Prosodic Marking of Information Status in Adults with Autism Spectrum Disorders

Martina Krüger[1], *Francesco Cangemi*[1], *Kai Vogeley*[2,3], *Martine Grice*[1]

[1]I*f*L Phonetik, University of Cologne, Germany
[2]Department of Psychiatry, University Hospital Cologne, Germany
[3]Research Center Juelich, Institute of Neuroscience and Medicine, Cognitive Neuroscience (INM-3), Germany

m.krueger@uni-koeln.de, fcangemi@uni-koeln.de, kai.vogeley@uk-koeln.de, martine.grice@uni-koeln.de

Abstract

When referring to an object or person, speakers select a referring expression along with an appropriate prosody. This choice is a highly context-dependent, listener-oriented aspect of language that has been reported to be difficult for individuals with Autism Spectrum Disorders (ASD) and associated mentalizing deficits [1, 2]. In a picture-based story-telling task, we investigated the encoding of a referent's givenness, focusing on prosodic choices. When new referents were introduced (or reintroduced) into the discourse, adults with ASD were similar to typically developed adults in their pitch accent *placement*, but differed in their choice of accent *type*. On new referents, the ASD group produced accents which are less prominent and which have a non-committal nature (H*), while the control group made greater use of more prominent accents (L+H*, L*+H). Thus, selecting the appropriate pitch accent type to mark a newly introduced referent is problematic for individuals with ASD.

Index Terms: prosody, information status, givenness, referring expressions, Autism Spectrum Disorders

1. Introduction

One of the main goals of conversation is to exchange information. Interlocuters achieve this goal by formulating propositions about individuals, events, things, and other entities. In doing so, they constantly *refer* to these entities to provide information about them.

In order to make a reference successful, the speaker has to make sure that the listener becomes or is familiar with the entity that the speaker is talking about. Furthermore, the speaker has to assess the degree of activation of already identifiable referents in the mind of the listener in order to avoid ambiguity among potential referents [3, 4]. On the other hand, redundant information is to be avoided and shorter and less specific expressions (e.g. pronouns) are preferable in order to make communication efficient and economical [5, 6]. To manage this balancing act, speakers constantly assess the current state of the entities in the mind of their listeners and adapt what they say and how they say it accordingly.

Prosody plays a central role in this organization of information. Words referring to new items are typically prosodically highlighted and receive a pitch accent, while words referring to given (old/known) items are generally attenuated or unaccented (e.g. [7, 8]). Furthermore, different pitch accent *types* have been shown to convey information about the degree of givenness of discourse referents [9, 10, 11]. Pitch accent types differ with regard to their perceived prominence [12] such that rising pitch accents are more prominent than falling pitch accents and higher and later pitch peaks are more prominent than lower and earlier ones.

Prosodic prominence correlates with the activation cost that a speaker has to apply to activate a referent in the mind of their listener (see Figure 1). If a referent has an active status, it is usually formally expressed through the absence of pitch accents/prosodic prominence and, morphosyntactically, through pronominalization. Thus, the activation cost is considered to be very low. If a referent is inactive or new, it is usually marked with a prominent pitch accent and expressed as a full noun phrase or name that unambiguously identifies it. In this case, the activation cost is very high. A referent that has been mentioned before but is currently not active, hence *semi*active, exhibits an intermediate information status with an associated intermediate activation cost that can, for example, manifest in less prominent pitch accent types when reactivated.

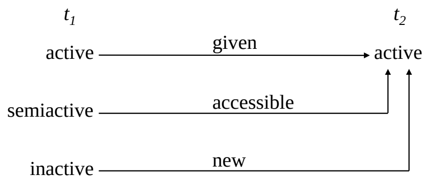

Figure 1: *Information status and associated activation cost (represented as arrows) as proposed in Chafe's cognitive approach (see [3, p. 73])*

Both highlighting and attenuation mechanisms require an assessment of whether items are new or given to the listener. Individuals with impaired mentalizing skills, such as individuals with Autism Spectrum Disorders (ASD), are thus expected to show different patterns of highlighting and attenuation from typically developed individuals.

Autism is a neurodevelopmental disorder characterized by persistent deficits in both social interaction and communication, associated with repetitive, restricted behavior, interests, and/or activities (DSM-V [13], ICD-10 [14]). According to the most recent revision of the DSM-V, autism has to be understood as a spectrum disorder, with the occurence and severity of deficits varying greatly from one individual to another. The spectrum ranges from individuals with "low-functioning" autism and associated learning difficulties who need support throughout life, to "high-functioning" autonomous individu-

als with normal or above average intelligence (IQ>70), corresponding to Asperger's Syndrome.
One of the core symptoms of ASD is related to deficits in social communication. One aspect that heavily impacts social communication is impaired prosody in general and impaired prosodic marking of pragmatic aspects of language in particular. Prosody has been found to be atypical in adults with ASD. This includes inappropriate accent placement [15] and a failure to use pitch to mark information structure [16]. In general, research suggests that unmarked accentuation patterns that can be predicted by grammatical rules such as those used for broad focus structures (nuclear accent on the last lexical item of a sentence, see, e.g., [17]) are found to be intact in high-functioning individuals with ASD, while accentuation that is dependent on contextual information, including contexts evoking contrastive meaning, is impaired [18, 19, 20, 21, 22, 23].

[24] investigated the prosodic marking of referential expressions encoding new and old (given) information in children with ASD. They found that children with ASD were able to produce utterances that demonstrated the use of accentuation to signal new information.

Although the studies mentioned above investigated prosody, they report on the presence or absence of accents rather than on the type of accents produced. As [25] pointed out, the investigation of prosody is difficult, especially when looking at fine-grained distinctions such as accent type. Nonetheless it is important to embark on such an analysis if the goal is to find out where exactly the differences between subjects with ASD and typically developed control persons lie.

In the following study, a cooperative story-telling task [26] has been used in order to elicit spontaneous but structured speech. This method retains the benefits of natural conversation while still providing controlled, quantifiable and comparable samples of referential expressions.

2. Methods

All aspects of this study have been approved by the local ethics committee of the Medical Faculty at the University of Cologne and were performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

All participants gave their written informed consent prior to participating in the experiment.

2.1. Participants

The ASD group comprised 16 native German speakers with ASD (11 male, aged 25-55), who had all been diagnosed with ICD-10: F84.5 (Asperger's Syndrome) and were recruited in the Autism Outpatient Clinic at the Department of Psychiatry at the University of Cologne (Germany). As part of a systematic assessment implemented in the clinic, diagnoses were made independently by two different specialized clinicians corresponding to ICD-10 criteria and supplemented by an extensive neuropsychological assessment.

The control group comprised 16 typically developed native German speakers (4 male, aged 19-27) who were recruited from the student pool of the University of Cologne (Germany).

All participants completed the German version of the *Autism-Spectrum Quotient* (AQ) questionnaire, an instrument that has been developed by Baron-Cohen and colleagues [27] to measure autistic traits in adults with normal intelligence. AQ scores range from 0 to 50, with higher scores indicating more autistic traits.

All participants completed the *Wortschatztest* (WST, [28]), a standardized German vocabulary test that not only exhibits high correlation to verbal intelligence, but also to general intelligence [29] and therefore serves as a measure of both verbal and general IQ.

Statistical tests revealed differences between groups for AQ scores (two-sample T-test, $p<0.001$), age (two-sample T-Test: $p<0.0001$) and sex (Pearson Chi-square test: $p=0.03$). Crucially, there was no significant difference between the groups for verbal IQ as measured by WST (two-sample T-test, $p= 0.137$).

2.2. Materials

The story-telling task was adapted from [26], who originally designed the task to explore effects of story coherence and complexity in referential choices of speakers with and without neurodegenerative and psychotic disorders. Furthermore, the task was built to investigate different discourse stages (introduction, maintenance, and shift of topic).

The task consisted of 18 stories, organized in three levels of increasing referential complexity: The level 1 stories featured one referent only, while the level 2 and level 3 stories featured two referents of different (level 2) or same (level 3) gender. Furthermore, half of the stories followed a logical chronological development (logical stories), while the other half did not (non-logical stories).

While the level 1 stories focused on the introduction and subsequent maintenance of one referent only, the level 2 and level 3 stories involved a more complex structure. Fossard and colleagues [26] defined three discourse stages for level 2 and level 3 stories, based on the six images composing each sequence, namely *introduction* of a new referent, *maintaining* this referent, and *shift* of the referent in focus in the image (see Figure 2 for an example of a level 2 story). Speech materials collected from these different discourse stages were analysed in terms of information status (new, reactivated, and given information [30]), considering referents as *new* when they were introduced for the first time (e.g. the man in picture 1 of Figure 2), as given when they were maintained from one picture to the next (e.g. the man in picture 2), and as reactivated when mentioned after another character had taken the focus of the narration (e.g. the man in picture 5, after the woman had been the focus of the narration in pictures 3 and 4).

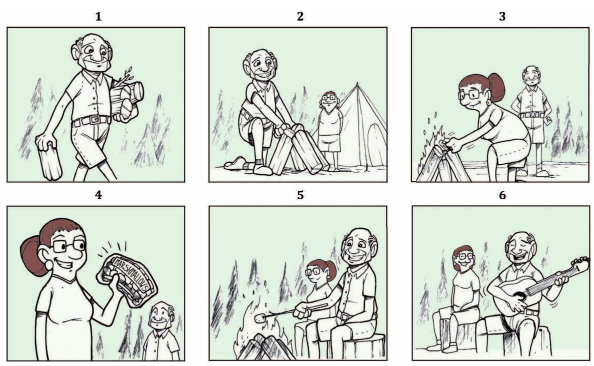

Figure 2: *Example of a level 2 story. Adapted for German from Fossard et al. [26].*

2.3. Task and Recordings

Participants were seated at a table opposite the confederate. A screen was placed between the participant and the confederate to prevent eye contact and any other non-verbal communication. The participants were asked to narrate stories based on sequences of six images (see Figure 2). They were led to believe that they would be helping the confederate, who had the same six images in random order, to arrange them in the correct order. Participants were also led to believe that the confederate was seeing the narrative sequences for the first time.

All stories were presented in randomized order. The first story was kept the same for all participants to familiarize themselves with the task. Participants were instructed not to describe the pictures one by one or give the numbers of the pictures but to try and narrate a cohesive story. Also, participants were instructed not to use proper names, nicknames or direct speech between the characters in order to ensure an appropriate elicitation of referential expressions.

2.4. Analyses

The recorded speech was segmented and annotated in Praat [31] by transcribers who were blind to the participants' diagnoses. Each story was segmented and transcribed at the word level. The positions of realized pitch accents and boundary tones were marked and their tonal configuration was categorized according to GToBI [32] on a tonal tier. Referential expressions that referred to one or both of the characters of the story were analysed with a simplified version of the RefLex annotation scheme [33, 34] that was complemented by an analysis of topichood in order to determine instances of reactivation. This was necessary, because a mere RefLex annotation would not adequately capture instances of reactivation, since the referent would be given, but there would be an activation cost involved in bringing this referent from the background to the foreground.

3. Results

A statistical analysis generated p-values using likelihood ratio tests to compare full models (Poisson Regression) with *group* as a fixed effect to the corresponding null models without the fixed effect. The total amount of produced referents was taken as an offset to adjust for differences in total numbers of mentions between the groups.

Logical and non-logical stories were taken together for the analyses. This decision was based upon the facts that a) all participants narrated non-logical stories in a coherent way, as if the stories were logical, and b) visual inspections of all generated graphs separated into logical versus non-logical stories exhibited no differences in accent distributions or referential choices.

In a first analysis, the distribution of nuclear and prenuclear accents as well as cases where no accent was realized on a referent was compared in both groups. The accent placement patterns for each type of information status in each group are shown in Figure 3.

In general, the patterns across reactivated and new referents were similar in both groups (see Figure 3), while the pattern for given referents differed between groups. There were no significant differences between the groups for the reactivated and new conditions, while there was a significant group difference for the no accent / given condition (see p-values in Table 1). Given referents predominantly received no accent in both groups, but more often in the control group (87% in the Control group versus 78% in the ASD group). There was an increased use of

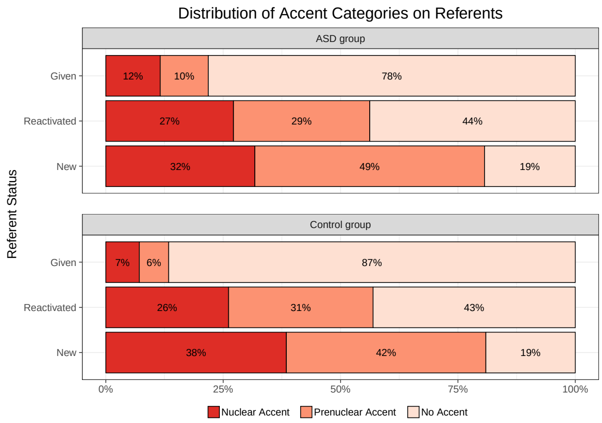

Figure 3: *Relative distribution of accent categories on referents in all stories as produced by the ASD group (top) and the control group (bottom). Percentages are rounded.*

both prenuclear and nuclear accents for reactivated referents in both groups. New referents predominantly received an accent in both groups, the accent being either prenuclear (49% and 42%) or nuclear (32% and 38%). Thus, an increase in (nuclear) accentuation can be observed from given through reactivated to new in both groups, reflecting the increased activation cost for referents that either have to be reactivated or introduced. However, the ASD group did not use deaccentuation as reliably as the control group to encode given referents.

Table 1: *The effect of group on Accent Placement. Italic type and asterisks indicate significant group differences*

	No Accent	Prenuc. Accent	Nuc. Accent
		p-values	
Given	*0.021**	0.079	0.304
Reactivated	0.853	0.760	0.683
New	0.850	0.512	0.224

In a second analysis, the different accent types that were used by speakers when producing nuclear accents on referents were compared. The nuclear accent types for each level of information status in each group are shown in Figure 4.

The two groups differ with regard to their choice of nuclear accent type on referents (see p-values in Table 2; accent types that are not included in this table did not occur often enough to be included in the statistical analysis).

In general, there was a significantly higher distribution of H* nuclear accents in the ASD group in comparison to the control group ($\chi^2(1) = 9.09$, p = 0.003). Furthermore, there were significant group differences for !H* accents on given referents, as well as for rising accents (L*+H/L+H*) and H* accents on new referents.

When producing nuclear accents on referents, both groups predominantly used H* accents on given referents. However, there was a significant group difference regarding the use of !H* accents on given referents, with more occurences of !H* in the control group than in the ASD group.

There were no group differences regarding choice of accent type in cases of reactivation. Both groups predominantly used H* accents on referents that were reactivated, with an increased

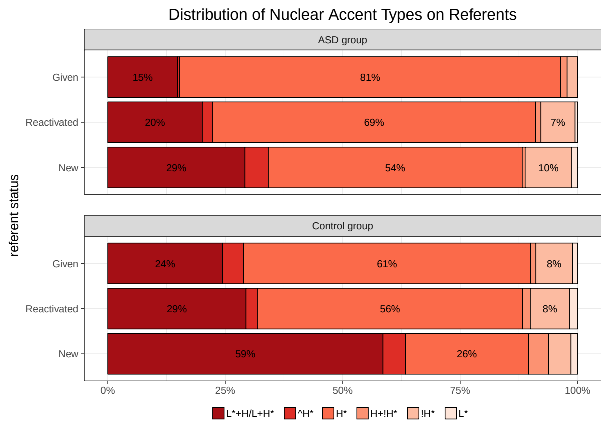

Figure 4: *Relative distribution of accent types on referents in all stories as produced by the ASD group (top) and the control group (bottom). Percentages are rounded. Accent types are ordered according to their perceived prominence from right (least prominent) to left (most prominent), see [12].*

Table 2: *The effect of group on choice of Accent Type. Italic type and asterisks indicate significant group differences*

	L+H*/L*+H	H*	!H*
		p-values	
Given	0.359	0.107	*0.028**
Reactivated	0.645	0.356	0.779
New	*0.011**	*0.007**	0.155

use of rising accents compared to cases where referents were given.

On new referents, the control group used H* accents significantly less often (only 26%) than the ASD group. Instead, the control group used rising accent types (59% L*+H/L+H*) significantly more often than the ASD group. Speakers from the ASD group, on the other hand, predominantly used H* accents (54% of cases) to mark new referents, while resorting to rising accents significantly less frequently.

In sum, an increase in prominence of accent type was observed from given through reactivated to new in both groups. However, new referents were more frequently marked with prominent rising accents in the control group than in the ASD group, who resorted to less prominent accents (H*, see [12] for prominence ratings of different pitch accents) instead. Furthermore, given referents were more frequently marked with less prominent !H* accents in the control group than in the ASD group, which again showed a tendency towards using H* accents (in 81% of cases versus 61% of cases in the control group).

4. Discussion

When telling stories, individuals with and without ASD used similar accent placement patterns to encode new or reactivated referents. Both groups placed more accents on new referents than on given referents, with reactivated referents falling inbetween the two. The only difference in accent placement between the two groups was in given information. Here, the control group deaccented more than the ASD group.

The two groups also differed in their choice of nuclear accent *type*. Crucially, every time a speaker chose to mark a referent of a story with a nuclear accent, the choice of H* was prevalent in the ASD group, unlike in the control group, where rising accents prevailed. H* accents are characterized by a smaller pitch excursion than the rising accents, L+H* and L*+H, that were predominantly used by the control group to signal new information. While H* may be appropriate in certain contexts, persistent production of this type of accent suggests that a speaker is using prosody in an atypical manner, or is inattentive to pragmatic context [35].

The extensive use of H* in the ASD group - the most striking aspect of their prosody - may in fact be a compensation strategy: Using a prosodic pattern that makes the referent neither too prominent nor too attenuated could be seen as a hedging, non-committal way of communicating.

While prosodic attenuation of previously mentioned information is sensitive to the speaker's assumption of the listener's knowledge state, it also seems to be susceptible to egocentric speaker-only knowledge [36]. This might explain the fact that the ASD group managed to attenuate given referents to some extent (although less than controls), since they may have been doing so based on their own knowledge state.

Taken together, these findings point towards a reduced ability in individuals with ASD to produce informative, committed, listener-oriented prosody and may offer an additional explanation as to why individuals with ASD are often faced with problems in social communication.

5. Acknowledgements

This work was funded by the German Research Foundation (DFG), CRC1252, A2 (Grice/Vogeley).

6. References

[1] J. E. Arnold, "Reference production: Production-internal and addressee-oriented processes," *Language and Cognitive Processes*, vol. 23, no. 4, pp. 495–527, 2008.

[2] L. Colle, S. Baron-Cohen, S. Wheelwright, and H. K. J. Van Der Lely, "Narrative discourse in adults with high-functioning autism or Asperger syndrome," *Journal of Autism and Developmental Disorders*, vol. 38, no. 1, pp. 28–40, 2008.

[3] W. Chafe, *Discourse, Consciousness, and Time*. Chicago: University of Chicago Press, 1994.

[4] K. Lambrecht, *Information Structure and Sentence Form*. Cambridge: Cambridge University Press, 1994.

[5] J. Arnold, "Reference Form and Discourse Patterns," PhD thesis, Stanford University:, 1998.

[6] H. P. Grice, "Logic and Conversation," in *Syntax and Semantics*, P. Cole and J. L. Morgan, Eds. New York & San Francisco & London: Academic Press, 1975, vol. 3, no. Speech Acts, pp. 41–58.

[7] A. Cruttenden, "The De-accenting of Given Information: A Cognitive Universal?" in *Pragmatic Organization of Discourse in the Language of Europe*, G. Bernini and M. L. Schwartz, Eds. The Hague: Mouton De Gruyter, 2006, pp. 311–355.

[8] J. Terken and J. Hirschberg, "Deaccentuation of Words Representing 'Given' Information: Effects of Persistence of Grammatical Function and Surface Position," *Language and Speech*, vol. 37, no. 2, pp. 125–145, 1994.

[9] J. B. Pierrehumbert and J. Hirschberg, "The Meaning of Intonational Contours in the Interpretation of Discourse," in *Intentions in Communication*, P. R. Cohen, J. Morgan, and M. E. Pollak, Eds. Cambridge: MIT Press, 1990, pp. 271–311.

[10] S. Baumann and M. Grice, "The Intonation of Accessibility," *Journal of Pragmatics*, vol. 38, no. 10, pp. 1636–1657, 2006.

[11] A. Chen, E. Den Os, and J. P. De Ruiter, "Pitch Accent Type Matters for Online Processing of Information Status: Evidence from Natural and Synthetic Speech," *The Linguistic Review*, vol. 24, no. 2-3, pp. 317–344, 2007.

[12] S. Baumann and C. T. Röhr, "The Perceptual Prominence of Pitch Accent Types in German," in *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, T. S. C. f. I. 2015, Ed. Glasgow, UK: The University of Glasgow, 2015, vol. 298, pp. 1–5.

[13] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Washington, DC: American Psychiatric Association, 5 2013.

[14] World Health Organization, *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization, 1992.

[15] L. D. Shriberg, R. Paul, J. L. McSweeny, A. Klin, D. J. Cohen, and F. R. Volkmar, "Speech and Prosody Characteristics of Adolescents and Adults With High-Functioning Autism and Asperger Syndrome," *Journal of Speech, Language, and Hearing Research*, vol. 4401, no. October, pp. 1097–1115, 2001.

[16] A.-M. R. DePape, A. Chen, G. B. C. Hall, and L. J. Trainor, "Use of prosody and information structure in high functioning adults with autism in relation to language ability." *Frontiers in Psychology*, vol. 3, p. 72, 2012.

[17] N. Chomsky and M. Halle, *The Sound Pattern of English*. New York: Harper and Row, 1968.

[18] J. Fine, G. Bartolucci, G. Ginsberg, and P. Szatmari, "The Use of Intonation to Communicate in Pervasive Developmental Disorders," *Journal of Child Psychology and Psychiatry*, vol. 32, no. 5, pp. 771–782, 1991.

[19] C. A. M. Baltaxe, "Use of contrastive stress in normal, aphasic, and autistic children," *Journal of Speech and Hearing Research*, vol. 27, no. 1, pp. 97–105, 1984.

[20] R. Paul, A. Augustyn, A. Klin, and F. R. Volkmar, "Perception and Production of Prosody by Speakers with Autism Spectrum Disorders," *Journal of Autism and Developmental Disorders*, vol. 35, no. 2, pp. 205–20, 4 2005.

[21] A. Nadig and H. Shaw, "Acoustic marking of prominence: how do preadolescent speakers with and without high-functioning autism mark contrast in an interactive task?" *Language and Cognitive Processes*, vol. 30, 2015.

[22] J. McCann, S. Peppé, F. E. Gibbon, A. O'Hare, and M. Rutherford, "Prosody and its relationship to language in school-aged children with high-functioning autism," *International Journal of Language and Communication Disorders*, vol. 42, no. 6, pp. 682–702, 2007.

[23] S. Peppé, J. McCann, F. Gibbon, A. O'Hare, and M. Rutherford, "Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism," *Journal of Speech Language and Hearing Research*, vol. 50, no. 4, pp. 1015–1028, 8 2007.

[24] P. McCaleb and B. M. Prizant, "Encoding of new versus old information by autistic children," *Journal of Speech and Hearing Disorders*, vol. 50, no. 3, pp. 230–240, 1985.

[25] H. Green and Y. Tobin, "Prosodic analysis is difficult...but worth it: A study in high functioning autism," *International Journal of Speech-Language Pathology*, vol. 11, no. 4, pp. 308–315, 2009.

[26] M. Fossard, A. M. Achim, L. Rousier-Vercruyssen, S. Gonzalez, A. Bureau, and M. Champagne-Lavau, "Referential choices in a collaborative storytelling task: Discourse stages and referential complexity matter," *Frontiers in Psychology*, in revision.

[27] S. Baron-Cohen, S. Wheelwright, R. Skinner, J. Martin, and E. Clubley, "The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians," *Journal of Autism and Developmental Disorders*, vol. 31, no. 1, pp. 5–17, 2001.

[28] K.-H. Schmidt and P. Metzler, *Wortschatztest: WST*. Weinheim: Beltz, 1992.

[29] W. Satzger, H. Fessmann, and R. R. Engel, "Liefern HAWIE-R, WST und MWT-B vergleichbare IQ-Werte?" *Zeitschrift für Differentielle und Diagnostische Psychologie*, vol. 23, no. 2, pp. 159–170, 2002.

[30] P. Hendriks, C. Koster, and J. C. Hoeks, "Referential choice across the lifespan: Why children and elderly adults produce ambiguous pronouns," *Language, Cognition and Neuroscience*, vol. 29, no. 4, pp. 391–407, 2014.

[31] P. Boersma, "Praat a System for Doing Phonetics by Computer," *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2001.

[32] M. Grice, S. Baumann, and R. Benzmüller, "German Intonation in Autosegmental-Metrical Phonology," in *Prosodic Typology: The Phonology of Intonation and Phrasing*, S.-A. Jun, Ed. Oxford: Oxford University Press, 2005, pp. 55–83.

[33] A. Riester and S. Baumann, *The RefLex scheme - annotation guidelines*. University of Stuttgart, 2017, vol. 14. [Online]. Available: https://elib.uni-stuttgart.de/handle/11682/9028

[34] S. Baumann and A. Riester, "Referential and Lexical Givenness: Semantic, Prosodic and Cognitive Aspects," in *Prosody and Meaning*, ser. Interface Explorations 25, G. Elordieta and P. Prieto, Eds. Berlin & New York: Mouton De Gruyter, 2012, pp. 119–162.

[35] J. J. Diehl, D. Watson, L. Bennetto, J. McDonough, and C. Gunlogson, "An acoustic analysis of prosody in high-functioning autism," *Applied Psycholinguistics*, vol. 30, no. 03, pp. 385–404, 2009.

[36] E. G. Bard, A. H. Anderson, C. Sotillo, M. Aylett, G. Doherty-Sneddon, and A. Newlands, "Controlling the Intelligibility of Referring Expressions in Dialogue," *Journal of Memory and Language*, vol. 42, no. 1, pp. 1–22, 2000.