Intelligibility of Noise-Adapted and Clear Speech in Child, Young Adult, and Older Adult Talkers

Rajka Smiljanic and Rachael C. Gilbert

Purpose: This study examined intelligibility of conversational and clear speech sentences produced in quiet and in noise by children, young adults, and older adults. Relative talker intelligibility was assessed across speaking styles.

Method: Sixty-one young adult participants listened to sentences mixed with speech-shaped noise at −5 dB signal-to-noise ratio. The analyses examined percent correct scores across conversational, clear, and noise-adapted conditions and the three talker groups. Correlation analyses examined whether talker intelligibility is consistent across speaking style adaptations.

Results: Noise-adapted and clear speech significantly enhanced intelligibility for young adult listeners. The intelligibility improvement varied across the three talker groups. Notably, intelligibility benefit was smallest for children’s speaking style modifications. Listeners also perceived speech produced in noise by older adults to be less intelligible compared to the younger talkers. Talker intelligibility was correlated strongly between conversational and clear speech in quiet, but not for conversational speech produced in quiet and in noise.

Conclusions: Results provide evidence that intelligibility variation related to age and communicative barrier has the potential to aid clinical decision making for individuals with speech disorders, particularly dysarthria.

In a companion article (Smiljanic & Gilbert, 2017), we examined acoustic–articulatory adjustments that children, young adults, and older adults implemented when speaking in response to noise (noise-adapted speech [NAS]) and when instructed to speak clearly (clear speech [CS]), both separately and in conjunction. The results showed some overlapping CS and NAS strategies (boost in spectral energy, increase in sound pressure levels [SPLs], longer voiced segments, slower speaking rate, and increased vowel space area [VSA]) as well as some divergent ones (increased F0 mean, jitter, and shimmer in NAS vs. increased pausing in CS). We also found age-related differences in conversational-to-clear and quiet-to-noise adjustments. Older adults produced overall slowest speaking rate, longest pauses, and smallest increase in F0 mean, 1–3 kHz energy, and SPL when speaking clearly. They produced smallest increases in VSA in both NAS and CS. When speaking clearly, children slowed down less than adults, increased the VSA least, increased harmonic-to-noise ratio, and decreased jitter and shimmer most. Children also increased mean F0 and F1 most in noise. In this article, we aim to determine whether intelligibility for young adult listeners is enhanced through the adaptations produced in response to noise and when instructed to speak clearly by the three talker groups.

Although the two speaking responses differ in the conditions under which talkers produced them, speaking clearly when talking to someone who is hard of hearing or not a native speaker or in response to noise, both adaptations characterize everyday communication. Examining how intelligibility is affected by talkers’ adaptation on the hypo- to hyperarticulated continuum in response to different communicative situations has implications for speech production theories (H&H theory; Lindblom, 1990). H&H theory is typically discussed in the context of healthy young adults. Here, we consider whether this theory can be extended to account for intelligibility variation in child and older adult talkers. Furthermore, CS, loud speech, slow speech, and speech produced in response to noise have been used as behavioral treatment techniques with the goal of enhancing intelligibility in clinical populations (Beukelman, Fager, Ullman, Hanson, & Logemann, 2002; Hustad & Weismer, 2007; Sadagopan & Huber, 2007). The results showed that...
Intelligibility of NAS and CS

Previous work demonstrated improved intelligibility of CS and NAS over conversational (CO) and quiet speech for a variety of listener groups and degraded listening conditions (Cooke, King, Garnier, & Aubanel, 2013; Godoy et al., 2013; Pichora-Fuller, Goy, & van Lieshout, 2010; Smiljanic & Bradlow, 2009). Only two studies to date compared intelligibility benefit of NAS and CS directly, reporting somewhat conflicting results. Goy, Pichora-Fuller, and van Lieshout (2013) found that NAS significantly improved intelligibility for older adult listeners in difficult listening conditions (with more background noise), but not for younger adult listeners. They found no intelligibility benefit for CS. Gilbert, Chandrasekaran, and Smiljanic (2014), on the other hand, demonstrated that both NAS and CS significantly increased intelligibility for young adult listeners and that this benefit extended to sentence recognition memory as well. Their results also showed that the combined acoustic–articulatory modifications produced by one talker, that is, speaking clearly in noise (CS + NAS), resulted in the largest intelligibility benefit.

Previous CS and NAS studies also found a striking variability in intelligibility benefit across and within talker populations. Unsurprisingly, late learners of English provided smaller CS intelligibility benefit compared to early learners and native speakers of English (Rogers, DeMasi, & Krause, 2010; Smiljanic & Bradlow, 2011). Even when comparing young adult native speakers of English, across-talker variability in intelligibility benefit is found. Ferguson (2004) reported that female talkers provided larger CS intelligibility gains compared to male talkers. In that study, an average intelligibility difference between CS and CO speech of 8.5 percentage points was found (which, when examined across all 41 talkers, ranged from −12 to 33 percentage points). Although the Ferguson study examined vowels, a similar intelligibility range was found for sentences in Smiljanic and Bradlow (2005) and Lam and Tjaden (2013). For NAS, Pittman and Wiley (2001) found an intelligibility benefit of 15 percentage points (when vocal levels were equated across conditions), though no individual talker data were reported. Similarly, Summers, Pisoni, Bernacki, Pedlow, and Stokes (1988) and Lu and Cooke (2008) reported an average benefit of 8 percentage points (across various noise levels) and a 9– to 25–percentage point increase in intelligibility, respectively. Green, Katiri, Faulkner, and Rosen (2007) found a 2–percentage point intelligibility benefit for NAS compared to speech produced in quiet. The current study builds on this work by examining intelligibility benefit for CS and NAS within the same study with consistent elicitation methods for 30 talkers across three age groups.

Intelligibility of Speaking Style Adaptations Across the Life Span

Children’s and older adults’ ability to enhance speech intelligibility in response to adverse communicative situations remains largely unexplored. Schum (1996) and Smiljanic (2013) both found that older adults’ intelligibility was enhanced through conversational-to-clear speech modifications. The first study found that CS gain in intelligibility was similar for older and young adults, whereas the second study found that older adults provided smaller gain in intelligibility compared to young adults. A few studies examined NAS production in older adults as a comparison group for individuals with Parkinson’s disease (e.g., Darling & Huber, 2011; Huber & Darling, 2011). They report that all talkers modified their speech in response to noise, but differences in the acoustic–articulatory modifications were also found between talker groups. Children as young as 3–5 years old produced perceptibly different speaking styles, demonstrating that even very young children have some ability to modify their speech when instructed to produce listener-oriented speech (Redford & Gildersleeve-Neumann, 2009; Syrett & Kawahara, 2014). However, acoustic analyses revealed that even by ages 9–10 and 13–14, children differed in precise CS modifications (e.g., vowel hyperarticulation) compared to adults (Pettinato & Hazan, 2013; Pettinato, Tuomainen, Granlund, & Hazan, 2016). This suggests that many features of adultlike speaking style adaptations continue to develop into adolescence. It remains to be determined whether these listener-oriented changes produced by children and older adults indeed result in enhanced intelligibility—a question none of the above studies addressed. The findings will enhance our understanding of developmental and aging effects on speech intelligibility variation.

Relative Talker Intelligibility

Finally, the current article addresses the question of relative talker intelligibility in the two adaptations. Previous
work found that, despite large variation in intelligibility levels, the relative intelligibility across talkers is remarkably consistent for different listener populations and environments. The ranking of most and least intelligible talkers is consistent for adult and child talkers and listeners (Hazan & Markham, 2004), for typically hearing and cochlear-implanted listeners (Green et al., 2007), for native and nonnative listeners (Bent, Kewley-Port, & Ferguson, 2010; Van Dommelen & Hazan, 2012), and for vocoded speech versus speech masked by multitalker babble (Bent, Buchwald, & Pisoni, 2009). This suggests that a talker’s speech clarity is an inherent quality independent of environment or listener characteristics (Bradlow, Torretta, & Pisoni, 1996). Whether such relative talker intelligibility holds across different intelligibility-enhancing adaptations produced by child and older adult talkers remains to be determined. This examination can provide further insight into the relationship between talker clarity and listener characteristics. Young adults are likely to have greater experience talking with other young adults and less experience talking with children and older adults. Episodic memory theories suggest that word recognition is enhanced if incoming speech is similar to the stored memory traces (Goldinger, 1996, 1998; Palmeri, Goldinger, & Pisoni, 1993). If such listener-driven perception holds here, young adult listeners should find young adult talkers most intelligible. If, on the other hand, inherent talker clarity determines intelligibility, listeners should find high- and low-intelligibility talkers in each of the three age groups. Furthermore, relative talker intelligibility across NAS and CS would be determined by the talker’s ability to produce intelligibility-enhancing modifications. It is thus plausible that relative talker intelligibility would change across different speaking style conditions. Ferguson (2004, 2012) found that vowel identification in noise for CO speech was correlated with CS across young adult talkers. However, the extent to which the relative intelligibility of individual talkers holds for sentences across different speaking style adaptations and across age groups is unknown.

The current perception study aims to determine whether young adult listeners benefit from NAS and CS modifications produced by children, young adult, and older adult talkers. We hypothesized that listeners will find NAS and CS more intelligible compared to the baseline CO and quiet speech. Furthermore, given that differences in the acoustic adjustments across speaking styles and talker groups reported in Smiljanic and Gilbert (2017) may explain intelligibility variation (Godoy et al., 2013; Kent et al., 1989; Turner, Tjaden, & Weismer, 1995; Weismer, Jeng, Laurens, Kent, & Kent, 2001), we also predicted that the intelligibility benefit would be greater for NAS compared to CS and for young adults compared to children and older adults. Finally, we hypothesized that the relative talker intelligibility across speaking styles will vary because talkers differ in implementing CS and NAS acoustic–articulatory adjustments.

**Method**

**Talkers**

Speech samples from 10 monolingual speakers of English from each of the three following groups were used in the intelligibility test: children (11–13 years old, mean = 12.3 years), young adults (18–29 years old, mean = 21.0 years), and older adults (60–84 years old, mean = 70.2 years). Children and young adults passed hearing screening bilaterally at 25 dB HL at octave frequencies between 500 and 4000 Hz. Four out of 10 older adult talkers had hearing within normal limits through 4000 Hz. Three out of 10 older adults had hearing within normal limits through 2000 Hz with 4000 Hz thresholds ranging between 40 and 45 dB HL. Three talkers had some degree of high-frequency hearing loss typical for presbycusis (ranging between 35–45 dB HL at 2000 Hz and 40–75 dB HL at 4000 Hz). None of the older adults were hearing aid users. Younger adults were all University of Texas at Austin undergraduate students. Children and older adults were recruited from the Austin, Texas, community.

**Materials**

Sixty meaningful sentences (e.g., *Farm animals stay in a barn*; Fallon, Trehub, & Schneider, 2002) were produced in CO speech first and then in CS. The CO and CS sentences were first elicited in quiet. Next, both styles were elicited in response to speech-shaped noise interference (NAS) presented via headphones (80 dB SPL). Speech-shaped noise was generated by obtaining the long-term average spectrum of six-talker babble (three male, three female; Van Engen & Bradlow, 2007) and shaping white noise to match that spectrum. For CO speech, talkers were instructed to speak in a casual manner as if they were talking to a friend or a family member. For CS, they were instructed to speak as if they were talking to someone who has a low proficiency in English and has difficulty following them conversationally (following Pichora-Fuller et al., 2010; Smiljanic & Bradlow, 2009). When eliciting NAS, the talkers read on a slide: “Now, you will hear some background noise. Pretend that you are in a noisy place, talking to your friends or family.” All sentences were normalized for root-mean-square amplitude. Forty unique sentences produced by each of the 30 talkers in all four speaking styles, 4,800 sentences total (40 sentences × 4 speaking styles × 30 talkers), were used in the intelligibility test. Sentences were mixed with portions of the noise with the same long-term average spectrum as the noise used to elicit NAS in production at −5 dB signal-to-noise ratio.

**Listeners**

Sixty-one young adult listeners (18–39 years old, mean = 20.2 years) participated in the speech intelligibility test. The remaining 20 sentences were used in another task, the results of which are not reported here.

---

1Reprinted with permission from Fallon, Trehub, & Schneider. Copyright © 2002, Acoustic Society of America.

2The remaining 20 sentences were used in another task, the results of which are not reported here.
task. All were native monolingual speakers of English. All passed a hearing screening, administered bilaterally at 25 dB HL at 500, 1000, 2000, and 4000 Hz. They were all University of Texas at Austin students recruited from the Linguistics Department subject pool.

Procedure

Listeners took part in one of three listening conditions consisting of sentences spoken by talkers from one age group. The test began with five practice sentences to familiarize the listener with the task; practice sentences consisted of talkers and stimuli not included in the test. Each listener then heard four unique sentences from each of 10 speakers (for a total of 40 sentences). The four sentences from one talker were each in a different speaking style (e.g., Sentence 1 in CO speech produced in quiet, Sentence 12 in CS produced in quiet, Sentence 30 in CO speech produced in noise, and Sentence 39 in CS produced in noise, all by young adult talker 3). Each listener heard a different combination of four unique sentences produced by 10 talkers and distributed across speaking styles. Sentences were distributed evenly across talkers and speaking styles. The order in which sentences were presented was randomized. Talker–style–sentence pairing was different for each listener. This was done to ensure that every listener heard every talker within one age group in every style without sentence repetition. In addition, the randomized study design, which was counter-balanced across talkers and speaking style conditions, ensured that perceptual learning could not have disproportionately contributed to better intelligibility in some conditions than others. Each talker’s speech intelligibility was assessed by 20–21 listeners in total. The experiment was presented in MATLAB. Listeners were instructed to write what they heard, typing one sentence at a time on the keyboard after stimulus presentation. The test was self-paced. Each sentence contained one target word (the final word in the sentence), which was scored as 1 (correct) or 0 (incorrect).

Results

Speech Intelligibility in Noise

A mixed-effects logistic regression was conducted on key word correct identification (lme4 package; Bates, Maechler, & Bolker, 2012). Talker, sentence, and listener were included in the model as random factors, and talker age group (children, young adults, or older adults), noise (produced in quiet or in noise), style (CO or CS), and their interactions were included as fixed effects. Random slopes were included in the model for both noise and style at the level of talker, because this level showed the greatest variance (Barr, Levy, Scheepers, & Tily, 2013). This determined the effect of age and communicative barrier on speech intelligibility. Figure 1 shows intelligibility scores in CO and CS produced in noise and in quiet for each talker group.

The probability of correct key word identification was significantly affected by all three fixed effects, talker age group, noise, and style (p < .001). There were significant two-way interactions between all three fixed effects (p values ranging between p = .013 and p < .001). The three-way interaction was not significant (p = .165). With regard to the main effect of talker age group, F(2, 2428) = 13.964, p < .01, pairwise comparisons showed that children were significantly less intelligible than young adults (p = .004), but not older adults (p = .234). Older adults were not significantly less intelligible than young adults, although there was a trend in this direction (p = .085). NAS was significantly more intelligible than speech produced in quiet, F(2, 2428) = 163.840, p < .001. Likewise, CS was overall significantly more intelligible than CO, F(2, 2428) = 101.838, p < .001.

The significant interaction between talker age group and noise, F(2, 2428) = 8.003, p < .001, revealed that intelligibility in quiet and in noise varied between children, young adults, and older adults. Listeners perceived speech produced in quiet by children to be significantly less intelligible than speech produced by older adults (p = .020), but not by young adults (p = .210). Older and young adults’ intelligibility was not significantly different (p = .275) in quiet. When speaking in noise, children were significantly less intelligible than young adults (p = .003). Older adults did not significantly differ from either group, although intelligibility difference between older and young adults approached significance (p = .072). In order to further shed light on intelligibility in noise for the three talker groups, we calculated NAS intelligibility gain (collapsed over CO and CS) as a proportion of the quiet intelligibility score (noise – quiet/quiet). The proportional gain in intelligibility was similar for children and younger adults (117% vs. 108%). This suggests that intelligibility was improved through children’s NAS modifications. However, these modifications may be different from the ones made by the young adults given that children were overall still less intelligible than young adults both in quiet (children’s intelligibility was 22% correct vs. 30% correct for young adults) and in noise (children’s intelligibility was 47% correct vs. 62% correct for young adults). Proportional NAS gain in intelligibility for older adults was substantially smaller than for both children and younger adults (38%). Their overall intelligibility levels for speech in quiet (37% correct) was the highest, whereas they were intermediate to the children and young adults in noise (51% correct). This suggests that listeners benefited least from older adults’ NAS modifications.

The significant interaction between talker age group and style revealed that the CS benefit was smaller for children (p = .002) compared to young and older adults (p < .001). Pairwise comparisons further showed that intelligibility across talker groups varied in each of the two speaking styles. There were no significant intelligibility differences between talker groups in CO (p values ranged between .528 and .835). In CS, however, listeners found young adults significantly more intelligible than children (p = .004). Older adults did not significantly differ from either group, although, as in previous analyses, the difference
between older adults' and younger adults' intelligibility approached significance ($p = .078$). Looking at the CS intelligibility gain (collapsed over quiet and noise conditions) as a proportion of CO intelligibility provides further insights. CS intelligibility benefit was smallest for children (34%) and substantially greater for both older and young adults (67% vs. 73%). Listeners benefited more from the conversational-to-clear speech changes produced by adult talkers compared to child talkers.

Finally, there was a significant interaction between noise and style, $F(2, 2428) = 6.201, p = .013$. Pairwise comparisons indicate that both CS and NAS resulted in significant intelligibility benefit ($p < .001$). However, examining proportional gain of each enhancement—CS and NAS—as a proportion of the baseline intelligibility reveals different intelligibility increases for CS and NAS. CS intelligibility benefit was larger for speech produced in quiet (133%) than for speech produced in noise (30%). Similarly, the intelligibility benefit for CO produced in noise relative to the CO produced in quiet was larger (162%) compared to the intelligibility benefit for CS produced in noise relative to CS produced in quiet (46%). Furthermore, even though the NAS CS was the most intelligible speaking style, it was less intelligible than if the individual NAS and CS enhancements were summed together. If the intelligibility benefit was additive, we would expect that the increase for CS in noise (N) would be equivalent to the increase we observed in quiet (Q). However, the CS increase on top of already speaking in noise was smaller (133% for QCS – QCO/QCO vs. 30% for NCS – NCO/NCO).

Even though the three-way interaction between talker age group, noise, and style was not significant, examining proportional gain in intelligibility for NAS, CS, and three talker groups separately provides further support for the results discussed above, namely, that intelligibility benefit was smaller for older adults' and children's acoustic–articulatory adjustments. Most notably, children’s proportional gain in intelligibility for CS relative to CO produced in noise is only 5% compared to 50% and 34% for young and older adults, respectively. Similarly, older adult talkers’ proportional gain in intelligibility for CO in noise relative to CO in quiet is 93% compared to 253% and 181% for children and young adults, respectively. Older adults’ proportional gain in intelligibility for CS in noise relative to CS in quiet is only 13% compared to 59% and 77% for children and young adults, respectively. These results further highlight the trends in intelligibility variation across talker groups and speaking style accommodations. The lack of statistical differences for some of these comparisons may be obscured by large intelligibility variation across individual talkers and speaking styles and should be examined further with greater numbers of talkers per age group.

**Relative Talker Intelligibility**

Individual talker mean intelligibility and ranking in each speaking style can be found in the Appendix. The ratio of words correctly transcribed for each talker across speaking styles ranged from 0.00 to 0.95, with a mean of 0.41 (Figure 2). Examination of the individual intelligibility rankings reveals that young adult listeners found talkers of all three age groups to be most/least intelligible in the three age groups to be most/least intelligible across speaking styles. In CO speech in quiet, for instance, OA02, YA10, and CH06 were the three most intelligible talkers and YA03, CH10, and YA01 were the three least intelligible talkers. Similarly, in NAS CO, YA10, YA09, and CH02 were the three most intelligible talkers and YA05, CH07, and CH09 were the three least intelligible talkers. As predicted,
the individual talker intelligibility rankings varied across speaking styles. For example, talker YA10 and talker OA02 were two of the most intelligible talkers in quiet. YA10 maintained this high level of intelligibility in both noise conditions. OA02, on the other hand, dropped to the 11th most intelligible talker in NAS CO and to the 24th most intelligible talker in NAS CS. Similarly, talker CH06 was ranked third in CO speech in quiet and 20th in NAS CS. Conversely, some talkers with low CO intelligibility enhanced their intelligibility significantly through speaking style changes. Talker YA02, for instance, was the least intelligible talker in CO speech in quiet (30th most intelligible) but became fifth most intelligible talker in NAS CS. Talker CH02 was ranked 25th in CO speech in quiet but became third most intelligible talker in NAS CO.

In order to examine the extent to which the relative intelligibility of individual talkers is consistent across speaking styles, four correlations were carried out on aggregated
intelligibility scores for each talker in each speaking style \((n = 30\) for each correlation). The four correlations were between (a) CO produced in quiet and CS produced in quiet, (b) CO produced in noise and CS produced in noise, (c) CO produced in quiet and CO produced in noise, and (d) CS produced in quiet and CS produced in noise. The strongest correlation was found for CO and CS produced in quiet \((r = .620, p < .001)\). Significantly correlated, but more weakly, were CS intelligibility scores produced in quiet and in noise \((r = .432, p = .017)\) and CO and CS intelligibility scores in noise \((r = .374, p = .042)\). Finally, no correlation was found between CO intelligibility scores produced in quiet and in noise \((r = .307, p = .099)\).

Discussion

Intelligibility of NAS and CS

Consistent with our predictions, the results revealed that intelligibility was increased when talkers were instructed to speak clearly and in response to noise. The two adaptations resulted in different gains: The intelligibility gain as a proportion of the baseline intelligibility was larger for NAS than for CS (88% vs. 58%). The largest intelligibility gain was found for CO in noise relative to CO in quiet (162%). The next largest improvement was provided when speaking clearly in quiet relative to CO in quiet (162%). Smallest intelligibility benefit was found for CS in noise both relative to CS in quiet (46%) and relative to CO in noise (30%). It is interesting to note that the largest intelligibility benefit was found for CO speech produced in noise even though the talkers were instructed to speak casually. This suggests that speaking in noise induced a level of automatic adaptation through increased vocal effort, which may facilitate signal transmission and which has also been found in birds and mammals, including bats (Brumm & Todt, 2002; Brumm, Voss, Köllmer, & Todt, 2004; Hage, Jiang, Berquist, Feng, & Metzner, 2013). The slight advantage of NAS over CS modifications found here could be related to the different acoustic–articulatory modifications adopted by talkers in these two types of communication situations (Smiljanic & Gilbert, 2017). Although responses to both barriers induced a decrease in speaking rate and an increase in 1–3 kHz energy, SPL, VSA, and harmonics-to-noise ratio, NAS was additionally characterized by an increased F0 mean and decreased jitter and shimmer. Whether these additional modifications significantly contributed to the NAS intelligibility advantage over CS remains to be determined.

The particular listening environment used here likely also played a role in NAS intelligibility advantage. That is, the larger gain in intelligibility for quiet-to-noise changes could result from the fact that the same masker was used in the speech intelligibility task as in eliciting NAS. Previous

\[3\]Correlation analyses examining relative talker intelligibility for different age groups would require more than 10 talkers per age group; thus, age-related differences in relative talker intelligibility are only indirectly discussed.

work showed that talkers modify their speech differently in response to different communicative barriers and adversity of the background (Cooke & Lu, 2010; Hazan & Baker, 2011; Lu & Cooke, 2008, 2009). Lu and Cooke (2008) demonstrated that different types of noise resulted in different adaptations and, furthermore, that the intelligibility benefit depended on the masker in the listening condition. They showed that, in a background of noise, NAS was more intelligible than speech produced in quiet (and then mixed with noise). The authors suggest that talkers monitor the background and tailor their modifications to counter energetic masking at the ears of the listener. The data reported here lend further support to this idea, as NAS modifications seem to be more effective in overcoming the masking effect of the noise in the listening test compared to the CS modifications produced in quiet and then mixed with noise. Listening environment thus may be an important factor when assessing intelligibility of talkers with dysarthria using CS or NAS as treatment techniques.

This study also revealed that intelligibility was highest for the two adaptations in conjunction, that is, listeners found speech produced clearly while in noise to be most intelligible. Increasing vocal effort in response to noise in combination with deliberate clear productions led to substantially greater intelligibility improvements than when each modification was produced separately. However, the combined NAS CS intelligibility benefit was smaller than if the individual NAS and CS enhancements were additive (e.g., the CS intelligibility gain was larger in quiet than in noise). This is not likely a result of a ceiling effect because the average performance was 60% correct in this most intelligible condition. Rather, this suggests that CS and NAS involve a number of shared modifications, which are used for each adaptation separately as well as when the two communicative barriers are combined. This notion is supported by the acoustic–articulatory analyses describing a number of similar modifications across the two speaking styles (Smiljanic & Gilbert, 2017). In contrast, Gilbert et al. (2014) found that intelligibility was higher for NAS CS than for the two adaptations separately. Note though that the findings here are based on the productions of 30 different talkers, including children and older adults, whereas Gilbert et al. examined only a single young adult talker. Results reported here highlight the need to further determine signal-, talker-, and listener-related factors that contribute to intelligibility variation.

Intelligibility of Speaking Style Adaptations Across the Life Span

Although listeners found children to be overall less intelligible, significant interactions revealed that this was true only in some speaking style conditions and in some talker group comparisons. Children were overall significantly less intelligible than older adults in quiet. However, intelligibility of all three talker groups was similar for CO speech produced in quiet. This result is in line with Hazan and Markham (2004), who found that 13-year-old children’s
casual speech produced in quiet was as intelligible as that of young adult men (but not women). Word intelligibility of child talkers in that study was assessed in 20-talker babble noise at an easier signal-to-noise ratio (+6 dB) compared to the current study because they also included a group of 7-year-old listeners. The relatively more difficult listening condition here did not reveal any overall differences in intelligibility for child talkers in their baseline speech. The present finding that older adults were also as intelligible as younger adults in CO is in accord with McAuliffe, Wilding, Rickard, and O’Beirne (2012), who found similar intelligibility for four young and four older adult talkers’ CO sentences presented in quiet.

Children, however, were significantly less intelligible in CS and NAS compared to young adult talkers. The smaller CS and NAS intelligibility benefit is likely not due to where on the psychometric function children’s and adults’ measures were obtained because the average intelligibility for the baseline CO speech in quiet was not significantly different across the three talker groups (13% vs. 18% vs. 23% correct for children, young adults, and older adults, respectively). The range of intelligibility scores for children (0–35%) and young adults (0–38%) also revealed similar variation in the baseline condition within these talker groups. Furthermore, child talkers were as intelligible as older adults; in, for instance, CO speech produced in noise. Children’s lower CS and NAS intelligibility could, in part, be attributed to their speech perception, for both phonemes and sentences, being more affected by speech degradation up to adolescence, therefore making communication in noise more detrimental to children compared to adults (Johnson, 2000; Wroblewski, Lewis, Valente, & Stelmachowicz, 2012). The smaller CS and NAS intelligibility benefit likely also reflects lack of fully adultlike productions in both temporal and spectral domains (Lee, Potamianos, & Narayanan, 1999), as well as inexperience in producing appropriate intelligibility-enhancing strategies in order to overcome adverse communicative conditions (Hazen, Tuomainen, & Pettinato, 2016; Pettinato & Hazan, 2013). In the companion article, we reported that children, compared to adult talkers, slowed down less, increased the VSA least, increased harmonic-to-noise ratio, and decreased jitter and shimmer most in CS. They also increased mean F0 and F1 most in noise. The combined production and intelligibility findings demonstrate that, although children were able to produce acoustic-articulatory modifications leading to intelligibility enhancement, these qualitatively and quantitatively different adaptations on average did not result in the adultlike intelligibility benefit by 11–13 years of age.

Intelligibility benefit for CS speech in quiet (as a proportion of the CO intelligibility) was similar for older and young adults (128% vs. 137%). These results are in line with that of Schum (1996), who reported equal CS gain in intelligibility for young and older talkers. The present results are different from that of Smiljanic (2013), in which older adults exhibited a smaller CS intelligibility benefit than young adults. Note that Schum reported only gain in intelligibility rather than the absolute intelligibility levels; thus, it is unclear if older and young adults were equally intelligible in CO speech in that study. The difference between the findings here and those of Smiljanic can be attributed in part to the stimuli (meaningful vs. semantically anomalous sentences). Furthermore, Schum’s study and this study examined speech from 10 younger and 10 older adult talkers, whereas Smiljanic’s study examined speech of only five older and five younger adults, which could have led to a less representative sample in the latter study.

The current results indicated that listeners found NAS produced by older adults somewhat less intelligible compared to that produced by young adults. Although these differences were not statistically significant, they approached significance. Intelligibility gain was smaller for CO and CS speech produced in noise as a proportion of the intelligibility scores in quiet by older adults than by young adults (93% vs. 181% for NAS CO and 34% vs. 50% for NAS CS). The smaller intelligibility benefit for older adults’ NAS could be related to the elevated hearing thresholds that naturally occur with aging. Research has shown that the magnitude of the noise adaptation depends on the intensity of the noise (Lu & Cooke, 2008), suggesting that talkers with diminished acoustic sensitivity might adapt their speech to noise to a lesser degree. There is evidence that this may be true for some of the talkers in this study. Two of the talkers who were identified as having some degree of high-frequency hearing loss typical of presbycusis (ranging between 35–45 dB HL at 2000 Hz and 40–75 dB HL at 4000 Hz) were ranked low in some noise conditions (OA04 was 22nd in both NAS CO and NAS CS; OA08 was 29th in NAS CS). However, the third talker, OA05, who also had some degree of high-frequency hearing loss, was ranked 18th in NAS CO. Talker OA07, who had hearing within normal limits through 4000 Hz, ranked 20th in NAS CO. This suggests that, although hearing loss in high frequencies may shape speech production in noise to some degree, it may not be the only factor determining intelligibility levels for older adults. It is important to note that Smiljanic and Gilbert (2017) reported that, compared to young adults, older adults produced smaller increase in F0 mean, 1–3 kHz energy, and SPL when speaking clearly. They also produced the smallest increases in VSA in both NAS and CS. Despite these different CS and NAS adaptation strategies, CS intelligibility benefit was similar for young and older adults, but NAS intelligibility was lower for older adults. These results highlight a lack of one-to-one mapping between acoustic-articulatory modifications and intelligibility benefit.

The lower intelligibility of NAS speech for older adults may also reflect levels of familiarity and frequency of communicating in particular settings, such as in noise. Older adults may accommodate to noise less overall due to avoiding noisy environments and lack of recent experience talking with others in such conditions. Although we did not ask about recent communicative experiences, future work should take this information into account. In contrast, listeners perceived speech produced in noise by both children and young adults to be more intelligible compared to when they were instructed to speak clearly. Intelligibility
Relative Talker Intelligibility

Looking at the average intelligibility across talker groups, it may seem that young adult listeners preferred young adults’ speech. However, young adult listeners here found a variety of talkers from all three age groups to be most/least intelligible. These findings lend support to the notion that acoustic–phonetic characteristics of the talker’s speech determine, to a large extent, intelligibility variation (Bradlow et al., 1996; Hazan & Markham, 2004; Smiljanic & Bradlow, 2005). Furthermore, these results suggest that an interaction between talker- and listener-related factors plays a smaller role in determining intelligibility variation at least when listening to native talkers of varying ages. With regard to the relative intelligibility ranking of talkers, a strong correlation was found between CO and CS produced in quiet. This extends the results from Ferguson (2004), who found that relative intelligibility for 41 talkers was consistent across the two speaking styles for vowels. Here, 30 talkers across three age groups showed similar intelligibility rankings across CO and CS for monosyllabic target words embedded in high-context sentences. This finding suggests that talkers who are inherently clear also employ effective CS intelligibility-enhancing strategies. It is not the case that a relatively intelligible talker’s baseline level precludes them from implementing acoustic–artulatory modifications that enhance their intelligibility further. Weaker correlations held between CS produced in quiet and in noise and between CO and CS produced in noise. No significant correlation was found for CO speech produced in quiet and in noise. This was evident even for talkers with comparable intelligibility levels in quiet. The more variable intelligibility of speech produced in response to noise compared to CS found for healthy talkers should be taken into consideration when deciding on treatment options, CS versus NAS for instance, for speakers with dysarthria.

General Discussion

The current results revealed that young adult listeners significantly benefited from NAS and CS adaptations produced by child, young adult, and older adult talkers. Results also showed several age-related differences in intelligibility. Listeners found the CS and NAS produced by children less intelligible compared to the CS and NAS produced by adult talkers. Intelligibility of NAS produced by older adults was marginally less intelligible compared to the two younger talker groups. These findings highlight difficulties beyond perceptual problems that these populations encounter when communicating in adverse conditions. Their own speech may not be understood well in environments such as noisy classrooms and doctors’ offices, which may lead to more effortful processing on the part of their communication partners. The findings also suggest that not all neurologically typical talkers can produce gradient hyperarticulations in response to communicative barriers that enhance intelligibility for the listeners (Lindblom, 1990).

The intelligibility variation reported here is related to the different acoustic–artulatory modifications implemented by the three talker groups in response to the two communication challenges (Smiljanic & Gilbert, 2017). The current study provides evidence that listeners were able to utilize some of these modifications to enhance speech recognition in noise. It is, however, important to keep in mind that a direct link between any one acoustic–artulatory modification and increased intelligibility remains rather tenuous (Godoy et al., 2013; Krause, 2001; Krause & Braida, 2002, Liu & Zeng, 2006; Picheny, Durlach, & Braida, 1989; Tjaden, Kain, et al., 2014; Uchanski, Choi, Braida, Reed, & Durlach, 1996). It is likely that a combination of a number of cue modifications that individual talkers made gave rise to the intelligibility benefit. The current study was not designed to assess how much each of these individual adaptations contributes to intelligibility, but rather to illustrate holistically how conversational-to-clear and quiet-to-noise strategies affect intelligibility levels. The results suggest that certain modifications, such as a decrease in speaking rate and vocalic durations, increased pausing, and increased energy in 1- to 3-kHz region, may contribute to enhanced intelligibility. Combined results of the current study and those reported in the companion article suggest, however, that using CS and NAS as global, speech-oriented techniques, rather than targeting individual speech parameters, may be a better way of eliciting a number of these modifications simultaneously. CS and NAS thus hold great potential as therapy and training techniques for speakers with dysarthria or second language learners.

There are other ways in which this research should be enhanced in the future. First, it should consider age-matched listener groups. It is possible that other age-matched talker–listener groups (older adult and child listeners) would reveal different intelligibility results (Ferguson & Kewley-Port, 2002). We argued above that talker characteristics, rather than listener experience, shape intelligibility variation. It is important to examine, however, to what extent older adult and child listeners’ speech recognition in noise is determined by talker clarity versus their experience with different talkers. Second, future work should include more talkers within each age group and extending the age of the older adult group to include older individuals. This would allow us to detect group differences more easily as well as to establish a degree of intelligibility variation for neurologically typical learners.
talkers. Finally, a pressing goal for future work is to identify which cognitive, linguistic, and physiological factors may contribute to the age-related intelligibility variation. The present work is an important step toward understanding the effects of inter- and intratalker variability and environmental factors on intelligibility. Although work reported here has limitations, it constitutes a meaningful step toward a more comprehensive examination of intelligibility issues arising from developmental and aging changes. Furthermore, it provides support for speech-oriented, behavioral therapy techniques, such as NAS and CS, as valid approaches for maximizing speech intelligibility.

References


Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. Retrieved from http://CRAN.R-project.org/package=lme4


Palmeri, T., Goldinger, S., & Pisoni, D. 

Lindblom, B. 


Appendix

Individual talker intelligibility scores and rank order for conversational (CO) and clear speech (CS) produced in quiet and in noise for young adults (YA), older adults (OA), and children (CH).

<table>
<thead>
<tr>
<th>Talker</th>
<th>Quiet</th>
<th></th>
<th></th>
<th>Noise</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CO</td>
<td>Rank</td>
<td>CS</td>
<td>Rank</td>
<td>CO</td>
<td>Rank</td>
</tr>
<tr>
<td>YA01</td>
<td>4.8</td>
<td>27</td>
<td>23.8</td>
<td>25</td>
<td>61.9</td>
<td>7</td>
</tr>
<tr>
<td>YA02</td>
<td>0.0</td>
<td>29</td>
<td>38.1</td>
<td>13</td>
<td>47.6</td>
<td>14</td>
</tr>
<tr>
<td>YA03</td>
<td>28.6</td>
<td>5</td>
<td>66.7</td>
<td>5</td>
<td>42.9</td>
<td>17</td>
</tr>
<tr>
<td>YA04</td>
<td>19.0</td>
<td>13</td>
<td>38.1</td>
<td>13</td>
<td>57.1</td>
<td>8</td>
</tr>
<tr>
<td>YA05</td>
<td>14.3</td>
<td>20</td>
<td>38.1</td>
<td>13</td>
<td>9.5</td>
<td>30</td>
</tr>
<tr>
<td>YA06</td>
<td>23.8</td>
<td>8</td>
<td>28.6</td>
<td>21</td>
<td>57.1</td>
<td>8</td>
</tr>
<tr>
<td>YA07</td>
<td>23.8</td>
<td>8</td>
<td>47.6</td>
<td>9</td>
<td>23.8</td>
<td>26</td>
</tr>
<tr>
<td>YA08</td>
<td>4.8</td>
<td>27</td>
<td>19.0</td>
<td>29</td>
<td>23.8</td>
<td>26</td>
</tr>
<tr>
<td>YA09</td>
<td>19.0</td>
<td>14</td>
<td>47.6</td>
<td>9</td>
<td>76.2</td>
<td>2</td>
</tr>
<tr>
<td>YA10</td>
<td>38.1</td>
<td>2</td>
<td>71.4</td>
<td>3</td>
<td>95.2</td>
<td>1</td>
</tr>
<tr>
<td>OA01</td>
<td>15.0</td>
<td>15</td>
<td>60.0</td>
<td>7</td>
<td>45.0</td>
<td>15</td>
</tr>
<tr>
<td>OA02</td>
<td>60.0</td>
<td>1</td>
<td>95.0</td>
<td>1</td>
<td>50.0</td>
<td>11</td>
</tr>
<tr>
<td>OA03</td>
<td>20.0</td>
<td>10</td>
<td>75.0</td>
<td>2</td>
<td>65.0</td>
<td>5</td>
</tr>
<tr>
<td>OA04</td>
<td>30.0</td>
<td>4</td>
<td>50.0</td>
<td>8</td>
<td>35.0</td>
<td>22</td>
</tr>
<tr>
<td>OA05</td>
<td>25.0</td>
<td>6</td>
<td>20.0</td>
<td>26</td>
<td>40.0</td>
<td>18</td>
</tr>
<tr>
<td>OA06</td>
<td>5.0</td>
<td>24</td>
<td>65.0</td>
<td>6</td>
<td>40.0</td>
<td>18</td>
</tr>
<tr>
<td>OA07</td>
<td>15.0</td>
<td>15</td>
<td>70.0</td>
<td>4</td>
<td>40.0</td>
<td>18</td>
</tr>
<tr>
<td>OA08</td>
<td>20.0</td>
<td>10</td>
<td>20.0</td>
<td>26</td>
<td>45.0</td>
<td>15</td>
</tr>
<tr>
<td>OA09</td>
<td>15.0</td>
<td>15</td>
<td>25.0</td>
<td>22</td>
<td>35.0</td>
<td>22</td>
</tr>
<tr>
<td>OA10</td>
<td>20.0</td>
<td>10</td>
<td>35.0</td>
<td>16</td>
<td>40.0</td>
<td>18</td>
</tr>
<tr>
<td>CH01</td>
<td>25.0</td>
<td>7</td>
<td>45.0</td>
<td>11</td>
<td>55.0</td>
<td>10</td>
</tr>
<tr>
<td>CH02</td>
<td>5.0</td>
<td>24</td>
<td>25.0</td>
<td>22</td>
<td>75.0</td>
<td>3</td>
</tr>
<tr>
<td>CH03</td>
<td>10.0</td>
<td>21</td>
<td>30.0</td>
<td>20</td>
<td>50.0</td>
<td>11</td>
</tr>
<tr>
<td>CH04</td>
<td>10.0</td>
<td>21</td>
<td>35.0</td>
<td>16</td>
<td>50.0</td>
<td>11</td>
</tr>
<tr>
<td>CH05</td>
<td>10.0</td>
<td>21</td>
<td>35.0</td>
<td>16</td>
<td>35.0</td>
<td>22</td>
</tr>
<tr>
<td>CH06</td>
<td>35.0</td>
<td>3</td>
<td>45.0</td>
<td>11</td>
<td>75.0</td>
<td>3</td>
</tr>
<tr>
<td>CH07</td>
<td>5.0</td>
<td>24</td>
<td>10.0</td>
<td>30</td>
<td>10.0</td>
<td>29</td>
</tr>
<tr>
<td>CH08</td>
<td>15.0</td>
<td>15</td>
<td>25.0</td>
<td>22</td>
<td>65.0</td>
<td>5</td>
</tr>
<tr>
<td>CH09</td>
<td>15.0</td>
<td>15</td>
<td>35.0</td>
<td>16</td>
<td>15.0</td>
<td>28</td>
</tr>
<tr>
<td>CH10</td>
<td>0.0</td>
<td>29</td>
<td>20.0</td>
<td>26</td>
<td>30.0</td>
<td>25</td>
</tr>
</tbody>
</table>