

1
2
3
4 **Title: The roles of bottom-up and top-down information in the recognition of reduced**
5 **speech: evidence from listeners with normal and impaired hearing**
6
7
8
9

10 Esther Janse^{a,b*}, Mirjam Ernestus^{c,b}
11

12
13 *Corresponding author: Esther.Janse@mpi.nl, Tel: +31 24 3521432, Fax: +31 24 3521213
14

15 ^a Utrecht institute of Linguistics OTS, Utrecht University, Janskerkhof 13, 3512 BL Utrecht, The
16 Netherlands
17

18 ^b Max Planck institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands
19

20 ^c Centre for Language Studies, Radboud University, PO Box 310, 6500 HD Nijmegen
21
22

23 **Abstract [word count: 200]**

24 Highly reduced pronunciation variants, such as something like 'yeshay' for *yesterday*, are abundant in
25 conversational speech. Previous research has shown that listeners understand such pronunciation
26 variants only in their sentential contexts. This study further investigates the roles of the acoustic
27 properties of reduced words themselves and of semantic/syntactic and acoustic information in their
28 contexts. We report five experiments in which participants were tested on the recognition of reduced
29 words originating from a corpus of conversational Dutch. Experiment 1, a visual cloze test,
30 demonstrates that our set of reduced words cannot be guessed just on the basis of their
31 semantic/syntactic context. Experiment 2 replicates the earlier finding that reduced words can only be
32 recognised in their contexts. Experiment 3 shows that the reduced words were less well recognised if
33 the context is presented visually in the form of orthographic transcriptions. These experiments suggest
34 that listeners need some acoustic properties of reduced words themselves, together with
35 semantic/syntactic *and* acoustic information in their contexts, to recognise reduced words.
36 Experiments 4 and 5 confirm the importance of acoustic information for word recognition by showing
37 that high-frequency hearing loss hinders both the interpretation of the target words' acoustic properties,
38 and the use of neighbouring context.
39
40
41
42
43
44
45
46
47
48

49 **Keywords**

50 Speech perception; reduced word forms; acoustic cues; hearing impairment; context effects,
51 spontaneous speech
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **The roles of bottom-up and top-down information in the recognition of reduced speech:**
5 **evidence from listeners with normal and impaired hearing**
6
7

8
9 **1.0 Introduction**

10
11
12 In conversational speech, words can be reduced quite drastically, for instance, the four-syllable word
13 *hilarious* can be pronounced as /hlɪrɪs/ (Johnson, 2004). Importantly, such reductions occur very
14 frequently, especially in highly frequent words or fixed expressions, where three-to-five syllables can
15 be reduced to only one (Ernestus, 2000). Johnson (2004) concluded that over 60% of the words in his
16 corpus of casual American English speech deviate from their citation forms on at least one segment,
17 and that several segments are absent in 28% of the word tokens.
18

19
20
21
22 Many people are not aware of such reduced pronunciation variants, either in their own speech
23 or in speech of others. They do not notice that segments are missing and tend to report that utterances
24 sound intact. In Kemps, Ernestus, Schreuder, and Baayen (2004) Dutch participants listened to
25 sentences and were asked to monitor for the phoneme /l/. Listeners often responded wrongly to
26 reduced forms such as 'eik' /ɛik/ (a reduced form of Dutch /ɛixələk/ *eigenlijk* 'actually'), thereby filling
27 in segments that were not part of the speech signal. Importantly, this happened only if the reduced
28 form was presented in sentence context.
29
30
31
32

33
34 Even though most listeners are not aware of reduced pronunciation variants, there is evidence
35 that reduced speech is more difficult to process than clear speech. Relatively mildly reduced forms
36 (e.g., variants with missing /t/s or schwas, or variants with flapped /t/s) produced in isolation are
37 recognised more slowly than their unreduced counterparts (Janse, 2004; Janse, Nooteboom, & Quené,
38 2007; Ranbom & Connine, 2007; Tucker & Warner, 2007; Ernestus & Baayen, 2007).
39
40

41
42 The presence of context has been shown to facilitate the perception of reduced speech
43 substantially. Ernestus, Baayen, and Schreuder (2002) showed that highly reduced word forms, cut
44 from a speech corpus of casual Dutch, were poorly recognised by young normal-hearing listeners
45 when presented as isolated fragments (52% correct). They also tested identification of the reduced
46 words if they were presented in their 'Limited Context', being the target word's adjacent vowels and
47 intervening consonants. Third, identification of the same target forms was also investigated when
48 presented in their full context. The results showed that identification of the highly reduced target
49 words went from 52% correct in the Isolation condition to 70% in the Limited context condition, to
50 92% in the Full context condition. Similarly, Pickett and Pollack (1963) showed that very short
51 excerpts from connected speech were often entirely unintelligible, but the same excerpts became
52 intelligible when presented re-embedded in the original context.
53
54
55
56
57

58
59 The question now arises which properties of the context are beneficial. There is no doubt that
60 semantic or syntactic information can aid spoken-word recognition (e.g., Swinney, 1979; McClelland
61
62
63
64
65

1
2
3
4 & Elman, 1986; Norris, 1994; Norris, McQueen & Cutler, 2000). Further, several studies have shown
5 that acoustic context cues facilitate recognition of words. Some of these contextual cues are relatively
6 local in the speech signal: if a word's acoustic features spread through coarticulation or assimilation,
7 this will mainly affect its immediately neighbouring segments (Gaskell & Marslen-Wilson, 1998;
8 Nolan, 1992, Gow 2001, 2002, and Ellis & Hardcastle, 2002). The same holds for durational cues
9 indicating the location of word boundaries (Spinelli, McQueen, & Cutler, 2003; Salverda, Dahan, &
10 McQueen, 2003; Shatzman & McQueen, 2006), or cues indicating grammatical function (Local,
11 2003).

12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
There are also acoustic context cues that are located at a greater distance from their source. Cues to the presence of /r/ in a word can be found a number of syllables before or after the word containing this phoneme (Kelly & Local, 1986; Tunley, 1999; Heid & Hawkins, 2000; West, 2000; Local, 2003). Similar long-distance effects were found for the feature [anterior] (Coleman, 2003). Hawkins (2003) argues that this spreading of features over a target word's neighbouring syllables may make the speech signal perceptually coherent.

Crucially, most studies on the role of context, regardless of whether this context mainly contained semantic/syntactic constraints or acoustic cues to the target word, have used carefully lab-recorded speech. The present study investigates the importance of acoustic properties of the word itself, and of information contained in a word's context in the perception of *spontaneous* rather than careful *read* speech. Though context plays a crucial role in understanding conversational speech, the very reduction of the context as well as the target that is characteristic of conversational speech may cause specific contextual influences to be less powerful in conversational than in read speech. Conversational speech is often faster than read speech, particularly if we focus on speech fragments containing reduced pronunciation variants. Listeners may not have the time to fully process all information contained in these speech stretches. This may hold for processing of both semantic/syntactic and acoustic information. A second issue is that acoustic cues signaling a target word in conversational speech could be reduced, relative to carefully read speech. Acoustic context cues in casual speech could then be too subtle to be picked up by listeners. The results of Pickett & Pollack (1963) suggest that listeners are able to pick up contextual information in speech that is read aloud fast, but their study did not specify how reduced the fast articulated target words or their contexts were, or which type of information was contained in the context. We specifically focus on stretches of conversational speech containing (strongly) reduced speech, and on the roles of different sources of information in the recognition of reduced words. First, we investigate which information in the words' contexts is used to facilitate target word recognition. Second, we investigate the importance of acoustic properties of the reduced words themselves for spoken-word recognition.

In order to investigate the roles of acoustic (bottom-up) information and semantic/syntactic context (top-down) information in the recognition of reduced words, we followed up the study of Ernestus, Baayen, and Schreuder (2002). We first set out to investigate how semantically/syntactically

1
2
3
4 constraining the context fragments of their (2002) study were to see which type of context in the
5 (2002) study was particularly beneficial to recognition of the target words. The results of a cloze test
6 on written materials (our Experiment 1) suggest that their “full context” did not contain many
7 syntactic/semantic cues. In Experiment 2, we replicated the Ernestus, Baayen and Schreuder (2002)
8 context benefit results with a number of improvements in the experimental design, including only
9 context fragments that are not strongly semantically or syntactically biasing. In Experiment 3, we
10 presented the auditory target words together with a (visual) orthographic transcription of their
11 sentential contexts. This was done to investigate whether the combination of the acoustic properties of
12 the reduced words and the semantic/syntactic information contained in their contexts would facilitate
13 word recognition as much as presenting the targets in their auditory contexts (as was done in
14 Experiment 2). The results showed that providing the visual context also boosted recognition of the
15 (auditorily presented) reduced words, but to a lesser extent than presenting the reduced words in their
16 auditory contexts. Together, these three experiments show that the context effects on target word
17 recognition in Ernestus, Baayen, and Schreuder (2002) should be attributed to the interaction between
18 the acoustic properties of the target word and information contained in the word's context (being both
19 semantic/syntactic constraints and acoustic cues to the target word).
20
21
22
23
24
25
26
27
28

29 Experiments 4 and 5 provide initial data about the relevance of the high-frequency part of the
30 spectrum for recognition of target words in and out of their natural contexts. Several studies have
31 shown that while young normal-hearing listeners are hardly aware of the massive reduction in
32 conversational speech style, reduced articulation is much more evident to people with hearing loss,
33 which often affects hearing in the high frequencies in particular (Picheny, Durlach, & Braida, 1985,
34 1986, 1989; Payton, Uchanski, & Braida, 1994; Uchanski, Choi, Braida, Reed, & Durlach, 1996). This
35 suggests that high-frequency acoustic information may be important for the recognition of
36 conversational speech. Experiment 4 tests this hypothesis directly. In Experiment 4, young normal-
37 hearing listeners were presented with the speech materials as in Experiment 2, but these materials were
38 filtered with a downward sloping filter, attenuating high frequencies more than low frequencies.
39
40
41
42
43
44

45 Experiment 5 investigates the recognition of reduced target words in and out of their contexts
46 by older adults. Many older adults suffer from (age-related) hearing impairment that particularly
47 affects the high frequencies. Unlike the young adults tested in Experiment 4, who suddenly had to deal
48 with a speech signal with attenuated high frequencies, older adults may have gradually developed
49 strategies to cope with a more impoverished speech signal. Such strategies may involve a stronger
50 reliance on low-frequency information (e.g., Stelmachowicz et al., 1990), or on semantic context
51 (Nittrouer & Boothroyd, 1990; Pichora-Fuller, Schneider, & Daneman, 1995). Experiment 5,
52 therefore, shows how indispensable high-frequency information is for the recognition of our reduced
53 speech materials.
54
55
56
57
58

59 In short, the present study investigates how bottom-up acoustic information in a target word
60 and in its context and top-down semantic/syntactic information in its context contribute to the
61
62
63
64
65

1
2
3
4 recognition of that word in conversational speech. These issues are addressed in five experiments: one
5 visual cloze test to test whether our targets' contexts were semantically/syntactically constraining, one
6 study in which auditory presentation of the reduced words was supplemented with a visual
7 presentation of their contexts, and three auditory recognition studies testing how well different groups
8 of listeners recognise the words in and out of their contexts.
9

10 **2.0 Experiment 1: Cloze procedure on sentence contexts**

11
12
13
14
15
16
17
18 Experiment 1 investigated whether the context effects on target word identification reported in
19 Ernestus, Baayen, and Schreuder (2002) should be attributed purely to strongly biasing semantic or
20 syntactic information in the context. We transcribed the stretches of speech used in Ernestus, Baayen,
21 and Schreuder (2002) orthographically and we presented the context fragments in a visual cloze task.
22 Since participants only received visual information, their guesses on what the target could have been
23 must be based on semantic and/or syntactic information in the words' contexts and cannot be based on
24 acoustic cues to the target word.
25
26
27
28
29
30
31

32 *2.1 Materials and methods*

33 *2.1.1 Materials*

34
35
36 The test stimuli of the Ernestus, Baayen, and Schreuder (2002) study were drawn from a corpus of
37 Dutch spontaneous casual conversations recorded in a soundproof room (Ernestus, 2000). Each of the
38 53 stretches incorporated in our study¹ contains a (reduced) token of 20 different target word types
39 (words or fixed expressions, such as, *in ieder geval* 'in any case'). The word tokens differed in degree
40 of reduction: degree of reduction was classified as High, Medium, or Low (cf. section 3.1.1 or
41 Ernestus, Baayen, & Schreuder, 2002, for a more elaborate description). An example of the sentence
42 context for the word *eigenlijk* 'actually' is *nou eigenlijk alleen voor deze keer niet* 'well actually only
43 for this time not'.
44
45
46
47
48
49
50

51 *2.1.2 Participants*

52 Eighteen young adults (all students at Utrecht University and all native speakers of Dutch) participated
53 in the experiment. The students were paid a small fee for their participation.
54
55
56

57 *2.1.3 Procedure*

58
59
60
61

¹ The Ernestus et al. study had 54, rather than 53, word tokens, but one token could not be retrieved anymore.
62
63
64
65

1
2
3
4 An internet application was made that presented the orthographic transcriptions of all sentence
5 fragments. Target words were replaced by dots. Participants read a sentence fragment and were asked
6 to guess three times which word they thought had been left out. They were encouraged to come up
7 with three different guesses. Only after filling in their three answers for that trial, they would proceed
8 to the next trial. The order in which the 53 sentence fragments were presented was randomised for
9 each participant.
10
11
12

13 14 15 16 2.2 Results 17

18
19 Figure 1 gives two histograms of the cloze results. Figure 1a represents how many (out of 18) students
20 guessed the target word correctly as their first guess. Figure 1b represents how many (again out of 18)
21 participants guessed the word correctly as either the first, second or third guess.
22
23

24
25
26 ##INSERT FIGURES 1A AND 1B ABOUT HERE##
27

28
29 Out of the 53 target words, only 7 items (3 items in the High reduction group, 3 items in the Low
30 reduction group, and 1 item in the Medium reduction group) were guessed correctly (either as first,
31 second or third guess) by at least a third (more than 5 out of 18) of the participants. This suggests that
32 the context effects in Ernestus, Baayen, and Schreuder (2002) cannot be attributed to strong syntactic
33 and/or semantic constraints.
34
35

36
37 In the following recognition studies, we analysed only the 46 stimuli that were guessed
38 correctly by less than a third of the participants (and considered the other 7 as fillers). We thus
39 concentrated on the stimuli with minimal semantic and syntactic cues in the context. The remaining
40 items had a mean cloze value of 1.11 ($SD=1.46$) (i.e., on average 1.11 out of 18 participants guessed
41 the word correctly).
42
43
44

45 46 47 48 3.0 Experiment 2: Young normal-hearing listeners 49

50
51 In the original Ernestus, Baayen, and Schreuder (2002) study, the effect of context on target word
52 recognition may have been confounded with, and possibly inflated by, a number of material and
53 design issues. Therefore, before we follow up on their study and address our experimental questions,
54 we first replicated their study with a number of important changes to see whether we would obtain
55 similar effects. These changes are discussed below. Minor changes are discussed in section 3.1.
56
57
58

59
60 *Blocking stimuli by speaker and familiarisation fragment*
61
62
63
64
65

1
2
3
4 The stretches of speech in Ernestus, Baayen, and Schreuder (2002) were from 17 different speakers
5 and were presented in random order. Identification of words that are presented in lists has been shown
6 to be worse if the words are from multiple speakers (and are presented in random order) compared to a
7 single speaker, presumably because listeners need time to adapt to each new speaker's characteristics
8 (Sommers, Nygaard, & Pisoni, 1994). The context benefit in Ernestus, Baayen, and Schreuder (2002)
9 may have been inflated by the continuous switching between speakers, which affects recognition of
10 short fragments more than that of longer fragments. Therefore, in the present study, fragments were
11 blocked by speaker which resulted in 17 speaker blocks with an average length of 3 items (target and
12 filler items). Furthermore, each block of fragments of one speaker was now preceded by a 1-second
13 long familiarisation fragment of this same speaker, drawn from the corpus that the target and filler
14 fragments were drawn from as well. These speaker familiarisation fragments were included to allow
15 speaker adaptation to take place before the test stimuli (targets and fillers) were presented.
16
17
18
19
20
21
22
23

24 *Equalised Intensity*

25
26 In the original (2002) study, intensity differences between speech stimuli were maintained. This could
27 have been partly responsible for the differences in intelligibility between word tokens of different
28 degrees of reduction, since the degree of reduction may naturally covary with signal intensity (the less
29 reduced stimuli being generally louder than the more reduced stimuli). In the present study, we wanted
30 to investigate intelligibility of our target items based on spectral properties of the signal, and on
31 information in their contexts. Mean intensity was therefore equalised over all fragments.
32
33
34
35
36

37 *Within-subjects design instead of between-subjects design*

38
39 In the Ernestus, Baayen, and Schreuder (2002) study, target word tokens were presented in three
40 contexts: in isolation, in limited context (together with adjacent vowels and intervening consonants),
41 and in the sentence fragment they were taken from. Context manipulation in their study was a
42 between-subjects factor: listeners were assigned to one of three context conditions (for all stimuli). In
43 the present study, context was a *within-subjects* factor to rule out the possibility that differences
44 between conditions are actually differences between participant groups. This is especially important in
45 Experiment 5 where we tested older listeners, since they generally form a much less homogeneous
46 group than young student populations. Furthermore, a within-subjects design also allowed us to relate
47 context use to individual characteristics (such as hearing acuity) of the older participants in
48 Experiment 5.
49
50
51
52
53
54
55
56

57 *3.1 Materials and methods*

58 59 60 *3.1.1 Materials*

1
2
3
4 Each of the 46 stretches of speech, originating from the Ernestus, Baayen, and Schreuder (2002) study,
5 contains a (reduced) token of 20 different target word types (words or fixed expressions, such as, e.g.,
6 *in ieder geval* 'in any case'). As mentioned before, the word tokens differed in degree of reduction.
7 There were tokens with Low (little or no) reduction. Secondly, there were tokens with Medium
8 reduction, missing several segments, but consisting of more segments than the initial, the final, and the
9 stressed segments of their unreduced counterparts, e.g., two-syllabic /'moxək/ instead of the three-
10 syllabic /'moxələk/ *mogelijk* 'possible.' Third, there were tokens with High reduction, consisting
11 maximally of the initial, final, and stressed segments of their unreduced counterparts, e.g.,
12 monosyllabic /'mok/ *mogelijk* 'possible.' Word tokens were classified on the basis of unanimous
13 phonemic transcriptions of three trained phoneticians who listened to the stretches of speech
14 containing the word forms as well as four preceding and four following syllables. The phoneticians
15 could replay these stretches as often as they thought was necessary for a reliable transcription. For the
16 transcriptions, a segment counted as present if there was an auditory trace (regardless of whether there
17 was clear evidence for this trace in the spectrogram).
18

19
20 In addition to the 46 target items, 62 filler items, which also varied in degree of reduction,
21 were interspersed with the target items. Note that, particularly for the target items, we included
22 multiple tokens of the same word type, e.g., two speakers' realisations of *mogelijk* 'possible'. Thus,
23 listeners would hear more than one token of the same word type, but always in different degrees of
24 reduction and always from different speakers. The filler items were from the same speakers as the
25 target items and were included to increase the diversity in correct answers from 20 different answers
26 (to the 46 targets) to 75 different answers (to the 46 targets and 62 fillers). Thus, all in all, 108 test
27 fragments (targets and fillers) were presented that elicited 75 different answers.
28

29
30 We tested identification of the test words (targets and fillers) only in sentential context and in
31 isolation. Since there were not sufficient stimuli to present every subject with at least 10 different
32 stimuli in three context conditions per degree of reduction, we restricted the experiment to the two
33 most extreme context conditions in order to determine the maximal effect of context: minimum
34 amount of context (i.e., none) and maximum amount of context (sentence fragment).
35

3.1.2 Participants

36
37 Twenty-four young adults, all students at the Radboud University of Nijmegen, participated in this
38 experiment. They were 3 men and 21 women, and were all native speakers of Dutch. They had a mean
39 age of 21 years ($SD=2.1$; range 18 to 25). Their hearing acuity in both ears was tested in a soundproof
40 booth with a Maico ST20 audiometer. Averaged over 1, 2 and 4 kHz, mean pure-tone threshold in the
41 better ear was 5.6 dB HL ($SD=3.7$; range 0 to 13 dB). Mean hearing acuity at several frequencies is
42 graphed in Figure 6 in section 6.1.2.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

3.1.3 Design and Procedure

Using a Latin-square design, we rotated context over the 108 test items on two experimental lists. Of each experimental list, three variants were made that varied randomly in the order in which the speaker blocks were presented. This made six different experimental lists to which the participants were randomly assigned.

Participants listened to stretches of speech over headphones (Sennheiser HD 215) in a soundproof booth, with each stretch of speech being presented only once. All speech fragments were presented diotically. Participants were asked to repeat back the speech fragment they heard as accurately and as fast as possible (note again that a sentence fragment either consisted of a word/ fixed expression, e.g., *eigenlijk* 'actually', or of the word/fixed expression as part of a longer fragment, e.g., *nou eigenlijk alleen voor deze keer niet* 'well actually only for this time not'). Participants' responses were recorded and scored by the test assistant immediately. The test assistant only scored whether the target word had been reproduced correctly, and ignored whether the context had been repeated correctly. If necessary, the recording could be used later to check a (unclear) response. We chose to elicit spoken responses, since we believe they are easier to provide than written responses, especially for the older participants in Experiment 5. Furthermore, we considered spoken responses to better reflect 'online' processing than written responses because they can be given faster.

All participants were first presented with the sentence condition, then the isolation condition. We know that restoration naturally occurs when listeners are presented with reduced forms in their appropriate contexts (Kemps, Ernestus, Schreuder, & Baayen, 2004). In the first part of the experiment the participants were therefore probably unaware of the goal of this study. If the isolated fragments were to be presented first, listeners might know what our research was about and might focus on reduction during the entire study which might result in less natural listening behaviour. Before each context condition, participants were presented with a practice familiarisation fragment and four practice items to get used to the task. In the actual test blocks, each familiarisation fragment was always followed by at least one filler item. Presentation level of the speech stimuli was kept constant for all participants at a mean of 80 dB SPL (a level which is still comfortable for those with normal hearing).

Finally, in order to have a measure of cognitive performance of the participants, we included a variant of the Stroop colour naming task in our procedure. The Stroop experiment was performed after the listening experiment. Participants were asked to name the colour of a rectangle, which either had the letter string "XXXX" in it in a large white font, or an incongruent colour name (e.g., the word RED in a green or blue rectangle). For each of the three rectangle colours (red, blue, green), there were three pictures (e.g., the red rectangle appeared with the letter strings "XXXX", "BLUE" or "GREEN"). The neutral picture (with the string "XXXX") was presented twice to balance the number of observations

1
2
3
4 per condition (neutral vs. incongruent). These 4 presentations per rectangle colour, times three colours,
5 made 12 trials. These were repeated 10 times in random order, so that there were 120 observations for
6 each participant in total (60 per condition). We measured by means of a voice key how fast people
7 named the colour of the rectangle. The young participants' results on this colour-naming test will be
8 presented together with those of the older adults (in Experiment 5, section 6.1.2).
9

10 11 12 13 14 15 3.2 Results and Discussion 16

17
18 A response was counted as correct if the target word token was reproduced correctly (whether the
19 surrounding words were reproduced correctly was not relevant here). In scoring the responses, we
20 noticed that some listeners would repeat reduced fragments in a reduced manner like the speaker. If
21 the response contained as few or even fewer segments than the presented fragment, we could not be
22 sure that the participant had recognised the fragment correctly or was just repeating what he or she
23 heard without recognising it. We decided to discard these responses, which formed 5% of the
24 responses, most of which were given in the sentence condition (e.g., the participant's response /eik/
25 was rejected for the target /eilk/ *eigenlijk* 'actually', and so was the response /hemə/ for /heməl/
26 *helemaal* 'entirely').
27
28
29
30
31
32

33 In Figure 2, accuracy rates are given for the word tokens in the two context conditions and the
34 three degrees of reduction (computed as the ratio of the correct responses, relative to the sum of the
35 correct and incorrect responses).
36
37
38

39 ## INSERT FIGURE 2 ABOUT HERE ##
40
41

42 The results (1051 observations) were analysed with a multi-level regression analysis for binomial data
43 (the dependent variable being whether a response was correct or incorrect) with participant and word
44 as crossed grouping factors (Pinheiro & Bates, 2000; Quené & van den Bergh, 2004; Baayen,
45 Davidson, & Bates, 2008). The factors were the Degree of Reduction of the word token (Low,
46 Medium, or High) and whether or not it was presented in Context. In the text below, regression
47 coefficients will be provided for the fixed effects with standard errors in brackets, as well as z-values
48 and significance values. The isolation condition with High degree of reduction was mapped on the
49 intercept. Relative to High reduction, items with a Medium degree of reduction elicited higher
50 accuracy scores ($\beta=1.69$ ($SE=0.27$), $z=4.79$, $p<0.001$). Similarly, items with Low degree of reduction
51 also elicited higher accuracy scores than the High reduction items ($\beta=1.69$ ($SE=0.27$), $z=6.30$,
52 $p<0.001$). Further, target words were better recognised in sentence context than in isolation ($\beta=2.69$
53 ($SE=0.40$), $z=6.68$, $p<0.001$). The statistical model gave a significantly better fit if Context was added
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 to the random item part, indicating that context was more helpful for some word types than for others
5 (loglikelihood ratio test: $\chi^2(2)=9.756, p<0.01$).
6

7 The present results look rather similar to those obtained in the Ernestus, Baayen, and
8 Schreuder (2002) study. Most importantly, we verified that presenting target words in their context
9 considerably improves identification performance over presenting them in isolation, even though we
10 only included contexts that are not semantically/syntactically constraining. Furthermore, we now took
11 care that participants had had the chance to adapt to the speaker's characteristics by blocking the target
12 words by speaker and by having every block start with fillers. The current results do not show an
13 interaction between context benefit and degree of reduction, unlike the results in Ernestus, Baayen,
14 and Schreuder (2002). Further analyses showed that this was not due to the exclusion of the items with
15 constraining semantic or syntactic context, but simply to the decreased power as a result of the change
16 from a between-subjects to a within-subject design (which resulted in half the number of stimuli per
17 condition). We will show that the interaction is statistically significant if we include the other two
18 listener groups (Experiments 4 and 5).
19
20
21
22
23
24
25

26 The identification results obtained in Experiment 2 show a robust context benefit effect.
27 Clearly, the context provided by the sentence fragments contains important information to aid
28 identification of the target words: for the word tokens in the High reduction category, identification
29 accuracy went from 53% in isolation to 90% in context. This raises the question what type of
30 information is present in the contexts. Experiment 1 showed that the contexts themselves were not
31 highly semantically/syntactically constraining. If the auditory context contains *acoustic* cues to the
32 target word's identity, then the combination of the auditory presentation of the target words with visual
33 orthographic transcripts of their contexts should yield a smaller context benefit than the auditory
34 context benefit observed here in Experiment 2. The effect of *visual* context on target word recognition
35 was tested in Experiment 3.
36
37
38
39
40
41
42

43 **4.0 Experiment 3: Auditory target words and visual context**

44 *4.1 Materials and methods*

45 *4.1.1 Materials*

46 The auditory target and filler materials were the isolation-condition materials of Experiment 2:
47 participants only heard test items excised out of their contexts. The context manipulation in this
48 experiment entailed whether or not the short auditory test fragments were presented together with a
49 visual orthographic transcript of their sentence contexts.
50
51
52
53
54
55
56
57
58

59 *4.1.2 Participants*

1
2
3
4 Twenty-four young adults, all students at the Radboud University of Nijmegen, participated in this
5 experiment. They were 4 men and 20 women, and were all native speakers of Dutch. They had a mean
6 age of 20 years ($SD=2.6$; range 17 to 26). None of them participated in Experiments 1 or 2. Their
7 hearing acuity in both ears was tested in a soundproof booth with an Oscilla USB-300 screening
8 audiometer. Averaged over 1, 2 and 4 kHz, mean pure-tone threshold in their better ear was 0.8 dB HL
9 ($SD=5.4$; range -8 to 12 dB).²
10
11
12

13 14 15 *4.1.3 Design and Procedure*

16 As in Experiment 2, we rotated context (with or without visual presentation of the sentence context)
17 over the 108 test items on two experimental lists (46 targets and 62 fillers). As in Experiment 2,
18 stimuli were blocked by speaker and each speaker block was preceded by a speaker familiarisation
19 fragment. Further, before each context condition, participants were presented with a practice speaker
20 familiarisation fragment and four practice items to get used to the task. Of each experimental list, three
21 variants were made that varied randomly in the order in which the speaker blocks were presented. This
22 made six different experimental lists to which the participants were randomly assigned.
23
24
25
26

27 Participants listened to the short stretches of speech over headphones (Sennheiser HD 215) in
28 a soundproof booth, with each stretch of speech being presented only once. All speech fragments were
29 presented diotically. Participants were asked to repeat back the speech fragment they heard as
30 accurately and as fast as possible. Participants' responses were recorded and scored by the test assistant
31 immediately, following the same procedure as in Experiment 2.
32
33
34

35 All participants were first presented with the visual-context condition in which they had to
36 combine the auditory presentation of the target with the orthographic presentation of its sentence
37 context displayed on the computer screen. In the last half of the trials, they were presented with the
38 auditory targets in isolation (no context provided). Thus, as in Experiment 2, participants were first
39 presented with the context condition, and then with the isolation condition. The target word's context
40 was presented visually in orthographic form 3000 ms prior to auditory presentation of the target. As in
41 the visual cloze procedure of Experiment 1, the target word itself was left out of this transcription and
42 was replaced by dots. The orthographic context only disappeared from the screen 2000 ms after target
43 onset and was therefore visible during presentation of the auditory target. Then, the next trial started.
44
45
46
47
48

49 Presentation level of the speech stimuli was kept constant for all participants at a mean of 80
50 dB SPL.
51
52
53

54 *4.2 Results and Discussion*

55
56
57
58
59 ² Hearing sensitivity in this group was slightly better than in the young adult group of Experiment 2 ($t(46)=3.66$,
60 $p<0.001$). It is not clear whether this is a true hearing sensitivity difference between groups, or whether this is
61 due to the use of a different screening audiometer.
62
63
64
65

1
2
3
4 Scoring of the responses was the same as in Experiment 2. In our scoring procedure, a response to a
5 reduced stimulus was discarded if it contained as few or fewer phonemes than the presented target
6 fragment. This led to the exclusion of 6% of the responses.
7

8
9 The remaining 1031 observations of Experiment 3 were analysed with a multi-level regression
10 analysis for binomial data (the dependent variable being whether a response was correct or incorrect)
11 with participant and word as crossed grouping factors. The factors were Degree of Reduction of the
12 word token (Low, Medium, or High) and whether or not the token was presented with its visual
13 context. The isolation condition with High degree of reduction was mapped on the intercept. Relative
14 to High reduction items, items with Low degree of reduction elicited higher accuracy scores ($\beta=0.85$
15 ($SE=0.27$), $z=3.19$, $p<0.01$), but there was no significant difference between items of High and
16 Medium reduction. Further, target words were better recognised when their auditory presentation was
17 supplemented with a visual presentation of their context ($\beta=0.72$ ($SE=0.19$), $z=3.74$, $p<0.001$). There
18 was no interaction between Degree of Reduction and Context.
19
20
21
22
23

24 The results were also compared to those of Experiment 2. Figure 3 shows the results of the
25 visual-context condition of Experiment 3, together with the results of Experiment 2 (also presented in
26 Figure 2). This is done for ease of comparison of the auditory and visual context conditions. Accuracy
27 rates are given for the target word tokens in the three context conditions (Experiment 2's isolation and
28 auditory context conditions and Experiment 3's visual context condition), broken down by the three
29 degrees of reduction of the targets. Accuracy is computed as the ratio of the correct responses, relative
30 to the sum of the correct and incorrect responses.
31
32
33
34

35
36
37 ### INSERT FIGURE 3 ABOUT HERE ###
38
39

40 The results from Experiments 2 and 3 (from both context conditions from both experiments)
41 were compared with a multi-level regression analysis for binomial data. The factors in this analysis
42 were Degree of Reduction of the word token, whether or not it was presented in context (Context
43 Presence), and Experiment (Experiment 2 or 3). The best-fitting model showed that there was a
44 significant effect of Context Presence, indicating that words were better recognised when their
45 contexts were available ($\beta=2.68$ ($SE=0.27$), $z=10.09$, $p<0.001$). Relative to High reduction items (on
46 the intercept), tokens of Low reduction were better identified ($\beta=1.54$ ($SE=0.25$), $z=6.06$, $p<0.001$).
47 Likewise, tokens of Medium reduction were better identified than those of High reduction ($\beta=0.91$
48 ($SE=0.26$), $z=3.45$, $p<0.001$). Further, in the experiment with visual context (Experiment 3), the
49 difference between items of High and Low reduction was smaller than in the experiment with auditory
50 context ($\beta=-0.66$ ($SE=0.32$), $z=-2.05$, $p<0.05$), as well as the difference in recognition accuracy
51 between items of Medium and High reduction ($\beta=-0.69$ ($SE=0.33$), $z=-2.09$, $p<0.05$). Furthermore,
52 recognition accuracy was overall higher in Experiment 3 than in Experiment 2 ($\beta=1.00$ ($SE=0.26$),
53 $z=3.83$, $p<0.001$). This was particularly due to performance differences in the isolation condition
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 across experiments (see Figure 4). Crucially, the Experiment by Context Presence interaction was
5 significant as well ($\beta=-1.98$ ($SE=0.32$), $z=-6.13$, $p<0.001$). We investigated this latter interaction in
6 more detail by performing separate subset analyses: first on the trials *without* context (see Figure 4),
7 and then on trials *with* context (see Figure 3) across the two experiments.
8
9

10
11 ## INSERT FIGURE 4 ABOUT HERE ##
12
13
14

15 The subset analysis on the isolation-condition trials across experiments (in both experiments
16 the last half of the trials) showed that items of Low reduction were more accurately identified than
17 items of High reduction ($\beta=1.11$ ($SE=0.23$), $z=4.92$, $p<0.001$). Perhaps more surprisingly, we also
18 found that recognition accuracy in Experiment 3 was *higher* than in Experiment 2 ($\beta=0.62$ ($SE=0.19$),
19 $z=3.31$, $p<0.001$). One possible explanation for this generally better isolation-condition performance in
20 Experiment 3 may be the slightly better hearing sensitivity of the Experiment 3 participants (cf.
21 section 4.1.2). This explanation seems unlikely, however, as we checked that individual differences in
22 hearing sensitivity (entered as a covariate) did not predict performance in either Experiment 2 or 3.
23
24

25
26 Alternatively, the better performance may be due to the fact that participants in Experiment 3
27 were presented with the reduced tokens excised out of their auditory context throughout the entire
28 study. Even in the condition *with* visual context (first half of the trials), they only heard the isolated
29 tokens, such that once they were halfway through the experiment (when the isolation condition
30 started), they may have become trained in 'analytical' listening. Furthermore, since the Experiment 3
31 participants only repeated the targets (without their context) in both context conditions, they may have
32 learned that there was repetition of word forms (remember that the 46 target tokens represented 20
33 different word types, such that there was repetition of word types across the two context conditions).
34 We investigated learning by adding Trial number to this subset analysis of the isolation-condition
35 trials of Experiments 2 and 3. We expected listeners to improve their recognition performance over
36 trials in both experiments. However, if the Experiment 3 participants have become trained in analytical
37 listening before they started with the isolation-condition trials, they should improve less over trials
38 than the Experiment 2 participants. The results of this analysis (which did not alter the overall result
39 pattern) showed an effect of Trial number, indicating that recognition accuracy indeed improved over
40 trials ($\beta=0.08$ ($SE=0.02$), $z=3.72$, $p<0.001$). Trial number also interacted with Degree of Reduction:
41 compared to their improvement on Highly reduced items, participants' performance improved less
42 over trials for the items of Low reduction ($\beta=-0.10$, ($SE=0.03$), $z=-3.78$, $p<0.001$). Importantly, the
43 Experiment 3 participants tended to improve less over trials than the Experiment 2 participants ($\beta=-$
44 0.04 , ($SE=0.02$), $z=-1.78$, $p=0.07$), supporting our 'training account' for the better performance of the
45 Experiment 3 participants.
46
47
48
49
50
51
52
53
54
55
56
57
58

59 The second subset analysis concerned the results of the auditory and visual context conditions
60 (both presented to participants as the first experiment block). As can be seen in Figure 3, recognition
61
62
63
64
65

1
2
3
4 accuracy was generally lower in Experiment 3's visual context condition than in Experiment 2's
5 auditory context condition ($\beta=-1.39$ ($SE=0.40$), $z=-3.46$, $p<0.001$). This analysis also replicated the
6 effect of Degree of Reduction: both items of Low reduction ($\beta=1.37$ ($SE=0.34$), $z=4.00$, $p<0.001$), and
7 items of Medium reduction ($\beta=1.35$ ($SE=0.39$), $z=3.43$, $p<0.001$) were better identified than items of
8 High reduction.
9

10
11 The higher accuracy rates in the auditory than visual context condition suggest that the
12 auditory context information helps identify the target word. The contextual acoustic cues may involve
13 spreading of features over the target word's neighbouring syllables or may be prosodic cues to the
14 target word's identity. If acoustic information contained in a target word itself and in its neighbouring
15 context is important for target word recognition, the question arises which parts of the spectrum are
16 relevant. We investigated whether decreased sensitivity to high frequencies interferes with
17 identification of the target words both in and out of their auditory contexts. In Experiments 4 and 5 we
18 tested auditory recognition of our target words with two groups of "hearing-impaired" listeners.
19
20
21
22
23
24

25 **5.0 Experiment 4: Young adults with a simulated hearing loss**

26
27

28
29 As mentioned in the Introduction, hearing-impaired adults with decreased sensitivity for the high
30 frequencies are known to have problems understanding conversational speech, which suggests that
31 high frequencies are relevant for the recognition of reduced speech.
32

33 Experiment 4 investigated the relative importance of the high-frequency spectrum for the
34 recognition of reduced speech. Participants in this experiment were presented with filtered speech in
35 which the higher frequencies are particularly attenuated.
36
37

38 We did not test actually hearing-impaired young adults because of the risk of potential
39 confounds, such as a somewhat lower education level in the group with hearing problems (than our
40 normal-hearing students of Study 2), or a more limited vocabulary knowledge or receptive command
41 of spoken Dutch.
42
43
44

45 *5.1 Materials and methods*

46
47

48 *5.1.1 Materials*

49

50 The materials were the same as those used in Experiment 2 (cf. section 3.1.1), but the speech stimuli
51 were filtered with a low-pass filter. This filter left frequencies below 1 kHz unaffected and had a -20
52 dB/octave dampening above 1 kHz (see filter function in Figure 6 in Section 6)³. This filtering was a
53 simplified approximation or simulation of age-related hearing loss, and does not mimic the distortion
54 caused by sensorineural hearing loss. (cf. e.g., Humes et al., 1987 and Hargus & Gordon-Salant, 1995
55
56
57
58

59
60 ³ PRAAT software was used (www.praat.org) and Paul Boersma is gratefully acknowledged for his help with the
61 PRAAT filter script: Filter (formula)... self * if x < 1000 then 1 else (x / 1000) ^ - log2 (10) fi
62
63
64
65

1
2
3
4 for noise-masking as a more accurate method to simulate hearing impairment in normal-hearing
5 listeners). Since we were mainly interested in the effect of attenuation of the high frequencies, we
6 opted for this simple filtering approach (we will come back to this in section 6).
7
8
9

10 5.1.2 Participants

11 Twenty-four students from the Radboud University Nijmegen, 23 women and 1 man, were paid to
12 participate in the experiment. They were all native speakers of Dutch and their mean age was 21 years
13 ($SD=2.0$, range 18 to 27). None of them participated in Experiments 1, 2 or 3. Their mean pure-tone
14 average (PTA) over 1, 2, and 4 kHz in the better ear was 6.0 dB ($SD=4.5$). There was one student with
15 a PTA of 25 dB, and one with a PTA of 12 dB, but all other students had mean PTAs below 7 dB (we
16 verified that inclusion of these two participants did not significantly lower the group's recognition
17 performance). Hearing sensitivity in this group did not differ significantly from that in the young adult
18 group of Experiment 2 ($t(46)<1$, *n.s.*). Mean hearing thresholds in the better ear of participants in this
19 group are given in Figure 6. The filter itself is also graphed in Figure 6.
20
21
22
23
24
25
26

27 5.1.3 Procedure

28 The procedure was the same as in Experiment 2. Filtering had brought the mean intensity of the clear
29 stimuli, equalised at 70.0 dB, to 69.8 dB. As in the previous experiments, stimuli were presented at 80
30 dB SPL.
31
32
33
34

35 5.2 Results and discussion

36
37
38 As in Experiment 2, we only analysed responses that we knew were either correct or incorrect. In
39 Figure 5, accuracy rates of these young listeners with simulated hearing impairment are given for the
40 word tokens in the two context conditions (as ratios of the correct responses, relative to the sum of the
41 correct and incorrect responses), broken down by degree of reduction.
42
43
44
45

46 ## INSERT FIGURE 5 ABOUT HERE ##
47
48

49 The results of the young listeners with simulated hearing impairment were compared to the
50 results obtained in Experiment 2 with normal-hearing young adults. Accuracy results of the two young
51 adult groups were analysed as in the previous study (with Filtering of the stimuli as a between-subjects
52 factor). Again, word tokens were identified better in sentence context than in isolation ($\beta=2.79$
53 ($SE=0.41$), $z=6.88$, $p<0.001$). Overall, the less reduced the target word token, the higher the
54 identification rates: relative to the items of High reduction, items with Medium degree of reduction
55 had higher accuracy ($\beta=1.91$ ($SE=0.28$), $z=6.79$, $p<0.001$), and so had items with Low degree of
56 reduction ($\beta=2.56$ ($SE=0.29$), $z=8.92$, $p<0.001$). The effect of Filtering was significant, which shows
57
58
59
60
61
62
63
64
65

1
2
3
4 that overall performance was better in the group who got the clear speech signal ($\beta=1.19$ ($SE=0.27$),
5 $z=4.41$, $p<0.001$), relative to the filtered group (on the intercept). The interaction between Context and
6 Degree of Reduction was now significant, indicating that context was more beneficial the more
7 reduced the word, replicating the findings by Ernestus, Baayen, and Schreuder (2002). Highly reduced
8 items and items with medium reduction benefited equally from context, but items with a Low degree
9 of reduction benefited less from context ($\beta=-2.19$ ($SE=0.43$), $z=5.07$, $p<0.001$). Most importantly,
10 there was a three-way interaction between Context, Degree of Reduction and Filtering. The young
11 listeners who got the clear signal benefited *more* from context for the words with Medium reduction
12 ($\beta=2.72$ ($SE=1.19$), $z=2.28$, $p<0.05$), and for the words with Low reduction ($\beta=2.49$ ($SE=0.82$), $z=3.04$,
13 $p<0.01$), than the young participants who listened to the filtered signal.

14
15 This three-way interaction was confirmed in item subset analyses (testing whether there was a
16 Filtering by Context interaction for the three degrees of reduction separately). For the items of Low
17 and Medium reduction, Filtering and Context showed a significant interaction: for both subsets, the
18 young adults presented with filtered speech benefited *less* from context than the young adults
19 presented with the unfiltered signal. For the items in the High reduction group, two groups of young
20 adults benefited equally from context. The statistical model gave a significantly better fit if Context
21 was added to the random item part, indicating that context was not equally helpful for all word types
22 (loglikelihood ratio test $X^2(2)=34.374$, $p<0.001$).

23
24 These results are in line with the hypothesis that high-frequency acoustic information plays an
25 important role in the recognition of reduced speech. Clearly, bottom-up interpretation of the acoustic
26 signal is affected by attenuation of the high frequencies: recognition accuracy in the most difficult
27 condition (Highly reduced words presented in isolation) dropped from 53% correct in Experiment 2 to
28 31% correct for the young adults who were presented with the filtered speech. Furthermore, these
29 listeners are generally less able to benefit from information contained in a word's context. Attenuation
30 of the high frequencies may degrade specific cues to the target word, and/or may simply weaken the
31 contextual semantic/syntactic constraints through lowered intelligibility of the neighbouring words. It
32 is surprising, however, that the context benefit effect is equal in size for the participants listening to
33 the filtered speech and those listening to the unfiltered speech for the highly reduced words. We will
34 come back to this in the General Discussion.

35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 **6.0 Experiment 5: Older listeners with hearing impairment**

52
53
54 In Experiment 4, participants were not used to listening to this type of 'muffled'-sounding speech. As a
55 consequence, they may not have been able to compensate for the attenuation of the high frequencies
56 with acoustic information present in the low frequencies or with semantic information. In Experiment
57 5, recognition of reduced words by older adults with varying degrees of hearing impairment was
58 tested. Age-related hearing impairment is gradual and progressive, such that older adults may develop
59
60
61
62
63
64
65

1
2
3
4 strategies to cope with a more impoverished speech signal. Such strategies may involve a stronger
5 reliance on acoustic information that is not so much affected by age-related hearing loss (such as low-
6 frequency information) and on semantic context (Nittrouer & Boothroyd, 1990; Pichora-Fuller,
7 Schneider, & Daneman, 1995). This experiment thus shows whether high-frequency information is as
8 indispensable for target recognition as suggested by Experiment 4.
9

10
11 Older adults, however, may differ from young adults in more aspects than their hearing
12 sensitivity. Age-related hearing problems involve more than loss of sensitivity, such as decreased
13 temporal and frequency selectivity. The filtering we applied to our materials in Experiment 4 does not
14 mimic the distortion caused by a sensorineural hearing loss. Such signal-distortion aspects of auditory
15 aging may be responsible for the finding that, even when older adults still have relatively good
16 hearing, they require a greater amount of word onset for correct identification of carefully pronounced
17 words than young adults in a gating procedure (Craig, 1992; Craig, Kim, Rhyner, & Chirillo, 1993).
18 Further, speech comprehension may be affected by age-related cognitive decline (Jerger, Jerger, &
19 Pirozzolo, 1991; van Rooij, Plomp, & Orlebeke, 1989; van Rooij & Plomp, 1990, 1992; Janse, 2009).
20 Among older adults, performance on speech perception tasks has been related to processing speed
21 (van Rooij, Plomp, & Orlebeke, 1989; Janse, 2009), measures of verbal working memory (e.g., Humes
22 et al., 2006) and executive control (Tun, O'Kane, & Wingfield, 2002). Given these differences
23 between young and older adults in auditory processing and in cognitive performance, equal
24 recognition performance for the older adults and the young adults tested in Experiment 4 may still
25 indicate that the older adults use compensation strategies to arrive at this performance level.
26
27

28 Since age group differences in information processing speed (Salthouse, 1996) or attention
29 capabilities may be partially responsible for potential performance differences between the older
30 adults of Experiment 5 and the young adults of Experiments 2 and 4, we investigated how the older
31 group compared to the young adults in terms of cognitive performance. Further, among the older
32 adults, we investigated how individual recognition performance and context use relate to individual
33 listener characteristics, such as hearing acuity, age and measures of cognitive performance.
34

35 36 37 38 39 40 41 42 43 44 45 46 *6.1 Materials and methods*

47 48 49 *6.1.1 Materials*

50 The materials were the same as those used in Experiment 2 (cf. section 3.1.1).
51
52

53 54 *6.1.2 Participants*

55 Thirty older adults, all native speakers of Dutch, were recruited via *Hoger Onderwijs voor Ouderen*
56 ('Higher education for older adults') at the Radboud University of Nijmegen. This organisation
57 provides academic courses on various topics at universities in the Netherlands for people over 55 years
58 of age. People who had subscribed for a course received an information letter in which they were
59
60
61
62
63
64
65

1
2
3
4 asked to participate in the study. Given the academic level of the courses, the older group was a good
5 match to the young (student) groups of Experiments 1-4 in terms of education level. The older
6 participants were 13 men and 17 women, and had a mean age of 72 years ($SD=5.3$; range 61 to 84).
7

8
9 Mean pure-tone thresholds (with standard error bars) for the older participants' better ear are
10 given in Figure 6. Averaged over 1, 2 and 4 kHz in the better ear, mean pure-tone average of the older
11 adults was 26.8 dB HL ($SD=15$; range 5 to 66 dB). Average hearing sensitivity in the older adults
12 sample was poorer than the hearing loss simulated in the young adults in Experiment 4: the filtering
13 caused a simulated hearing loss (averaged over 1, 2, and 4 kHz) of 20 dB HL. We will come back to
14 this difference in hearing sensitivity in the discussion of the combined results (section 6.2). Hearing
15 loss, if present, in this participant group was most likely mainly due to age effects: participants had not
16 encountered hearing problems before the age of 50. Nevertheless, on the basis of the audiogram alone,
17 we cannot rule out that some adults may have encountered some noise-induced hearing loss earlier in
18 life or other pathologies affecting their hearing. Out of the 30 participants, two had a hearing aid in
19 one ear, and two had hearing aids in both ears. These four participants were asked to participate in the
20 listening study without their hearing aids to see how they performed in an 'unaided' situation. By-
21 subject analysis of performance showed that these four older adults did not behave differently from the
22 other older participants.
23
24
25
26
27
28
29
30
31

32 ### INSERT FIGURE 6 ABOUT HERE ###
33
34

35 As mentioned, we included a variant of the Stroop colour naming task in our procedure to
36 have a measure of cognitive performance. Our choice for this task was based on the study by Sommers
37 and Danielson (1999). These authors investigated the identification of words that varied in
38 neighbourhood density in young and older adults. Importantly, they found that older adults'
39 susceptibility to interference in two Stroop-like tasks predicted their ability to identify 'hard' words,
40 that is, words with many similar-sounding neighbours. The correlation was particularly strong for
41 'hard words' embedded in low-predictability, rather than high-predictability, sentences. Our speech
42 materials have relatively weak semantic constraints and poor intelligibility, and therefore also present
43 a situation with strong lexical competition. Unlike Sommers and Danielson (1999), who opted for
44 auditory variants of interference tasks (their 'inhibition index' was based on a modified version of the
45 selective attention paradigm by Garner (1974) and on an auditory Stroop paradigm), we investigated
46 whether a measure of interference obtained in the visual domain would be predictive of speech
47 perception in our study. If an interference measure obtained in a non-auditory modality is predictive of
48 speech perception performance, it is more likely to be related to general cognitive, rather than
49 auditory, performance.
50
51
52
53
54
55
56
57
58

59 In our variant of the classic Stroop task (Stroop, 1935), we measured by means of a voice key
60 how fast participants named the colour of the rectangle in a neutral condition (the rectangle contained
61
62
63
64
65

1
2
3
4 the letter string 'XXXX') and in an incongruent condition in which the coloured rectangle contained an
5 incongruent colour name (e.g., a blue rectangle containing the word 'RED' or 'GREEN'). There were
6 60 observations for every participant in each of the two conditions. In the neutral condition, mean
7 colour naming time for the older adults was 608 ms ($SE=12$) and 599 for the young adults ($SE=12$) in
8 Experiments 2 and 4. In the incongruent condition, mean naming time (of the correct colour) was 719
9 ms ($SE=13$) for the older adults and 672 ms for the young adults ($SE=12$). Analysis of naming times
10 showed a significant effect of Congruency ($F_1(1,76)=356.3, p<0.001$), but no effect of Age Group, and
11 an interaction between Congruency and Age Group ($F_1(1,76)=11.43, p<0.001$). This suggests that the
12 older adults experienced greater interference from the incongruent colour names than the young adults.
13
14

15 We investigated whether differences among the older adults on this interference measure are
16 associated with their individual ability to understand reduced words out of context or to make efficient
17 use of contextual cues. A participant's colour naming speed in the neutral condition was considered a
18 measure of information processing speed. Naming speed in the incongruent condition, a measure of
19 interference, was considered as a predictor as well, and so was a combined measure of the two
20 conditions of a participant's performance (naming speed in incongruent condition divided by this
21 participant's naming speed in neutral condition, as a normalised measure of interference). Because
22 these three measures are highly correlated, the relevance of each of them was tested in separate
23 models.
24
25
26
27
28
29
30
31
32

33 *6.1.3 Procedure*

34 The procedure was the same as in Experiments 2 and 4. Note again that all stimuli were presented at
35 80 dB SPL for all listeners.
36
37
38
39

40 *6.2 Results and discussion*

41
42
43 As in Experiments 2, 3 and 4, we only analysed responses that we knew were either correct or
44 incorrect (this led to the exclusion of 5% of the responses). In Figure 7, accuracy rates of the older
45 listeners are given for the target words in the two context conditions, broken down by degree of
46 reduction. Figure 8 shows the combined results of the auditory isolation and auditory context
47 conditions for the three listener groups of Experiments 2, 4 and 5.
48
49
50
51

52 ## INSERT FIGURE 7 ABOUT HERE ##

53 ## INSERT FIGURE 8 ABOUT HERE ##
54
55
56

57 We first compared the results of the older listeners to those of the young normal-hearing listeners
58 (Experiment 2). As before, the responses (2358 observations altogether) were analysed with a multi-
59 level regression analysis for binomial data (the dependent variable being whether a response was
60
61
62
63
64
65

correct or incorrect) with participant and word as crossed grouping factors. Factors were Age group, Degree of Reduction of the word and the Context in which it was presented. All three factors showed significant effects. The significant effects and interactions are given in Table 1.

Table 1. Overview of significant effects and interactions in the different result analyses. The condition mapped on the intercept is the isolation condition for items with a High degree of reduction (young adults).

Effect	β	SE	z	p
<i>Older adults vs. young listeners (Experiment 5 vs. Experiment 2)</i>				
Context	2.08	0.46	4.50	<0.001
Degree of Reduction (Medium)	0.86	0.28	3.13	<0.01
Degree of Reduction (Low)	1.61	0.29	5.56	<0.001
Participant Group (older adults)	-1.82	0.32	-5.69	<0.001
Context x Degree of Reduction (M)	-2.63	1.11	-2.36	<0.05
Participant Group x Context x Degree of Reduction (Medium)	-3.23	1.17	2.77	<0.01
Participant Group x Context x Degree of Reduction (Low)	-2.56	0.82	-3.13	<0.01
<i>Older adults vs. young listeners who heard filtered speech (Experiment 5 vs. Experiment 4)</i>				
Context	2.85	0.37	7.78	<0.001
Degree of Reduction (Medium)	1.86	0.21	8.91	<0.001
Degree of Reduction (Low)	2.84	0.21	13.59	<0.001
Context x Degree of Reduction (Low)	-2.08	0.31	-6.78	<0.001
<i>Individual differences among the older adults</i>				
Context	3.57	0.53	6.79	<0.001
Degree of Reduction (Medium)	1.86	0.28	6.62	<0.001
Degree of Reduction (Low)	2.95	0.28	10.42	<0.001
Context x Degree of Reduction (Low)	-1.82	0.42	-4.31	<0.001
Hearing loss	-0.04	0.01	-4.49	<0.001
Residual-effect-of-Age	-0.07	0.02	-2.95	<0.01
Context x Hearing loss	-0.02	0.01	-2.26	<0.05

Older participants showed poorer identification performance than the young normal-hearing listeners. Furthermore, as in the analysis of Experiment 2, words were better recognised in their (auditory) sentential contexts than in isolation; and, relative to items of High reduction, the items of Medium reduction and those of Low reduction were better identified. We again found a significant interaction between Context and Degree of Reduction. Importantly, whereas the Context by Age Group interaction was not significant, the three-way interaction between Context, Degree of Reduction and Age Group was. While the two listener groups did not differ in the context effect for highly reduced words, the context benefit on recognition was smaller for the older adults than for the young adults for words with Medium reduction and words with Low reduction. We confirmed our

1
2
3
4 interpretation of this three-way interaction by separately testing item subsets (breaking items down by
5 the three Degrees of Reduction) for the Age Group by Context interaction. For the items of Low and
6 Medium reduction, Age Group and Context showed a significant interaction: for both item subsets, the
7 older adults benefited less from context than the young adults (Low reduction items: $\beta=-1.83$
8 ($SE=0.82$), $z=-2.23$, $p<0.05$; and Medium reduction items: $\beta=-2.48$ ($SE=1.13$), $z=-2.20$, $p<0.05$). For
9 the items in the High reduction group, older adults and young benefited equally from context. These
10 subset analyses replicated the pattern we found in the comparison of normal-hearing young adults with
11 the young adults with simulated hearing loss (Experiments 2 and 4). As before, the statistical model
12 gave a significantly better fit if Context was added to the random item part (loglikelihood ratio test
13 $\chi^2(2)=62.154$, $p<0.001$).

14
15 We then compared the results of the older listeners to those of the young adults who had been
16 presented with the filtered speech (Experiment 4). Apart from Context and Degree of Reduction, we
17 also tested whether Age Group had an effect. The significant effects and interactions in this
18 comparison are also given in Table 1. This model replicated the effects of Context and Degree of
19 Reduction, and their interaction. As before, the statistical model gave a significantly better fit if
20 Context was added to the random item part (loglikelihood ratio test $\chi^2(2)=64.489$, $p<0.001$). There
21 was no overall effect of Age Group, and Age Group did not interact with either of the two factors.
22 This means that simulation of an age-related hearing loss brought young participants to about the same
23 performance level as the older participants. Furthermore, the pattern of context use across the different
24 degrees of reduction is similar for the two hearing-impaired age groups, as can be seen from Figure 8.

25
26 Finally, we also analysed the results of the older adults separately to investigate whether
27 individual characteristics, such as hearing loss, would predict performance and context use among the
28 older adults. Apart from the factors manipulated in the study (Degree of Reduction and Context),
29 several other predictors were entered into the statistical model of the older listeners' data: the
30 participant's gender, hearing loss, age, and the measures derived from the Stroop task (average neutral
31 naming speed, average naming speed in the incongruent condition, and an interference value, being
32 the ratio of these two). Because age and hearing loss are correlated in this participant group (Pearson
33 correlation coefficient for this group was $R=0.463$, $p=0.010$), we orthogonalised the two predictors by
34 replacing Age by the residuals of a linear regression model predicting Age as a function of hearing
35 loss.

36
37 As can be seen from Table 1, this model replicated the earlier effects for Context, Degree of
38 Reduction, and their interaction. In addition, hearing loss and the residual-effect-of-Age affected
39 performance: the more hearing loss, the poorer overall identification performance, and also the older
40 the participant, the poorer performance. Importantly, Context interacted with hearing loss: the more
41 hearing loss the participant had, the *smaller* the beneficial effect of context.

42
43 A participant's average colour naming speed in the Stroop neutral (XXXX) condition, which
44 can be taken as a measure of information processing speed, did not predict performance in the
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 identification task, nor did average naming time in the incongruent condition, nor the interference
5 combination measure. Thus, the overall Age effect on performance (cf. also Jerger, Jerger, &
6 Pirozzolo, 1991; van Rooij, Plomp, & Orlebeke, 1989; van Rooij & Plomp, 1990, 1992) may result
7 from some type of age-related cognitive decline that is not captured by the Stroop task. Measures of
8 cognitive function other than a measure of inhibitory abilities might have been a better choice. Since
9 wider context is needed to resolve ambiguous words, age-related declines in verbal working memory
10 could turn out to be important (Wingfield, Alexander, & Cavigelli, 1994). These issues remain for
11 further investigation.

12
13
14
15
16 In conclusion, the pattern of results observed for the older adults was very similar to that
17 obtained with the young adults with a simulated hearing loss. The two "hearing-impaired" groups
18 showed equally poor recognition of reduced words presented out of context. Moreover, both groups
19 benefited less from information contained in a word's context than normal-hearing adults, except for
20 the most reduced forms. This effect of poorer hearing on context use is supported by our finding that
21 among the older participants, those with *more* hearing loss benefited *less* from context.

22
23
24
25
26 Interestingly, the difference in context benefit between the hearing-impaired and normal-
27 hearing listeners only manifests itself for the words with Low and Medium reduction. We will come
28 back to this issue in the Discussion.

31 32 **7.0 Discussion**

33
34
35 The aim of this study was twofold. First, we investigated the role of acoustic properties of reduced
36 words for word recognition. Second, we investigated the role of information contained in a word's
37 context for spoken word recognition, being semantic/syntactic (top-down) information, as well as
38 acoustic cues to the target word.

39
40
41 The results of Experiments 2 through 5 speak to the first issue. Reduced word forms presented
42 out of context were repeatedly shown to be more difficult to process than less reduced forms: the
43 degree of reduction of the target word affected target intelligibility in all experiments. Further, the
44 comparisons between listener groups of Experiments 2, 4 and 5 emphasised the importance of good
45 hearing in bottom-up processing of the speech forms.

46
47
48 Importantly, the results of Experiment 3 showed that participants can be trained to recognise
49 reduced target words excised out of their auditory context: target recognition accuracy in the isolation
50 condition was higher in Experiment 3, where participants listened to isolated reduced words
51 throughout the experiment, than in Experiment 2. The young listeners of Experiment 3 seem to get
52 more (acoustic) information out of these short fragments in the second half of the experiment because
53 they have had more training (which includes information about which words are likely to occur in the
54 experiment). This is in line with our finding that the older adults (Experiment 5) showed similar
55 performance to the young adults who heard the filtered speech (Experiment 4), despite the fact that the
56
57
58
59
60
61
62
63
64
65

1
2
3
4 hearing loss simulation in the young adults was an underestimation of the older adults' hearing
5 problems in terms of simulated sensitivity (cf. Figure 6), and because the filtering did not simulate the
6 distortion caused by sensorineural hearing loss. These results suggest that older adults might have
7 outperformed the young adults if the two groups had been accurately matched on hearing abilities. As
8 suggested by Stelmachowicz et al. (1990), older adults may perform relatively better because they
9 have learnt to pay more attention to low-frequency information.

10
11
12
13 The second issue, the role of context information for word recognition in conversational
14 speech, is addressed in several ways in this study. Experiment 1, a visual cloze test, showed that most
15 target words from the Ernestus, Baayen, and Schreuder (2002) study could not be guessed purely from
16 their contexts. Experiment 2 replicated the Ernestus, Baayen and Schreuder (2002) finding that
17 presentation of the target words in their auditory contexts strongly increased target recognition
18 accuracy (from 53% to 90% for highly reduced words). Importantly, Experiment 3, which combined
19 the auditory presentation of the excised target words with the visual (orthographic) presentation of
20 their sentence contexts, showed that the contexts did contain some semantic/syntactic constraints, as
21 the presentation of visual context also improved auditory target word recognition. These results
22 demonstrate that the semantic/syntactic cues are by themselves of little help, but they can constrain the
23 number of words that fit the acoustic signal. The weak constraints interact with the acoustic properties
24 of the reduced target words to facilitate their recognition. The results of Experiment 3 also provided
25 evidence that the auditory context results in our Experiments 2, 4 and 5 are to some extent due to
26 *acoustic* cues in the context, because presenting the target words in their auditory context boosted
27 recognition more than presenting them with their context in a visual form. In the Introduction we
28 suggested the possibility that contextual (acoustic) information in a less formal speech style could be
29 more difficult to pick up than in read speech because conversational speech can be articulated very
30 fast. The fact that participants got more information out of the auditory context in Experiment 2 than
31 out of the visual context in Experiment 3 to facilitate target word recognition is all the more interesting
32 considering that Experiment 3 provided a very clear orthographic representation of the target word's
33 semantic/syntactic context (whereas information concerning neighbouring words and thus
34 semantic/syntactic context was transient and less clear in Experiment 2). The present results therefore
35 clearly demonstrate that contextual *acoustic* cues play an important role for spoken word recognition
36 in conversational speech.

37
38
39
40
41
42
43
44
45
46
47
48
49
50
51 Further, Experiments 4 and 5 showed that a substantial part of the contextual information is
52 carried by the high-frequency part of the speech signal. Attenuation of high frequencies led to a
53 reduced context benefit on word recognition. The reduced context benefit was evidenced by three
54 findings: 1) compared to the normal-hearing young adults, the young adults in whom we simulated a
55 hearing impairment benefited less from context; 2) the older participants benefited less from context
56 than the young normal-hearing adults; and 3) among the older listeners, we found that the more
57 hearing-impaired participants were, the less benefit they gained from context. These context findings
58
59
60
61
62
63
64
65

1
2
3
4 complement the robust finding of older listeners' increased reliance on *semantic* context (Nittrouer &
5 Boothroyd, 1990; Wingfield, Aberdeen, & Stine, 1991; Pichora-Fuller, Schneider, & Daneman, 1995;
6 Sommers, & Danielson, 1999; Sheldon, Pichora-Fuller, & Schneider, 2008). In most of these semantic
7 context studies, standardised Sentence-Perception-In-Noise materials (SPIN-R) were used in which
8 the sentence-final word is either highly predictable from the preceding semantic context, or has a low
9 predictability (Kalikow, Stevens, & Elliott, 1977; Bilger, Neutzel, Rabinowitz, & Rzeczkowski, 1984).
10 When equated for identification performance in the low-probability condition (by having different
11 signal-to-noise ratios), subjects with age-related hearing loss benefited more from context in the high-
12 probability condition than old subjects with near-normal hearing, who in turn benefited more from
13 context than young (normal-hearing) subjects (Pichora-Fuller, Schneider, & Daneman, 1995). Pichora-
14 Fuller and colleagues found this "greater context benefit" pattern over a range of signal-to-noise ratios
15 and thus over a range of performance levels. In our study, recognition performance in the isolation
16 condition was not equated across groups. This allowed us to investigate which sources of information
17 are available if the input is exactly the same for all participants. The information contained in our
18 auditory context fragments, being semantic/syntactic and acoustic cues, may not be as helpful to
19 hearing-impaired listeners as strongly biasing semantic context. A reduced benefit from *acoustic*
20 context may in fact add to older listeners' increased reliance on *semantic* context and top-down
21 processing.
22

23
24
25
26
27
28
29
30
31
32 Interestingly, contextual benefit for normal-hearing and hearing-impaired listeners was equal
33 for the Highly reduced words. This first of all demonstrates that a smaller context benefit for adults
34 with hearing impairment for the words with Low and Medium reduction should not be attributed to a
35 floor effect (i.e., overall lower performance on our test materials). There are at least two possible
36 explanations for the finding of equal benefit for the Highly reduced words. First, Highly reduced
37 words are articulatorily weak which may imply a higher degree of coarticulation with neighbouring
38 words. Thus, Highly reduced words may have more acoustic cues in their context than Low/Medium
39 reduced words. The more acoustic cues to the target, the more likely it is that they are picked up by
40 hearing-impaired listeners. Second, hearing-impaired listeners may (unconsciously) refrain from using
41 acoustic contextual information, unless word identification without using acoustic context is almost
42 impossible. Acoustic cues in neighbouring words may be subtle and difficult to interpret, given the
43 high variability in conversational speech, particularly for hearing-impaired listeners. As a
44 consequence, hearing-impaired listeners mainly use acoustic context information for recognition of the
45 difficult Highly reduced words, which they identify correctly without context in only 23% (older
46 adults) or 31% (young adults with simulated hearing impairment) of the cases. Further research is
47 required to investigate these two possibilities.
48

49
50
51
52
53
54
55
56
57 The results of our study, particularly the different recognition scores of words in auditory
58 (Experiment 2) and visual (Experiment 3) contexts, suggest that acoustic cues in a word's context are
59 more useful than previously assumed in models of spoken word recognition (e.g., McClelland &
60
61
62
63
64
65

1
2
3
4 Elman, 1986; Norris, 1994; Norris, McQueen, & Cutler, 2000; Gaskell & Marslen-Wilson, 1998), and
5 fits in with the idea that the spreading of acoustic cues over a target word's neighbouring syllables
6 makes the speech signal perceptually coherent (Hawkins, 2003) and that this perceptual coherence is
7 the 'perceptual glue' of speech (Remez, 2003). It is also in line with the 'perceptual grouping
8 hypothesis' proposed in Dilley and McAuley (2008) that "prosodic (i.e., fundamental frequency and
9 duration) cues distal from the locus of segmentation or lexical access of a word affect the unfolding
10 process of perceiving prosodic constituents, thereby influencing word segmentation and lexical
11 recognition in a manner consistent with principles of perceptual organization for non-speech auditory
12 patterns" (p.296).
13
14
15
16
17

18 A first step towards modelling the role of near phonetic context for spoken-word recognition
19 was made in papers on place assimilation and deletion of segments (e.g., Gaskell & Marslen-Wilson,
20 1996; 1998; Gaskell, 2003; Gaskell & Snoeren, 2008; Mitterer, Yoneyama, & Ernestus, 2008).
21 Hawkins and Smith (2001), in their POLYSP model (for 'polysystemic speech understanding'), focus
22 on the role of systematic fine phonetic detail that may be spread throughout the speech signal and how
23 it facilitates analysis at all linguistic levels simultaneously. This integration of information over longer
24 time scales may not be easily captured in recurrent neural networks (Hawkins & Smith, 2001).
25 However, future models of speech perception cannot afford to ignore these wider-range bottom-up
26 acoustic effects.
27
28
29
30
31

32 In conclusion, the current results stress the importance of acoustic information in a reduced
33 word and in its context for recognition, and have provided some initial data on the importance of the
34 high-frequency spectrum. Our results also suggest that listeners can be trained to get more out of a
35 spectrally impoverished speech signal by paying more attention to information in other parts of the
36 spectrum. Finally, we have shown that semantic/syntactic information contributes to word recognition,
37 particularly in combination with acoustic information in the reduced word.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 **Reference List**

5 Baayen, R.H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modelling with crossed random
6 effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.

7
8
9
10 Bilger, R C., Neutzel, J. M., & Rzezowski, C. (1984). Standardisation of a test of speech perception
11 in noise. *Journal of Speech and Hearing Research*, 27, 32-48.

12
13
14
15 Coleman, J. (2003). Discovering the acoustic correlates of phonological contrasts. *Journal of*
16 *Phonetics*, 31, 351-372

17
18
19
20 Craig, C. H. (1992). Effects of aging on time-gated isolated word-recognition performance. *Journal of*
21 *Speech and Hearing Research*, 35, 234–238.

22
23
24
25 Craig, C. H., Kim, B. W., Rhyner, P. M. P., & Chirillo, T. K. B. (1993). Effects of word predictability,
26 child development, and aging on time-gated speech recognition performance. *Journal of Speech and*
27 *Hearing Research*, 36, 832–841.

28
29
30
31 Dille, L.C., & McAuley, J.D. (2008). Distal prosodic context affects word segmentation and lexical
32 processing. *Journal of Memory and Language*, 59, 294-311.

33
34
35
36 Ellis, L., & Hardcastle, W. J. (2002). Categorical and gradient properties of assimilation in alveolar to
37 velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics*, 30, 373–396.

38
39
40
41 Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch, a corpus-based study*
42 *of the phonology-phonetics interface*. Utrecht: LOT.

43
44
45
46 Ernestus, M., Baayen, H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and*
47 *Language*, 81, 162–173.

48
49
50
51 Ernestus, M., & Baayen, R.H. (2007). The comprehension of acoustically reduced morphologically
52 complex words: The roles of deletion, duration and frequency of occurrence. *Proceedings of the 16th*
53 *International Congress of Phonetic Sciences, Saarbruecken, Germany*, 773-776.

54
55
56
57 Garner, W.R. (1974). *The processing of information and structure*. Potomac, MD (USA): Erlbaum.

58
59
60
61 Gaskell, M. G. (2003). Modeling regressive and progressive effects of assimilation in speech
62 perception. *Journal of Phonetics*, 31, 447–463.

1
2
3
4
5 Gaskell, M.G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical
6 access. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 144–158.
7
8

9
10 Gaskell, M.G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech
11 perception. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 380–396.
12
13

14 Gaskell, M.G., & Snoeren, N. (2008). The impact of strong assimilation on the perception of
15 connected speech. *Journal of Experimental Psychology: Human Perception and Performance*, 34,
16 1632–1647.
17
18

19
20
21 Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of*
22 *Memory and Language*, 45, 133–159.
23
24

25
26 Gow, D. W. (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of*
27 *Experimental Psychology: Human Perception and Performance*, 28, 163-179.
28
29

30
31 Hargus, S. E., & Gordon-Salant, S. (1995). Accuracy of speech intelligibility index predictions for
32 noise-masked young listeners with normal hearing and for elderly listeners with hearing impairment.
33 *Journal of Speech and Hearing Research*, 38, 234–243.
34
35

36
37 Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech
38 understanding. *Journal of Phonetics*, 31, 373-405.
39
40

41
42 Hawkins, S., & Smith, R. (2001). Polysp: a polysystemic, phonetically rich approach to speech
43 understanding. *Italian Journal of Linguistics – Rivista di Linguistica*, 13, 99-188.
44
45

46
47 Heid, S., & Hawkins, S. (2000). An acoustical study of long-domain /r/ and /l/ coarticulation. In
48 Proceedings of the fifth seminar on speech production: models and data (ISCA). Kloster Seon,
49 Bavaria, Germany (pp. 77–80).
50
51

52
53 Humes, L.E., Dirks, D.D., Bell, T.S., & Kincaid, G.E. (1987). Recognition of nonsense syllables by
54 hearing-impaired listeners and by noise-masked normal hearers. *Journal of the Acoustical Society of*
55 *America*, 81, 765-773.
56
57
58
59
60
61
62
63
64
65

- 1
2
3
4 Humes, L.E., Lee, J.H., & Coughlin, M.P. (2006). Auditory measures of selective and divided
5 attention in young and older adults using single-talker competition. *Journal of the Acoustical Society*
6 *of America*, 120, 2926-2937.
7
8
9
10 Janse, E. (2004). Word perception in fast speech: artificially time-compressed vs. naturally produced
11 fast speech. *Speech Communication*, 42, 155-173.
12
13
14 Janse, E., Nootboom, S., & Quené, H. (2007). Coping with gradient forms of /t/-deletion and lexical
15 ambiguity in spoken word recognition. *Language and Cognitive Processes*, 22, 161-200.
16
17
18
19 Janse, E. (2009). Processing of fast speech by elderly listeners. *Journal of the Acoustical Society of*
20 *America*, 125, 2361-2373.
21
22
23
24 Jerger, J., Jerger, S., & Pirozzolo, F. (1991). Correlational analysis of speech audiometric scores,
25 hearing loss, age, and cognitive abilities in the elderly. *Ear and Hearing*, 12, 103-109.
26
27
28
29 Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama, & K.
30 Maekawa (Eds.), *Spontaneous speech: Data and analysis. Proceedings of the 1st Session of the 10th*
31 *International Symposium* (pp. 29–54). Tokyo, Japan: The National International Institute for Japanese
32 Language.
33
34
35
36
37 Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility
38 in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society*
39 *of America*, 61, 1337-1361.
40
41
42
43 Kelly, J., & Local, J. K. (1986). Longdomain resonance patterns in English. In *International*
44 *conference on speech input/output: Techniques and applications* (pp. 304–9). Conference Publication
45 No. 258. London: Institute of Electrical Engineers
46
47
48
49 Kemps, R., Ernestus, M., Schreuder, R., & Baayen, R.H. (2004). Processing reduced word forms: The
50 suffix restoration effect. *Brain and Language*, 19, 117-127.
51
52
53
54 Local, J. (2003). Variable domains and variable relevance: interpreting phonetic exponents. *Journal of*
55 *Phonetics*, 31, 321-339.
56
57
58
59 McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive*
60 *Psychology*, 18, 1-86.
61
62
63
64
65

1
2
3
4
5 Mitterer, H., Yoneyama, K., & Ernestus, M. (2008). How we hear what is hardly there: Mechanisms
6 underlying compensation for /t/-reduction in speech comprehension. *Journal of Memory and*
7 *Language, 59*, 133-152.
8
9

10
11 Nitttrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young
12 children and older adults. *Journal of the Acoustical Society of America, 87*, 2705-2715.
13
14

15
16 Nolan, F. (1992). The descriptive role of segments: evidence from assimilation. In G. J. Docherty, &
17 D. R. Ladd (Eds.), *Papers in laboratory phonology II* (pp. 261–280). Cambridge: Cambridge
18 University Press.
19
20

21
22 Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52*,
23 189-234.
24
25

26
27 Norris, D., McQueen, J.M. & Cutler, A. (2000). Merging information in speech recognition: Feedback
28 is never necessary. *Behavioral and Brain Sciences, 23*, 299-370.
29
30

31
32 Payton, K. L., Uchanski, R. M., & Braida, L. D. (1994). Intelligibility of conversational and clear
33 speech in noise and reverberation for listeners with normal and impaired hearing, *Journal of the*
34 *Acoustical Society of America, 95*, 1581–1592.
35
36

37
38 Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing. I.
39 Intelligibility differences between clear and conversational speech,' *Journal of Speech and Hearing*
40 *Research, 28*, 96–103.
41
42

43
44 Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II.
45 Acoustic characteristics of clear and conversational speech, *Journal of Speech and Hearing Research,*
46 *29*, 434–446.
47
48

49
50 Picheny, M. A., Durlach, N. I., & Braida, L. D. (1989). Speaking clearly for the hard of hearing. III.
51 An attempt to determine the contribution of speaking rate to differences in intelligibility between clear
52 and conversational speech, *Journal of Speech and Hearing Research, 32*, 600–603.
53
54

55
56 Pichora-Fuller, M. K., Schneider, B., & Daneman, M. (1995). How young and old listeners listen to
57 and remember speech in noise. *Journal of the Acoustical Society of America, 97*, 593-608.
58
59
60
61
62
63
64
65

- 1
2
3
4 Pinheiro, J.C., & Bates, D.M. (2000). *Mixed-Effects Models in S and S-Plus*. New York: Springer.
5
6
7 Pickett, J.M., & Pollack, I.(1963). Intelligibility of excerpts from fluent speech: Effects of rate of
8 utterance and duration of excerpt, *Language & Speech*, 3, 151-164.
9
10
11 Quené, H. & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures
12 designs: A tutorial. *Speech Communication*, 43, 103-121.
13
14
15
16 Ranbom, L.J. & Connine, C.M. (2007). Lexical representation of phonological variation in spoken
17 word recognition. *Journal of Memory and Language*, 57, 273-298.
18
19
20
21 Remez. R.E. (2003). Establishing and maintaining perceptual coherence: Unimodal and multimodal
22 evidence. *Journal of Phonetics*, 31, 293-304.
23
24
25
26 van Rooij, J.C.G.M., Plomp, R., & Orlebeke, J. F. (1989). Auditive and cognitive factors in speech
27 perception by elderly listeners. I. Development of test battery. *Journal of the Acoustical Society of*
28 *America*, 86, 1294–1309.
29
30
31
32 van Rooij, J.C.G.M., & Plomp, R. (1990). Auditive and cognitive factors in speech perception by
33 elderly listeners. II. Multivariate analyses. *Journal of the Acoustical Society of America*, 88,
34 2611–2624.
35
36
37
38 van Rooij, J.C.G.M., & Plomp, R. (1992). Auditive and cognitive factors in speech perception by
39 elderly listeners. III. Additional data and final discussion. *Journal of the Acoustical Society of*
40 *America*, 91, 1028–1033.
41
42
43
44 Salthouse, T. A. (1996). The processing speed theory of adult age differences in cognition.
45 *Psychological Review*, 103, 403-428.
46
47
48
49 Salverda, A.P., Dahan, D., & McQueen, J.M. (2003). The role of prosodic boundaries in the resolution
50 of lexical embedding in speech comprehension. *Cognition*, 90, 51-89.
51
52
53
54 Shatzman, K.B., & McQueen, J.M. (2006). Segment duration as a cue to word boundaries in spoken-
55 word recognition. *Perception & Psychophysics*, 68, 1-16.
56
57
58
59
60
61
62
63
64
65

1
2
3
4 Sheldon, S., Pichora-Fuller, M.K., & Schneider, B.A. (2008). Priming and sentence context support
5 listening to noise-vocoded speech by younger and older adults. *Journal of the Acoustical Society of*
6 *America*, 123, 489-499.
7
8
9

10 Sommers, M. S., Nygaard, L. C., & Pisoni, D. B. (1994). Stimulus variability and spoken word
11 recognition. I. Effects of variability in speaking rate and overall amplitude, *Journal of the Acoustical*
12 *Society of America*, 96, 1314-1324.
13
14
15

16 Sommers, M. S., & Danielson, S. E. (1999). Inhibitory processes and spoken word recognition in
17 young and older adults: the interaction of lexical competition and semantic context. *Psychology and*
18 *Aging*, 14, 458-472.
19
20
21

22 Spinelli, E., McQueen, J. M., & Cutler, A. (2003). Processing resyllabified words in French. *Journal*
23 *of Memory and Language*, 48, 233-254.
24
25
26

27 Stelmachowicz, P.G., Lewis, D.E., Kelly, W.J., & Jesteadt, W. (1990). Speech perception in low-pass
28 filtered noise for normal and hearing-impaired listeners. *Journal of Speech and Hearing Research*, 33,
29 290-297.
30
31
32

33 Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental*
34 *Psychology*, 18, 643-661.
35
36
37

38 Swinney, D. (1979). Lexical access during sentence comprehension: (re)consideration of context
39 effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645-659.
40
41
42

43 Tucker, B.V., & Warner, N. (2007). Inhibition of processing due to reduction of the American
44 English flap. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbruecken,
45 August 2007, 1949-1952.
46
47
48

49 Tun, P.A., O'Kane, G., & Wingfield, A. (2002). Distraction by competing speech in young and older
50 adult listeners. *Psychology and Aging*, 17, 453-467.
51
52
53

54 Tunley, A. (1999). Coarticulatory influences of liquids on vowels in English. Unpublished Ph.D.
55 dissertation, University of Cambridge.
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Uchanski, R. M., Choi, S., Braida, L. D., Reed, C. M., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing. IV. Further studies of the role of speaking rate, *Journal of Speech and Hearing Research*, 39, 494–509.

West, P. (2000). Perception of distributed coarticulatory properties of English /l/ and /ɛ̃/. *Journal of Phonetics*, 27, 405–425.

Wingfield, A., Aberdeen, J. S., & Stine, E. A. (1991). Word onset gating and linguistic context in spoken word recognition by young and elderly adults. *Journal of Gerontology*, 46, 127-129.

Wingfield, A., Alexander, A.H., & Cavigelli, S. (1994). Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging. *Language and Speech*, 37, 221-235.

1
2
3
4 FIGURE CAPTIONS
5
6

7 Figure 1. Number of participants (out of 18) that guessed the target word correctly from its context:
8 either as the first guess (Figure 1a), or as any of three guesses (Figure 1b). The horizontal line in the
9 lower graph indicates our cut-off criterion.
10

11
12 Figure 2. Results for the young normal-hearing listeners: Accuracy rates for the words of varying
13 degrees of reduction (High, Medium, Low) in the two context conditions (in isolation or in sentences).
14
15

16
17 Figure 3. Results for young normal-hearing listeners: Recognition accuracy rates for the words of
18 varying degrees of reduction (High, Medium, Low) in the three context conditions (in isolation
19 (Exp.2), with auditory context (Exp.2), and with visual context (Exp.3)).
20
21

22
23 Figure 4. Recognition accuracy rates for the words of varying degrees of reduction (High, Medium,
24 Low) in the isolation context conditions of Experiments 2 and 3.
25
26

27
28 Figure 5. Results of young adults with simulated hearing impairment: Accuracy rates for the word
29 tokens of varying degrees of reduction (High, Medium, Low) in the two context conditions.
30
31

32
33 Figure 6. Average hearing loss in the better ear (pure-tone thresholds in dB HL) for the three listener
34 groups in Experiments 2, 4 and 5 (error bars represent one standard error). The low-pass filter function
35 for the hearing loss simulation of Experiment 4 is also graphed.
36
37

38
39 Figure 7. Results of the older participants: Accuracy rates for the words of varying degrees of
40 reduction (High, Medium, Low) in the two context conditions (in isolation or in sentences).
41
42

43
44 Figure 8. Results of all three listener groups of Experiments 2, 4 and 5 (young adults, young adults
45 who were presented with filtered speech, and the older adults): Accuracy rates for the words of
46 varying degrees of reduction (High, Medium, Low) in the two context conditions (in isolation or in
47 their auditory contexts).
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure1a

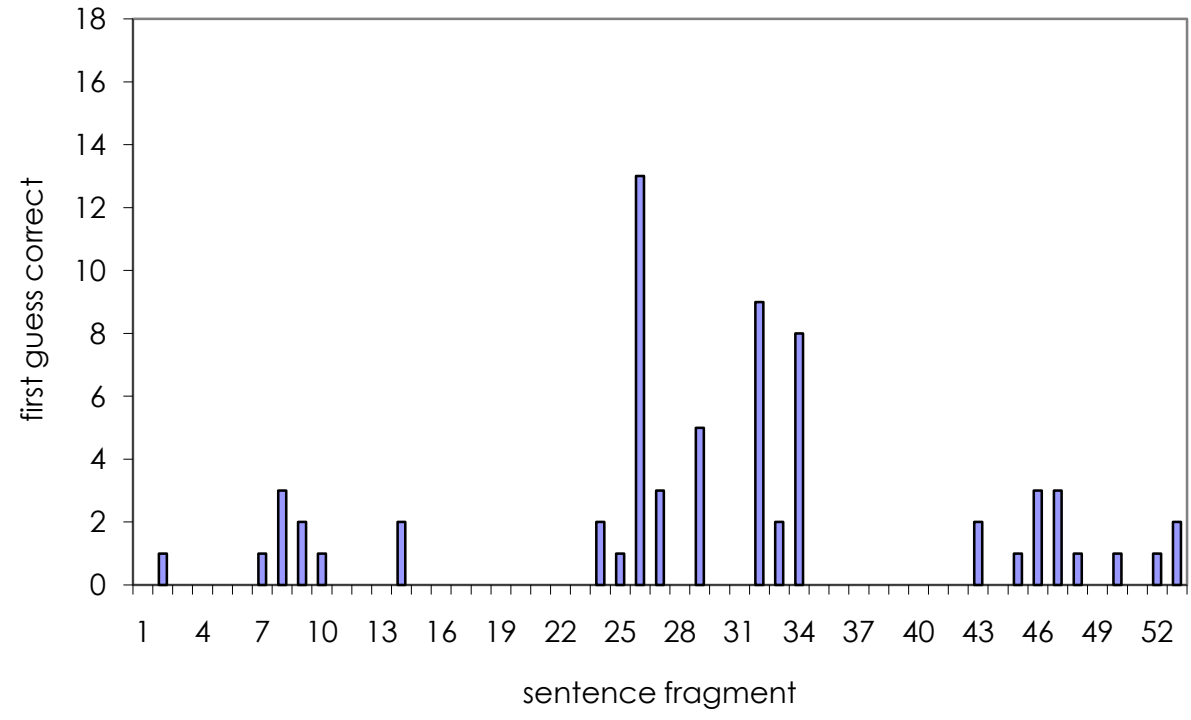


Figure1b

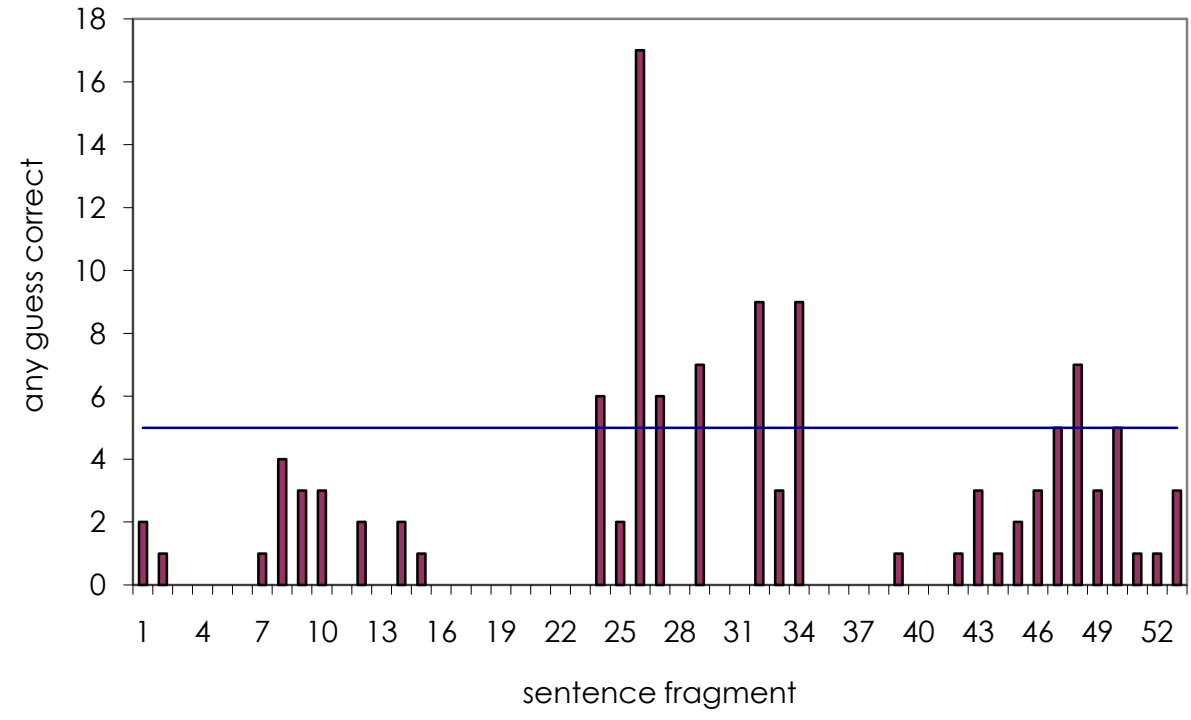


Figure2

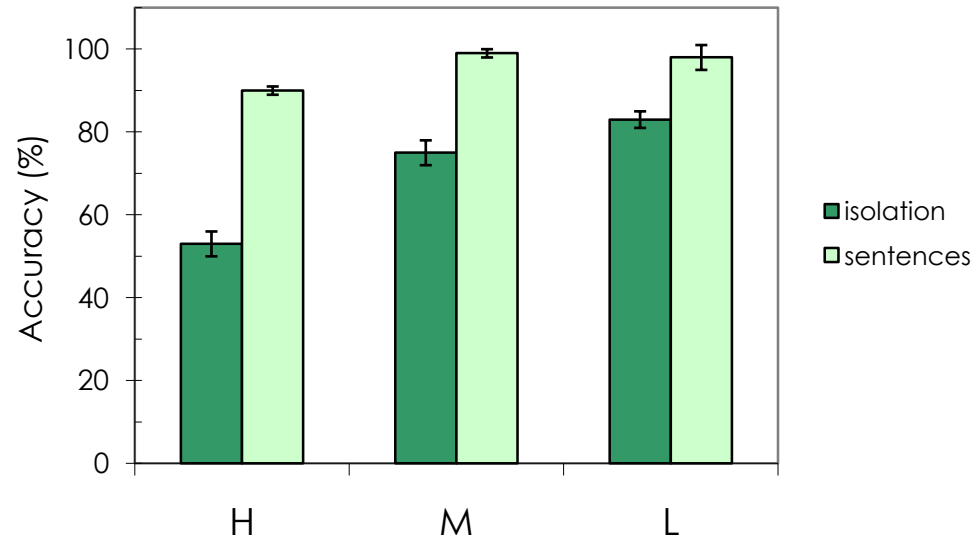


Figure3

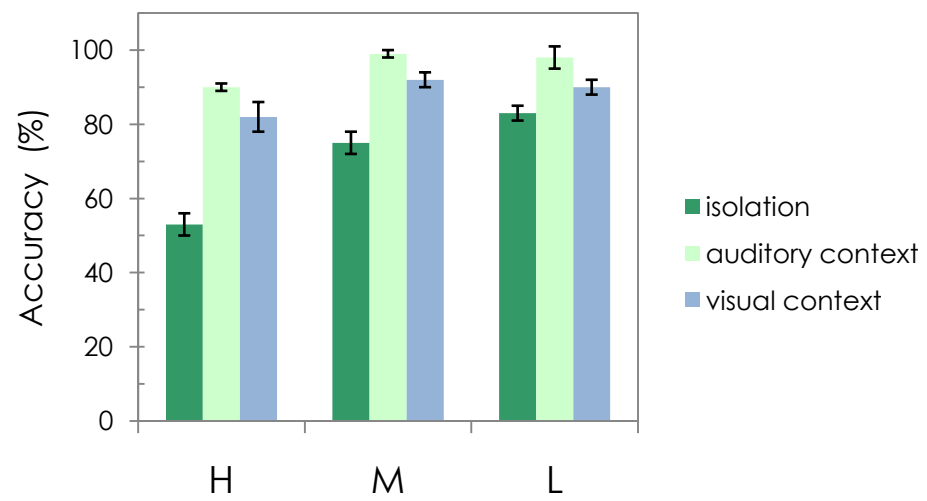


Figure4

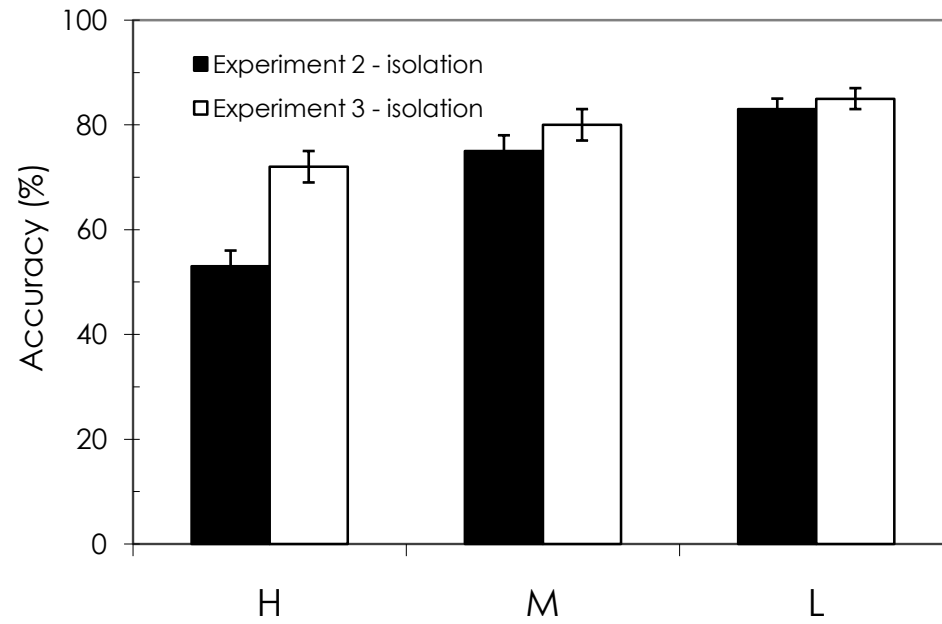


Figure5

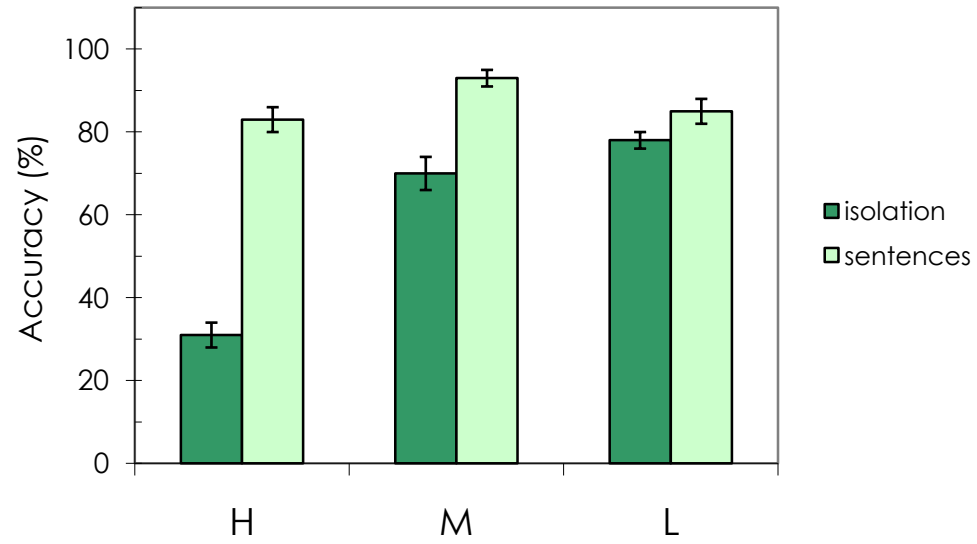


Figure6

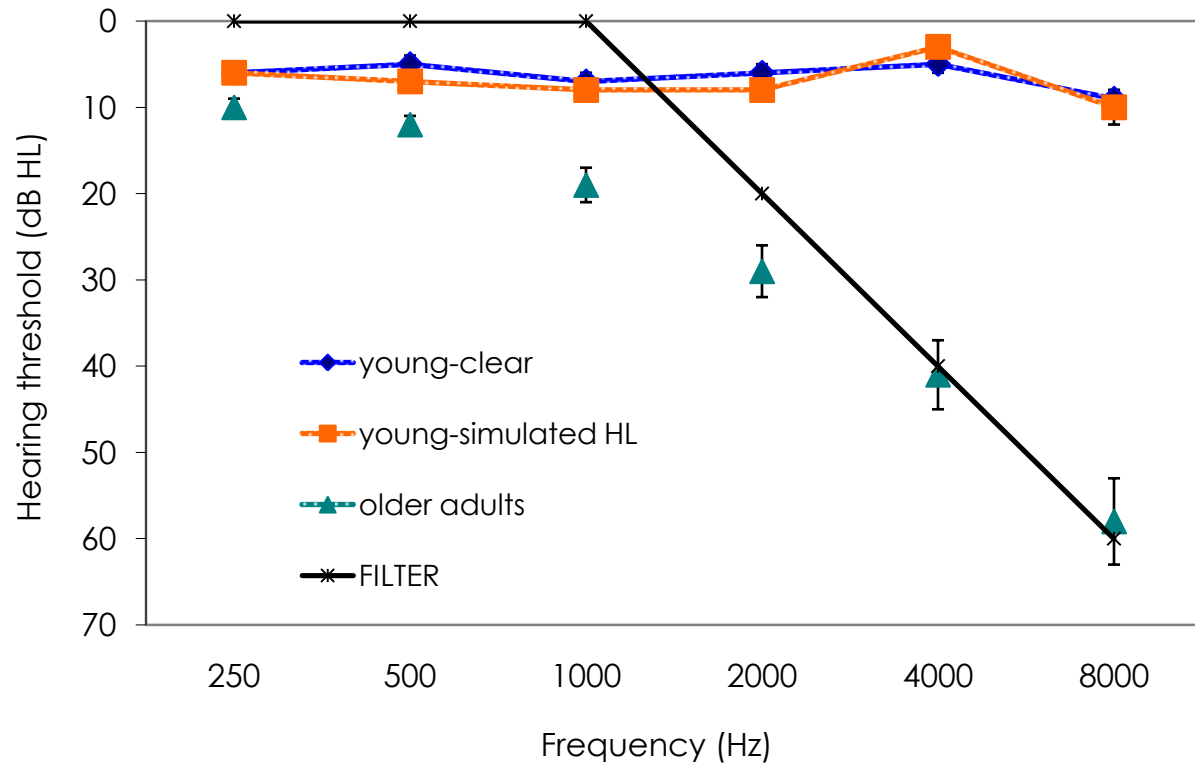


Figure7

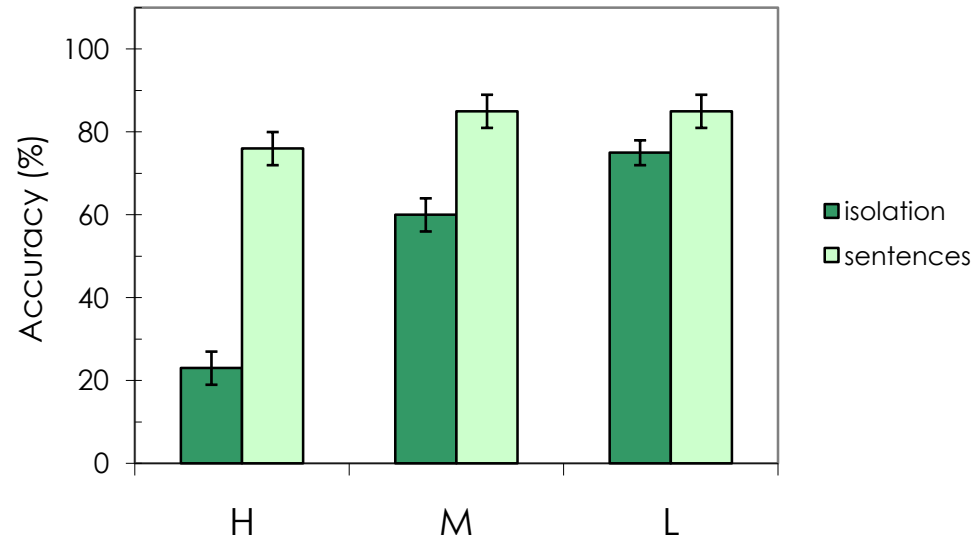


Figure8

