From Nonverbal Cues to Perception: Personality and Social Attractiveness

A.Vinciarelli[1,2], H.Salamin[1], A.Polychroniou[1], G.Mohammadi[2,3], and A.Origlia[4]

[1] University of Glasgow, Glasgow (UK)
{vincia,hsalamin,annap}@dcs.gla.ac.uk
[2] Idiap Research Institute, Martigny (Switzerland)
[3] EPFL, Lausanne (Switzerland)
gmohamma@idiap.ch
[4] University of Naples "Federico II", Naples (Italy)
antonio.origlia@unina.it

Abstract. Nonverbal behaviour influences to a significant extent our perception of others, especially during the earliest stages of an interaction. This article considers the phenomenon in two zero acquaintance scenarios: the first is the attribution of personality traits to speakers we listen to for the first time, the second is the social attractiveness of unacquainted people with whom we talk on the phone. In both cases, several nonverbal cues, both measurable and machine detectable, appear to be significantly correlated with quantitative assessments of personality traits and social attractiveness. This provides a promising basis for the development of computing approaches capable of predicting how people are perceived by others in social terms.

Keywords: Social Signal Processing, Nonverbal Behaviour, Personality, Social Attractiveness, Voice Quality, Laughter, Back-Channel.

1 Introduction

Social Cognition has shown that unconscious, automatic cognitive processes of which we are unaware have significant influence on our behaviour and attitudes towards others, especially in zero acquaintance scenarios and early stages of an interaction [25, 26]. The key aspect of the phenomenon is the *perception-action link* [5], namely the automatic and unmediated activation of behavioural patterns after the very simple perception of appropriate stimuli, whether these correspond to verbal messages (e.g., emotionally or idelogically oriented messages), context and environment characteristics (e.g., wheather and time of the day), or nonverbal behavioural cues (e.g., facial expressions and speaking style) [1, 2].

From a computing point of view, the perception-action link is interesting for two main reasons: the first is that it makes human behaviour potentially easier to predict. In fact, if a certain stimulus tends to elicit always the same behavioural pattern, the uncertainty about behaviour under observation can be reduced. For example, when a dyadic interaction participant displays certain nonverbal cues, it becomes easier to predict whether the other participant will back-channel or not [14]. The second reason is that human behaviour can possibly be changed by generating or displaying appropriate stimuli. For example, people that hold certain personality traits tend to spend more time with robots that simulate those same traits [24].

In both cases, the key issue is to understand how people perceive a given stimulus, i.e. what is the social meaning that people tend to attach to it [29]. The reason is that such a meaning seems to determine the behavioural patterns that the stimulus activates [1, 2]. In other words, once we have attached a certain meaning to a stimulus, we tend to automatically react to it always in the same way. Hence, this paper considers the way nonverbal behavioural cues typically used in conversations (speaking style, voice quality, prosody, laughter, back-channel, etc.) influence social perception in zero-acquaintance scenarios. In particular, the paper considers two problems: the first is how nonverbal vocal behaviour influences the perception of personality traits in people that listen to a speaker for the first time [6, 19, 20, 22]. The second is the role of laughter, back-channel and turn-taking in shaping the perception of social and task attractiveness when people talk together for the first time [11, 12].

According to the Social Signal Processing paradigm [27, 28], the ultimate goal of this investigation is the development of approaches capable of predicting automatically not only how individuals perceive one another, but also how they react to the nonverbal cues they mutually display. However, the perception of both personality and social attractiveness has been investigated extensively in psychology as well and the results of this work can provide indications on the role of nonverbal communication in social interactions.

The rest of the paper is organized as follows: Section 2 shows the results obtained in personality perception experiments, Section 3 shows how several nonverbal cues (laughter, back-channel, etc.) shape social and task attractiveness, and Section 4 draws some conclusions.

2 Personality Perception: from Speech to Traits

Personality is the latent construct that accounts for "*individuals' characteristic patterns of thought, emotion, and behaviour together with the psychological mechanisms - hidden or not - behind those patterns*" [6]. Whenever we enter in contact with another person, we quickly develop an impression about her that leads to the attribution of personality traits that, while not being necessarily accurate, still guide our social behaviour, especially in the earliest stages of an interaction [25, 26]. This section investigates the effect of nonverbal vocal behaviour (in particular prosody and voice quality) on such phenomenon.

2.1 Measuring Personality: the Big-Five Model

The personality model most commonly applied in the literature, known as the *Big-Five* Model (BF), relies on five broad dimensions that not only capture most

ID	Statement	ID	Statement
1	This person is reserved	6	This person is outgoing, sociable
2	This person is generally trusting	7	This person tends to find fault with others
3	This person tends to be lazy	8	This person does a thorough job
4	This person is relaxed, handles stress well	9	This person gets nervous easily
5	This person has a few artistic interests	10	This person has an active imagination

Table 1. BFI-10 questionnaire. The table reports the questions of the BFI-10 and the respective IDs.

of the observable differences between people, but also are stable across cultures and situations [18]:

- *Extraversion*: Active, Assertive, Energetic, Outgoing, Talkative, etc.
- *Agreeableness*: Appreciative, Kind, Generous, Forgiving, Sympathetic, etc.
- *Conscientiousness*: Efficient, Organized, Planful, Reliable, Responsible, etc.
- *Neuroticism*: Anxious, Self-pitying, Tense, Touchy, Unstable, Worrying, etc.
- *Openness*: Artistic, Curious, Imaginative, Insightful, Original, etc.

Following the lexical hypothesis, adjectives like those in the list above are the physical trace that personality leaves in language [18]. Hence, personality is *measured* by assigning an individual five numerical scores (one per dimension) that account for how well such adjectives describe the person.

The attribution of the scores is typically performed with questionnaires that consider observable behaviour and characteristics of an individual. In this work, we adopted a short version of the *Big-Five Inventory* (see Table 1) [16]. In particular, the assessors involved in the experiments have been asked to answer the questions of Table 1 after listening for 10 seconds to a person they have never heard before (see below for more details). The answers are selected out of a Likert scale with five possible values, from "*Strongly Disagree*" to "*Strongly Agree*", mapped into the inverval $[-2, 2]$. If Q_i is the answer to question i, the scores for the different dimensions are calculated as follows: Extraversion: $Q_6 - Q_1$, Agreeableness: $Q_2 - Q_7$, Conscientiousness: $Q_8 - Q_3$, Neuroticism: $Q_9 - Q_4$, Openness: $Q_{10} - Q_5$. The resuting range for each dimension is $[-4, 4]$.

2.2 Speech and Personality

The interplay between personality and speech takes two main forms. On one hand, our personality is likely to influence the way we speak and leave *markers* in it [20]. On the other hand, our voice quality and speaking style elicit the attribution of certain personality traits rather than others [19]. While the above has been proposed as a hypothesis roughly one century ago [17], quantitative studies have been performed only since the late seventies. The speech features used in this work are influenced by the results obtained since in the psychological literature and address nonverbal vocal behavioural cues most likely to influence personality perception [22].

The first step of the feature extraction process is the segmentation into *syllables* and *syllable nuclei*. The reason is that these units are less affected by noise and result into more reliable information when processed. In absence of an explicit syllable segmentation, we applied an automatic approach (see [15] for more details) based on the following definition [4, p.275]: "*[a syllable is] a continuous voiced segment of speech organized around one local loudness peak, and possibly preceeded and/or followed by voiceless segments*". The nucleus of the so obtained syllables is the region where the energy is in within $3dB$ from the loudness peak. The length of each syllable is taken as a feature while the ratio between the number of detected syllables and the total duration of an utterance is taken as a measure of speech rate.

Syllable nuclei have been used to extract voice quality related features. The first is *harmonicity*, a measure of the ratio between the energy in the periodic part of the speech signal and the noise (see [3] for the method applied). The second is the *spectral centroid*, perceptually correlated with voice brightness. The centroid is calculated as the average of the frequencies, weighted by their respective energies. The distribution of the energies is used to compute *spectral skewness* (how much energy is above the spectral centroid) and *spectral kurtosis* (how much the energy distribution is different from a Gaussian).

The spectral slope is a measure of the difference between the amount of energy found in the low frequency area and the high frequency area. Spectral slopeness measures have been shown to be effective in emotion discrimination [23]. In this work, spectral tilt is estimated by considering the Long-Term Average Spectrum (LTAS) and taking the slopeness of the trend line computed over the frequency bins. In the employed feature extraction algorithm, the width of the frequency bins is set to 1000 Hz, the low frequency area is comprised between 0 and 1000 Hz and the high frequency area is comprised between 1000 and 4000 Hz. These values are commonly found in the literature.

The frequency values of the first three formants specify the frequencies around which energy concentrates because of the examined vowel. Formant bandwidths describe the area of influence of the considered formants over the spectrum. The two last measurements provide information about syntax-related energy distribution in the syllable nucleus.

Other two voice quality related measures are extracted from syllable nuclei: *Jitter* and *Shimmer*. Jitter is defined as "*the average absolute difference between a period and the average of it and its four closest neighbours divided by the average period*". It is included in the feature set in order to describe the stability of the periodic component inside the syllable nucleus. Shimmer is defined as "*the average absolute difference between the amplitudes of consecutive periods divided by the average amplitude*" and it is included in the features set in order to describe the stability of the energetic component inside the syllable nucleus.

Features concerning the length of syllables and their nuclei are included in the features set to describe the amount of stress in the utterance while energy related features are included as they are a powerful indicator of arousal and dominance levels [21].

The likelihood of dynamic tones (glissando) in a syllable nucleus is estimated for each syllable as the ratio between the actual rate of change of the pitch movement crossing the syllable nucleus and the glissando perception threshold employed in [13]. If the observed rate of change exceeds the threshold, the value of the likelihood is set to 1. This parameter gives an account of whether the pitch movement crossing the syllable nucleus will be perceived as a dynamic tone or as a static one.

For an entire audio clip to be represented, it is necessary to estimate statistical properties of the features above that are extracted from each syllable and nucleus separately. In this work, the statisticals adopted are mean, standard deviation, minimum, maximum and entropy. Different statisticals are used for different features (see caption of Figure 2 for more details).

2.3 Experiments and Results

The goal of the experiments is to show whether the nonverbal cues described in the previous section actually influence the perception of personality or not. A pool of 11 assessors has listened to 640 audio clips of length 10 seconds. The total number of individuals talking in the corpus is 322, with the most represented person speaking in 16 clips and 61% of the individuals speaking only in one clip. For each clip, each of the assessors has filled the questionnaire of Table 1, for a total of 70400 questions answered. The assessments have been performed via an online application and each judge has worked independently of the others. The clips have been presented in a different order to each assessor to avoid tiredness effects. The clips have been assessed in sessions no longer than 30 minutes (no more than two sessions per day) to avoid the lack of concentration resulting from the prolonged repetition of a potentially tedious task. For a given clip, the score for each dimension is the average of the scores assigned by each assessor individually.

The clips have been randomly extracted from the news bulletins broadcasted in Switzerland in February 2005. Only assessors that do not speak the language of the clips (French) have been selected. In this way, the assessors should be influenced only by nonverbal behaviour. The data is emotionally neutral and attention has been paid to avoid words that might be accessible to non-French speakers and have a priming effect (e.g., names of famous people or places) [1, 2].

Figure 1 shows the distribution of the scores across the clips of the corpus. For certain traits (in particular Extraversion and Conscientiousness) the assessments cover a large fraction of the range with sufficient frequency. For others, the distribution is peaked around 0, the value corresponding to the answer "*Neither agree nor disagree*". These results are not surprising because Extraversion and Conscientiousness are well known to be perceived quickly and effectively in the

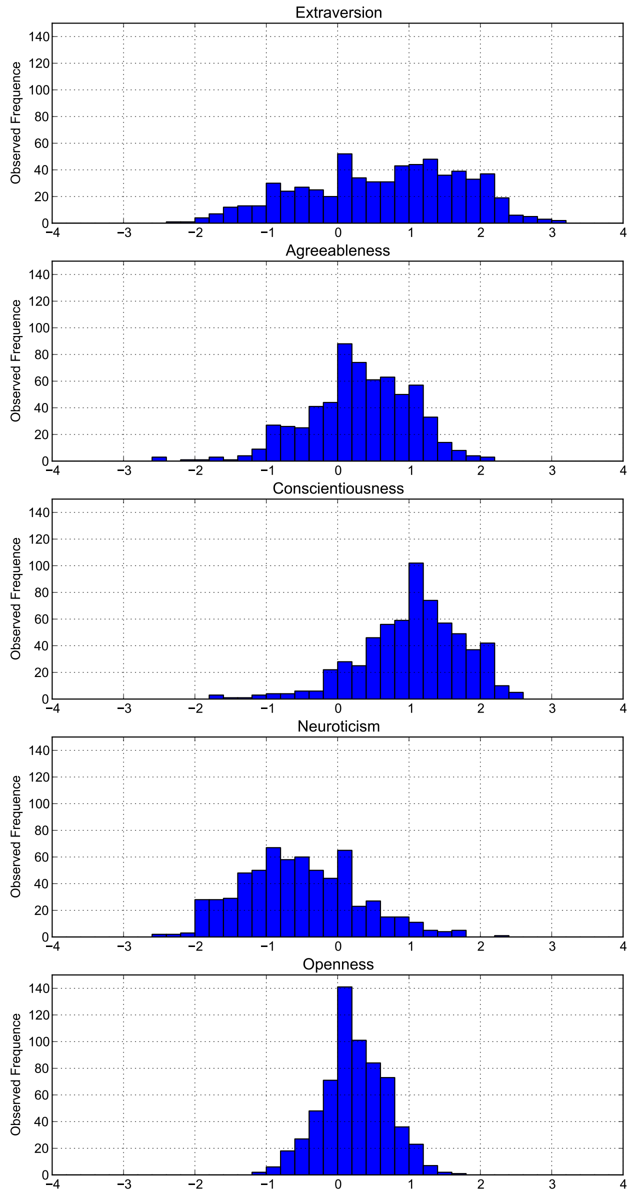

Fig. 1. Distribution of personality scores across different traits.

very first instants after a first encounter [8]. In contrast, the other dimensions are difficult to assess in zero acquaintance scenarios.

Figure 2 shows the correlation between speech cues (see previous section) and personality scores. The horizontal lines in the bar charts correspond to a

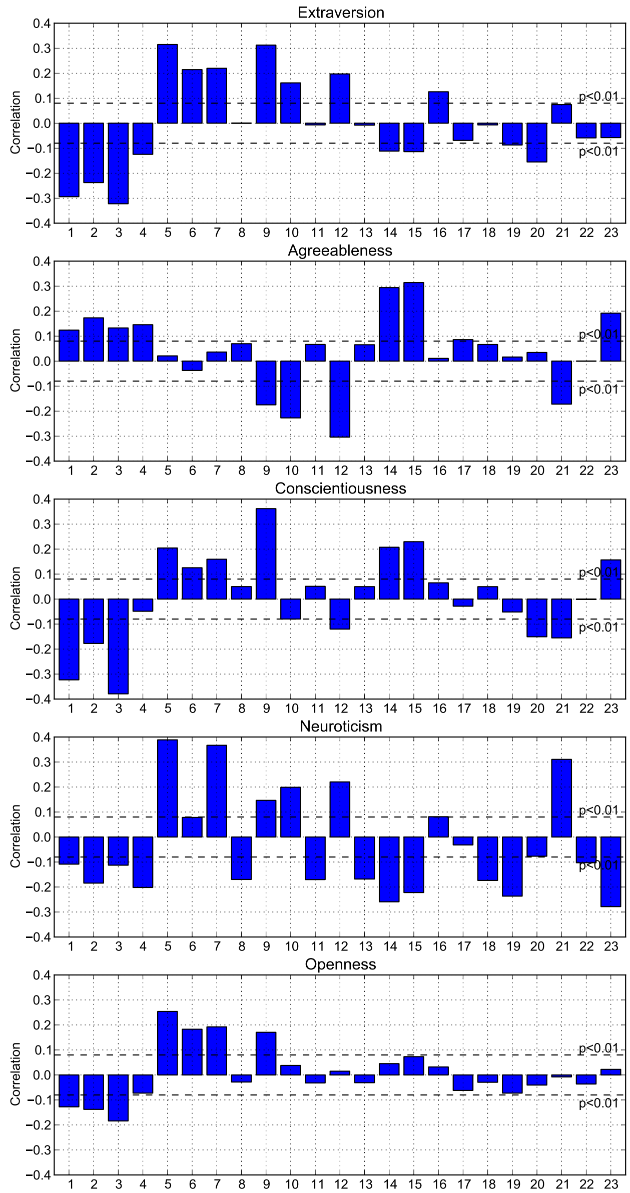

Fig. 2. Correlation between features and personality scores: Nuclei Len. Mean and Entropy (1,2), Syllables Len. Mean and Entropy (3,4), F0 Mean, Stdev., Minimum and Entropy (5,6,7,8), Speech Rate (9), Spectral Centroid and Entropy (10,11), Spectral Tilt and Entropy (12,13), Spectral Skewness and Kurtosis (14,15), Energy Mean, Maximum and Entropy (16,17,18), Jitter and Shimmer (19,20), Mean of F1, F2 and F3 (21,22,23).

significance level of 1%. For all traits, except Openness, at least half of the cues are correlated with p-value lower than 1%. However, the cues with higher correlation change depending on the trait. The perception of Extraversion seems to be dominated by length of syllables and vowels (the shorter syllables and vowels, the higher the assigned Extraversion score), pitch (the higher the pitch, the higher the perceived Extraversion) and speaking rate (the faster the speaker, the more extrovert she sounds). In contrast with the psychological literature, the energy seems not to have an effect, but this is probably due to the relatively small variability of loudness in the corpus.

In the case of Agreeableness, spectral cues dominate perception: brighter voices (higher center of mass in the power spectrum) and higher spectral tilt (higher energy fraction on the fundamental frequency) are perceived as less agreeable. In contrast, voices for which the power spectrum is peakier and tends to be skewed towards higher frequencies are perceived as more agreeable. These latter cues affect the perception of Conscientiousness in the same way, together with the speaking rate (people that talk faster look more competent). In the case of Neuroticism, the higher pitch and first formant means, the higher the score. No evident effects are observed for Openness and the reason is probably that this trait is difficult to assess in a scenario like the one considered in this work, as it is evident in the score distribution narrowly peaked in correspondence of the "*Neither agree nor disagree*" answer.

3 Nonverbal Communication and Social Attractiveness

No other technologies have been accepted as widely as cellular phones in everyday life. In 2005, a mere 15 years after their first apperance in the consumer electronics market, there was one mobile phone subscription every third person in the world, with 82 subscriptions per 100 persons in Europe (the most "mobile" continent) and 19 countries where the number of subscriptions exceeded the size of the population (see [7, 9] for up-to-date figures). The ubiquitous diffusion of mobile phones is a major change in the way we develop and maintain our social ties [10]. However, the impact on conversation, the primary site of human sociality, has not been investigated extensively. The results presented in this section try to address such a gap by showing how a number of nonverbal behavioural cues influence the perception of social and task attractiveness [11, 12] between unacquainted people talking on the phone.

3.1 Data and Scenario

The experiments of this section have been performed over a collection of 26 phone calls between unacquainted individuals. The total number of involved subjects is 52 (no person participates in more than one call). During the data collection, the subjects are invited to the laboratory, but they do not meet one another before the call. The conversations are centered around the *Winter Survival Scenario* (WSS): the two persons play the role of members of a rescue team that must

ID	Statement	ID	Statement
1	I think (s)he could be a friend of mine	1	I couldn't get anything accomplished with him (her)
2	I would like to have a friendly chat with him (her)	2	(S)he is a typical goof off when assigned a job to do
3	It would be difficult to meet and talk with him (her)	3	I have confidence in his (her) ability to get the job done
4	We could never establish a personal firednship with each other	4	If I wanted to get things done, I could probably depend on him (her)
5	(S)he just would not fit into my circle of friends	5	(S)he would be a poor problem solver
6	(S)he would be pleasant to be with	6	I think studying with him (her) would be impossible
7	I feel I know him (her) personally	7	You could count on him (her) getting the job done
8	(S)he is personally offensive to me	8	I have the feeling (s)he is a very slow worker
9	I do not care if I ever get to meet him (her)	9	If we put our heads together, i think we could come up with some good ideas
10	I sometimes wish I were more like him (her)	10	(S)he would be fun to work with

Table 2. Social (left column) and task (right column) attractiveness questionnaires.

support a group of people that have survived a plane crash in Northern Canada. It is winter (temperatures around -40^oC) and the survivors have extracted 12 items from the plane. However, they have to leave the place of the crash and they can bring only part of the 12 items. During the call, the rescue members are expected to identify the items that maximize the chances of survival. The subjects are paid 6 British Pounds for their participation. Furthermore, they get 3 extra Pounds each time they select a good item, but are penalized by the same amount each time they select a wrong one (in any case, a minimum payment of 6 Pounds is guaranteed).

The conversations have been captured with two cellular phones (Nokia N900) that record not only what the subjects say (via both microphone and speakers), but also the movement of the phones via accelerometers, gyroscopes and magnetoscopes. All signals are synchronized to allow a multimodal analysis of subjects behaviour. The conversations have been annotated manually in terms of *turns* (who speaks when), *silences* (when none of the subjects talks), *laughter* (both individual and common) and *back-channel* (short vocal bursts that react and/or accompany the speech of others without attempts of grabbing the floor).

Furthermore, the subjects have filled the questionnaires proposed in [11, 12], aimed at assessing both social (how much we enjoy interacting with another person) and task (how much we like to work with another person) attractiveness of

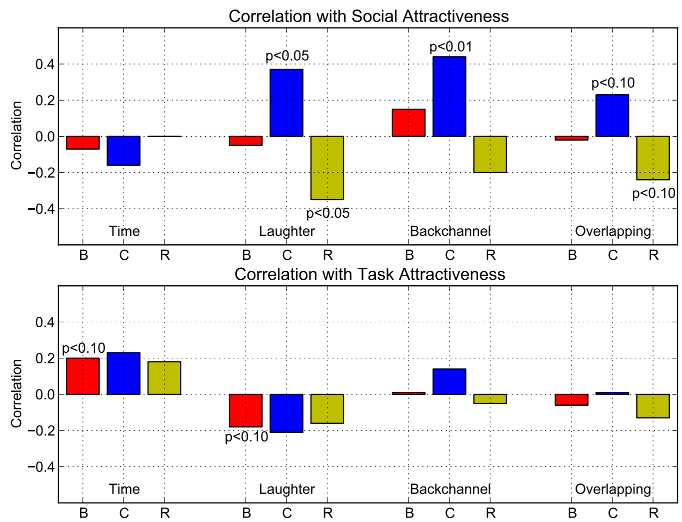

Fig. 3. Recognition performance as a function of the clips length. The right plot shows the results for the two classes separately.

their interlocutor (see Table 2). Like in the case of the BFI-10, the questions are associated to Likert scales with five possibile answers, from "*Strongly Disagree*" to "*Strongly Agree*", mapped into the interval $[-2, 2]$ (see Table 2 for more details). If Q_i is the answer to question i, the social attractiveness is calculated as $Q_1+Q_2-Q_3-Q_4-Q_5+Q_6+Q_7-Q_8-Q_9+Q_{10}$. In the case of the task attractiveness, the sum is as follows: $-Q_1-Q_2+Q_3+Q_4-Q_5-Q_6+Q_7-Q_8+Q_9+Q_{10}$.

3.2 Experiments and Results

Figure 3 shows the correlation between the nonverbal cues annotated in the data and the attractiveness scores obtained from the questionnaires. The cues considered are the fraction of time a person talks during a call, the number of times a person laughs, the number of times a person performs back-channel, and the number of times there is overlapping speech. For each cue, the correlation is calculated for all speakers (B), only for the subjects that call (C) and only for the subjects that receive (R).

The upper chart shows the results for the social attractiveness. When all speakers are taken into account, no correlation reaches a p-value lower than 10%. However, the situation changes when considering separately the subjects that call and those that receive. The former tend to be more socially attractive when they laugh more ($p < 5\%$), when they show more back-channel ($p < 1\%$) and when there is more overlapping speech ($p < 10\%$). In the case of the subjects that receive the situation is opposite, namely they are more appreciated if they laugh less ($p < 5\%$) and there is less overlapping speech ($p < 10\%$). With the exception of the fraction of time people talk, the results seem to suggest that the expectations are different depending on whether a person calls or is called. However, the effect might depend on the particular scenario adopted and it should be confirmed by collecting more data.

In the case of task attractiveness, no correlations reach a p-value lower than 10% and it is not possible to say whether there is a difference between being the person that calls or the one that receives. However, correlations with acceptance level lower than 10% are obtained when considering all of the subjects. In particular, people that talk more, but laugh less, seem to be more attractive when it is necessary to accomplish a task. In this case as well, the number of calls (26) and subjects (52) is relatively low and more solid evidence can be obtained only by collecting more data.

4 Conclusions

This paper has investigated how people perceive a number of nonverbal behavioural cues in social terms. Two problems have been considered: the first is how people attribute personality traits to speakers they have never heard before. The second is the perception of social and task attractiveness in phone calls between unacquainted individuals.

In the first case, the experiments have focused on voice and speaking style characteristics (voice quality, speaking rate, etc.) and show that a large number of nonverbal cues correlate to a statistically significant degree with personality assessments. In the second case, laughter and backchannel have been shown to influence significantly the perception of social attractiveness, but they do it in a different way depending on whether a person calls or is called. To the best of our knowledge, it is the first time that someone reports about such a lack of symmetry between two people involved in a phone conversation.

The results presented in this work are interesting under two main respects. The first is that they provide further information about nonverbal cues involved in important social phenomena. The second is that they provide useful indications about the cues to be detected in order to develop automatic approaches capable of predicting the way people perceive others.

Acknowldgments: The research that has led to this work has been supported in part by the European Community's Seventh Framework Programme (FP7/2007-2013), under grant agreement no. 231287 (SSPNet), in part by the Swiss National Science Foundation via the National Centre of Competence in Research IM2 (Information Multimodal Information management) and in part by the Finnish Funding Agency for Technology and Nokia Research Center Finland as part of Human Emotional Interaction project 2010-2011.

References

1. J.A. Bargh, M. Chen, and L. Burrows. Automaticity of Social Behaviour: Direct Effects of Trait Construct and Stereotype Activation on Action. *Journal of Personality and Social Psychology*, 71(2):230–244, 1996.
2. J.A. Bargh and E.L. Williams. The Automaticity of Social Life. *Current Directions in Psychological Science*, 15(1):1–4, 2006.
3. P. Boersma and D. Weenink. Praat: doing phonetics by computer [Computer program]. Version 5.2.40, 2011.
4. C. D'Alessandro and P. Mertens. Automatic pitch contour stylization using a model of tonal perception. *Computer Speech and Language*, 9(3):257–288, 1995.
5. A. Dijksterhuis and J.A. Bargh. The Perception-Behavior Expressway: Automatic Effects of Social Perception on Social Behavior. In M. P. Zanna, editor, *Advances in Experimental Social Psychology*, pages 1–40. 2001.
6. D. C. Funder. Personality. *Annual Review of Psychology*, 52:197–221, 2001.
7. ITU. The World in 2010: ICT Facts and Figures. Technical report, International Telecommunication Union.
8. C.M. Judd, L. James-Hawkins, V. Yzerbyt, and Y. Kashima. Fundamental Dimensions of Social Judgment: Unrdestanding the Relations Bteween Judgments of Competence and Warmth. *Journal of Personality and Social Psychology*, 89(6):899–913, 2005.
9. K. Kalba. The Global Adoption and Diffusion of Mobile Phones. Technical Report December, Center for Information Policy Research Harvard University, 2008.
10. R. Ling. *New Tech, New Ties. How Mobile Communication is Reshaping Social Cohesion*. MIT Press, 2008.
11. J.C. McCroskey and T.A. McCain. The measurement of interpersonal attraction. *Speech Monographs*, 41(3):261–266, 1974.
12. L. McCroskey, J. McCroskey, and V. Richmond. Analysis and Improvement of the Measurement of Interpersonal Attraction and Homophily. *Communication Quarterly*, 54(1):1–31, 2006.
13. P. Mertens. The Prosogram: Semi-Automatic Transcription of Prosody Based on a Tonal Perception Model. In *Proceedings of Speech Prosody*, 2004.
14. L. P. Morency, I. de Kok, and J. Gratch. A Probabilistic Multimodal Approach for Predicting Listener Backchannels. *Journal of Autonomous Agents and Multi-Agent Systems*, 20(1):70–84, 2010.
15. M. Petrillo and F. Cutugno. A syllable segmentation algorithm for english and italian. In *Proc. of Eurospeech*, pages 2913–2916, 2003.
16. B. Rammstedt and O.P. John. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41(1):203–212, 2007.
17. E. Sapir. Speech as a personality trait. *The American Journal of Sociology*, 32(6):892–905, 1927.
18. G. Saucier and L.R. Goldberg. The language of personality: Lexical Perspectives on the Five-Factor Model. In J.S. Wiggins, editor, *The Five-Factor Model of Personality*. 1996.
19. K.R. Scherer. Personality inference from voice quality: The loud voice of extroversion. *European Journal of Social Psychology*, 8:467–487, 1978.
20. K.R. Scherer. Personality markers in speech. In *Social markers in speech*, pages 147–209. Cambridge University Press, Cambridge, 1979.

21. K.R. Scherer, T. Johnstone, and G. Klasmeyer. Vocal expression of emotions. In R.J. Davidson, K.R. Scherer, and H.H. Goldsmith, editors, *Handbook of Affective Sciences*, pages 433–456. Oxford University Press, 2003.
22. K.R. Scherer and U. Scherer. Speech behavior and personality. In *Speech evaluation in psychiatry*, pages 115–135. Grune & Stratton, New York, 1981.
23. L. Tamarit, M. Goudbeek, and K. Scherer. Spectral slope measurements in emotionally expressive speech. In *Proceedings of Speech Analysis and Processing for Knowledge Discovery*, 2008.
24. A. Tapus, C. Tapus, and M. Matarić. User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, 1(2):169–183, 2008.
25. J. S. Uleman, L. S. Newman, and G. B. Moskowitz. *People as flexible interpreters: Evidence and issues from spontaneous trait inference*, volume 28, pages 211–279. Elsevier, 1996.
26. J. S. Uleman, S. A. Saribay, and C. M. Gonzalez. Spontaneous inferences, implicit impressions, and implicit theories. *Annual Reviews of Psychology*, 59:329–360, 2008.
27. A. Vinciarelli, M. Pantic, and H. Bourlard. Social Signal Processing: Survey of an emerging domain. *Image and Vision Computing Journal*, 27(12):1743–1759, 2009.
28. A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schroeder. Bridging the Gap Between Social Animal and Unsocial Machine: A Survey of Social Signal Processing. *IEEE Transactions on Affective Computing (to appear)*, 2012.
29. T. Wharton. *Pragmatics and Non-Verbal Communication.* Cambridge University Press, 2009.