

Adults with Asperger syndrome are less sensitive to intonation than control persons when listening to speech

To appear in *Culture and Brain*, DOI 10.1007/s40167-016-0035-6

Martine Grice<sup>a</sup>, Martina Krüger<sup>a</sup> and Kai Vogeley<sup>b,c</sup>

a Institute of Linguistics, Department of Phonetics, University of Cologne, Germany

b Department of Psychiatry and Psychotherapy, University Hospital Cologne, Germany

c Research Center Juelich, Institute of Neuroscience and Medicine, Cognitive Neuroscience (INM-3), Germany

Martine Grice (corresponding author), Institute for Linguistics, Department of Phonetics, University of Cologne, Herbert-Lewin-Str. 6, D-50931 Köln, Germany. Phone +49 221 - 470 4253, Fax +49 221 - 470 5938, Email: martine.grice@uni-koeln.de

Running title: Reduced sensitivity to intonation in AS

Grant sponsor: Deutsche Forschungsgemeinschaft, Excellent Research Programme; UoC

Emerging Groups (Grice: ZUK81/1) UoC Forum (Vogelely).

## Abstract

In conversation, speakers typically draw attention to items that are meant to be informative by pronouncing the words referring to these items in a particular way. These words have a distinct intonation, and are accented – typically involving a rise or fall in vocal pitch on the stressed syllable. Listeners use this information to know which part of the sentence is new, and therefore worthy of attention.

In a perception study, adults with Asperger Syndrome (AS) and a group of control persons were instructed to rate the informativeness of words, based on how they sounded. The AS group showed a reduced sensitivity to intonation and subsequently based their judgement less on the way the word was pronounced and more on word frequency and semantic features of the words themselves. This finding is in concordance with a general reduced sensitivity to non-verbal cues in social encounters and to a propensity towards literal interpretation in the group of persons with AS.

*Keywords:* Asperger syndrome (AS); Autism Spectrum Disorders (ASD); prosody; intonation; information status; informativeness; perception; word frequency.

## Introduction

Intonation is an essential component of conversation, conveying affective and grammatical functions simultaneously, making its interpretation highly context dependent. A key function of intonation is to mark out the informative parts of utterances. This is achieved by highlighting words that are new to the discourse, signalling that they refer to information that is considered to be more worthy of attention than words referring to information that is given in the discourse. In West Germanic languages such as English and German, highlighted words are not only louder, longer and more carefully pronounced than the surrounding words (Mücke & Grice 2014; Breen et al. 2010), but they are typically accented, involving a pitch movement on or in the vicinity of the stressed syllable of the highlighted word. The main (nuclear) pitch accent is structurally the most prominent in a phrase (Cruttenden 2006; Ladd 2008), and the placement of this accent is the focus of this study.

To illustrate how this accentuation expresses informativeness, the example in (1) shows a context, A, in which the concept of *banana* is introduced, followed by two alternative answers, B and B'. The main accent is in capital letters.

(1) A. Thomas bought a BANANA.

B. He LIKES bananas.

B'. \*He likes BANANAS.

Answer B, where *bananas* is unaccented, would sound appropriate in this context, whereas answer B', where *bananas* is accented, would not. As the concept of *banana* is already given in A, it needs less highlighting (Baumann et al. 2015; Cruttenden 2006; Pierrehumbert & Hirschberg 1990). This is because speakers account for what their listeners already know (Chafe 1976; Galati & Brennan 2010; Gregory et al. 2001), using intonation to draw the listeners' attention to the new parts of the message.

The ability to appreciate the mental state of the conversational partner, in other words: “mentalizing”, is required for the adequate use of intonation to express informativeness (Kaland et al. 2013; Clark & Murphy 1982; Galati & Brennan 2010). This is in line with reports that mentalizing difficulties affect pragmatic language abilities rather than more formal aspects of language, such as syntax (but see Tager-Flusberg (2000) for exceptions) and even other aspects of grammatical prosody (such as word stress and phrasing, Chevallier et al. 2009).

It has long been established that adults with Asperger Syndrome (AS) produce atypical intonation in their own speech (Kanner 1943; Asperger 1944; Tager-Flusberg et al. 2005) which is reported as sounding “robotic” in some studies and, conversely, “singsong” in others (Shriberg et al. 2001; Paul et al. 2005; see also Filipe et al. 2014 for a discussion of variability in F0 contours in AS). They also show inappropriate highlighting (Shriberg et al. 2001) and a failure to use pitch to mark information structure (DePape et al. 2012). This appears to be related to a failure in AS to react adequately to subtle non-verbal cues when navigating social encounters, while the ability to explicitly use verbal information is well preserved (Senju et al. 2009; Kuzmanovic et al. 2011). These nonverbal features include not only visually accessible signals, but also audible signals in the speech itself, including the intonation.

In comparison to the wealth of studies on the production of intonation in adults with AS, there are relatively few studies on perception. These studies have mainly been concerned with the recognition of emotion and affect (inter alia Chevallier et al. 2011; Golan et al. 2007; Kleinman et al. 2001; Korpilahti et al. 2007), reporting an over reliance on the lexical content, especially when information is conflicting (Lindner & Rosén 2006; Stewart et al. 2013; Kuzmanovic et al. 2011). Hobson et al. (1988) as cited in Kleinman et al. (2001) on the other hand showed that individuals with autism were better at identifying vocal emotional

expressions (humming, sighing) than verbal emotional expressions, which suggests that adding a verbal dimension to stimuli influences their judgements.

The few studies available on the perception of accentual highlighting in adults and adolescents with AS are concerned with *contrast* (Chevallier et al. 2010; Hesling et al. 2010; Paul et al. 2005), and *emphasis* or *focus* (Globerson et al. 2015). Hesling et al. (2010) investigated several aspects of prosody perception (and production) and found significantly lower scores in the autism group as compared to a control group. This included a task that assessed the comprehension of contrastive accentual highlighting, where subjects had to identify the item that the speaker was missing in sentences like “I wanted BREAD and apples”. Paul et al. (2005) found comparable results for the comprehension of contrastive accentual highlighting in sentences like “I want CHOCOLATE ice cream” vs. “I want chocolate ICE CREAM”. However, Chevallier et al. (2010) found no differences between groups in perceiving and interpreting contrastive accentual highlighting of the word “or” as an exclusive disjunction in sentences such as “There is a window OR a monkey”. Furthermore, Paul et al. (2005) also reported a problem with perceiving highlighting in general, including word stress, e.g. PERmit vs. perMIT, although two later studies with adults and children failed to confirm this (Chevallier et al. 2009 and Grossmann et al. 2010 respectively). Thus, perception studies to date provide a somewhat inconclusive picture, despite the fact that there is a reported effect of autistic traits on the perception of accentual highlighting in the general population (Bishop 2012).

In this study we focus on the perception and interpretation of accentual highlighting, but without this highlighting being necessarily contrastive. Instead we are concerned with the ability to gauge the informativeness of words in an utterance based on their intonation. We hypothesized an interaction between Group (AS vs. control) and Accentuation (Acc. vs.

NoAcc.) assuming that persons with AS would be able to read less informativeness into the stimuli on the basis of accentuation than normal non-affected control persons.

## Method

### *Participants*

The clinical group comprised 37 adults with AS (23 male, aged 23-54, mean age 42.5, SD=9.7; ICD-10: F84.5). They were diagnosed and recruited in the Autism Outpatient Clinic at the Department of Psychiatry at the University Hospital Cologne (Germany). Diagnoses were made independently by two specialized clinicians, employing ICD-10 criteria, supplemented by an extensive neuropsychological assessment. The control group comprised 37 typically developed adults (24 male, aged 21-65, mean age 38.4, SD=14.4) recruited from the general population (e-mail distribution list with an invitation to share the link with others). Statistical tests revealed no significant differences between groups regarding gender (Pearson Chi-square test:  $p=0.8092$ ) and age (two-sample T-Test:  $p=0.1552$ ).

All subjects reported normal hearing abilities with no known hearing impairments. On the basis of their professional qualifications and their current employment status, we could infer that 76% of subjects in the AS group and 65% in the control group passed the German *Abitur*. All remaining subjects in the AS group and all but four in the control group had at least successfully finished secondary school. On this basis we could estimate that no member of either group is likely to have a deficit corresponding to an IQ of 70 or lower according to ICD10.

The study has been approved by the local ethics committee of the Medical Faculty at the University of Cologne and has been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All persons gave their written informed consent prior to their inclusion in the study.

## *Materials*

Speech materials were taken from recordings made by ten native German speakers and comprised 28 read sentences of the type “I’ll take the banana with me”. The appendix illustrates the types of context provided for elicitation: in (a) the target word *banana* is new, in (b) it is given (Röhr & Baumann 2010; 2011). Contexts were varied so as to elicit unaccented and accented target words, and in the latter case, a range of accents. All target sentences had the same structure, comprising 5 or 6 words, with the target word in the penultimate position of the sentence, followed by a verbal particle. They were presented auditorily to the participants in isolation (without the context of the story used for elicitation).

Target words were common nouns with feminine gender (Ballade *ballad*, Banane *banana*, Dame *lady*, Lawine *avalanche*, Rosine *raisin*) and proper nouns (*Janina*, *Nina*, *Dr. Bahber*, *Dr. Bieber*, *Romana*). Of these, four nouns were inanimate, the remaining six were animate (persons). In total, 11 of the 28 target sentences contained inanimate nouns. The word frequency differed considerably across the target words as estimated using the SUBTLEX Corpus (Brysbaert 2011). Neither frequency nor animacy were specifically controlled for, as the material for the production study focused only on the phonetic properties of the words. All words had penultimate lexical stress (= a default stress on the second to last syllable, e.g. BaNAne; RoMAna) with a long vowel in the stressed syllable and voiced consonants in the stressed and post-stressed syllables.

## *Procedure*

The experiment was conducted online via SosciSurvey, a software package for scientific online questionnaires (Leiner 2014). Subjects were invited by email with a link to the survey and were able to complete the survey at home. In the preface of the survey, subjects were familiarized with the structure of the stimuli and were instructed to pay attention to how the noun in the sentence sounded. The preface also contained example test sentences that the

subjects could replay as often as needed in order to adjust sound settings and ensure a satisfactory sound quality. In case of questions, problems or doubts regarding any aspect of the experiment, participants were advised to email the principal investigator prior to participating in the experiment. In the main part of the survey, written instructions “The noun sounds as if it were...” were followed by two repetitions of each auditory stimulus. Subjects then were asked to rate informativeness by placing a roll bar on a continuous line between two end-points, labelled as ‘known’ and ‘new’. The responses were encoded as interval data (1= ‘known’, 100= ‘new’). Subjects rated each test sentence three times in randomized order, amounting to a total of 84 responses in the main part of the experiment.

#### *Statistical analysis*

The results were analysed using R (R Development Core Team 2011) and lme4 (Bates et al. 2012) to perform a linear mixed effects analysis with Response (ranging from 1 to 100) as a dependent measure. Visual inspection of residuals revealed some violation of normality and homoscedasticity. These violations can be disregarded for this analysis due to the size of this data set (for a discussion on this issue see Faraway 2004).

The factors Group (AS vs. control) and Accentuation (Acc. vs. NoAcc.) were included as fixed effects along with Repetition (1st, 2nd or 3rd repetition), Animacy of the Words and Word frequency (log values from the Subtlex Corpus (Brysbaert et al. 2011)). Interactions between Group x Accentuation, Group x Repetition, Group x Animacy and Group x Word Frequency were tested. We included Speaker, Subject and Target word as random intercepts. We report p-values based on Likelihood Ratio Tests comparing a reduced model (without the effect in question) against the full model (including all effects). The random effects structure was constant across these comparisons.

## Results

### *Interaction Group x Accentuation*

There was a significant interaction of Group (AS vs. control) and Accentuation (Acc. vs. NoAcc.) ( $\chi^2(1)=67.35$ ,  $p<0.001$ ). The ratings in the AS group increased (=interpreted as newer) from unaccented to accented target words only slightly by 2.02 (SE= 1.46), but increased substantially in the control group by 17.07 (SE=1.83)).

(FIGURE-1)

The mixed effect model comparison revealed that Accentuation had a different influence on the ratings in the two groups: When comparing accented and unaccented words, this difference was significantly larger in the control group than the AS group (Figure 1). Note that the scale in Figure 1 shows values up to 50%. This is because the target words were produced with a definite article (e.g. “the banana”), which signals that the information is not brand new. Singular nouns require an article, and in order to control for the effect of prosody alone, it was necessary to use the same one for all tokens. Since the target words consisted of both proper nouns and common nouns, the definite article was the only article available to us. The indefinite article cannot be used with proper nouns in German.

### *Interaction Group x Animacy*

There was a significant interaction of Group and Animacy ( $\chi^2(1)=7.86$ ,  $p=0.005$ ). The ratings of the AS group increased substantially by 12.44 from Inanimate to Animate (SE=1.3) while the ratings of the control group increased less strongly by only 7.7 (SE=1.69). This indicates that the category of Animacy of the target words plays a more important role in the rating of newness in the AS group as compared to the control group.

### *Interaction Group x Word Frequency*

There was also a significant interaction of Group and Word Frequency ( $\chi^2(1)=8.95$ ,  $p=0.003$ ). The ratings of the AS group slightly decreased by 2.34 per log frequency (SE=1.71), while

the ratings of the control group did not decrease substantially (decrease of 0.09 per log frequency (SE=0.76)). Thus, word frequency affects the rating differently across the two populations. The difference between frequent and infrequent words is larger in the AS group. Unlike the control group, the AS group rated words that are frequently used in speech (e.g. *banana*) as more given than words that are less frequent (e.g. *avalanche*), irrespective of their prosodic characteristics. Figure 2 plots mean ratings for each word against word frequency, as estimated from the Subtlex Corpus (Brysbaert 2011). The trend line for values of the control group is based on  $y = -0,2094x + 37,242$  ( $R^2 = 0,0006$ ). The trend line for values of the HFA/AS group is based on  $y = -2,8336x + 31,682$  ( $R^2 = 0,1512$ ).

(FIGURE-2)

#### *Interaction Group x Repetition*

No significant interaction of Group and Repetition was found ( $\chi^2(1)=2.78$ ,  $p=0.095$ ).

Repetition as a factor in itself was not significant ( $\chi^2(1)=1.61$ ,  $p=0.2052$ ).

Table 1 summarizes the results, informally reflecting the trends in the statistical analysis.

(TABLE 1)

## Discussion

Results revealed that when making judgements as to the informativeness of words and the concepts these words refer to, the two groups behaved substantially differently: While both groups rated accented words as sounding newer than unaccented words, the difference in ratings for the control group was considerably larger than for the AS group, indicating that persons with AS took intonation into account to a much lower degree during the processing of their judgment as compared to the control group.

Notably, as a by-product of the study and although that was not the explicit research question here, we are able to identify two different compensation strategies that are employed

by individuals with AS in order to make sense of the concept of givenness, these two candidates are related to word frequency and animacy. It is interesting that the two potential compensation mechanisms do not have anything to do with the “intersubjective” social encounter, here: other person as the producer of the utterance, but both are related to “objective” features of the word, either word frequency or the semantic category of animacy.

First, the AS group (but not the control group) rated objectively less frequent words (e.g. *avalanche*) as newer than objectively frequent ones (e.g. *banana*; as estimated using the SUBTLEX Corpus (Brysbaert 2011)), regardless of the way the word was pronounced. Thus, they appear to make use of their own generic knowledge of a word and its objective frequency, resulting in higher (i.e. newer) ratings for less frequent and therefore less familiar words, apparently paying less attention to the way the words were pronounced. It has to be pointed out that our online experimental setup and questionnaire did not permit us to completely rule out verbal intelligence differences across the two populations. However, since both groups show a sensitivity to aspects of language – intonation, word frequency and animacy all rely on language related competence – we consider it to be unlikely. Nonetheless, this issue needs to be considered as a limitation of the study and will be taken into account in future studies.

Second, although both groups showed to some extent sensitivity in their judgments to the category animacy – words referring to inanimate objects tending to be rated as more given than words with an animate referent – this effect was stronger in the AS group. This is another indication that the lexical properties of words could play a greater role than their intonation for persons with AS if they compete with intonation.

The over-reliance on lexical content in AS subjects can be linked to the deficit in the capacity to understand metaphors or irony, to “read between the lines” and the extreme literalism for which people with autism are known (Dewey & Everard 1974; Happé 1993;

Mitchell et al. 1997, Riedel et al. 2014), whereas their reduced sensitivity to intonation reflects a reduced sensitivity to non-verbal cues in general (Georgescu et al. 2013; Kuzmanovic et al. 2014). These deficits can be subsumed under the umbrella term of mentalizing, or perspective-taking, deficits in persons with AS (Frith 2012; Happé & Frith 2014; Chung et al. 2014), although it is still an open question whether the social cognitive deficit is really the core of the autistic phenomenology (Leekam 2016). Moreover, difficulties encountered by the AS subjects in our study may not only lie in the interpretation of the speaker's intentions but also in the complexity of the speech material, which has been shown to lead to impaired performance in perception experiments (Järvinen-Pasley et al. 2008; Lepistö et al. 2008; O'Connor 2012; Singh & Harrow 2014; Tager-Flusberg 2000): The speech material included a range of different accents, rather than one single accent type (as is often the case in experiments on accentual prominence, e.g. Globerson et al. 2015), such that the accentuation was not reflected in one acoustically identifiable pattern.

Since intonation, and in particular accentuation, varies in a context dependent way, whereas animacy and frequency of words is context independent, our results can be taken to support findings in other domains in terms of underlying cognitive mechanisms: It has been shown that people with AS have difficulties in perceiving signals that vary in a context dependent way, corroborating the concept of a *weakness of central coherence* (Frith 1989; 2012). It appears that instead of making sense of information from intonation “intersubjectively”, these subjects attempt to recruit intellectual or “objective” compensation mechanisms, using lexical properties of words that are invariable to calculate the degree of givenness of an object or person referred to with this word.

## Acknowledgements

We thank Christine Röhr for making available the materials used for experiments reported in Röhr & Baumann (2011) and Baumann, Röhr & Grice (2014). We would also like to thank Bodo Winter for help with the statistical analysis and Francesco Cangemi for discussion of the implications. All errors are the responsibility of the authors. This work was partly funded by the German Research Council as part of University of Cologne's Excellent Research support program awarded to Martine Grice (UoC Emerging Group, ZUK 81/1) and to Kai Vogeley (UoC Forum).

The authors declare that they have no conflict of interest.

## References

- Asperger, H. (1944). Die „Autistischen Psychopathen“ im Kindesalter. *European Archives of Psychiatry and Clinical Neuroscience*, 117(1), 76-136.
- Bates, D., Maechler, M. & Bolker, B. (2012). lme4: Linear mixed-effects models using Eigen and Eigenfaces. R package version 0.999999-0. <http://CRAN.R-project.org/package=lme4>
- Baumann, S., Röhr, C. T. & Grice, M. (2015). Prosodische (De-)Kodierung des Informationsstatus im Deutschen. *Zeitschrift für Sprachwissenschaft* 34(1), 1-42.
- Bishop, J. (2012). Focus, prosody, and individual differences in „autistic“ traits: Evidence from cross-modal semantic priming. UCLA Working Papers in Phonetics, No. 111, 1-26.
- Breen, M., Fedorenko, E., Wagner, M., Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, 25 (7), 1044-1098.
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bólte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology* 58, 412-424.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, C. (ed.), *Subject and Topic* (pp. 25-55). New York: Ac. Press.
- Chevallier, C., Noveck, I., Happé, F. & Wilson, D. (2009). From acoustics to grammar: Perceiving and interpreting grammatical prosody in adolescents with Asperger Syndrome. *Research in Autism Spectrum Disorders*, 3(2), 502-516.
- Chevallier, C., Wilson, D., Happé, F. & Noveck, I. (2010). Scalar inferences in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders* 40, 1104-1117.
- Chevallier, C., Noveck, I., Happé, F. & Wilson, D. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia* 49(3), 507-517.

- Chung, Y.S., Barch, D., Strube, M. (2014). A meta-analysis of mentalizing impairments in adults with schizophrenia and autism spectrum disorder. *Schizophrenia Bulletin* 40(3), 602-616.
- Clark, H. H. & Murphy, G. L. (1982). Audience design in meaning and reference. In Le Ny, J.-F. & Kintsch, W. (eds.), *Language and Comprehension* (pp. 287-299). North Holland, New York.
- Cruttenden, A. (2006). The De-accenting of Given Information: a Cognitive Universal? In Bernini, G. & Schwartz, M. L. (eds.), *Pragmatic Organization of Discourse in the Language of Europe* (pp. 311-355). The Hague: Mouton de Gruyter.
- DePape, A.-M. R., Chen, A., Hall, G. B. C. & Trainor, L. J. (2012). Use of prosody and information structure in high functioning adults with Autism in relation to language ability. *Frontiers in Psychology* 3, Article 72.
- Dewey, M. A. & Everard, M. P. (1974). The near-normal autistic adolescent. *Journal of autism and childhood schizophrenia* 4 (4), 348-356.
- Faraway, J. J. (2004). *Linear models with R*. CRC Press.
- Filipe, M., Branco, P., Frota, S. Castro S. L. & Vicente S. (2014). Atypical Prosody in Asperger Syndrome: Perceptual and Acoustic Measurements. *Journal of Autism and Developmental Disorders*. 44. 1972-1981.
- Frith, U. (2012). Why we need cognitive explanations of autism. *The Quarterly Journal of Experimental Psychology* 65(11), 2073-92.
- Galati, A. & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language* 61, 35-51.
- Georgescu A. L., Kuzmanovic B., Schilbach L., Tepest R., Kulbida R., Bente G., Vogeley K. (2013). Neural correlates of “social gaze” processing in high-functioning autism under systematic variation of gaze duration. *Neuroimage: Clinical* 3, 340-351.

- Globerson, E., Amir, N., Kishon-Rabin, L., & Golan, O. (2015). Prosody recognition in adults with high-functioning autism spectrum disorders: from psychoacoustics to cognition, *Autism Research*, 8(2):153-63.
- Golan, O., Baron-Cohen, S., Hill, J. J. & Rutherford, M. D. (2007). The 'Reading the Mind in the Voice' Test-Revised: A Study of Complex Emotion Recognition in Adults with and Without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders* 37, 1096–1106.
- Gregory, M. L., Jurafsky, D., & Healy, A. F. (2001). The role of the Hearer in Durational Shortening. Poster presented at *Architectures and Mechanisms for Language Processing (AMLaP) 7*. Saarbrücken, Germany.
- Happé, F. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition* 48, 101-119.
- Happé F., Frith, U. (2014). Annual research review: Towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry* 55(6), 553-557.
- Hesling, I., Dilharreguy, B., Peppé, S., Amirault, M., Bouvard, M. & Allard, M. (2010). The Integration of Prosodic Speech in High Functioning Autism: A Preliminary fMRI Study. *PLoS one* 5(7): e11571.
- Hobson, R. P., Ouston, J., & Lee, A. (1988). Emotion recognition in autism: Coordinating faces and voices. *Psychological Medicine* 18, 911–923.
- Järvinen-Pasley, A., Peppé, A., King-Smith, G. & Heaton, P. (2008). The relationship between form and function level receptive prosodic abilities in autism. *Journal of Autism and Developmental Disorders* 38 (7), 1328–1340.

- Kaland, C., Swerts, M. & Kraemer, E. (2013). Accounting for the listener: Comparing the production of contrastive intonation in typically-developing speakers and speakers with autism. *Journal of the Acoustical Society of America* 134(3), 2182-2196.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217-250.
- Kleinman J., Marciano P. L. & Ault, R. L. (2001). Advanced theory of mind in high functioning adults with autism. *Journal of Autism and Developmental Disorders* 31, 29–36.
- Korpilahti, P., Jansson-Verkasalo, E., Mattila, M.-L., Kuusikko, S., Suominen, K., Rytty, S., and Moilanen, I. (2007). Processing of affective speech prosody is impaired in Asperger syndrome. *Journal of Autism and Developmental Disorders*, 37(8), 1539-1549.
- Kuzmanovic, B., Schilbach, L., Lehnhardt, F.-G., Bente, G. & Vogeley, K. (2011). A matter of words: Impact of verbal and nonverbal information on impression formation in high-functioning autism. *Research in Autism Spectrum Disorders* 5, 604–613.
- Kuzmanovic, B. Schilbach, L., Georgescu, A. Kockler, H., Santos, N., Shah, N.J., Bente, G., Fink, G.R., Vogeley, K. (2014). Dissociating animacy processing in high-functioning autism: neural correlates of stimulus properties and subjective ratings. *Soc Neurosci* 9, 309-325.
- Ladd, B. (2008). *Intonational Phonology*, second edition. Cambridge: Cambridge University Press.
- Leiner, D. J. (2014). SoSci Survey (Version 2.4.00-i) [Computer Software]. Available from <https://www.soscisurvey.de>
- Leekam, S. (2016). Social cognitive impairment and autism: what are we trying to explain? *Philosophical Transactions of The Royal Society B Biological Sciences* 371: 20150082.

- Lepistö, T., Kajander, M., Vanhala, R., Alku, P., Huotilainen, M., Näätänen, R., & Kujala, T. (2008). The perception of invariant speech features in children with autism. *Biological Psychology*, 77(1), 25–31.
- Lindner J. L. & Rosén, L. A. (2006). Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. *Journal of Autism and Developmental Disorders* 36, 769–777.
- Mitchell, P. Saltmarsh, R., & Russell, H. (1997). Overly literal interpretations of speech in autism: Understanding that messages arise from minds. *Journal of Child Psychology and Psychiatry* 38, 685-691.
- Mücke, D. & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation - Is it mediated by accentuation? *Journal of Phonetics* 44, 47-61.
- O'Connor, K. (2012). Auditory processing in autism spectrum disorder: A review. *Neuroscience and Biobehavioral Reviews* 36, 836–854.
- Paul, R., Augustyn, A., Klin, A. & Volkmar, F. R. (2005). Perception and Production of Prosody by Speakers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders* 35(2), 205-220.
- Pierrehumbert, J. B. & Hirschberg, J. (1990). The Meaning of Intonational Contours in the Interpretation of Discourse. In: Cohen, P.R., Morgan, J. & M.E. Pollack (eds.), *Intentions in Communication* (pp. 271-311). Cambridge: MIT Press.
- R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistica Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Riedel, A., Heejung S., Haser, V., Hermann, I., Ebert, D., Riemann, D., Bubl, E., Tebartz van Elst, L., Hölzel, L. P. (2014). Freiburg Questionnaire of linguistic pragmatics (FQLP): psychometric properties based on a psychiatric sample. *BMC Psychiatry* 14, 14.

- Röhr, C. T., Baumann, S. (2010). Prosodic marking of information status in German. *Proceedings of the 5th International Conference on Speech Prosody* 100019 Chicago, 1-4.
- Röhr, C. T. & Baumann, S. (2011). Decoding Information Status by Type and Position of Accent in German. *Proceedings 17th International Congress of Phonetic Sciences*, Hong Kong, China, 1706-1709.
- Senju, A., Southgate, V. White, S. & Frith, U. (2009). Mindblind Eyes: An Absence of Spontaneous Theory of Mind in Asperger Syndrome. *Science* 325, 883 -885.
- Shriberg, L. D., Paul, R., McSweeney, J., Klin, A., Cohen, D. J. & Volkmar, F. R. (2001). Speech and Prosody Characteristics of Adolescents and Adults with High Functioning Autism and Asperger Syndrome. *Journal of Speech, Language, and Hearing Research* 44, 1097–1115.
- Singh, L. & Harrow, M. S. (2014). Influences of Semantic and Prosodic Cues on Word Repetition and Categorization in Autism. *Journal of Speech, Language, and Hearing Research*, released May 6, 2014.
- Stewart, M. E., McAdam, C., Ota, M., Peppe, S., & Cleland, J. (2013). Emotional recognition in autism spectrum conditions from voices and faces. *Autism* 17(1), 6-14.
- Tager-Flusberg, H. (2000). Language and understanding minds: connections in autism. In: Baron-Cohen, S., Tager-Flusberg, H. & D. Cohen (eds), *Understanding other minds. Perspectives from developmental cognitive neuroscience* (pp. 124-149). Oxford, Oxford University Press.
- Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. *Handbook of autism and pervasive developmental disorders, 1*, 335-364.

## Appendix

Examples of test sentences used to elicit the speech materials (in bold), test words underlined.

Subjects listened to the sentences in bold only.

(a1) Test word new: „Was hätten Sie gerne?“ „**Ich nehme die Banane mit**.“, antwortet

Thomas dem Obsthändler. Normalerweise ernährt er sich sehr ungesund und isst zwischendurch ständig Süßigkeiten. Außerdem treibt er fast nie Sport und wenn doch, dann am liebsten Minigolf.

*(„What would you like?“ **„I'll take the banana with me“**, says Thomas to the fruit merchant. He usually eats very unhealthily and he is always eating sweets between meals. He hardly ever plays sport, and if he does he prefers mini golf.)*

(a2) Test word new: „Was machen wir?“ „**Wir rufen die Romana an**.“, antwortet Tina ihrer

Freundin. Beide sind auf der Suche nach einer guten Russisch-Lehrerin. Weil sie die Kultur dieses Landes lieben, werden sie im nächsten Semester dort studieren. Englisch ist allerdings die einzige Fremdsprache, die die beiden bisher beherrschen.

*(„What are we going to do?“ **„We will give Romana a call“**, Tina answers her friend. They are looking for a good Russian teacher. They love the culture of this country, which is why they plan to spend a semester there. But so far, English is the only foreign language that both of them have a good command of.)*

(b1) Test word given: Thomas hat gerade auf dem Markt eine Banane gekauft. **Er steckt sich die Banane ein**. In Zukunft möchte er sich viel gesünder ernähren.

*(Thomas has just bought a banana at the market. **He puts the banana in his pocket**. In the future he wants to eat much more healthily.)*

(b2) Test word given: Tina und ihre Freundin möchten bei Romana Russisch lernen. **Sie sprechen die Romana an.** Mit ihrer Hilfe werden sie die Sprache sicher schnell lernen.

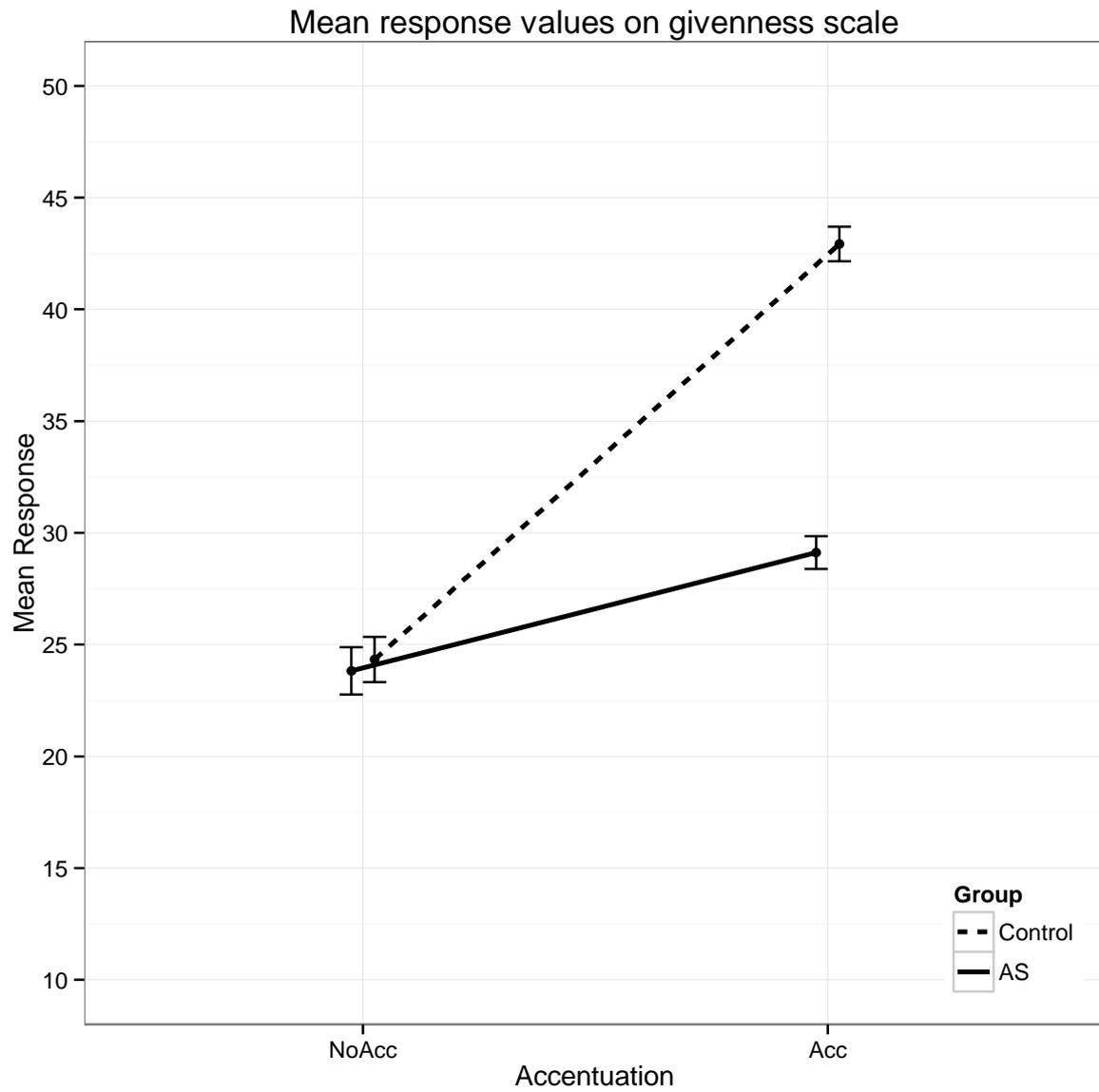
*(Tina and her friend would like to take Russian lessons with Romana. **They approach Romana.** With her help they will surely learn the language in no time.)*

**Table 1** Summary of sensitivity to different properties of the target word in the two groups.

The main effect (Accentuation) is in the heavy-lined box (– lack of sensitivity, + weak sensitivity, ++ strong sensitivity).

	Accentuation	Semantic category of animacy	Word frequency
Control	++	+	-
AS	+	++	+
Effect on rating	Accented is newer (Control > AS)	Animate is newer (AS > Control)	Infrequent is newer (AS only)

**Figure 1** Distribution of ratings for no accent=NoAcc and accent=Acc according to their mean response values, control and AS groups.



**Figure 2** Mean ratings of each target word and the corresponding Subtlex Word Frequency values

