a referential deictic. The semantic link with the word it refers to, however, is weak. The guest speaks of an abstract "this" in his utterance "well this is it". He does not refer to anything physically present in the room; he just points to a mental image. Therefore, the deictic is part of the background. This, in turn, is in line with the assigned DA, since both concepts are background expressions.

A similar symmetry between the nature of the DA and the NCI is observed in Examples 2 and 3. In Example 2, the NCI can be considered both as a background and a foreground expression (see **Figure 3**: 1). Within the concept of DAs, it can also be considered as occurring in the background and in the foreground. In the background, it is a turn-assigning DA with which the host hopes to receive a response to his nomenclature dilemma. In this case, the non-verbal behavior is perceived as spontaneous. Rather than to visualize the referential utterance *dajta* "come on" the speaker tries to emphasize his frustration with the interlocutors, i.e., if you know better then please explain, and thereby assigns the turn someone else. The 'open hand gesture' is observed to signal this. In the foreground, it has an instructing function, within the task dimension, since he demands a response. The host is referencing actual people in the room and due to the imperative use of the referential utterance *dajta* "come on you two", the observed conversational expression may be interpreted as instructing. *Dajta* is perceived as an explicit speech referent, and the non-verbal behavior seems to directly visualize it. Similarly, in Example 3, the NCI (see **Figure 4**: 3) is not a typical background referential deictic as the co-host physically leans towards the host while he nods towards him in order to prompt him to continue. The referent of the NCI is, therefore, an actual person (the host) in the room, and the NCI also occurs in the foreground. This duality is reflected in DAs as well. On the one hand, the co-host assigns the turn to the host (within the turn dimension); on the other, the co-host instructs the host to speak (within the task dimension). Again, this is an example of the difficulty in differentiating between background and foreground expressions.

Finally, as hypothesized, regulators are the group of NCIs that co-occur with unambiguous turn management DAs. In Example 1, the secondary turn management DA co-occurs with the group of regulators, specifically a communication regulator. Whereas the group of deictic NCIs crosses the foreground-background barrier, regulators are background expressions. The fact that the accompanying NCI is a regulator and not a referent from the deictic group, which would be more typical for feedback elicitation, further endorses the assignment of a turn dimension DA and highlights the turn management intent. Example 3 illustrates a similar unambiguity when regulators are used for turn-taking. There is an underlying agreement of the nature of the DAs and the NCI in the last utterance "yes" (see **Figure 4**: 4). Namely, both take place in the background. And clearly, they were well understood by the interlocutors as no one else tried to take the turn. The opposite phenomenon is observed at the beginning of the same example (see **Figure 4**: 1) with the host's filler "uh". It is an example of a failed turn-take attempt since the co-host interrupts the host. There was no noticeable non-verbal behavior accompanying the host's filler, which is why no NCI could be assigned. However, one might argue that the fact that there is no NCI accompanying his turn-take attempt, contributes to the reason of why the attempt failed. Consequently, this might be considered a supporting example of Birdwhistell's [22] findings that successful communication requires both verbal and non-verbal components.

We can therefore confirm our hypothesis that turn management DAs co-occur with regulators. The case analysis further supports the hypothesis that turn management DAs particularly co-occur with communication regulators. Moreover, we can observe that during propositional content, i.e., task-oriented DAs, use of illustrators (foreground NCIs) is more common (see **Figure 2**: B in Example 1). In accordance with Cooperrider's [27] characteristics of foreground-background gestures, we observed the spatial prominence of each type of gesture. Example 1 shows how non-verbal behavior changes in parallel with the change of DAs. As the DAs changed from task with the function of informing to turn management with the function of turn release (see **Figure 2**, B and 1), so did the NCI. It shifted from foreground behavior to background behavior. Moreover, foreground NCIs are far more prominent than the background NCIs. It seems that, as this simultaneous shift in DAs and foreground-background behavior occurs, body behavior is decelerated and minimized. Our findings, therefore, corroborate Cooperrider's [27] the special hallmarks of foreground-background gestures.

There are, however, border cases. For example, the background DA of providing feedback during active listening, such as uttering the supportive "yes" or "mm-hmm", can be accompanied by a slight nod of the head. Head nodding is generally considered a foreground gesture, if it signals a "yes" or "no" answer, since it can substitute speech altogether. However, in background use, one does not provide an answer, but merely signals to the interlocutor, that one is listening to them and wishes them to continue their turn. Hence, the act is clearly of background nature. Nevertheless, it is impossible to state that at the same time one does not also agree with what the interlocutor is saying. Agreement, however, is considered a foreground act. This is a typical case where the duality no longer applies. Hence, it is possible even for background DAs, such as feedback providing and eliciting, to co-occur with foreground NCIs. Moreover, even task-oriented DAs are often accompanied with batons, a representative background NCI, since they signal importance or set the rhythm but do not convey any propositional content. It is therefore difficult to extend the shared background-foreground nature hypothesis to other DAs. Despite this observation, the exploration of the shared nature in foreground DAs offers an interesting research question for future research.

A potential concept to further elaborate on the underlying nature is to observe whether the gesture is prominent (in its iteration or spatial dimensions) or subtle [27], as observed in Example 1. In accordance with this distinction, a subtle nod suggests background nature whereas a prominent nod suggest that the gesture is of foreground nature. Moreover, the relative timing may also, provide additional insight in the communicative intent. Although, not directly investigated in this research, is seems that when the stroke phase of the embodiment (especially a hand gesture) co-occurs with a specific speech referent (i.e. the gesture starts at the same time as the spoken articulation) the information provided is propositional, i.e., of foreground nature, whereas when the stroke phase occurs outside boundaries of the targeted referent (or without one) the information provided is of background nature. An example would be phrases "look over there!" and "what do you mean?". In general, deictics will accompany both phrases. On the one hand, the phrase "look over there" is clearly a task-oriented DA and will be accompanied by a pointer, the stroke of which will occur aligned with the verbal articulation of "there". On the other hand, the stroke phase of a similar gesture 'visualizing' the "you" in "what do you mean?" will co-occur with "mean" and will be recognized as a referential deictic in turn management (i.e., as turn offer). Thus, in our future investigations, we tend to analyze if the alignment of verbal structure with the prosody of non-verbal cues (i.e. the cues preceding verbal acts, cues following verbal acts, cues at the beginning or end of verbal cats) may shed further light on the true purpose of the shared nature.

5. Conclusions

In this chapter, we examined what kind of non-verbal behavior accompanies turn management DAs. For the annotation of turn management DAs, the ISO 24617-2 scheme's functions sufficed. Nevertheless, turn management DAs frequently overlap with other DAs, especially within the time management dimension. The fact that it is sometimes very difficult to decide which dimension and function is the most fitting shows the importance of multidimensional DA tagging. As a future endeavor, it would be more functional to create annotation schemes that, besides being multidimensional, denote the hierarchical order of the tags assigned, for example, the primary, secondary, tertiary, etc. dimensions and functions.

Cooperrider's [27] distinction between gestures that occur in the foreground or background proved an effective method within the concept of DAs. We hypothesized that there is an interlink between background NCI and background DAs. Since regulators, specifically, communication regulators, convey typical background NCI, we predicted their co-occurrence with turn management DAs. Indeed, the present case study confirms this hypothesis. Moreover, an interlink with deictic NCI was observed. As they can be of either background or foreground nature, the premise that background DAs co-occur with background NCI is maintained. This duality is not observed only within NCI but also within DAs. An utterance can have alternative expressions, one of background nature and one of foreground nature. However, the duality occurs simultaneously for NCI and for DAs. Hence, the fact that there is the same duality at the NCI level and at the DAs level strengthens the hypothesis of an interlink between the two concepts.

Acknowledgements

This paper is partially funded by the Slovenian Research Agency, project HUMANIPA (research core funding No. J2-1737 (B)). This paper is partially funded by European Union's Horizon 2020 research an innovation program, project PERSIST (grant agreement No. 875406).

Conflicts of interest

The authors declare no conflict of interest.

Author details

Izidor Mlakar*, Matej Rojc, Darinka Verdonik and Simona Majhenič
Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia

*Address all correspondence to: izidor.mlakar@um.si

References

[1] Austin J L. How to do things with words. Oxford: Clarendon Press; 1962.

[2] Searle, J R. Speech acts: An essay in the philosophy of language. Volume 626. Cambridge university press; 1969.

[3] Bunt, H. The DIT++ taxonomy for functional dialogue markup. In: Heylen D, Pelachaud C, Catizone R, & Traum D, editors. Towards a Standard Markup Language for Embodied Dialogue Acts. Proceedings. Budapest; 2009. p. 36-48.

[4] Schmitz B, Quantz J J. Dialogue Acts in Automatic Dialogue. In: Interpreting Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation, TMI-95. Leuven; 1996. p 33-47.

[5] Heeman P, Allen, J. The Trains 93 Dialogues: TRAINS Technical Note 94-2. Computer Science Department. The University of Rochester; 1995.

[6] Di Eugenio B, Jordan P W, Pylkkänen L. The COCONUT project: dialogue annotation manual. ISP Technical Report 98-1. University of Pittsburgh; 1998.

[7] Godfrey J J, Holliman E C, McDaniel J. SWITCHBOARD: Telephone speech corpus for research and development. In: Proceedings of the ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing; 23-26 March San Francisco, CA, USA; 1992. Volume 1. p. 517-520.

[8] Core M G, Allen J F. Coding dialogues with the DAMSL annotation scheme. In: Traum D R, editor. Working Notes: AAAI Fall Symposium on Communicative Action in Humans and Machines. Boston; 1997. p. 28-35.

[9] ISO 24617-2. Language resource management – Semantic annotation framework (SemAF): Part 2: Dialogue acts. Reference number ISO 24617-2:2012(E). Geneva; 2012.

[10] Bunt H. The Semantics of Dialogue Acts. In: Proceedings 9th International Conference on Computational Semantics (IWCS 2011); January 12-14; Oxford, UK, p. 1-14.

[11] Petukhova V, Bunt H. Introducing Communicative Function Qualifiers. In: Fang A, Ide N, Webster J, editors. Language Resources and Global Interoperability. Proceedings of the second International Conference on Global Interoperability for Language Resources (ICGL 2010); City University of Hing Kong; 2010. p.123-131.

[12] Yoshino K, Tanaka H, Sugiyama K, Kondo M, Nakamura S. Japanese Dialogue Corpus of Information Navigation and Attentive Listening Annotated with Extended ISO-24617-2 Dialogue Act Tags. LREC; 2018.

[13] Navarretta C, Paggio P. Dialogue Act Annotation in a Multimodal Corpus of First Encounter Dialogues. LREC; 2020.

[14] Petukhova V, Bunt H. The coding and annotation of multimodal dialogue acts. In: Proceedings 8th International Conference on Language Resources and Evaluation (LREC 2012); Istanbul: 2012.

[15] Petukhova V, Bunt H. A multidimensional approach to multimodal dialogue act annotation. In: Proceedings of the Seventh International Workshop on Computational Semantics (IWCS); 2007. p. 142-153.

[16] Hunyadi L, Váradi T, Kovács G, Szekrényes I, Kiss H, Takács K. Human-human, human-machine

communication: on the HuComTech multimodal corpus. In: CLARIN Annual Conference 2018; Pisa, Italy; 2018.

[17] Feyaerts K, Brône G, Oben B. Multimodality in interaction. In: Dancygier B, editor. The Cambridge Handbook of Cognitive Linguistics. Cambridge University Press: Cambridge; 2017. p. 135-156.

[18] Mlakar I, Kačič Z, Rojc M. A Corpus for Analyzing Linguistic and Paralinguistic Features in Multi-Speaker Spontaneous Conversations–EVA Corpus. International Journal of Computers, 2. 2017;136-145.

[19] McNeill D, Duncan S. Growth points in thinking-for-speaking. In: McNeill D, editor. Language and Gesture (Language Culture and Cognition). Cambridge: Cambridge University Press; 2000. p. 141-161.

[20] Peirce C S. Collected papers of Charles Sanders Peirce. Vol. 5. Harvard University Press; 1965.

[21] Ekman P, Friesen W. Constants across cultures in the face and emotion. Journal of Personality and Social Psychology. 17 (2). 1971;124-129.

[22] Birdwhistell R L. Kinesics and Context: Essays on Body Motion Communication. Philadelphia: University of Pennsylvania Press; 1970.

[23] Allwood J. A framework for studying human multimodal communication. In: Rojc M, Champbell N, editors. Coverbal Synchrony in Human-Machine Interaction. Ed. 1. Boca Raton: CRC Press; 2014. p. 17-39.

[24] McNeill D, Levy E, Duncan S D. Gesture in Discourse. In: Tannen D, Hamilton H E, Schiffrin D, editors. The Handbook of Discourse Analysis 2. Wiley-Blackwell; 2015. p. 262-289.

[25] Church R B, Goldin-Meadow S. So how does gesture function in speaking, communication, and thinking? In: Church R B, Alibali M W, Kelly S D, editors. Why Gesture? How the hands function in speaking, thinking and communicating. Gesture Studies 7. Philadelphia: John Benjamins Publishing Company; 2017. p. 397-412.

[26] Keevallik L. What does embodied interaction tell us about grammar? In: Research on Language and Social Interaction, 51(1). 2018. p. 1-21.

[27] Cooperrider K. Foreground gesture, background gesture. In: Gesture, 16(2). 2017. p. 176-202.

[28] Mlakar I, Verdonik D, Majhenič S, Rojc M. Towards Pragmatic Understanding of Conversational Intent: A Multimodal Annotation Approach to Multiparty Informal Interaction– The EVA Corpus. In: International Conference on Statistical Language and Speech Processing. Springer, Cham; 2019. p. 19-30.

[29] Paggio P, Navarretta C. The Danish NOMCO corpus: multimodal interaction in first acquaintance conversations. In: Language Resources and Evaluation, 51(2). 2017. p. 463-494.

[30] Cohen J. A coefficient of agreement for nominal scales. In: Educational and Psychological Measurement, 20(1). 1960. p. 37-46.

Chapter 7

The Most Powerful Thing You'd Say Is Nothing at All: The Power of Silence in Conversation

Bashir Ibrahim and Usman Ambu Muhammad

Abstract

After a long period of neglect, silence is currently receiving an increased amount of attention in the literature of sociolinguistics and pragmatics. Since the publication of Tannen and Saville-Troike and Jaworski, many international conferences, books, monographs, articles, PhD theses and book chapters continue to emerge. Many of those publications recognized silence as a powerful tool of communication; and that it is not peripheral to speech because any form of analysis that is applied to speech could also be applied to the analysis of silence. Silence has been broadly classified as communicative and non-communicative, and it serves both positive and negative functions. As silence performs two opposite functions, its interpretation depends on some factors such as the socio-cultural background of the actors involved in the use and the interpretation of the silence act, and the context of its use. This chapter starts with an introduction which covers review of related literature, and then proceeds with the classification of silence. It continues with discussing some functions of silence, and then talks about interpretation of silence in social contexts. Finally, the chapter examines some instances of the power of silence in conversation.

Keywords: Silence, conversation, communication, sociolinguistics, pragmatics

1. Introduction

Conversation is, of course, not a mere disorderly chunk of speech. There are protocols that guide exchange of talk as well as when not to say anything. The exchange protocols and the time of silence rely heavily on culture and context. Culture here refers to the norms of the society that guide how the talk or silence is used, and the context refers to the setting in which the conversation holds. Trudgill [1] reported that '[in] some Caribbean communities, as among certain groups of Black American adolescents, it is perfectly normal, at least in some situations, for everyone to talk at once. Schegloff [2], however, argues that the tendency of two speakers talking at the same time appeared to be a departure from what is basically known in conversation as one-speaker-at-a-time.

Various studies have looked at silence from communicative perspective such as some studies of some tribes in Africa as among the Akan people of Ghana [3], the Igbo people of Nigeria [4]; religious silence such as among British Quakerism and British Buddhism [5]; or silence in the lawyering process [6]. Others have looked at

silence in social interaction. Trudgil [7], for example, stated silence of longer than four seconds is not allowed especially among native speakers of English who are not close friends because such silence is considered as embarrassment. Similarly, Tannen [8] reported that silence was averted during conversation among some New Yorkers of East European Jewish background. These reports indicated that silence is perceived as a negative trait particularly among Europeans.

In addition, Jaworski [9] reported that one of the passengers in a 5-hour journey by train in Poland expressed disgust over their silence throughout the journey, describing the situation as if they were going to a funeral. It is of little surprise then that Scollon [10], who uses the theory of metaphor to study silence, describes silence as malfunction – 'If one assumes the engine should be running, the silences will indicate failures. Smooth talk is taken as a natural state of the smoothly running cognitive and interactional machine'. It has to be noted, however, in the situations cited above, silence is portrayed as part of the participants' preference for talk or avoidance of relationship. In all those cases, silence is non-communicative because the silence of the passengers in a train, for example, cannot be interpreted. Also, in the case of the conversation between New Yorkers and non-New Yorkers, silence was described as a boundary marking the beginning or end of speech, and the desire of the New Yorkers to continue talking. Such desire might be a tendency acquired since childhood or a personality trait of the conversational partners involved.

There is, however, another type of silence which occurs during conversation when the current speaker stops talking or when the next speaker takes the floor from the current speaker. The former is referred to as 'switching pause' while the latter as 'inturn pause' [10]. The turn exchanges is referred to as turn-taking. Some pioneering works on turn-taking in cross-cultural encounters include Basso [11] and Scollon and Scollon [12] who conducted studies on Western Apache and the Athabaskan Indians respectively. Their studies indicated variations in turn- taking habits of the Western Apache cultures and the Athabaskans with that of the 'West'. Reporting Scollon and Scollon's [12] research, Trudgill [7] stated that:

> *The Athabaskans go away from the conversation thinking that English speakers are rude, dominating, superior, garrulous, smug and self-centered. The English speakers, on the other hand, find the Athabaskans rude, superior, surly, taciturn and withdrawn.*

Such perception culminated from what one of conversation partners believed to be usurping his right to the floor while the other thought his partner has relinquished such a right. Turn-taking, therefore, is not just exchange of speakership from one person to the other but how such transfer is made without violating exchange etiquette. Members of speech community know how to participate in turn-taking exchanges and how to adhere to rules that are appropriate in their community, by allowing overlap (if acceptable) to occur between utterances, and by using pause lengths that are compatible with their particular socio-cultural norms.

2. Role of culture in the use and interpretation of silence

As various cultures differ in their use of silence, the interpretation of someone's silence can also be culture-dependent. Fast rate of speech, for example, is valued in European communities while in some communities such as the Navajo and Indians Athabaskans, longer silences are tolerated. Such silences have,

therefore, become the feature of the speech of those communities, and therefore, not a problem. Also among the Japanese, Nakane [13] reported that 'Japanese silence stands out not only in comparison with Southern Europeans or New Yorkers but also with East Asian neighbors such as the Koreans and Chinese as well'. Silence is used more often by the Japanese in highly emotional situations such as courtship. 'Young spouses who are deeply in love, for example, often express their affection for each other by nonverbal means and silence' [9]. There was no consensus among researchers on Japanese silence. Some of the researchers – Anderson [14] and Miller [15] criticized the view that Japanese are more reserved compared to other Asians because the claim, according to Anderson, was purely intuitive, lacking any empirical backing. Anderson added that Japanese do talk, and sometimes they talk a lot, but the context of the talk is culturally sanctioned. The above arguments indicate the need to explain further the power of silence in conversation.

To some other communities, however, particularly in the west, past rate of speech is the expected trait during conversation [8]. For this reason, 'the debate over whether silence or talk should take more priority will never be settled with reconciliation across all contexts' [16]. This is obvious because silent cultures might not discard their silent practice for the voluble ones or vice versa. The Japanese, for example, might not dispose of its 'quietness' because another culture somewhere is garrulous. This is because 'cross-cultural uses of silence are rooted in the observation of different types of taboo, practical magic, and in varying beliefs as to how much talk is necessary in a given situation' [9].

3. Classification of silence

Earlier, since 1771, Dinouart, cited in Perniola [17] classified silence into ten types:

i. Prudent silence: Silence use to avoid jeopardizing oneself.

ii. Artificial silence: Deliberate use of silence in order to benefit from what others may say.

iii. Courteous silence: The use of silence to show approval of something usually accompanied by nodding or gesture.

iv. Teasing silence: Using silence as a means of deception to show approval while the performer of the silence act is actually deceiving the perceiver of the silence act.

v. Spiritual silence: Silence used by those who believe that it (silence) is a means of spiritual connection between themselves and the divine being.

vi. Stupid silence: In this type of silence the performer remains mute because he has nothing to express, particularly when he was alone.

vii. Applauding silence is used to show approval in front of the person being heard, mostly accompanied by nodding one's head.

viii. Contempt silence is used to show disregard towards what is said.

ix. Capricious silence: This type of silence is used by those who want show their liking or disliking of what they hear, mostly accompanied by opening their mouths.

x. Political silence: This type of silence is used by those who use circumspection, not saying all that they think, not entirely explaining their conduct or purposes.

The above proposal broadly explains the types of silence that may likely occur in various situations. A close look at the classification would reveal the need for reclassification. The first one (prudent silence) for example can be merged with the second (artificial silence) and the last one (political silence) because the intent of using the silence in these three situations is similar – to avoid jeopardizing oneself. Thus, both can be classified under prudent silence. Also, as courteous silence and applauding silence are used to show approval, they can be merged under applauding or courteous silence.

Later, other researchers classified silence by considering some factors such as genre and the context in which the silence is practiced. Saville-Troike [18], for example, proposed *etic* categories: Institutionally-determined silence, Group-determined silence, and Individually-determined/negotiated silence. As for Kurzon [19], there is a *typology* of silence comprising conversational, thematic, textual and situational silences. In his classification of silence, Jaworski [9] proposed 'fuzzy categories'. He approached silence from a socio-pragmatic perspective where he discussed silence 'as a component of various communicative situations and a tool of communicative expression' [9]. He categorizes silence into two broad categories: Communicative and Non-communicative. Jaworski believes that not all types of silence are communicative. As such, 'the actual interpretation of someone's silence takes place only when the communication process is expected or perceived to be taking place'. It is of no use, for example, trying to infer meaning from some silent person who was found alone 'day-dreaming' in the classroom. In this situation, his silence can be termed as 'stupid silence' [17]. But if communication was triggered by two strangers, for example, and silence occurred during conversation, then the silence is perceived to be meaningful and, therefore, subject to interpretation. This latter case of the use of silence is communicative while the former is non-communicative.

As communicative silence is perceived to be meaningful and interpretable, Jaworski [9] expands it to include the following types:

i. Silence as State: This is the type of silence where communicative event is structured or framed. The information transmitted through this type of silence is mostly in the form of visual arts, music, literary works, kinesics or proxemics behaviors. Jaworski argues that in the aforementioned genres, silence is a substance presented by the communicator in nonverbal form.

ii. Silence as Formulaic: This is 'a customary act of saying nothing in reaction to specific stimuli' sometimes 'accompanied by other nonverbal behavior such as bowing, smiling, waving and so on'. Interpretation of such type of silence depends largely on cultural practices, and contexts. Such formulaic silences include instances where loss of face is perceived such as someone passing gas, belching or spitting in public. This type of silence is also practiced during funerals or some rituals.

iii. Silence as Activity: This type of silence includes refraining from speaking, and acts of failing to mention something. Two examples provided by Jaworski may suffice here:

a. Mother gave her approval in silence.

b. They moved to the next point of the agenda in silence

In examples 'a' and 'b' above, silence is conceptualized as part of an activity following certain speech acts. According to Jaworski [9] the three categories he proposed 'are the most prototypical ones and seem to provide a good starting point for studying silence in interpersonal communication'. He believes that the categories 'cover a wide range of forms and situations in which the concept of silence can be used to explain and account for problems of miscommunication and misunderstanding, indirectness, ritualized behavior, and cross-cultural communication'.

As Jaworski [9] approached silence from social and pragmatic perspectives, his classification centred mainly on the use of silence in social encounters, and the meaning that can be derived from using silence in the social context being talked about. His recognition of silence as communicative and non-communicative is, certainly, commendable because he acknowledges that not all silences are interpretable and meaningful. Silence of someone who is found alone, for example, is non-interpretable and meaningless unless preceded by talk or act which requires the silent person to talk but he chooses to remain silent. Silence is non-communicative when it serves linguistic function as where interlocutor pauses, interrupts or overlaps during conversation. Turn exchanges can occur in both local and foreign language use. In both the local and foreign language use interlocutors are expected to follow certain conventions depending on the situation.

4. The power of silence in conversation

In this section, an attempt has been made to show the power of silence in conversation from pragmatic perspective. The examples did not include the use of silence from conversation analysis perspective i.e. turn-taking. During conversation, it is pertinent for conversation partners to understand when to say something and when to remain silent. Also, in situations where silence is used instead of talk, conversation partners must try to infer meaning to the silence act in order to avoid misunderstanding and confrontation. The interpretation of the silence act, however, might not always be accurate due to its various nuances. The use of silence in similar situations, for example, may evoke different interpretations depending on culture, situation or setting. Saville-Troike [18] cited two examples as they occurred among the Japanese and the Igbo of Nigeria:

A: Please marry me
B: [Silence; head and eyes lowered] (Acceptance)

In Japanese culture, silence in the above context signifies acceptance because, according to Jaworski [9], young spouses who are deeply in love often express their affection for each other by nonverbal means and in silence. In the Igbo culture of Southern Nigeria, however, similar scenario can mean either rejection or acceptance as in the following example:

A: Are you still mad at me?
B: [Silence] (Affirmative)

In the preceding example, if the girl continues to stand there saying nothing it means denial, but acceptance if she ran away Nwoye [4] cited in Tannen [8]. The interpretation of silence in the two cultures cited above is rather convoluted because the scenarios are similar but the act of staying or running away interprets the silence act. This is parallel with what Davidson [20] cited in Jaworski [9] who states that '[s]ilences that occur immediately after the speech act of invitation, offer, request, or proposal are typically interpreted as rejections'. The silences in the above scenarios evoke different interpretations – acceptance in the case of the Japanese girl, and rejection and/or acceptance in the case of the Igbo girl. Apart from portraying the ambiguity of silence, the above examples show that silence is not an empty 'locution', but "a potent communicative weapon" [21] which is used in formal and social situations. Hence, silence often sends the most powerful message in a more safer and apt way than verbal communication.

In situations where communication is involved or is perceived to be taking place, there should be the sender and the receiver of the message. Such message can either in verbal or non-verbal forms. Silence, therefore, is a non-verbal communication which carries symbolic significance, and it is interpretable based on the intent, situation and context of its use. Consider again the following situation by Jaworski [9]:

Speaker A: How do you see this shirt?
Speaker B: [looks at the shirt, silence (0.1)]. Yeah, it's good.
Speaker A: No. You didn't like it.

The slight silence of speaker B sends a powerful message to speaker A who concluded that the intention of speaker B has been hidden, and therefore interpreted the silence of 0.1 seconds as dislike despite an utterance which likely conceals the real intent of B. Nakane [22], citing Crown and Feldstein [23] believes that lengths of pauses and tempo of speech can be associated with personal traits and can contribute to listener's impression of the speaker. The use of silence and the act of looking at the shirt in the above example created an impression in the mind of speaker A that his shirt was not liked despite the response of B in the affirmative.

Looking at silence as a stimulus, it can often be less demanding particularly on the part of the addresser and the addressee, and particularly when the context is clear to both of them. Consider this example provided by Jaworski [9]:

Peter: How much do you earn at this new place?
Mary: [Si lence]
Peter: Well, you don't have to tell me.

In the above example, Peter presumably asked a question which he should not have asked because one's wages is confidential and personal. Though the relationship of Peter to Mary is not stated, it can be said that the relationship was very strong to the extent that Peter feels that asking such a question might not lead to a confrontation. Peter, however, received a big surprise with silence which indicated the unwillingness of Mary to reveal her salary status to him. The communicators achieved both their communicative and informative intentions in a more optimal way. Peter inferred meaning from the silence of Mary; and on her part, Mary has passed information to Peter in a more solid and concise way possible. Peter immediately infers Mary's silence as unwillingness to expose her salary information to him, therefore the information [silence] is worth Peter's effort to process and interpret

it. On her part, it was Mary's choice and preference to express her intention with silence, which might be safer for her than to give a verbal response.

In another situation below, A wants to use B's umbrella but the silence of the latter sends a powerful message to B which compels the former to withdraw his earlier request. This situation confirms what Davidson [20] says about silences that come after requests, proposals and offers as rejections. Although B did not say anything, her silence is obvious – a resounding 'no'. Had it been B uses the word 'no', A might feel more offended as the answer might sound offensive and defiant:

A: Can I use your umbrella?
B: [silence]
A: Well, I can use Ben's.

In the above example, A interprets B's silence as a warning or order which implies 'no don't take it' (a warning), or simply 'I ordered you not use it'. As B did not use verbal response, the tendency of hot feeling is reduced, and A quickly changes his decision of using B's umbrella to Ben's.

However, not all similar situations like the above can end smoothly. Sometimes the silence act might end up embarrassing the conversational partner. For example, Jaworski [9] narrated that his neighbor's daughter was married out, and after the wedding the neighbor visited him and his wife. As their discussion unfolds, the neighbor asked him how much she owed him. This is what he said after the woman asked him that question:

I was genuinely appalled at that question
so I did not say anything and just looked
peeved at the woman. After a moment she said
"Do you want me to jump out of the window".
I said "Yes". [italics mine]

The above incidence shows how the author felt after the question, and therefore remained silent, but disgustingly, continued to look at the woman. The woman, on her side felt embarrassed and ashamed of the question she asked due the silence act that followed her question, and asked if she could jump out of the window. The reply of the author in the affirmative indicates how dismayed he was. He stated that:

"I knew she did not intend to be rude to me,
but I felt insulted. If I had decided to tell her
that I thought she was being rude to me at that moment,
I may have hurt her in turn" [9]

In the above example, silence serves as a repair mechanism of a seeming confrontation. On the side of the author, his silence serves a referential function which might be interpreted as "You shouldn't have asked this question"; or "Why do you asked me this question?" or "I didn't like what you asked". On her side, the woman felt very much embarrassed by the question she asked, and quickly changed the matter to a joke by asking whether she could jump out of the window. The author instantly accepted the change of the topic, thereby repairing the conversation. Any verbal response to that type of request which appears to be disgusting might mess up the whole situation and brings confrontation and dissonance. The most powerful thing that might repair the interaction is silence. From the foregoing examples, it can be suggested that most uses of conversational silence are negative. The interpretation of the silence act can either be rejection, order, warning or defiance.

5. Conclusion

This chapter discussed the power of silence in conversation by bringing an overview of the classification of silence, its function in communication, and how it is (mis)interpreted in various contexts. The result of this discussion shows that silence is ambiguous and its interpretation varies even within similar contexts. In many of the situations cited above silence was used as rejection, order, warning or defiance which are apparently unpleasant in human interactions. Use of verbal means to express such negative acts might increase the unpleasantness in the social interaction. The most suitable alternative could be the use of silence in order to reduce the pain of using verbal communication. Silence could, of course, be a 'reliever' of social tension that might occur in many human interactions, and an effective tool of sending a powerful message. Future studies may consider how silence accompanied with other non-verbal acts such as a grin or a smirk contributes to the interpretation of the silence act.

Author details

Bashir Ibrahim* and Usman Ambu Muhammad
School of General Studies, Kano University of Science and Technology, Wudil, Kano State, Nigeria

*Address all correspondence to: bashirdukawa@gmail.com

References

[1] Trudgill, P. (1974). Sociolinguistics: An introduction to language and society. Middlesex, England: Penguin books

[2] Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in society*, *29*(1), 1-63.

[3] Agyekum, K. (2002). The communicative role of silence in Akan. *Pragmatics*, *12*(1), 31-32.

[4] Nwoye, G. O. (1978) Unpublished manuscript. Georgetown University. Ref. in Saville- Troike, M. (1985).

[5] Bell, S., & Collins, P. (2014). Religious silence: British Quakerism and British Buddhism compared. *Quaker Studies*,*3*(1), pp. 1-26.

[6] Kurzon, D. (1995). The right of silence: A socio-pragmatic model of interpretation. *Journal of pragmatics*, *23*(1), 55-69.

[7] Trudgill, P. (1983). *On dialect: Social and geographical perspectives*. Wiley-Blackwell.

[8] Tannen, D. (1985). Silence: Anything but. In D. Tannen, & M. Saville-Trioke, (Eds.), *Perspective on Silence.* Norwood, NJ: Ablex PublishingCorporation.

[9] Jaworski, A. (1993). *The power of silence.*Newbury Park, London: Sage Publications

[10] Scollon, R. (1985). 'The Machine Stops: Silence in the Metaphor of Malfunction'. In D. Tannen, & M. Saville-Trioke, (Eds.), *Perspective on Silence*.Norwood, NJ: Ablex PublishingCorporation.

[11] Basso, K. H. (1970). " To give up on words": Silence in western apache culture. *Southwestern Journal of Anthropology*, *26*(3), 213-230.

[12] Scollon, R., & S. W. Scollon (1983) Face in interethnic communication. In Richards, J., & R. Schmidt (eds), *Language and Communication*. London: Longman, pp. 156-188.

[13] Nakane, I. (2003). *Silence in Japanese-Australian Classroom Interaction: Perceptions and Performance*. (Unpublished PhD thesis). University of Sydney

[14] Anderson, F. E. (1992) The enigma of the college classroom: nails that don't stick up. InWadden, P. (ed.), *A Handbook for Teaching English at Japanese Colleges andUniversities*. Oxford: Oxford University Press, pp. 101-110.

[15] Miller, L. (2000) Negative assessments in Japanese-American workplace interaction. InSpencer-Oatley, H. (ed.), *Culturally Speaking: managing rapport in talk acrosscultures*. London: Continuum, pp. 240-254.

[16] Bao, D. (2014). *Understanding silence and reticence: Ways of participating in second language acquisition*. United Kingdom: Bloomsbury Publishing PLC.

[17] Perniola, M. (2010). Silence, the utmost in ambiguity. *CLCWeb: Comparative Literature and Culture*, *12*(4), 2.http://dx.doi.org/10.7771/1481-4374.1670

[18] Saville-Troike, M. (1985) The place of silence in an integrated theory of communication. In Tannen, D. & M. Saville-Troike (eds), *Perspectives on Silence* . Norwood, NJ: Alex, pp. 3-18.

[19] Kurzon, D. (2007). Towards a typology of silence. *Journal of Pragmatics*, *39*(10), 1673-1688. doi:10.1016/j.pragma.2007.07.003

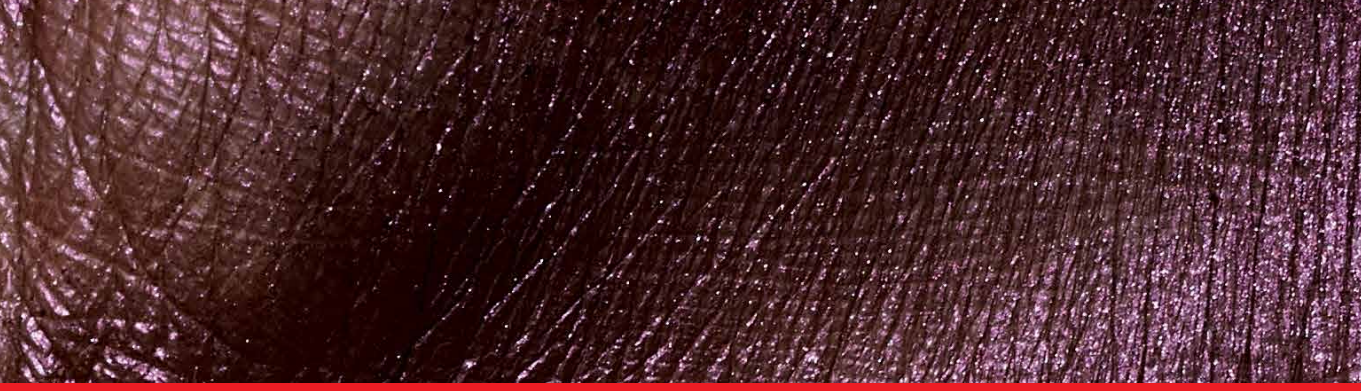
[20] Davison, J. (1984). Subsequent versions of invitations, offers, requests, and proposals dealing with potential or

actual rejection. In J. M. Atkinson & J. M. Heritage (Eds.), Structures of social action: Studies in conversational analysis (pp. 102-128). Cambridge, UK: Cambridge University Press.

[21] Wardhaugh, R. (1985). *How Conversation works.* New York, USA:Basil Blackwell

[22] Nakane, I. (2012). Silence. In Paulston, C.B, Kiesling, S.F., & Rangel, E.S. (Eds)., *The Handbook of Intercultural Discourse and Communication.* UK: Wiley-Blackwell

[23] Crown, C. L., & S. Feldstein (1985) Psychological correlates of silence and sound in conversational interaction. In Tannen, D., & M. Saville-Troike (eds), *Perspectives on Silence*. Norwood, NJ: Ablex, pp. 31-54.



Edited by Xiaoming Jiang

The use of nonverbal cues in social activities is essential for human daily activities. Successful nonverbal communication relies on the acquisition of rules of using cues from body movement, eye contact, facial expression, tone of voice, and more. As such, this book adds to our understanding of nonverbal behavior by examining state-of-the-art research efforts in the field. The book addresses the classification and training of nonverbal communication with advanced technologies, gives an overview on factors underlying the learning and evaluating of nonverbal communications in educational settings and in digital worlds, and characterizes the latest advancement that uncovers the psychological nature underlying nonverbal communication in conversations. We hope the book will reach a large audience for a variety of purposes, including students and professors in academic institutions for teaching and research activities as well as researchers in industries for the development of communication-related products, benefiting both healthy individuals and special populations.

Published in London, UK

IntechOpen

ISBN 978-1-83962-710-1