University of Groningen

Listening Effort

Pals, Carina

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Pals, C. (2016). *Listening Effort: The hidden costs and benefits of cochlear implants*. [Thesis fully internal (DIV), University of Groningen]. Rijksuniversiteit Groningen.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 09-09-2024

CHAPTER 6

General discussion

Carina Pals

The title of this thesis is "Listening effort, the hidden costs and benefits of cochlear implant hearing." So what are these *hidden* costs and benefits of hearing with a cochlear implant (CI)? The obvious *visible* benefit of a CI is the restored hearing ability, which allows the CI user to participate more comfortably in our predominantly hearing society. Some of the better performing CI users are even able to communicate over the telephone, without the visual aid of lip reading. These visible benefits are measurable in terms of speech understanding, and can be easily observed by friends and coworkers. The *hidden* costs and benefits, on the other hand, are internal to the listener and less easily measured or observed. Even if speech understanding performance is similar, there may still be differences between normal hearing (NH) listeners and CI users, between individual CI users, or within a single CI user for different device settings, configurations, or different listening situations. Differences, for example, in *listening effort*.

Listening effort refers to the cognitive processing load associated with listening. In the context of this thesis the focus is specifically on the effort related to speech understanding. For NH listeners, speech understanding in ideal listening conditions seems to be effortless and automatic (Mattys, Davis, Bradlow, & Scott, 2012; Wild et al., 2012). Speech that is degraded, whether due to factors internal or external to the listener, however, does not match the listener's phonological representations in long-term memory, and requires increased cognitive processing for the interpretation of the message (e.g. Lunner & Sundewall-Thorén, 2007; Rönnberg, Rudner, Foo, & Lunner, 2008; Wild et al., 2012; Wingfield & Tun, 2007). CI mediated speech, due to both technical and neural limitations of electric stimulation, results in perceptual representations that are spectro-temporally degraded, i.e. sparser in information regarding the frequency content and timing of the signal, and distinctly different from NH (e.g. Blamey et al., 1992; Stickney et al., 2006). Especially for postlingually deafened CI users, who formed their speech representations in long-term memory based on normal acoustic speech input before losing their hearing, this degraded input from electric stimulation can lead to a mismatch between the incoming speech and representations in long-term memory. Speech understanding in otherwise favorable listening conditions may therefore already be more effortful for CI users than for NH listeners, which is supported by research using CI simulated speech in NH listeners (Wagner, Pals, de Blecourt, Sarampalis, & Başkent, 2016). When the speech signal is additionally degraded, for example due to interfering background noise, resolving the mismatch may be even more effortful.

Increased listening effort serves the purpose of maintaining speech understanding, but at what cost? The additional recruitment of limited cognitive resources for effortful listening reduces the spare capacity available and limits, therefore, the performance on simultaneous tasks and cognitive processes. Even between listening conditions that result in similar intelligibility, differences in listening effort have been shown to affect performance on a secondary task (e.g. Broadbent, 1958; Sarampalis, Kalluri, Edwards, & Hafter, 2009), short-term memory for the correctly heard speech (e.g. McCoy et al., 2005; Pichora-Fuller, Schneider, & Daneman, 1995; Rabbitt, 1966, 1991), which in turn affects discourse comprehension (Pichora-Fuller, 2003), and long-term episodic memory of the speech (Sörqvist & Rönnberg, 2012). In other words, while the listener appears to understand what is said at that moment, the actual comprehension and later memory for the message may be compromised. One can imagine, then, effortful listening could potentially affect academic or professional performance of the listener (Ljung, Sörqvist, Kjellberg, & Green, 2009; Van Engen & Peelle, 2014). In addition to these cognitive consequences, sustained periods of effortful listening, for example in a noisy work environment, can lead to mental fatigue (Hornsby, 2013; McGarrigle et al., 2014), and has been shown to correlate with stress-related sick-leave from work (Kramer, Kapteyn, & Houtgast, 2006). In short, effortful listening can lead to a broad array of negative consequences for the listener and their active participation in society (Hua et al., 2014).

To come back to the population of interest in this thesis: how much is actually known about listening effort in CI users? Although a large body of scientific work had traditionally explored effects of CI processing and device configurations on speech intelligibility (e.g. Fu, Shannon, & Wang, 1998; Spriet et al., 2007; Wilson et al., 1991), at the outset of this project, little had been published on *listening effort* and CIs. Measures of speech intelligibility, such as the percentage of correctly repeated words or sentences, or the signal-to-noise ratio (SNR) that results in a certain level of intelligibility, reflect the end result of all the perceptual and cognitive processes involved in speech understanding. Yet, they do not reveal the nature or magnitude of cognitive processing, or listening effort, that was invested to reach this level of speech understanding. Measures of listening effort can complement intelligibility measures to reveal the otherwise hidden cost of increased cognitive processing load for speech understanding in challenging conditions (e.g. Gatehouse & Gordon, 1990; Larsby, Hällgren, Lyxell, & Arlinger, 2005; McGarrigle et al., 2014), and can therefore, potentially, shed new light on how the process of speech perception is affected by different aspects of hearing with a CI, specifically when intelligibility measures fail to show an effect.

To summarize, when the incoming auditory signal does not match the phonological representations in long-term memory, as it may well be the case for CI listeners, speech understanding requires increased cognitive processing, i.e. listening effort. This increased processing load does not necessarily lead to a loss of intelligibility, and may therefore go undetected. However, even if the end result intelligibility is not affected, increased listening effort can have detrimental consequences for the listener, such as reduced memory for the heard speech, which can, for example, affect discourse comprehension, and for example, lead to poor academic performance. Listening effort can, therefore, have a significant impact on the lives of CI users. Yet relatively little is known about how CI processing affects listening effort. This gives rise to the following questions. At a scientific level; how do intelligibility and listening effort interrelate for different CI configurations, individuals, or listening situations? At a more practical level; can a measure of listening effort uncover hidden benefits of CI processing that are not revealed by measures of intelligibility? Finally at a clinical level: does any such measure of listening effort seem promising for clinical applications?

THIS THESIS

The aim of this thesis was to systematically explore how differences in CI processing or device configurations can affect intelligibility, which was extensively studied before, as well as listening effort, which had not been studied before. More specifically, this thesis explores whether and how different CI settings might affect listening effort when intelligibility shows no change. The studies progressed from simple and more controlled, e.g. testing NH listeners using CI simulations quiet, to more complex and closer to real-life, e.g., including speech perception in background noise, and eventually examining effects of actual CI settings in actual CI users. The first two studies were conducted using normal-hearing participants presented with CI simulated speech, to control for much of the between-CI-user variability. Thus, these studies could focus strictly on the effects of spectral resolution on listening effort for speech in otherwise optimal listening conditions (Chapter 2) or the effect of added low-frequency acoustic speech to simulated CI speech, thus simulating electric-acoustic stimulation (EAS), on speech perception in quiet and in noise (Chapter 3). The next study explored the effect of background noise and noise-type on the perception of clear, unprocessed speech for NH listeners (Chapter 4). In the final chapter, the insights gained in the first three studies were combined and applied to investigate how spectral resolution affects listening effort in CI users, manipulating spectral resolution by varying the number of active electrodes of the CI (Chapter 5).

The main method chosen to objectively quantify listening effort was a dual-task paradigm combining a primary intelligibility task with a secondary visual response-time task. In a dual-task paradigm two tasks compete for cognitive resources and performance on the secondary task therefore reveals the cognitive processing load for the primary listening task (e.g. Baddeley & Hitch, 1974; Kahneman, 1973; Tyler, Hertel, McCallum, & Ellis, 1979), much in the same way effortful listening can interfere with concurrent tasks in real-life situations. The dual-task paradigm's long history of use as a measure for cognitive effort in hearing research (e.g. Broadbent, 1958), relative ease of implementation, the fact that it requires no expensive equipment, and its ability to measure intelligibility and listening effort simultaneously, made it a good starting point in the search for a suitable scientific, as well as clinical, tool. In the earlier studies described in this thesis (Chapter 2 and 3) the dual-task was complemented with a subjective self-report measure of listening effort to compare the perceived effort with the objectively measured cognitive processing load. In the later studies (Chapter 3 and 4), simple response-time measures, such as verbal response-times, reflecting cognitive processing time (Gatehouse & Gordon, 1990) and the sentence verification task, reflecting comprehension and processing time (Baer, Moore, & Gatehouse, 1993), compared with the dual-task paradigm as potential measures of listening effort. Such simple response time measures, due to the simplicity of the tasks, might be more widely applicable in clinical settings.

As the main interest of this thesis was to investigate 'hidden' effects, i.e. effects that are not revealed by traditional intelligibility measures, the studies were designed such that the effects of experimental parameters on listening effort could be compared at equal levels of intelligibility. To recapitulate, the parameters investigated in this thesis were spectral resolution in NH listeners (Chapter 2) and in CI users (Chapter 5), simulated EAS compared to CI alone in NH listeners (Chapter 3), and background noise and noise type in NH listeners (Chapter 4).

THE FINDINGS

In general, the *intelligibility results* confirmed our expectations based on the literature. Spectral resolution, in both NH listeners and CI users, has been shown to improve sentence intelligibility in quiet until reaching ceiling performance at a relatively low number of spectral channels (Fishman, Shannon, & Slattery, 1997; Friesen, Shannon, Başkent, & Wang, 2001). In this thesis, for NH listeners increased spectral resolution improved sentence intelligibility in quiet up to 6 spectral channels (Chapter 2), and in CI users, 7 active electrodes produced near-ceiling intelligibility, showing no improvement for further increased spectral resolution (Chapter 5). As for the effect of simulated EAS compared to CI, the literature shows improved speech perception in noisy listening conditions (e.g. Brown & Bacon, 2009; Dorman, Spahr, Loizou, Dana, & Schmidt, 2005; Kong & Carlyon, 2007). In this thesis, adding low-frequency acoustic speech to CI simulated speech to simulate EAS improved speech perception in background noise, thus allowing for equal intelligibility at lower SNRs than for CI alone (Chapter 3). In Chapter 4, the effects of background noise and noise type were investigated in NH listeners for speech masked by steady-state, speech-shaped noise (SSN) and 8-talker babble. For NH listeners, understanding speech masked by 8-talker babble required a higher SNR than speech masked by steady-state, speech-shaped noise (SSN) to achieve the same level of target speech intelligibility, in line with previous research that shows better speech intelligibility in SSN compared to 8-talker babble (Chapter 4; Lecumberri & Cooke, 2006). In summary, in each of the studies in this thesis the intelligibility results were in line with expectations based on the literature, and intelligibility was successfully fixed at the desired levels for the conditions of interest.

In the first two chapters a *subjective self-report measure of listening effort* was administered alongside with the objective dual-task measure. In clinical settings the fit of hearing devices is usually mainly evaluated using intelligibility measures and for additional information, clinicians rely on subjective reports made by the patient. Therefore, if a subjective self-report measure could reliably reveal ‘hidden’ effects, effects not reflected by the traditional intelligibility measures, this could be useful in clinical settings. In this thesis, however, the results of the subjective measures were observed to closely follow the intelligibility results and revealed no additional effects (Chapter 2 and 3), while, as will be further described below, the objective measures of listening effort did reveal additional effects. Prior research often shows that subjective effort measures appear not to correlate with objective measures (e.g. Anderson Gosselin & Gagné, 2010; Feuerstein, 1992; Fraser, Gagné, Alepins, & Dubois, 2010). In fact, subjective estimates of effort often appear to reflect perceived accuracy (e.g. Feuerstein, 1992), as was also the case in this thesis. Subjective self-report, while convenient, may therefore not be the most suitable measure for revealing effects that are not already reflected in intelligibility measures.

THE DUAL-TASK MEASURE OF LISTENING EFFORT, RESULTS AND IMPLICATIONS

Unlike the subjective measure of listening effort, the *objective, dual-task measure of listening effort* did successfully reveal effects of simulated CI processing and device configurations on listening effort in NH listeners, even at equal intelligibility. In Chapter 2, increased spectral resolution for speech in quiet resulted in significantly improved secondary task performance for NH listeners up to 8 spectral channels while intelligibility reached ceiling at 6 channels. In Chapter 3, simulated EAS compared to CI alone, improved listening effort in quiet listening conditions at near-ceiling intelligibility and in SSN at 50% intelligibility. In a dual-task paradigm, improved performance on the secondary task implies that the primary task required fewer cognitive resources, thus leaving more resources available to allocate to the execution of the secondary task (e.g. Baddeley & Hitch, 1974; Kahneman, 1973; Tyler, Hertel, McCallum, & Ellis, 1979). These dual-task results thus show that, compared to high spectral resolution, poor spectral resolution requires increased cognitive processing, even if speech intelligibility is at ceiling. Furthermore, the results show that both increasing spectral resolution, or providing additional low frequency acoustic speech, as in EAS, can potentially reduce the cognitive processing load of speech understanding, even when these improvements do not appear to further benefit intelligibility.

These findings may be explained using the ease-of-language-understanding (ELU) model (Rönnberg et al., 2013, 2008). The ELU predicts little or no interference with concurrent tasks for the perception of well-formed speech. However, when the auditory signal does not match representations in long-term memory, explicit cognitive processing is required to resolve this mismatch (Rönnberg et al., 2013, 2008). The more mismatches occur, the more cognitive processing is required, thus depleting resources for concurrent tasks or cognitive processes. Empirical evidence does indeed suggest that only those portions of the bottom-up speech signal that cannot be matched with top-down predictions are passed on for further, higher-order cognitive processing (Sohoglu, Peelle, Carlyon, & Davis, 2012; Van Engen & Peelle, 2014). While intelligibility for 6-channel CI simulated speech is at ceiling, the interaction with the secondary task thus suggests that the reduced spectral resolution of the speech signal does still result in mismatches with the speech representations in the listener's long-term memory, and therefore requires increased cognitive processing for interpretation. Increasing spectral resolution of the CI simulated speech up to 8 spectral channels significantly reduced interfere between the speech task and the secondary task (Chapter 2), suggesting the 8-channel CI simulated speech is more similar to the listener's representations. Providing low frequency acoustic speech in addition to the CI simulated speech signal similarly improved secondary task performance (Chapter 3).

Improving spectral resolution beyond 8 spectral channels, however, did not lead to further increase in secondary task performance (Chapter 2), and while EAS improved secondary task performance compared to simulated CI alone, no distinction could be made between the different EAS simulations combining 300 Hz or 600 Hz low-pass filtered speech with CI simulated speech (Chapter 3). However, this absence of dual-task interaction with the secondary task does not necessarily imply 'effortless' listening. The secondary-task performance for listening conditions with 8 spectral channels and up resulted in similar performance as when the task was performed without a simultaneous speech task suggesting ceiling performance. Interpreting the speech may still require increased cognitive processing compared to NH. As long as the combined processing load of the primary listening task and the secondary task do not exceed the limit of available processing resources, this increased processing load of the listening task will not affect the secondary task.

While the dual-task paradigm successfully revealed effects of spectral resolution on listening effort in NH participants, in CI users it was not as successful. In CI users, changing the spectral resolution by varying the number of active electrodes between 7 and 15 did not affect secondary task performance (Chapter 5). While this could be due to a number of limitations of the dual-task paradigm, which were discussed in chapter 5, as well as the same ceiling effect discussed above, the ELU might provide a different explanation. For the CI users, due to their experience listening with a CI for at least one year, and some participants even many more years, their phonological representations may have adapted to the input they hear every day.

The change from their normal, full electrode array to a limited subset of those electrodes may have been less drastic than the change for NH participants from normal acoustic hearing to the 6-channel, noise-band vocoded CI simulations. The 7-active-electrode speech input may therefore not have required increased cognitive processing to resolve the mismatch with representations in long-term memory to the degree that this processing requirement interfered with secondary task performance.

However, as will be further detailed below, the results of the sentence verification task, did suggest the decreased spectral resolution for CI users does affect speech comprehension and the required processing time. This need not contradict the explanation above, since as long as the limit of the total available resources has not been reached, increases in listening effort do not affect the secondary task.

SIMPLE RESPONSE TIME MEASURES OF LISTENING EFFORT, RESULTS AND IMPLICATIONS

In Chapters 4 and 5, two simple response time measures of listening effort were compared with the dual-task paradigm. Effortful listening has been suggested to increase not only the need for cognitive processing resources, but also lead to increased processing time (e.g. Francis & Nusbaum, 2009; Rönnberg et al., 2013; A. E. Wagner, Toffanin, & Başkent, 2016). Previous research has successfully used response times to speech as a measure that complements intelligibility measure, and reflects 'ease of listening' (Baer et al., 1993; Gatehouse & Gordon, 1990).

In Chapter 4, 'verbal response times' were used, i.e. the time required to start repeating a sentence after it was heard (Gatehouse & Gordon, 1990). This measure can be collected during a traditional speech intelligibility task by recording the verbal responses and scoring manually or using speech detection software. The effects of noise and noise type at different intelligibility levels on listening effort were measured using both the dual-task paradigm and the verbal response times. The dual-task results showed reduced secondary task performance when noise was present, but did not differentiate between the noise types or between 79% and near-ceiling intelligibility. Similar to the dual-task paradigm, the verbal response times showed an effect of the presence of noise; responses were slower for speech in noise than for speech in quiet. Unlike the dual-task paradigm, the verbal response times also showed a significant difference between the 79% and near-ceiling intelligibility. Neither measure showed an effect of noise type at fixed intelligibility levels (Chapter 4). These results suggest that verbal response times may be at least as suitable for measuring listening effort as the dual-task paradigm employed in this thesis, and its ease of implementation could make it a practical method for clinical use.

In Chapter 5, a sentence verification task was used. In this task the participant is presented with sentences that are either unmistakably true or nonsense, and they respond by pressing a key indicating whether the sentence was true or false (Adank & Janse, 2009; Baer et al., 1993). The sentence verification task accuracy and response times reflect both comprehension and processing speed, respectively. The task is simple, easy to explain and perform, and the response collection is easily automated, making it an appealing candidate for a clinical tool. In CI users, increased spectral resolution from 7 active electrodes up did not affect the dual-task measures of intelligibility or listening effort. The sentence verification task, on the other hand, revealed a clear improvement in both speech comprehension and speed-of-processing for increasing the number of active electrodes from 7 to 11 (Chapter 5). The same trend of improved comprehension and speed-of-processing was found in NH listeners for improved spectral resolution up to 16 spectral channels (Chapter 5), while in a different group of (similar) NH participants the dual-task measure showed improvement only up to 6 channels for intelligibility and 8 channels for listening effort (Chapter 2). In part, these differences in results could be attributed to the different speech materials used for the sentence verification task and the dual-task paradigm. Nevertheless, these results for both NH and CI listeners establish the sentence verification task as a likely candidate for a clinical measure of listening effort.

One challenge arises from the sentence verification task results: not only the response times improved with increased spectral resolution when the dual-task measures of listening effort and intelligibility had not, but so did the accuracy scores. The accuracy scores can be interpreted to indirectly reflect intelligibility; when the sentence is not heard correctly and the answer must be guessed, there is a 50% chance of getting a wrong answer. Thus, does this mean that for the speech materials used in the sentence verification task, intelligibility was affected by the decreased spectral resolution even when intelligibility in the dual-task was at ceiling? Possibly. However, the accuracy scores on the sentence verification task do not purely reflect intelligibility. In order to correctly answer whether the sentence was true or false/nonsense, the listener has to *comprehend* the sentence. Comprehension goes beyond just perception, in that the meaning has to have been understood and the participant has to reason about the meaning of the sentence (Ralston, Pisoni, Lively, Greene, & Mullennix, 1991; Wingfield & Tun, 2007). This requires further cognitive processing of the speech than only individually perceived phonemes or words, and therefore likely relies more on cognitive capacity (Just & Carpenter, 1992). The reduction of cognitive resources due to increased listening effort may therefore also affect comprehension.

Regardless of whether the accuracy scores on the sentence verification task reflect intelligibility of the sentence materials, listening effort, or both, the results show the value of additional measures. Based on the intelligibility or dual-task listening effort results, one would have had to conclude increased spectral resolution beyond 7 active electrodes in CI users, or 8 spectral channels for NH listeners, no longer improves speech perception. The sentence verification task revealed both improved comprehension and processing speed for further increased spectral resolution in both participant groups.

SUMMARY AND CONCLUSIONS

The results for the dual-task paradigm and the simple response-time measures of listening effort illustrate the added value of measures that tap into effort and cognitive processes involved in speech understanding to complement the traditional intelligibility measures. The studies described in this thesis were designed such that the effects of CI processing and device configurations on listening effort could be compared at equal intelligibility. Subjective self-report closely followed intelligibility and revealed no additional effects. The dual-task did show that secondary task performance could be affected differently for conditions that did produce similar intelligibility, thus illustrating the value of a measure of listening effort to complement intelligibility measures. The sentence verification task further revealed effects where the dual-task did not. This may be due to a limitation of the dual-task; when the listening task and secondary task combined do not require all available cognitive resources to be performed simultaneously, the tasks will not interact, and changes in listening effort will not affect secondary task performance. All in all, the sentence verification task appears to be a useful behavioral measure to investigate speech comprehension and listening effort in both NH listeners and CI users, and may prove to be a suitable tool for clinical use.

As for CI processing and device configurations, our studies in NH listeners suggest that reduced spectral resolution, results in increased processing load for speech perception. Both increasing spectral resolution (Chapter 2) and providing low pass filtered acoustic speech to simulate EAS (Chapter 3) can improve listening effort. No distinction could be made between the different EAS conditions, however, for future research it might be interesting to revisit this question using the sentence verification task to see if further hidden effects can be uncovered. While these strategies appear to work well in simulations and NH listeners, they may not be easily realizable in CI users. The source of reduced spectral resolution in CI users is not only related to the device, but also due to factors such as dead regions in the cochlea and current spread, which limits the CI users' effective use of spectral information contained in the signal (Fu et al., 1998; Stickney et al., 2006). Similarly, EAS may not be an option for many CI users, as they may not have any residual hearing. The results described in Chapter 5, however, are hopeful; the sentence verification task did show improved comprehension for increased spectral resolution from 7 up to 11 active electrodes in CI users.