References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


Summary

Understanding a single person talking is effortless in quiet surrounding. However, the task becomes much more challenging when the talker is surrounded by a loud crowd. The listener has to focus on the voice of the talker (target) to follow what they say, and not pay attention to the surrounding background (masker). For individuals with hearing impairment, who might already experience problems in understanding speech in quiet, the challenge of speech in noise becomes even more difficult. In the present dissertation, I investigated how characteristics from the voice of a talker – with focus on the pitch and the timbre – contribute to understand speech in difficult listening situations, such as noisy environments. Pitch is the height of the voice of a talker (low-pitched, high-pitched voices), whereas timbre is its character or color (clear, brassy, strident, warm voices). The overarching goal of this dissertation being to identify the voice characteristics necessary for better speech-in-noise understanding for hearing impaired (HI) listeners, especially users of cochlear implants (CIs). A CI is provided to individuals with severe or profound sensorineural hearing loss, when a hearing aid does not produce satisfying outcomes. However, percept of speech is degraded through a CI, and the conveyed pitch cues are weak.

How did I conduct my research?

I conducted behavioral experiments first on normal-hearing (NH) listeners to examine how they performed listening to voices with different characteristics (Chapters 2 and 3). Combined with these voice manipulations, in order to make speech perception more challenging, I presented sentences interrupted with silent gaps, or with noise filling in the gaps. The participants had to repeat what they could understand, by perceptually restoring the parts of the sentences replaced with silence or noise. This ability of the brain to reconstruct an incomplete sentence is called top-down restoration of speech. NH listeners usually show a restoration benefit, that is, as silent gaps are filled with noise, their intelligibility of the sentences increases. This method of measuring intelligibility of the interrupted sentences and the restoration benefit allowed me to test the importance of some voice characteristics at the level of the ear (the bottom-up cues) and at the level of the brain (the top-down cues). NH participants also listened to simulations of CI (Chapter 4), before actual CI users were tested (Chapter 5). In the latter two cases, speech (i.e. the bottom-up cues)
is degraded and is potentially more difficult to understand. The assumption is that the degradation of the bottom-up cues makes the brain (i.e. top-down restoration mechanisms) act differently than for NH listeners listening to speech that is not degraded. I intended to investigate whether voice pitch helped to better understand and restore degraded speech.

**The role of voice characteristics for NH listeners**

A man’s voice was manipulated (via its pitch and timbre) to make it sound like a woman’s voice (Chapter 2). Then, the man’s and the woman’s voices were made to alternate around the gaps (silent or filled with noise), disrupting the continuity of the voice within a sentence. Participants had to link the speech segments from the man’s voice with those from the woman’s voice in order to reconstruct the complete sentence. In this case, their speech perception decreased. I also showed that this perception decreased when only the timbre cues were manipulated but not when only the pitch cues were manipulated. These results suggest that the continuity of the timbre cues was more important for the intelligibility of interrupted speech than that of pitch cues, at least for the voice manipulations reported in the present dissertation. However, as far as restoration of interrupted speech is concerned, none of the disruptions in voice continuity had an effect on the restoration benefit, even when the two alternating voices were consistently reported to be from different talkers. Thus, voice continuity does not seem to be a prerequisite for top-down restoration of interrupted speech.

Next, I focused on the intonation of the sentence, as it is thought to help assembling successive speech segments into a coherent sentence. I manipulated the pitch contours to modify the sentence intonation (Chapter 3). Only the most atypical contours (inverted intonation) hindered interrupted speech intelligibility. No effect of the pitch contour manipulations was observed on top-down restoration.

Taken together, the results from these studies (Chapter 2 and 3) suggest that interrupted-speech perception and restoration are robust to most voice manipulations. This may be because the linguistic context of the sentence helps to understand the sentence despite the voice manipulations.
The role of access to pitch: from NH listeners to CI users

CIs convey weak pitch cues. I investigated if the poor performance of speech in noise by CI users could be due to this weak pitch percept. For this I tested how speech perception was affected in sentences with pitch as compared to sentences without or at best, weak pitch. I started with NH listeners listening to CI simulations either with original pitch cues or without any pitch cue (Chapter 4). To serve this purpose, I developed a new method to simulate CI with the possibility to add pitch cues. This CI simulation shows that the addition of pitch cues improves speech understanding at different degrees of speech degradation (N.B.: the lower the spectral resolution, the more degraded speech is). The results also confirm the importance of pitch in top-down repair of degraded speech for spectral resolutions in the range of actual CI users. This suggests that the combination of low spectral resolution and weak pitch representation may contribute to the poor top-down restoration of speech observed in CIs.

Translating the results from NH listeners to actual CI users was done by testing CI users with residual hearing in their non-implanted ear (bimodal CI users). Bimodal CI users wear a CI in one ear and a hearing aid (HA) in the other ear. It is possible for them to have access to pitch cues via the low-frequency residual hearing in their HA. I thus tested them twice, with and without their HA along their CI, to investigate if they benefited from the pitch cues from the HA (Chapter 5). However, at the group level no benefit from the HA was observed for top-down restoration of speech, although individual analysis showed that some bimodal users could benefit from adding the low-frequency acoustic residual hearing to trigger top-down repair mechanisms. Moreover, a larger number of participants could benefit from the acoustic cues delivered by their HA along their CI for top-down restoration of interrupted speech as the gaps were shortened. This suggests that every improvement in quality of bottom-up cues (in frequency and in time) may contribute with an additive effect to better top-down restoration, showing a powerful interaction between bottom-up cues and top-down repair mechanisms. However, for individuals showing no bimodal benefit, it is possible that access to pitch via the low frequency residual hearing was either not a sufficient cue or was not combined properly with cues from the CI.
Conclusion

In summary, this dissertation shows that speech perception can be robust to voice manipulations when linguistic context can prevail over these manipulations. Moreover, it also shows the importance of pitch for interrupted-speech perception as well as for triggering the top-down restoration mechanisms, especially when bottom-up cues are degraded. However, the promising results from CI simulations were not replicated with actual CI users at the group level. That is, speech perception improvement expected with access to pitch cues was not captured. Nevertheless, on a more positive note, individual analysis suggests that the additional HA can provide better sound quality. Further research is thus needed to supplement the contribution of these findings in improving speech perception in noise for CI users.
Samenvatting

Het kost geen moeite om te begrijpen wat een spreker zegt als het stil is in de omgeving. Het wordt echter veel lastiger zodra de spreker zich bevindt in een lawaaiige groep mensen. De luisteraar moet zich concentreren op de stem van de spreker (target) om goed te kunnen volgen wat er gezegd wordt, terwijl het achtergrondgeluid (masker) zoveel mogelijk genegeerd moet worden. Voor mensen met slechthorendheid, die vaak al problemen hebben bij het verstaan van spraak in stille omgevingen, wordt deze uitdaging nog moeilijker. In dit proefschrift heb ik onderzocht hoe de eigenschappen van de stem van een spreker, met nadruk op toonhoogte en timbre, invloed hebben op het verstaan van spraak in lastige luistersituaties. Met de toonhoogte wordt de hoogte van de stem bedoeld (lage stem, hoge stem) en de timbre is zijn karakter (heldere, schelle, brassige of warme stem). Het hoofddoel van dit proefschrift is om aan te tonen welke verschillende karakteristieke eigenschappen van de stem nodig zijn voor een beter begrip van spraak in achtergrondgeluid bij slechthorende luisteraars, met name gebruikers van het cochleaire implantaat (CI). Een CI wordt voorzien aan mensen met ernstige of volledige slechthorendheid, wanneer een gehoorapparaat niet voldoende helpt. Het spraakbegrip wordt echter gedegradeerd als gevolg van het CI en toonhoogte cues worden zwak overgedragen.

Hoe heb ik mijn onderzoek uitgevoerd?

Eerst heb ik gedragsexperimenten gedaan bij normaal horende (NH) luisteraars om te onderzoeken hoe ze presteerden bij het luisteren naar stemmen met verschillende karakteristieke eigenschappen (hoofdstuk 2 en 3). Ik heb hen laten luisteren naar zinnen die onderbroken worden door delen weg te laten en ze te vervangen met stille periodes of periodes van achtergrondgeluid, om hun spraakperceptie te testen. De deelnemers moesten vervolgens herhalen wat ze konden begrijpen van de zin door goed te luisteren en de delen van de zin op te vullen die ik vervangen had met stilte of achtergrondgeluid. Het vermogen van de hersenen om een incomplete zin zoals deze te herstellen noemt men ook wel top-down restauratie van spraak. NH luisteraars ondervinden hier vaak voordelen van. Ze begrijpen namelijk de zinnen beter naarmate de stille periodes worden opgevuld met achtergrondgeluid. Dankzij deze methode, waarbij het begrip van de onderbroken zinnen en de mate van restauratie wordt gemeten, kon ik het belang van zintuiglijke waarneming
SAMENVATTING

(d.w.z. wat het oor verwerkt, bottom-up cues) en cognitieve waarneming (d.w.z. wat het brein verwerkt, top-down cues) van sommige karakteristieke eigenschappen van een stem bepalen. NH luisteraars luisterden tevens naar simulaties van CI (hoofdstuk 4), voordat werkelijke CI gebruikers werden getest (hoofdstuk 5). In de laatste twee gevallen wordt spraak (bottom-up cues) gedegradeerd en is het daardoor potentieel lastiger om te begrijpen. De aannemer is dat bij de degradatie van de bottom-up cues het brein (top-down restauratie mechanismen) anders functioneert dan wanneer NH luisteraars luisteren naar niet gedegradeerde spraak. Mijn doel was om te onderzoeken of de toonhoogte helpt bij het spraakbegrip en het herstel van gedegradeerde spraak.

De rol van de karakteristieke eigenschappen van een stem bij NH luisteraars

De stem van een man werd bewerkt (met behulp van toonhoogte en timbre) om het meer op een vrouwenspreken te laten lijken (hoofdstuk 2). Vervolgens werd de mannenstem en de verkregen vrouwenspreken zo aangepast dat de onderbrekingen (stilte of achtergrondgeluid) afgewisseld werden tussen de twee verschillende stemmen, met als doel de continuïteit van de zin te versturen. Deelnemers moesten vervolgens despraaksegmenten van man-envrouwstemonderlingverbinden om de volledige zin te kunnen reconstrueren. In dit geval nam hun spraakperceptie af. Ik heb tevens aangetoond dat de perceptie afneemt wanneer alleen de timbre veranderd werd, maar niet wanneer alleen de toonhoogte veranderd werd. Deze resultaten doen suggereren dat de continuïteit van timbre cues belangrijker is voor onderbroken spraakbegrip dan dat van toonhoogte cues, althans voor de stem die gebruikt werd in dit proefschrift. Echter, als men kijkt naar herstel van onderbroken spraak, had geen van de onderbrekingen in de continuïteit van de stem een effect op het voordeel van spraakherstel, zelfs niet wanneer consistent werd aangegeven dat de twee afwisselende stemmen van verschillende sprekers waren. Zodoende lijkt dus de continuïteit van de stem niet van invloed te zijn op top-down herstel van onderbroken spraak.

Vervolgens heb ik me gericht op de timbre van de zin, aangezien die helpt bij het opbouwen van een coherente zin uit spraaksegmenten. Ik heb de contouren van de toonhoogte gemanipuleerd om de timbre van de zin aan te passen (hoofdstuk 3). Alleen de meest atypische contouren (geïnverteerde timbre) belemmerden onderbroken
spraakbegrip. Er is dus geen invloed van toonhoogte contour manipulaties op top-down spraakherstel gevonden.

Als de resultaten van deze studies samen worden genomen, doet dat suggereren dat onderbroken spraakperceptie en onderbroken spraakherstel niet beïnvloed worden door manipulaties van de stem. Dit kan wellicht komen doordat de taalkundige context helpt bij het begrijpen van de zin, ondanks de stemmanipulaties.

**De rol van toegang tot toonhoogte: van NH luisteraars naar CI gebruikers**

CIs dragen slecht toonhoogte over. Ik heb onderzocht of het slechte begrip van spraak in achtergrondgeluid bij CI gebruikers komt door slechte overdracht van toonhoogte. Hiervoor heb ik getest hoe spraakperceptie beïnvloed werd in zinnen met aanwezige toonhoogte ten opzichte van zinnen zonder toonhoogte of, op zijn minst, een zwakke toonhoogte. Ik ben begonnen met NH luisteraars die luisterden naar simulaties van CI-spraak met de originele toonhoogte of CI-spraak zonder toonhoogte (hoofdstuk 4). Om dit doel te behalen heb ik een nieuwe methode ontwikkeld om CI te simuleren, waarbij het mogelijk is om toonhoogte cues in te voegen. Deze CI simulatie demonstreert dat de toevoeging van toonhoogte cues spraakbegrip bevordert op verschillende niveaus van spraakdegradatie. (N.B.: hoe lager de spectrale resolutie, hoe verder gedegradeerd de spraak is). De resultaten bevestigen tevens het belang van toonhoogte in top-down herstel van gedegradeerde spraak voor spectrale resoluties die men bij werkelijke CI gebruikers kan verwachten. Dit doet suggereren dat de combinatie van lage spectrale resolutie met zwakke toonhoogte representatie bij kan dragen aan slechte top-down spraakherstel, zoals dat waargenomen wordt bij CI gebruikers.

Het omzetten van de resultaten van NH luisteraars naar werkelijke CI gebruikers werd gedaan door CI gebruikers met restgehoor in hun niet-geïmplanteerde oor te testen (bimodale CI gebruikers). Bimodale CI gebruikers dragen hun CI in een oor en het gehoorapparaat in het andere. Het is voor hen alsnog mogelijk om toegang te hebben tot toonhoogte cues via het lage frequentie restgehoor via hun gehoorapparaat. Hierom heb ik ze tweemaal getest. Eenmaal met en eenmaal zonder hun gehoorapparaat, maar beide samen met hun CI. Zo heb ik onderzocht of ze voordeel hadden dankzij de toonhoogte cues in hun gehoorapparaat (hoofdstuk 5).
SAMENVATTING

Gemiddeld genomen over de hele groep was er echter geen voordeel aangetoond van het gehoorapparaat in het geval van top-down herstel van spraak. Individuele analyse liet echter toch zien dat sommige bimodale gebruikers wel degelijk voordeel konden hebben door de toevoeging van lage frequentie restgehoor om top-down restauratiemechanismen te activeren. Bovendien kon een grotere hoeveelheid deelnemers voordeel hebben dankzij de akoestische cues van hun gehoorapparaat in combinatie met hun CI voor top-down herstel van onderbroken spraak wanneer de pauzes in de zinnen verkort werden. Dit doet suggereren dat elke verbetering in kwaliteit van bottom-up cues (zowel in frequentie als in tijd) bij zou kunnen dragen om nog betere top-down spraakherstel te verkrijgen. Dit demonstreert een sterke wisselwerking tussen bottom-up cues en top-down restauratiemechanismen. Echter, voor individuen die geen bimodaal voordeel lieten zien, is het mogelijk dat de toegang tot toonhoogte cues via het lage frequentie restgehoor niet voldoende was of dat de signalen van de CI er niet adequaat mee werden gecombineerd.

Conclusie

Samengevat laat dit proefschrift zien dat spraakperceptie robuust kan zijn tegen stemmanipulaties wanneer taalkundige context van meer belang is dan deze manipulaties. Bovendien is ook het belang van de aanwezigheid van toonhoogte gedemonstreerd voor zowel onderbroken spraakperceptie als top-down herstelmechanismen, met name wanneer bottom-up cues werden gedegradeerd. De veelbelovende resultaten van de CI simulaties konden echter niet gerepliceerd worden bij de werkelijke CI gebruikers op het niveau van de hele groep. Ofwel, de verbeteringen in spraakperceptie die men verwacht bij toegang tot toonhoogte cues konden niet aangetoond worden. Anderzijds is het wel positief dat individuele analyse toch laat zien dat een toegevoegd gehoorapparaat kan resulteren in betere geluidskwaliteit bij de gebruiker. Meer onderzoek is dus nodig om de werkelijke toegevoegde waarde aan te tonen bij verbeterde spraakperceptie in achtergrondgeluid voor CI gebruikers.
About the author

This is a very short summary about the life of Jeanne Clarke, as she likes to talk about herself in the third person (she has the same initials as Julius Caesar after all).

After growing up in the countryside in the middle of France, Jeanne moved to the “big city” of St-Etienne to study physics at the university. With her Bachelor’s degree in hands, she moved to Grenoble with Bastien for her Master’s degree in signal processing. She was interested in research and did her internships in the local laboratories. But after her graduation, she found a job as a consulting engineer (which, in her humble opinion, is the equivalent of being a piece of meat at the butcher, sold to the first costumer who wants it). She moved to the urban area of Paris to program printers, a job she didn’t enjoy much. She still carries the scars of this experience as she gets really upset when a printer does not work, thus getting fired was a very positive thing! She moved back with Bastien in Grenoble and went back to university for a Master’s degree in Cognitive sciences and to get back to research. Just after graduation, she travelled to Groningen for an interview for the PhD position she would get. It was the beginning of the summer and the weather was nice in the Netherlands. She didn’t know what she was really getting into when she took the job, starting in October, when the days are already much shorter and the rain is a daily routine. Bastien joined her in a heartbeat in this new adventure, she had just enough time to buy her green hoody. They soon realized they had something in common with the local people: complaining about the weather, even if there’s nothing much you can do about it. What they could not understand was why Dutch people did not complain about the food, because something could definitively be done to improve it! However, one thing they couldn’t complain about for sure was the standard of living. And after 4 years in this country (without getting pneumonia, I thank my warm green hoody), they had an amazing daughter. Almost another year went by during which Jeanne finished writing up the present dissertation. The whole family is back to France for the post-doc position she found in Marseille. Sure the weather is different, so she’s not wearing her green hoody anymore!