Hearing Impairment and Perceived Clarity of Predictable Speech

Carine Signoret and Mary Rudner

Objectives: The precision of stimulus-driven information is less critical for comprehension when accurate knowledge-based predictions of the upcoming stimulus can be generated. A recent study in listeners without hearing impairment (HI) has shown that form- and meaning-based predictability independently and cumulatively enhance perceived clarity of degraded speech. In the present study, we investigated whether form- and meaning-based predictability enhanced the perceptual clarity of degraded speech for individuals with moderate to severe sensorineural HI, a group for whom such enhancement may be particularly important.

Design: Spoken sentences with high or low semantic coherence were degraded by noise-vocoding and preceded by matching or nonmatching text primes. Matching text primes allowed generation of form-based predictions while semantic coherence allowed generation of meaning-based predictions.

Results: The results showed that both form- and meaning-based predictions make degraded speech seem clearer to individuals with HI. The benefit of form-based predictions was seen across levels of speech quality and was greater for individuals with HI in the present study than for individuals without HI in previous studies. However, for individuals with HI, the benefit of meaning-based predictions was only apparent when the speech was slightly degraded. When it was more severely degraded, the benefit of meaning-based predictions was only seen when matching text primes preceded the degraded speech. The benefit in terms of perceptual clarity of meaning-based predictions was positively related to verbal fluency but not working memory performance.

Conclusions: Taken together, these results demonstrate that, for individuals with HI, form-based predictability has a robust effect on perceptual clarity that is greater than the effect previously shown for individuals without HI. However, when speech quality is moderately or severely degraded, meaning-based predictability is contingent on form-based predictability. Further, the ability to mobilize the lexicon seems to contribute to the strength of meaning-based predictions. Whereas individuals without HI may be able to activate explicit working memory capacity for storing meaning-based predictions, individuals with HI may already be using all available explicit capacity to process the degraded speech and thus become reliant on explicit skills such as the temporal fine structure of speech signal into a specific number of frequency bands in which the temporal amplitude envelope is preserved. Therefore, the potential support of such predictability could be invaluable.

Sensorineural HI is a common form of internal degradation of the auditory signal and is associated with aging. According to the definition of disabling hearing loss proposed by the World Health Organization (i.e., more than 40 dB loss in the better ear for adults), one-third of people aged 65+ has a disabling hearing loss (World Health Organization 2018). This also means that 95% of people aged 80+ has a disabling hearing loss (see median values of expected hearing threshold deviations presented in International Organisation for Standardization DN/EN ISO 7029:2017). Sensorineural HI challenges communication in everyday life situations. Not only is the speech signal degraded by internal factors but sensorineural HI also reduces spectrotemporal resolution. This results in difficulties to segregate target speech from other external auditory signals, such as surrounding music or background noise that occurs in many social situations, especially when they take place in public places including restaurants, train stations, supermarkets, or city centers. Thus, it seems likely that predictability of upcoming speech would be of particular value for individuals with sensorineural HI. The main aim of the present study is to determine whether form- and meaning-based predictions can improve the perceived clarity of externally degraded speech for individuals with HI.

External degradation of speech can be manipulated with noise-vocoding (Shannon et al. 1995). This involves decomposition of the speech signal into a specific number of frequency bands in which the temporal amplitude envelope is preserved, but the temporal fine structure of speech signal within each frequency band is removed, allowing us to precisely control the level of sound quality. Intelligibility of noise-vocoded speech (NVS) varies as a function of the number of frequency bands and the number of bands leading to better intelligibility (Shannon et al. 1995) and consequently better perceptual clarity (Sohoglu et al. 2014). Indeed, NVS perceptual clarity has been shown to be highly associated with NVS intelligibility (Obleser et al. 2008). NVS has been used in several studies investigating degraded speech perception for individuals without HI (Davis et al. 2011; Hervais-Adelman et al. 2011), as well as in studies comparing speech perception performance in younger (e.g., below 35 years old) and older (e.g., above 55 years old) adults without HI (Thiel et al. 2016; Rosemann et al. 2017). Some studies have shown that older adults need more frequency bands to achieve the same level of perception performance as younger adults (Sheldon et al. 2008b; Experiment 3 in Sheldon et al. 2008a). Others have shown no difference in NVS perceptual learning between older and young adults (Peelle & Wingfield 2018).

Conclusions: Taken together, these results demonstrate that, for individuals with HI, form-based predictability has a robust effect on perceptual clarity that is greater than the effect previously shown for individuals without HI. However, when speech quality is moderately or severely degraded, meaning-based predictability is contingent on form-based predictability. Further, the ability to mobilize the lexicon seems to contribute to the strength of meaning-based predictions. Whereas individuals without HI may be able to activate explicit working memory capacity for storing meaning-based predictions, individuals with HI may already be using all available explicit capacity to process the degraded speech and thus become reliant on explicit skills such as the temporal fine structure of speech signal into a specific number of frequency bands in which the temporal amplitude envelope is preserved. Therefore, the potential support of such predictability could be invaluable.

Sensorineural HI is a common form of internal degradation of the auditory signal and is associated with aging. According to the definition of disabling hearing loss proposed by the World Health Organization (i.e., more than 40 dB loss in the better ear for adults), one-third of people aged 65+ has a disabling hearing loss (World Health Organization 2018). This also means that 95% of people aged 80+ has a disabling hearing loss (see median values of expected hearing threshold deviations presented in International Organisation for Standardization DN/EN ISO 7029:2017). Sensorineural HI challenges communication in everyday life situations. Not only is the speech signal degraded by internal factors but sensorineural HI also reduces spectrotemporal resolution. This results in difficulties to segregate target speech from other external auditory signals, such as surrounding music or background noise that occurs in many social situations, especially when they take place in public places including restaurants, train stations, supermarkets, or city centers. Thus, it seems likely that predictability of upcoming speech would be of particular value for individuals with sensorineural HI. The main aim of the present study is to determine whether form- and meaning-based predictions can improve the perceived clarity of externally degraded speech for individuals with HI.

External degradation of speech can be manipulated with noise-vocoding (Shannon et al. 1995). This involves decomposition of the speech signal into a specific number of frequency bands in which the temporal amplitude envelope is preserved, but the temporal fine structure of speech signal within each frequency band is removed, allowing us to precisely control the level of sound quality. Intelligibility of noise-vocoded speech (NVS) varies as a function of the number of frequency bands and the number of bands leading to better intelligibility (Shannon et al. 1995) and consequently better perceptual clarity (Sohoglu et al. 2014). Indeed, NVS perceptual clarity has been shown to be highly associated with NVS intelligibility (Obleser et al. 2008). NVS has been used in several studies investigating degraded speech perception for individuals without HI (Davis et al. 2011; Hervais-Adelman et al. 2011), as well as in studies comparing speech perception performance in younger (e.g., below 35 years old) and older (e.g., above 55 years old) adults without HI (Thiel et al. 2016; Rosemann et al. 2017). Some studies have shown that older adults need more frequency bands to achieve the same level of perception performance as younger adults (Sheldon et al. 2008b; Experiment 3 in Sheldon et al. 2008a). Others have shown no difference in NVS perceptual learning between older and young adults (Peelle & Wingfield 2018).

INTRODUCTION

We tend to hear what we expect to hear—this is a common observation and recently supported by data showing that the perceived clarity of degraded speech is enhanced by both predictability based on form (i.e., phonological/lexical) at the word level and on meaning (i.e., semantic) at the sentence level (Signoret et al. 2018). For individuals with hearing impairment (HI) who may struggle particularly hard to understand degraded speech, the potential support of such predictability could be invaluable.

Sensorineural HI is a common form of internal degradation of the auditory signal and is associated with aging. According to the definition of disabling hearing loss proposed by the World Health Organization (i.e., more than 40 dB loss in the better ear for adults), one-third of people aged 65+ has a disabling hearing loss (World Health Organization 2018). This also means that 95% of people aged 80+ has a disabling hearing loss (see median values of expected hearing threshold deviations presented in International Organisation for Standardization DN/EN ISO 7029:2017). Sensorineural HI challenges communication in everyday life situations. Not only is the speech signal degraded by internal factors but sensorineural HI also reduces spectrotemporal resolution. This results in difficulties to segregate target speech from other external auditory signals, such as surrounding music or background noise that occurs in many social situations, especially when they take place in public places including restaurants, train stations, supermarkets, or city centers. Thus, it seems likely that predictability of upcoming speech would be of particular value for individuals with sensorineural HI. The main aim of the present study is to determine whether form- and meaning-based predictions can improve the perceived clarity of externally degraded speech for individuals with HI.

External degradation of speech can be manipulated with noise-vocoding (Shannon et al. 1995). This involves decomposition of the speech signal into a specific number of frequency bands in which the temporal amplitude envelope is preserved, but the temporal fine structure of speech signal within each frequency band is removed, allowing us to precisely control the level of sound quality. Intelligibility of noise-vocoded speech (NVS) varies as a function of the number of frequency bands and the number of bands leading to better intelligibility (Shannon et al. 1995) and consequently better perceptual clarity (Sohoglu et al. 2014). Indeed, NVS perceptual clarity has been shown to be highly associated with NVS intelligibility (Obleser et al. 2008). NVS has been used in several studies investigating degraded speech perception for individuals without HI (Davis et al. 2011; Hervais-Adelman et al. 2011), as well as in studies comparing speech perception performance in younger (e.g., below 35 years old) and older (e.g., above 55 years old) adults without HI (Thiel et al. 2016; Rosemann et al. 2017). Some studies have shown that older adults need more frequency bands to achieve the same level of perception performance as younger adults (Sheldon et al. 2008b; Experiment 3 in Sheldon et al. 2008a). Others have shown no difference in NVS perceptual learning between older and young adults (Peelle & Wingfield 2018).
Only one study to our knowledge has previously investigated NVS perception for individuals with HI (Souza & Boike 2006) and reported lower speech recognition performance for older individuals with HI than younger adults without HI at NVS 4 and 8 bands (as well as clear speech), but similar speech recognition performance at NVS 1 and 2 bands. However, age but not the degree of hearing loss (mild to severe) was a significant predictor of NVS recognition performance, possibly due to greater lexical knowledge associated with age (Wingfield et al. 2005; Sheldon et al. 2008a) or slower processing. In the study of Neger et al. (2014) comparing younger and older adults (of whom one-third had untreated HI), processing speed was a significant predictor of NVS perception performance for older adults only. To our knowledge, no study has yet investigated NVS perception for individuals with HI. However, as the envelope of the speech signal is preserved in both sensorineural HI and NVS, we assume that individuals with HI and hearing aids will be able to perceive NVS in much the same way as individuals with normal hearing. Because no difference in NVS perceptual learning has been reported between younger and older adults without HI, NVS perception of individuals with HI with hearing aids should be similar to NVS perception of older adults without HI.

NVS is a key tool for investigating the influence of knowledge-based support during degraded speech perception. For example, giving access to the content of the upcoming auditory sentence before hearing the speech signal leads to a clearer perception of the sentence (Davis et al. 2005). For individuals without HI, this effect, known as the “pop-out effect,” has been investigated extensively (Davis et al. 2005; Davis & Johnsrude 2007; Hervais-Adelman et al. 2011; Wild et al. 2012), whereas it has not, to our knowledge, been reported whether individuals with HI experience the pop-out effect. For individuals without HI, presenting matching text only 200 msec before presenting a degraded spoken sentence, improves the perceived clarity of the sentence more than meaningless consonant strings (Wild et al. 2012; Signoret et al. 2018). This improvement in perceptual clarity is thought to be related to knowledge-based predictions about the phonological/lexical form of the upcoming word influencing upcoming speech processing (Sohoglu et al. 2014; Signoret et al. 2018). Perceptual clarity of degraded spoken sentences is also greater for sentences with high compared with low semantic coherence (Signoret et al. 2018). Such an improvement may be explained by the fact that knowledge-based predictions about the semantic meaning of the upcoming sentence facilitate processing by reducing the number of lexical candidates (Federmeier 2007). The facilitative effects of phonological/lexical form- and semantic meaning-based predictions are independent but also additive, suggesting that several sources of knowledge-based predictions (i.e., at different linguistic levels, see Pickering & Garrod 2013) could combine to improve the perceptual clarity of degraded speech (Signoret et al. 2018). Although the independent and additive facilitative effects of form- and meaning-based predictions have been highlighted for individuals without HI (Signoret et al. 2018), it is unclear whether the pattern of effects would be similar for individuals with HI.

The ability to understand degraded speech is associated with working memory (Rudner et al. 2011), especially for individuals with HI. For individuals without HI, the perceptual clarity benefit associated with form- and meaning-based predictability is also related to working memory (Signoret et al. 2018). The Ease of Language Understanding model (Rönnberg et al. 2013) proposes two particular roles of working memory in speech processing: predictive and postdictive. Predictively, working memory is likely to be involved in the storage of knowledge-based predictions while postdictively, working memory is likely to be involved during the explicit processing of degraded speech when a mismatch occurs between knowledge-based predictions and the auditory signal. As the perceived auditory signal is always degraded for individuals with HI, it is likely that they use linguistic and cognitive abilities differently than individuals without HI (Wingfield et al. 2005; Rogers et al. 2012). In particular, individuals with HI are probably used to utilizing working memory capacity in its predictive function to minimize slow and effortful explicit processing (Rudner et al. 2011). If the speech signal matches knowledge-based predictions, lexical access will be faster. For individuals with HI using hearing aids, there is evidence that degraded speech recognition is more dependent on working memory when speech is more highly degraded and no meaning-based predictions are available (Rudner et al. 2011). Because knowledge-based predictions are dependent on the lexicon, it could be argued that the larger the lexicon, the better the quality of the meaning-based predictions, and the higher the probability of implicit processing of the incoming signal (Benard et al. 2014). It is important to evaluate whether the benefit of form- and meaning-based predictions during perception of degraded speech is associated with these cognitive and linguistic factors. This knowledge could be used in the design of technological supports such as written text assistance for individuals with HI and consequently improve their experience of everyday life in social situations.

In the present study, we investigated how form- and meaning-based predictions influence the perceptual clarity of degraded speech for individuals with HI. Based on our previous study investigating this question for individuals without HI (Signoret et al. 2018), we compared perceptual clarity ratings of spoken sentences with high or low semantic coherence that were degraded with noise-vocoding and preceded by matching or nonmatching (nM) text primes. Matching text primes allowed generation of form-based predictions while coherence allowed generation of meaning-based predictions (Signoret et al. 2018). Individuals with moderate to severe sensorineural hearing loss and wearing hearing aids were recruited to participate in the present study. Because NVS seems to be perceived similarly by individuals with and without HI, we expected to find main effects of both form- and meaning-based predictions as well as a significant interaction showing that meaning-based predictions enhance perceptual clarity even when form-based predictions can be made based on the text primes provided. We also investigated how linguistic and cognitive factors are associated with the enhancement of the perceptual clarity of degraded speech by form- and meaning-based predictions. We hypothesized that (1) greater working memory capacity will correlate with higher clarity ratings for heavily degraded spoken sentences with low semantic coherence (Rudner et al. 2011) and (2) greater verbal fluency (indexing a larger mental lexicon) will correlate with higher clarity ratings for slightly degraded spoken sentences with high semantic coherence (Benard et al. 2014).
MATERIALS AND METHODS

Participants
Twenty-two native Swedish speakers (11 women; age range: 52 to 75, M_age = 69.41, SD_age = 5.62) participated in this study. All were experienced (at least 1 year) bilateral hearing aid users (M = 5.86 years, SD = 3.03) with moderate-to-severe symmetrical sensorineural hearing loss (see pure-tone audiometry [PTA] thresholds in Fig. 1 at 8 frequencies ranging from 125 to 8000 Hz). All participants were recruited through the audiological department at Linköping Hospital and reported normal or corrected-to-normal vision and had no history of psychological disorder or neurological disease. After reading an information letter, all participants provided written informed consent to the study, which was conducted in accordance with the guidelines of the Declaration of Helsinki and approved by the regional Ethics Committee in Linköping (2011/403–31).

Materials
The details of the materials are described in Signoret et al. (2018). In brief, we used a set of 200 syntactically correct Swedish sentences of which 100 with high semantic coherence (HC, e.g., “Hennes dotter var för ung för att gå på disco” (“Her daughter was too young for the disco”)) and 100 with low semantic coherence (LC, e.g., “Hennes hockey var för tät för att gå på bomull” (“Her hockey was too tight for the cotton”)). Sentences were matched in terms of length, words, number of words, syllables, and letters between the two subsets of sentences.

The sound quality of the 200 sentences was degraded with noise-vocoding at 1 (NV1), 3 (NV3), 6 (NV6), or 12 (NV12) contiguous frequency bands (cf. Signoret et al. 2018). NV1 is unintelligible to listeners but intelligibility increases with increasing number of bands. Words degraded with noise-vocoding up to 4 bands (i.e., severely degraded) are difficult to perceive, while those degraded with noise vocoding 5 to 8 bands are around 50% intelligible, and those degraded with noise vocoding at more than 10 bands are easy to perceive (see Shannon et al. 2004). After processing, all stimuli were normalized to reach the same total RMS power. The onset time of each word in each sentence was determined with the NALIGN automatic segmentation algorithm (Sjölander & Heldner 2004). Stimulus presentation and behavioral responses were controlled using E-Prime 2.0 software (Psychology Software Tools, Inc. www.pstnet.com/eprime) with a DELL Latitude 6420 laptop computer connected with an external 20-inch screen, a Swedish keyboard, and an external soundcard (M-Audio FireWire 410). Auditory stimuli were binaurally presented to the participants through headphones (Sennheiser HD 600) connected to the external soundcard. Participants did not wear their own hearing aids, but the presentation levels of the sentences were individually adjusted according to the Cambridge formula (Moore & Glasberg 1998), with frequency-dependent amplification based on the pure-tone thresholds of the best ear. For the experienced hearing aid users included in the study, this ensured audibility of the clear sentences and reasonable intelligibility of NVS sentences while retaining differences in perceived clarity between sound quality levels. Participants were tested in a sound-attenuated room (standard ISO 8253 and ISO BS EN6189).

Perceptual Clarity Experiment
The perceptual clarity experiment conformed to a repeated-measures design with sound quality (clear, NV12, NV6, NV3, NV1), prime (M, nM), and coherence (HC, LC) as within-participant factors. Participants were instructed to keep their eyes on the computer screen, and to rate the clarity of each sentence they heard on a seven-point Likert scale (from one for very unclear to seven for totally clear). They practiced the task on four trials in which they had to recognize these two sound quality levels (NV1 and clear) as the endpoints of the rating scale (1 and 7, respectively). No or negligible influence from the presence of primes or semantic coherence is expected at these two sound quality levels, which were used as an indication of task compliance (Signoret et al. 2018). Each trial began with a white fixation cross in the middle of a black computer screen for 500 msec. Twenty LC and 20 HC sentences were presented at each level of sound quality (clear, NV12, NV6, NV3, NV1); sentence assignment to sound quality level was counterbalanced across participants. Ten of the LC and HC items at each sound quality level were presented with corresponding text primes (M primes) and the other 10 with length-matched strings of random consonants (nM primes). This ensured that there was no repetition of any sentence within participants while all sentences were presented in every condition across participants. A total of 200 sentences were randomly presented. Primes were presented as white text (courier new, 32 points) on a black screen, replacing the fixation cross in the middle of the screen. Primes appeared word-by-word such that each word preceded the homologous auditory segment by 200 msec (Wild et al. 2012). Each text prime (M or nM) was displayed for a period equal to the time required to produce the primed word in the corresponding audio file. At the end of the trial, a question mark appeared, and participants rated the clarity of the spoken sentence. Immediately after the response of the participant, a white fixation cross appeared on the black screen for 500 msec before the next trial.

Cognitive/Linguistic Test Battery
To determine whether individual differences in cognitive/linguistic abilities could predict rated perceptual clarity, an independent test battery described below was administered.

Cognitive Abilities • Block design: For determining cognitive status and nonverbal intelligence, the BD subtest of the Wechsler Adult Intelligence Scale (WAIS-IV test, Hartman 2009) was used. This test has previously been shown to be correlated with the Wechsler Adult Intelligence Scale total score.
(Wechsler 1981). The test was administered following the standard procedure and scoring was calculated with time bonus. **Size-Comparison Span test:** The Size-Comparison Span test (SiCSpan; Sörqvist et al. 2010) is sensitive to working memory capacity. In the SiCSpan test, questions in Swedish about the relative size of two objects in the same category were presented on a screen (e.g., is a cow larger than a cat?). After a yes/no button press response (right hand for yes/left hand for no), another word designating an object belonging to the same object category was presented (e.g., crocodile). The participants were instructed to maintain this intermediate word in memory for later recall. Participants began the test with two questions (i.e., two intermediate words to recall) and the difficulty increased progressively until six questions (i.e., intermediate words to recall). Immediately after the final intermediate word had been presented, the participants were cued to recall. To do this successfully, the participants had to recall all the intermediate words, while inhibiting report of the words included in the sizecomparison questions. The score was calculated as the number of correct responses (CRs) during the recall phase. The maximum score is 40.

**Linguistic Abilities • Verbal Fluency test:** Verbal fluency was tested using an adaptation of the first part of the FAS test (Spren & Strauss 1998). The participants were given 90 sec to write down as many Swedish words as possible beginning with the letter F. They were instructed to avoid repetitions, nonexistent words, variations of the same word and proper nouns and not to care about the orthography. The participants were informed that after the 90 sec were up, they would be given the opportunity to correct their spelling if they wished, so that would not spend time attending to orthography during the test. The score corresponded to the number of correct words given by the participant.

**Verbal Information Processing Speed test:** Verbal abilities were tested using the Verbal Information Processing Speed (VIPS) test (Hällgren et al. 2001). Participants responded by pressing the yes button with the right hand and the no button with the left hand. Scores were calculated by averaging reaction times (RTs) for CRs.

**Warm-up task:** This test was designed to obtain baseline RT. The participants were asked to respond by pressing the appropriate button as fast as possible when “yes” or “no” (i.e., “ja” or “nej” in Swedish) appeared on the screen. Ten trials were presented in total to train the participants to use the response buttons. A success rate of 60% is mandatory to allow the participant to perform the other tasks of the VIPS test. The RTs to this task will be subtracted to the RTs on the other tasks in order to obtain an unbiased measure of access speed (i.e., excluding motor reaction in pressing a button or general processing speed, see Carroll et al. 2016).

**Lexical-decision task:** Participants were asked to judge whether three-letter strings were legitimate Swedish words. Forty items were used in the test: 20 real familiar Swedish words and 10 pseudowords (i.e., pronounceable letter strings) and 10 nonwords (i.e., unpronounceable consonant strings).

**Phonological decision task:** Participants were asked to decide whether two visually presented words rhymed or not. Thirty-two pairs of words were presented; half of the pairs rhymed, half did not. In each case, half of the pairs were orthographically similar. For example, rhyming pairs included orthographically similar items, e.g., KATT/HATT (i.e., cat/hat) and orthographically dissimilar items, e.g., DAGS/LAX, (i.e., day/salmon pronounced [daks/laks]). Nonrhyming pairs included orthographically similar items, e.g., STÅL/STEL (i.e., steel/stiff), and orthographically dissimilar items, e.g., CYKEL/SPADE (i.e., bike/shovel). Thirty-two pairs, eight of each type, were presented.

**Semantic decision task:** Participants were asked to decide whether the presented word belonged to a predefined semantic category. Four categories (animals, lands, sports, and vegetables) were used. Within each of 4 blocks (1 per category), 24 word items were presented, of which 16 belonged to the category and 8 did not.

**Overall Procedure**

After informed consent was obtained, PTA and a handedness questionnaire (Oldfield 1971) were administered. The cognitive-linguistic test battery was administered in the following order: BD, FAS, VIPS (warm-up, lexical decision task, phonological decision task, and semantic decision task), and SiCSpan. The perceptual clarity experiment was administered last. The total duration of the test session was approximately 75 min.

**Statistical Analysis**

As variance was expected to be too low at NV1 and clear sound quality levels, clarity ratings for these two levels were not planned to be included in the analysis of variance (ANOVA). Mean clarity ratings were then entered in a 3*Sound Quality (NV12, NV6, NV3) * 2*Prime (M, nM) * 2*Coherence (HC, LC) repeated-measures ANOVAs in Statistica 13 (Hill & Lewicki 2005). The Student–Newman–Keuls test was applied for post hoc comparisons (Howell 1998). An alpha level of 0.05 was used after Greenhouse-Geisser correction where appropriate. Due to the small sample size and the comparison between different measurement scale units, Spearman rank correlations were chosen for calculating correlations between clarity ratings (Likert-scale) and scores (CRs or RTs) on the cognitive-linguistic test battery, to confirm (1) the involvement of working-memory capacity in the most degraded condition (NV3) of the experiment and (2) the involvement of linguistic abilities at low levels of sound quality degradation. Spearman rank correlations were also calculated to determine the association between cognitive-linguistic abilities working-memory and verbal fluency and the perceptual clarity benefits contingent on matching text primes and semantic coherence. Bonferroni correction for multiple comparisons was applied where appropriate.

**RESULTS**

**Perceptual Clarity Experiment**

The clarity ratings at the five sound quality levels (clear, NV12, NV6, NV3, and NV1) with and without text primes (M, nM, respectively) and for sentences of high and low semantic coherence (HC, LC, respectively) are shown in Table 1.

Repeated-measure ANOVAs conducted on the 3 * 2*Sound Quality conditions revealed a main effect of sound quality \[F_{(1.68,35.37)} = 83.31, \ p < 0.001; \ \eta^2 = 0.799\]. Posthoc comparisons (Student–Newman–Keuls) demonstrated that clarity was rated higher at NV12 (M = 3.76; SE = 0.48) than at NV6 (M = 3.04; SE = 0.46), and higher at NV6 than at NV3 (M = 2.41; SE = 0.35; all \( p < 0.001\)). Clarity was rated higher in the
M (M = 3.69; SE = 0.58) than in the nM (M = 2.45; SE = 0.51) condition as revealed by a main effect of prime \( F_{(1,21)} = 61.69; p < 0.001; \eta^2_p = 0.746 \) demonstrating that matching text primes enhance perceptual clarity. Clarity was rated higher in HC (M = 3.32; SE = 0.56) than in LC (M = 2.82; SE = 0.49) sentences as revealed by a main effect of coherence \( F_{(1,21)} = 18.68; p < 0.001; \eta^2_p = 0.471 \) demonstrating that semantic coherence enhances perceptual clarity.

There was a significant interaction between Prime and Coherence \( F_{(1.27,26.61)} = 3.89; p = 0.050 \), suggesting an increasing effect of prime with decreasing sound quality level. The interaction between Sound Quality and Prime \( F_{(1.27,26.61)} = 3.89; p = 0.050 \) was due to the difference in perceptual clarity between M and nM conditions (i.e., the effect of prime) being smaller at NV12 (M = 4.24 and 3.28, respectively), than at NV6 (M = 3.73 and 2.35, respectively) and NV3 (M = 3.09 and 1.73, respectively) sound quality level, although all differences were statistically significant (all ps < 0.006), suggesting an increasing effect of prime with decreasing sound quality level. The interaction between Sound Quality and Coherence was not significant.

Little variance was expected at NV1 and clear sound quality levels, but the inspection of the results indicated different findings. Detailed results for these two sound quality levels indicated that, at NV1, 3 participants rated perceptual clarity higher for M than nM sentences, suggesting that they perceived greater clarity with matching text primes even when speech was unintelligible. One participant gave low ratings overall and rated clear sentences as one. This participant had the worst hearing thresholds on intermediate frequencies (0.5 to 2 kHz). When we removed these participants from the analysis (N = 18, 9 women; Mage = 69.22, SD age = 6.13), variance decreased at the NV1 sound quality level (M = 4.06 and 3.32, respectively; p < 0.001) but not in the nM condition (M = 2.57 and 2.33, respectively; p > 0.05, n.s.). This finding indicates that semantic coherence facilitates perceptual clarity only when providing full phonological/lexical information up front.

The two-way interaction between Prime and Coherence was tempered by a three-way interaction among sound quality, prime, and coherence \( F_{(2,42)} = 3.37; p = 0.050; \eta^2_p = 0.138 \) suggesting that the effects of prime and coherence both depend on the sound quality level (Fig. 3). The effect of prime (M versus nM) was observed at all sound quality levels for all sentences. In other words, for individuals with HI, matching text primes increase the perceived clarity of speech irrespective of the semantic coherence of the sentences. The effect of coherence (HC versus LC) was observed at all degraded sound quality levels in M condition (all ps < 0.001) but only at the NV12 sound quality level in nM condition (p < 0.001). In other words, the perceived clarity of degraded sentences increases with semantic coherence if matching text primes are available but semantic coherence gives no benefit in perceived clarity when text primes are nM, except at the NV12 sound quality level.

Similarly, a significant interaction between Sound Quality and Prime \( F_{(1.27,26.61)} = 3.89; p = 0.050 \) was due to the difference in perceptual clarity between M and nM conditions (i.e., the effect of prime) being smaller at NV12 (M = 4.24 and 3.28, respectively), than at NV6 (M = 3.73 and 2.35, respectively) and NV3 (M = 3.09 and 1.73, respectively) sound quality level, although all differences were statistically significant (all ps < 0.006), suggesting an increasing effect of prime with decreasing sound quality level. The interaction between Sound Quality and Coherence was not significant.

Little variance was expected at NV1 and clear sound quality levels, but the inspection of the results indicated different findings. Detailed results for these two sound quality levels indicated that, at NV1, 3 participants rated perceptual clarity higher for M than nM sentences, suggesting that they perceived greater clarity with matching text primes even when speech was unintelligible. One participant gave low ratings overall and rated clear sentences as one. This participant had the worst hearing thresholds on intermediate frequencies (0.5 to 2 kHz). When we removed these participants from the analysis (N = 18, 9 women; Mage = 69.22, SD age = 6.13), variance decreased at the NV1 sound quality level (M = 1.05; SD = 0.155), but we still observed substantial variance at the clear sound quality level (M = 6.50; SD = 0.767). This could be explained by the fact that even at the clear sound quality level, individuals with HI still do not perceive clear speech as clear. Indeed, at this sound quality level, the benefit of prime (0.328) but not coherence (0.092) is still observable. Supplementary analysis including clear and NV1 sound quality levels to obtain a 5*Sound Quality (clear, NV12, NV6, NV3, NV1) × 2*Prime (M, nM) × 2*Coherence (HC, LC)
repeated-measures ANOVAs confirmed that the effect of coherence was not significant at the clear sound quality level ($p = 0.31$). However, excluding these participants from the analysis had a negligible impact on the results and thus we retained them in the analysis.

**Comparison of Individuals With or Without HI**

To examine differences in perceptual clarity ratings between listeners with and without HI (Signoret et al., 2018), a repeated-measures ANCOVA was conducted. As no variance was observed at the clear and NV1 Sound quality levels for listeners without HI, mean clarity ratings were entered in a 3*Sound Quality (NV12, NV6, NV3) × 2*Prime (M, nM) × 2*Coherence (HC, LC). As the two groups are not matched on age (M = 69.41, SD = 5.62 and M = 31.34, SD = 10.32, respectively, for listeners with and without HI; $t(49) = 14.81, p < 0.001$), this variable was entered as a covariate. Results revealed a main effect of sound quality [$F(2,78) = 21.23, p < 0.001$], showing better perceptual clarity rating with higher sound quality level, as well as a main effect of prime [$F(1,39) = 9.58, p < 0.004$] and a marginal interaction between Prime and Hearing status [$F(1,39) = 3.97; p = 0.053$]. Posthoc analysis (unequal N Tukey HSD) suggested that benefit in perceptual clarity from matching text primes was larger for listeners with HI than listeners without HI (mean benefit = 1.24 and 0.83, respectively), while no significant difference in perceptual clarity ratings was observed between the two groups when matching text primes were presented ($p = 0.458$).

**Cognitive/Linguistic Test Battery**

Results obtained in the different tests of the cognitive/linguistic battery are displayed in Table 2 and are comparable with performances observed in previous literature with similar age groups (for BD test see, Troyer et al. 1994; for FAS test, see Tombaugh et al. 1999; and for VIPS test, see Häflgren et al. 2001), except for SiCSpan scores that are lower than previously reported (M = 29.0, SD = 5.01) for a group of hearing aid users of 67.4 years old with mild to moderate hearing loss (Heinrich et al. 2016). Three participants elected not to complete the VIPS test and one elected not to complete the SiCSpan test.

Intercorrelations between tests of the cognitive/linguistic battery are shown in Table 3. VIPS RTs were negatively associated with working memory abilities (i.e., SiCSpan scores) but this correlation did not survive Bonferroni correction. SiCSpan scores were however positively associated with BD scores, which is not surprising as BD is used as a general measure of nonverbal intelligence.

**TABLE 2. Cognitive/linguistic test battery results**

| Test        | CRs (%)   | SD (%)  | RTs (msec)   | SD (msec) |
|-------------|-----------|---------|--------------|-----------|
| BD          | 35.32     | ±10.83  | 1156.54      | ±261.26   |
| FAS         | 13.55     | ±5.54   |              |           |
| SiCSpan     | 14.76     | ±6.42   |              |           |
| VIPS        | 91.25     | ±6.16   | 1156.54      | ±261.26   |
| Warm-up     | 86.84     | ±12.5   | 825.83       | ±244.42   |
| Lexical     | 96.05     | ±6.49   | 1631.14      | ±466.74   |
| Phonological| 81.32     | ±15.44  | 961.72       | ±216.45   |
| Semantic    | 96.37     | ±2.53   | 834.19       | ±124.71   |

Mean CRs and SD are reported for BD, SiCSpan, and verbal fluency (FAS) tests. Mean RTs and SD for CRs are reported for the tasks (warm-up, lexical decision task, phonological decision task, and semantic decision tasks) included in the VIPS test.

**TABLE 3. Spearman’s rho correlation coefficients between the tests in the cognitive/linguistic test battery: mean correct responses for nonverbal ability with BD test, working memory with SiCSpan test, and verbal fluency with FAS test and mean reaction times for the decision tasks (lexical, phonological, and semantic) included in the VIPS test**

| Spearman's Rho | BD   | FAS  | SiCSpan |
|----------------|------|------|---------|
| FAS            | 0.205|      |         |
| SiCSpan        | 0.670*| 0.018|         |
| VIPS           | −0.316| −0.336| −0.546  |
| Phonological   | −0.272| −0.300| −0.518  |
| Lexical        | −0.453| −0.238| −0.530  |
| Semantic       | −0.324| −0.308| −0.551  |

*p < 0.001.*

BD, block design; SiCSpan, size-comparison span; VIPS, verbal information processing speed.

**Auditory Skills**

We also examined with Spearman’s rho correlations the relation among auditory skills (i.e., PTA thresholds) and age. The results showed that, in the sample studied, age was not significantly associated with HI on low frequencies thresholds (averaged across 125, 250, and 500 Hz; $r = 0.35$, $p = 0.11$), speech frequencies thresholds (averaged across 1000, 2000, and 4000 Hz; $r = 0.12$, $p = 0.61$) or high frequencies thresholds (averaged across 6000 and 8000 Hz; $r = −0.05$, $p = 0.84$). This is likely due to the fact that the studied sample was recruited from a clinical population.

**DISCUSSION**

The present study aimed to investigate whether meaning-based predictions (driven by semantic coherence) and
form-based predictions (driven by the phonological/lexical form of words) could enhance perceived clarity of degraded speech for individuals with HI and how this knowledge-based predictability is associated with linguistic and cognitive factors. The results showed that form-based predictability improved the perceived clarity of degraded speech for individuals with HI, irrespective of meaning-based predictability. However, in contrast to individuals without HI (Signoret et al. 2018), in the absence of form-based predictability, meaning-based predictions seem to be only possible when speech is only slightly degraded (NV speech 12 bands), i.e., where the participants could probably hear well enough without the matching text primes to benefit from sentence coherence. When it is more severely degraded (NV speech 6 and 3 bands), the benefit of meaning-based prediction was only seen when form-based predictions could also be made. The benefit in terms of perceptual clarity of meaning-based predictions was positively related to verbal fluency, as measured by the FAS test (Spreen & Strauss 1998) in an adapted written procedure, but was not related to working memory capacity as measured with SiCSpan test (Sörqvist et al. 2010).

The present study is the first to our knowledge to investigate both form- and meaning-based predictability on the perceptual clarity of speech for individuals with HI. The results demonstrated that the presentation of matching text primes before the degraded spoken sentences improved the perceived clarity of NV speech clarity at all levels of sound quality for individuals with HI by allowing form-based predictions. This result demonstrates that predictions based on the phonological/lexical form at a word level improve the perceived clarity of degraded speech for individuals with HI. Because perceived clarity and intelligibility of NV speech are highly correlated for individuals without HI (Obleser et al. 2008), we speculate that matching text primes will also enhance the intelligibility of degraded speech for individuals with HI. The results of the present study extend previous results observed for younger individuals without HI (Obleser et al. 2012) and extending these studies should investigate if other types of predictability, such as emotion-based predictability, would affect degraded speech perception in a similar manner.

Form-based predictability could be manipulated for example with the congruency between the prosody and the content of a sentence, or by comparing neutral versus emotional words or sentences. The present findings demonstrated that meaning- and form-based predictability have cumulative, but not independent, facilitative effects on degraded speech perception for individuals with HI. This supports the idea that more precise predictions (at different linguistic levels) about upcoming spoken words can help compensate for sensory degradation. It could be the case that emotion-based predictability has also an additive facilitative effect on degraded speech perception.

The effect of form-based predictability was also observed, although smaller, in clear sentences. The clear level of sound quality was originally included as a control of good task compliance. However, it is not surprising to observe a facilitative effect of form-based predictability at this clear sound quality level. Indeed, participants included in this study are individuals with moderate to severe symmetrical sensorineural HI. This means that even if the signal is perfectly clear in the environment, it is still perceived with some internal degradations that are not entirely compensated by hearing aid amplification. At a sound quality level that is too low to provide intelligibility even for persons with no HI (i.e., NVS with one band), 3 participants showed an effect of form-based predictability, by reporting better perceptual clarity ratings for sentences preceded by matching text primes than nM text primes. This may indicate that these participants were particularly influenced by the presentation of prior matching text primes. It is possible that these participants perceived auditory illusions due to the presence of matching text primes (Rogers et al. 2012) and may have relied more on the form-based predictions as the stimulus-driven information decreased. Future studies should further investigate the importance of form-based prediction for individuals with HI, as has been done for individuals without HI (Sohoglu et al. 2014). Typical questions to be answered relate to the duration and optimization of this facilitative effect. Using this knowledge, we can improve the design of social environments, especially where important information must reach the listeners (as in airports or train stations for example).

In the present study, the benefit in terms of perceptual clarity of meaning-based predictions was positively related to verbal fluency, suggesting that for individuals with HI, the ability to mobilize the lexicon contributes to the strength of meaning-based predictions. This is in line with previous results showing that lexicon is related to knowledge-based predictions (Benard et al. 2014). This benefit of meaning-based predictability was particularly related to verbal fluency when matching text primes were presented beforehand, suggesting that individuals with HI use their lexical knowledge to benefit from meaning-based predictability particularly when form-based predictability is available. As we have already discussed, meaning-based prediction
of moderately to severely degraded NVS only seems to be possible for individuals with HI when form-based predictions can be made. Thus, for individuals with HI, verbal fluency seems to contribute to the enhancement of perceived clarity of NVS through meaning-based predictions when prior lexical knowledge is available, possibly by facilitating application of that knowledge.

The benefit of meaning-based predictions was not, however, related to working memory capacity. This result is in line with previous research showing that speech comprehension in difficult listening conditions is related to verbal fluency but not to working-memory for older listeners (Schneider et al. 2016) and extends it to perceptual clarity. It could then be suggested that individuals with HI do not generally use working memory to generate meaning-based predictions, likely because all available capacity is being used to process the degraded speech. However, when it became too difficult to process the degraded speech, i.e., in the most challenging listening situations, working memory capacity was positively associated with perceptual clarity ratings. This suggests that when working memory is not engaged in solving the primary task of speech perception, available capacity is used to generate and store meaning-based predictions. It is interesting to note that both in the present study and in our previous study in individuals without HI (Signoret et al. 2018), the benefit of meaning- rather than form-based predictability is associated with individual differences. Whereas individuals without HI may be able to devote explicit working memory capacity to store meaning-based predictions, individuals with HI may become reliant on explicit skills such as their verbal fluency to generate useful meaning-based predictions when listening to the degraded speech. This interpretation in terms of limited working-memory capacity may also apply to results showing that informing listeners beforehand about the content of a meaningless degraded spoken sentence significantly improves recognition of its last word (Freyman et al. 2004). Presumably, semantic priming relieves working memory load, making more cognitive resources available to process the last word.

CONCLUSIONS

The results of the present study showed that both form- and meaning-based predictability have a facilitative influence on the perceived clarity of degraded speech for individuals with HI. In other words, predictability makes degraded speech sound clearer for this group. The effect of form-based predictability was greater for individuals with HI in the present study than for individuals without HI in our previous study (Signoret et al. 2018). Indeed, when form-based predictions were available, there was no significant difference in ratings of perceived clarity between groups. However, when speech quality was moderately or severely degraded, meaning-based predictability was contingent on form-based predictability. The benefit in terms of perceptual clarity of meaning-based predictions was positively related to verbal fluency but not working memory. This suggests that the ability to mobilize the lexicon contributes to the strength of meaning-based predictions. Individuals with HI may already be using available working memory capacity to process the degraded speech and thus become reliant on explicit skills such as their verbal fluency to generate useful meaning-based predictions.
Rogers, C. S., Jacoby, L. L., Sommers, M. S. (2012). Frequent false hearing by older adults: The role of age differences in metacognition. *Psychol Aging, 27*, 33–45.

Rönnberg, J., Lunner, T., Zekveld, A., et al. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Front Syst Neurosci, 7*, 31.

Roth, T. N., Hanebuth, D., Probst, R. (2011). Prevalence of age-related hearing loss in Europe: A review. *Eur Arch Otorhinolaryngol, 268*, 1101–1107.

Rudner, M., Rönnberg, J., Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *J Am Acad Audiol, 22*, 156–167.

Schneider, B. A., Avivi-Reich, M., Daneman, M. (2016). How spoken language comprehension is achieved by older listeners in difficult listening situations. *Exp Aging Res, 42*, 31–49.

Shannon, R. V., Fu, Q.-J., Galvin, J. (2004). The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Oto-Laryngol Suppl, 50–54*.

Shannon, R. V., Zeng, F. G., Kamath, V., et al. (1995). Speech recognition with primarily temporal cues. *Science, 270*, 303–304.

Sheldon, S., Pichora-Fuller, M. K., Schneider, B. A. (2008a). Effect of age, presentation method, and learning on identification of noise-vocoded words. *J Acoust Soc Am, 123*, 476–488.

Thiel, C. M., Özyurt, J., Nogueira, W., et al. (2016). Effects of age on long term memory for degraded speech. *Front Hum Neurosci, 10*, 473.

Tombaugh, T. N., Kozak, J., Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Arch Clin Neuropsychol, 14*, 167–177.

Troyer, A. K., Cullum, C. M., Smernoff, E. N., et al. (1994). Age effects on block design: Qualitative performance features and extended-time effects. *Neuropsychology, 8*, 95–99.

Wechsler, D. (1981). The psychometric tradition: Developing the Wechsler adult intelligence scale. *Contemp Educ Psychol, 6*, 82–85.

Wingfield, A., Tun, P. A., McCoy, S. L. (2005). Human auditory cortex is sensitive to the perceived clarity of speech. *Neuroimage, 60*, 1490–1502.

World Health Organization. (2018). WHO global estimates on prevalence of hearing loss. Retrieved October 8, 2018, from http://www.who.int/deafness/estimates/en/.