Spectral contrast effects and auditory enhancement under normal and impaired hearing

Andrew J. Oxenham*

Department of Psychology, University of Minnesota – Twin Cities, Elliott Hall N218, 75 East River Road, Minneapolis, Minnesota 55455, USA

Abstract: We are generally able to identify sounds and understand speech with ease, despite the large variations in the acoustics of each sound, which occur due to factors such as different talkers, background noise, and room acoustics. This form of perceptual constancy is likely to be mediated in part by the auditory system’s ability to adapt to the ongoing environment or context in which sounds are presented. Auditory context effects have been studied under different names, such as spectral contrast effects in speech and auditory enhancement effects in psychoacoustics, but they share some important properties and may be mediated by similar underlying neural mechanisms. This review provides a survey of recent studies from our laboratory that investigate the mechanisms of speech spectral contrast effects and auditory enhancement in people with normal hearing, hearing loss, and cochlear implants. We argue that a better understanding of such context effects in people with normal hearing may allow us to restore some of these important effects for people with hearing loss via signal processing in hearing aids and cochlear implants, thereby potentially improving auditory and speech perception in the complex and variable everyday acoustic backgrounds that surround us.

Keywords: Hearing loss, Cochlear implants, Context effects

PACS number: 43.66.Ba  [doi:10.1250/ast.41.108]

1. INTRODUCTION

Sounds are rarely if ever heard in complete isolation, devoid of any context or surrounding environment. The context in which a sound is presented can often affect the physical acoustics of the sound, such as when a particular speech sound is spoken by different talkers, in different rooms, or in the presence of different background noises. Nevertheless, in most cases we are able to accurately identify specific speech sounds, despite the wide range of different acoustic conditions in which we encounter them. This phenomenon is a form of perceptual constancy, whereby different physical waveforms are perceived as having the same identity or belonging to the same category [1].

The neural mechanisms underlying the various forms of auditory context effects remain uncertain. However, many are thought to involve some form of adaptation and consequent contrast enhancement. For instance, after a broadband sound with a spectral peak or notch, a following spectrally flat broadband sound is often perceived to have a spectral notch or peak, respectively, at the corresponding frequency, in a way that is analogous to visual contrast aftereffects, such as seeing a dark spot after looking at a bright light and then looking away.

Prior context has been known to affect speech perception for over 60 years [2]. Although early explanations tended to focus on speech-specific mechanisms, such as motor theories of speech production [3], many of the phenomena that do not involve semantic or linguistic knowledge can be understood in terms of more general auditory phenomena, such as spectral contrast effects [4]. Spectral contrast may also underlie a well-known auditory context effect known as auditory enhancement [5–7].

Because of their role in maintaining perceptual constancy, context effects are likely to play an important role in everyday auditory and speech perception. It is therefore critical to understand the mechanisms behind these effects and to determine whether any of these effects are altered by hearing loss or by cochlear implants. In particular, any alteration or reduction in context effects induced by hearing loss may result in a reduced ability to adapt to new acoustic environments and new talkers, which in turn may affect everyday communication in ways that are not reflected in clinical tests of auditory and speech perception. More importantly, by mapping out differences in context effects between normal-hearing listeners and listeners with
2. SPEECH CONTEXT EFFECTS

2.1. Vowel Enhancement in Normal-hearing Listeners and Cochlear-implant Users

Early work using artificial vowels had shown that the percept of a vowel could be induced by a broadband harmonic complex when it was preceded by a broadband stimulus that had spectral gaps where the formant frequencies should be [8]. In other words, the precursor with the spectral gaps produced the equivalent of a negative afterimage on the subsequent stimulus, even if the formants in the subsequent stimulus were not elevated in level, relative to the surrounding harmonics.

One possible mechanism by which such adaptation and contrast enhancement could occur is via the medial olivocochlear (MOC) efferent system [10], which could decrease the gain of the outer hair cells in frequency regions that are being stimulated, while leaving the gain in non-stimulated frequency regions intact. If the MOC system was involved, then we would not expect this vowel enhancement to occur in cochlear-implant users, as the cochlear implant bypasses the cochlea and stimulates the auditory nerve directly.

We tested this hypothesis by measuring the ability of listeners to identify artificial vowels with and without a 1-second precursor, as a function of the relative level of the harmonics at formant frequencies [9]; see Fig. 1. We found that cochlear-implant users showed evidence of the effect of a precursor, but that the effect was reduced relative to what was found with normal-hearing listeners. However, when normal-hearing listeners were presented with the stimuli via a noise-excited envelope vocoder, designed to simulate the effects of poorer spectral resolution, the enhancement effects were also reduced to the levels found with the cochlear-implant users.

Overall, the results suggested that vowel enhancement effects were reduced in cochlear-implant users, but that the reduction was not due to the loss of efferent activation, as the same reduction in enhancement effects was observed in normal-hearing listeners when spectral resolution was appropriately reduced. These results ruled out the MOC system as the sole source of vowel enhancement effects, but also illustrated how a loss of spectral resolution could result in a reduction in speech context effects.

2.2. Changes in Perceived Vowel Identity due to Preceding Context in Normal-hearing Listeners

Another widely studied context effect in speech perception involves the effects of a preceding sentence (or other stimulus) on the identity of a target vowel [11–13]. Specifically, if the precursor sentence is filtered to emphasize the formant frequencies of one vowel and deemphasize the formant frequencies of a second vowel, then the final target vowel is more likely to be perceived as the second vowel, in line with expectations based on spectral contrast enhancement. Some work had shown that the effect is reduced when the precursor and target are presented to the same ear than when they are presented to opposite ears, consistent with a peripheral explanation [13]. On the other hand, the effect is also reduced (but not by as much) when interaural time differences were used to lateralize the precursor and target to opposite sides, even though the stimulus intensity at the two ears remained the same [13]. These results suggest that the effect may have both peripheral and central components.

These early conclusions were supported by a recent study from our lab that also found a stronger effect of ear than of perceived direction, using two different methods of generating an acoustic continuum between two vowels (/i/ and /æ/, as in “bit” and “bet”) [14]. Our study also found evidence for effects of attention on the context effects. Different precursor sentences were presented to each ear (or lateralized to each side), and listeners were instructed to attend to either one side or the other before judging the identity of a target vowel that was generated on the continuum between /i/ and /æ/. When the target vowel was presented diotically, the sentence to which the listener had attended had a greater effect on the perceived identity of the vowel. However, when the target vowel was presented to one side or the other, the side of presentation of the precursor had a stronger effect than the side to which attention was directed [14].

![Fig. 1 Schematic diagram of the stimuli used in the vowel identification experiment [9]. Vowel identification was measured with and without the precursor as a function of formant level in both normal-hearing listeners and cochlear-implant users. Reproduced from with permission of the American Institute of Physics.](image)

109
The results are not consistent with a purely peripheral mechanism underlying the vowel context effect. Instead they suggest a complex interaction between peripheral and central mechanisms, modulated by subtle but measurable effects of attention.

### 2.3. Changes in Perceived Vowel Identity by Preceding Context in Cochlear-implant Users

The same spectral contrast context effect on vowel identity was recently studied in cochlear-implant users [15]. In contrast to the results found for vowel enhancement in implant users [9], here the cochlear-implant users actually showed larger context effect sizes than were found with normal-hearing listeners, by a factor of two or more, even when accounting differences in overall sensitivity to the phoneme boundary [15]. The reason for this increase in effect size remains somewhat unclear. However, one hypothesis suggested by Feng and Oxenham [15] is that the cochlear-implant users are relying on different spectral cues to distinguish the two vowels [16]. In particular, the cochlear-implant users may be relying on gross differences in the overall spectrum (such as overall spectral tilt, or shifts in the spectral centroid), whereas the normal-hearing listeners are relying on more fine-grained spectral differences, such as the different peaks of the formant frequencies. If the gross differences produce larger spectral changes, then the cochlear-implant users may exhibit larger context effects than the normal-hearing listeners, who are relying on more subtle cues that are less affected by the context filtering.

The results are important because they show that cochlear implants produce different patterns of results than normal hearing, but that the direction of the change (increase or decrease in context effects) is not always the same.

### 3. AUDITORY ENHANCEMENT

Auditory enhancement refers to the increase in audibility of a target sound in the presence of a masking sound, produced by a copy of the masking sound prior to the presentation of the target and masker. The effects of auditory enhancement have been known for some time, as early as 1940 [5]. However, it was only in the 1980s that Viemeister [6] and Viemeister and Bacon [7] developed a method with which to quantify the effects of enhancement. They found that the threshold for detecting a pure tone target in the presence of a spectrally flanking masker was reduced (improved) by adding a precursor that was a copy of the masker [6]. In addition, the enhanced component produced more forward masking on a subsequent probe, suggesting that enhancement was not just a reduction in the effectiveness of the masking components (perhaps via adaptation), but instead was an effective increase in the internal representation of the target, which caused it to produce more masking to a subsequent tone [7]. This pattern of results led to the hypothesis of “adaptation of inhibition,” whereby the mutual inhibition of spectrally adjacent components is reduced or adapted by the precursor, meaning that the adjacent masker components now inhibit the response to the target less than they would have in the absence of the precursor [7].

### 3.1. Mechanisms of Auditory Enhancement, Studied in Normal-hearing Listeners

A recent study from our lab provided a more direct test of the adaptation of inhibition hypothesis by estimating the perceived loudness of the target and masker components in the presence and absence of the precursor [17]. Normal-hearing listeners were asked to judge the loudness of these components relative to a comparison stimulus that was presented 2 seconds later. Interestingly, the precursor had no effect on the loudness of the masker components, or on the loudness of the target when it was presented on its own. However, when the target and masker were presented together, the precursor produced an increase in the perceived loudness of the target by an effective 10 dB [17], which is similar to what has been observed in masking studies.

The results are consistent with the adaptation-of-inhibition hypothesis, in that the effective level of the target is only affected when the masker components are present at the same time, and that the effective level of the target is greater, in absolute terms, when the precursor is present.

### 3.2. Effects of Hearing Loss and Cochlear Implants on Auditory Enhancement

Two recent studies in our lab have explored auditory enhancement in both cochlear-implant users [18] and listeners with sensorineural hearing loss of presumed cochlear origin [19]. In both cases, detection thresholds for a pure tone in the presence of spectrally flanking masker tones were reduced by the introduction of a precursor that was a copy of the masker. In both cases the absolute amount of enhancement was less than was found in normal-hearing listeners, although the amount of enhancement was actually similar across all three groups, if enhancement was calculated as a proportion of the total dynamic range for each participant.

One interesting finding is that enhancement is observed under simultaneous masking, but not under forward masking for both cochlear-implant users [18] and hearing-impaired listeners [19,20]. This intriguing difference suggests a potential difference in mechanism underlying enhancement in simultaneous vs. forward masking, which has yet to be fully elucidated. It may be that the within-
channel adaptation mechanisms are functional, whereas the across-channel inhibition mechanisms are not, perhaps due to a loss of peripheral frequency selectivity.

Finally, it has become clear that not all forms of auditory enhancement are the same. Although the detection threshold for a simultaneously masked target tone was enhanced by the presence of a precursor in cochlear-implant users [18], a more recent study has shown that supra-threshold effects of enhancement are not necessarily found in cochlear-implant users [21]. In that study, listeners were asked to judge whether a probe tone had been presented in the preceding mixture of a target tone and spectrally flanking masker tones. A previous study had shown that a precursor could decrease the level of the target needed to perform the task by as much as 20 dB, and that the time-constants associated with this effect required only a relatively short onset in the tens of milliseconds, and an offset that extended out to as far as one second [22]. In contrast, cochlear-implant users showed essentially no enhancement under such conditions, resulting in a 20-dB or more difference between the performance of the two groups. As with the work on vowel enhancement [9], much of the difference could be ascribed to the loss of spectral resolution, because normal-hearing listeners showed a similar lack of enhancement when spectral resolution was limited by passing the stimuli through a noise-excited envelope vocoder with shallow filter slopes. It remains unknown whether hearing-impaired listeners exhibit such suprathreshold enhancement, or whether the amount of enhancement is related to the degree or type of hearing loss.

Overall, the results so far suggest that the mechanisms underlying auditory enhancement are present to some extent in people with hearing loss and cochlear implants, but that some large, and presumably important, effects are not experienced in these populations, leaving room for interventions via signal processing that may reintroduce some enhancement effects and potentially improve acoustic communication in noisy and variable situations.

4. SUMMARY

Context effects in auditory and speech perception are the rule rather than the exception. Nevertheless, the nature of these effects and their neural bases remain poorly understood. By carrying out empirical tests of such context effects using speech and non-speech auditory stimuli in normal-hearing listeners, it is possible to develop a model of the underlying mechanisms. In addition, by studying these context effects in clinical populations with hearing loss and cochlear implants, it becomes possible to determine the extent to which such effects are altered by changes in cochlear processing or a different stimulation mode (electric as opposed to acoustic).

Understanding how these effects occur and how they are affected by hearing loss could lead to improved signal processing algorithms for both hearing aids and cochlear implants that will ideally restore the auditory and speech context effects observed in the normal-hearing population. Approaches as simple as channel-based automatic gain control could be used to restore some context effects, so long as we know from studies in normal-hearing listeners what time constants for both attack and release would be most appropriate. These approaches should in turn make it easier for people with hearing loss to function in everyday acoustic environments, where perceptual constancy is required.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (Grant R01 DC012262). Current and past lab members involved in the work described include Lei Feng, Heather Kreft, and Ningyuan Wang.

REFERENCES

[1] P. K. Kuhl, “Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories,” J. Acoust. Soc. Am., 66, 1668–1679 (1979).
[2] D. E. Broadbent, P. Ladefoged and W. Lawrence, “Vowel sounds and perceptual constancy,” Nature, 178(4537), 815–816 (1956).
[3] A. M. Liberman and I. G. Mattingly, “The motor theory of speech perception revised,” Cognition, 21, 1–36 (1985).
[4] L. L. Holt and A. J. Lotto, “Speech perception within an auditory cognitive science framework,” Curr. Dir. Psychol. Sci., 17, 42–46 (2008).
[5] J. F. Schouten, “The residue, a new component in subjective sound analysis,” Proc. Kon. Ned. Akad. Wetensch., 43, 356–365 (1940).
[6] N. F. Viemeister, “Adaptation of masking,” in Psychophysical, Physiological and Behavioural Studies in Hearing, G. van den Brink and F. A. Bilsen, Eds. (Delft University Press, Delft, Netherlands, 1980), pp. 190–198.
[7] N. F. Viemeister and S. P. Bacon, “Forward masking by enhanced components in harmonic complexes,” J. Acoust. Soc. Am., 71, 1502–1507 (1982).
[8] Q. Summerfield, M. P. Haggard, J. Foster and S. Gray, “Perceiving vowels from uniform spectra: Phonetic exploration of an auditory after-effect,” Percept. Psychophys., 35, 203–213 (1984).
[9] N. Wang, H. Kreft and A. J. Oxenham, “Vowel enhancement effects in cochlear-implant users,” J. Acoust. Soc. Am., 131, EL421–426 (2012).
[10] J. J. Guinan, Jr., “Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans,” Ear Hear., 27, 589–607 (2006).
[11] P. Ladefoged and D. E. Broadbent, “Information conveyed by vowels,” J. Acoust. Soc. Am., 29, 98–104 (1957).
[12] C. E. Stilp, P. W. Anderson and M. B. Winn, “Predicting contrast effects following reliable spectral properties in speech perception,” J. Acoust. Soc. Am., 137, 3466–3476 (2015).
[13] A. J. Watkins, “Central, auditory mechanisms of perceptual compensation for spectral-envelope distortion,” J. Acoust. Soc. Am., 90, 2942–2955 (1991).
[14] L. Feng and A. J. Oxenham, “Spectral contrast effects produced by competing speech contexts,” J. Exp. Psychol. Hum. Percept. Perform., 44, 1447–1457 (2018).

[15] L. Feng and A. J. Oxenham, “Effects of spectral resolution on spectral contrast effects in cochlear-implant users,” J. Acoust. Soc. Am., 143, EL468 (2018).

[16] M. B. Winn and R. Y. Litovsky, “Using speech sounds to test functional spectral resolution in listeners with cochlear implants,” J. Acoust. Soc. Am., 137, 1430–1442 (2015).

[17] N. Wang and A. J. Oxenham, “Effects of auditory enhancement on the loudness of masker and target components,” Hear. Res., 333, 150–156 (2016).

[18] H. A. Kreft and A. J. Oxenham, “Auditory enhancement in cochlear-implant users under simultaneous and forward masking,” J. Assoc. Res. Otolaryngol., 18, 483–493 (2017).

[19] H. A. Kreft, M. Wojtczak and A. J. Oxenham, “Auditory enhancement under simultaneous masking in normal-hearing and hearing-impaired listeners,” J. Acoust. Soc. Am., 143, 901–910 (2018).

[20] L. M. Thibodeau, “Performance of hearing-impaired persons on auditory enhancement tasks,” J. Acoust. Soc. Am., 89, 2843–2850 (1991).

[21] L. Feng and A. J. Oxenham, “Auditory enhancement and the role of spectral resolution in normal-hearing listeners and cochlear-implant users,” J. Acoust. Soc. Am., 144, 552 (2018).

[22] L. Feng and A. J. Oxenham, “New perspectives on the measurement and time course of auditory enhancement,” J. Exp. Psychol. Hum. Percept. Perform., 41, 1696–1708 (2015).