European Archives of Oto-Rhino-Laryngology (2019) 276:2673–2680
https://doi.org/10.1007/s00405-019-05502-9

OTOLOGY

Perception of lexical stress cued by low-frequency pitch and insights into speech perception in noise for cochlear implant users and normal hearing adults

Hilal Dincer D'Alessandro[1,2] · **Patrizia Mancini**[1]

Received: 9 February 2019 / Accepted: 3 June 2019 / Published online: 8 June 2019

Abstract

Purpose Cochlear implant (CI) users show great difficulty for understanding speech in noise and this fact may partly stem from their poor low-frequency (LF) pitch perception and temporal fine structure (TFS) processing. Clinical assessment of pitch perception is usually based on non-speech tasks. However, linguistically relevant contexts such as lexical stress may better reflect the role of pitch in speech perception, especially for everyday speech where background noise is inevitable. Hence, the study aimed to assess perception of lexical stress cued by LF pitch and TFS cues for CI and normal hearing (NH) listeners, and to investigate relationships with speech perception in noise.

Methods The low-pass-filtered Word Stress Pattern (WSP-LPF) test was used to evaluate perception of lexical stress cued by LF pitch. Speech perception was assessed with the sentence test with adaptive randomized roving level (STARR) test which presented everyday sentences at low, medium, and high levels in a fluctuating noise to estimate a Speech Reception Threshold. This new test intended to give a realistic estimate for real-world listening.

Results Median WSP-LPF scores in NH ($N=18$) and CI listeners ($N=18$) were 12.0 Hz and 67.0 Hz, respectively. The corresponding STARR scores were − 9.1 dB and 17.3 dB. Group differences were statistically significant ($p<0.001$). Analysis showed significant positive correlations for NH ($r_s=0.50$) and CI listeners ($r_s=0.60$).

Conclusions Present findings reveal stronger correlations than previous studies using non-speech materials, supporting that CI listeners' poor speech perception in noise might be strongly associated with their inability for LF pitch perception and TFS processing.

Keywords Lexical stress · Pitch perception · Speech perception in noise · Cochlear implants · Temporal fine structure

Introduction

The cochlear implant (CI) has opened up a new therapeutic option to people with severe to profound sensorineural hearing loss for restoring functional hearing and achieving better spoken language skills. As a result of recent advances in CI technology, many children and adolescents implanted at an early age show age-appropriate language and communication skills [1–4] while most adult CI users achieve a very good performance for open-set speech perception in quiet and are even able to converse on the phone [5–7]. However, current cochlear implants have a reduced capacity to convey pitch cues that are fundamental for auditory perception [8]. Thus, CI users, in comparison to people with normal hearing (NH) or hearing aids (HA), receive degraded temporal and spectral information especially in the low frequency (LF) domain. This fact becomes more evident when CI users face challenging speech perception tasks such as listening in the presence of background noise which is a natural part of daily life [6, 7, 9, 10].

Lexical stress is defined as the relative prominence on a syllable in a multisyllabic word. Perceptually, a stressed syllable is characterized by changes in amplitude, duration and pitch or Fundamental Frequency (F0) which is the prime acoustic correlate of pitch in a speech signal [11,

✉ Hilal Dincer D'Alessandro
hilal.dincer@uniroma1.it; dincer.hilal@gmail.com

1 Department of Sense Organs, Sapienza University of Rome, Viale dell'Università 31, 00161 Rome, Italy

2 Department of Audiology, Faculty of Health Sciences, Istanbul Aydın University, Beşyol Mah. İnönü Cad. No: 38 Küçükçekmece, Istanbul, Turkey

12]. When lexical stress is produced through pitch alone, it is also called pitch accent. Languages vary in their stress assignment pattern within words. Despite of presence of certain linguistic rules such as stress dominance (i.e., the most frequent stress pattern) and stress neighborhood (i.e., the number of words sharing the same final sequence and stress pattern), lexical stress placement might be unpredictable. Italian is such a language where spelling-sound correspondences at the segmental level are regular, but lexical stress placement is largely variable [13, 14]. The most frequent pattern in Italian is the stress assignment on the penultimate syllable with about 80% of multisyllabic words bearing such stress (e.g., bam-BI-no 'child') [15, 16] while about 20% bear antepenultimate stress (e.g., A-ni-ma 'soul') [15, 17].

Lexical stress may have profound consequences for speech perception and language acquisition, since the stressed syllable may signal the meaning of the word (e.g., là 'there' versus la 'the') [18]. Indeed, sensitivity to overall stress pattern of the native language seems to be present from very early on. Newborns with normal hearing show native stress pattern in their cries [19]. Electrophysiological responses reveal that infants at 4–5 months of age already distinguish between trochaic (i.e., strong–weak) and iambic (i.e., weak–strong) stress patterns in a language-specific manner [20, 21]. Likewise, congenitally deaf children implanted at an early age show the same developing pattern within the first 6 months of CI use with event-related potentials reflecting habituation to the native stress pattern and a negative mismatch response to foreign stimulus [21, 22]. However, findings in older children with cochlear implants indicate perception of lexical stress as the most difficult one among other tasks such as perception of word pattern and intonation [23]. Similarly, severely to profoundly hearing-impaired adults are reported to have difficulties in perceiving the stress pattern and to be able to perceive the stress pattern only when amplitude cue is available in the stimuli and even exaggerating F0 or durational cues does not improve performance [24, 25].

The Word Stress Pattern (WSP) test, integrated in the Auditory Speech Sounds Evaluation (ASSE) suite, is a newly developed test to evaluate listeners' perception of lexical stress by changes in pitch, i.e., listeners' perception of pitch in a linguistic context [26, 27]. The test exists in two modes: non-filtered WSP and low-pass-filtered WSP (WSP-LPF). In the non-filtered mode, high-frequency (HF) cues are available whilst in the alternative mode, HF cues are removed by low-pass filtering the speech stimuli and high-pass filtered white noise is added. Hence, low-pass filtering maintains only frequencies in which the F0 is contained and the higher harmonics are suppressed. The WSP test has been validated for a Germanic (Dutch) and two Romance languages (Italian and Romanian) and normative data are reported in 90 adults. NH findings show a statistically significant difference between LPF versus non-filtered versions of the WSP test with a tendency towards better outcomes on the LPF mode, especially for Italian participants. Not surprisingly, results in hearing-impaired listeners are significantly worse than NH people; participants with LF hearing loss perform significantly poorer than those with HF hearing loss, especially for WSP-LPF testing; and among all hearing-impaired listeners, CI users have more difficulty in performing the test and to obtain scores that are within normal range [26, 27]. WSP outcomes in relation to other speech perception measures in CI users, however, are so far not reported.

Pitch perception is known to contribute to the processing of linguistic information and to be indicative of speech perception skill. On the other hand, clinical assessment of pitch perception is usually performed using non-speech tasks such as discrimination of tone changes. However, non-speech tests based on tone complexes are shown to overestimate listeners' performance and it is believed that the use of linguistically relevant contexts such as lexical stress may reflect a more realistic assessment for the role of pitch at a functional level, probing everyday listening situations where background noise is inevitable and speech level may fluctuate together with the noise [26, 27]. Hence, the objectives of the present study have been to evaluate perception of lexical stress cued by LF pitch for Italian-speaking adults with cochlear implants and normal hearing as well as to investigate the relationships with outcomes obtained from a speech perception test presented in fluctuating noise. For this purpose, the LF pitch perception is assessed using the LPF mode of WSP test. The outcomes are thought to be indicative of the ear's processing capacity of Temporal Fine Structure (TFS) which is known to be important for the F0 information, for frequencies lower than 1000 Hz, and for understanding speech in a fluctuating background. The speech perception performance in noise is tested with the Sentence Test with Adaptive Randomized Roving Level (STARR) test where sentences are presented at a roving level whilst the noise is adapted to obtain signal-to-noise ratio (SNR) at which the subject reaches 50% correct level referred as the Speech Reception Threshold (SRT) level. This new test approach intends to give a more realistic estimate for real-world listening conditions where speech and noise levels vary together [7, 9, 10, 26–28].

Materials and methods

Participants

Eighteen CI users (7 females and 11 males) and eighteen NH listeners (10 females and 8 males) participated in this study. All participants were native Italian speakers. The mean ages for CI and NH participants were 63 years (range

40–81 years, SD = 10.8) and 26 years (range 20–41 years, SD = 5.7) respectively. All CI users were post-lingually deafened, with an average duration of 51.3 months profound deafness before implantation (range 3–360 months, SD = 85.2). The minimum duration of CI experience was 6 months (mean = 46.7 months, SD = 53.0).

NH participants were persons with no otologic history and their hearing thresholds in both ears were ≤ 15 dB HL (mean = 9.8 dB HL, SD = 3.0) at octave frequencies from 125 to 8000 Hz. All CI participants, in either ear, had unaided hearing thresholds above 80 dB HL for frequencies 125–1000 Hz, hence did not show any degree of LF residual hearing that may have interfered with pitch perception outcomes. The mean CI threshold at octave frequencies from 125 to 8000 Hz was 35.8 dB HL (SD = 4.1) and the mean recognition score when listening to bi-syllabic words in quiet was 78.4% (range 50–100%, SD = 15.7). Nine CI participants were Advanced Bionics HiRes users and nine participants were Med-El FS4 users. All but one CI participants were unilaterally implanted. The bilateral listener had a simultaneous implantation with Med-El devices in both ears. Demographic information and audiological outcomes for CI participants are given in Table 1.

Procedure

All testing was performed in a sound-proofed room. Unaided hearing thresholds were assessed presenting a warble tone through an Aurical audiometer and TDH39 headphones whilst aided thresholds were obtained in the sound field. Pitch and speech stimuli were presented via a computer and a preamplifier connected directly to a single loudspeaker positioned 1 m in front of the participants. CI participants were asked to set their sound processor to a comfortable listening level. For bilateral listener, pitch and speech perception performance were assessed ear by ear. A test session including the time for an audiogram, instructions, a practice list for STARR and a training module for WSP-LPF was always completed within 1 h.

The WSP-LPF test

The WSP-LPF test was used to evaluate LF pitch perception skills in a linguistic context. The stimuli were based on pseudo words. The recordings were done with a female phonetician who was asked to speak with a flat intonation. Each word was modeled by a three-syllable sequence. The syllables /mi, ma, mu, ni, na, nu/ were used to generate all three-syllable consonant vowel (CV) sequences. The selection was done following a comparison of phoneme inventories and syllable forms across Italian, Romanian and Dutch. The analysis indicated CV as the choice of syllable type which occurred more frequently in all three languages. Moreover, voiced, sonorant speech sounds were preferred as they were more robust in terms of cross-language variations and allowed stimuli to carry pitch continuously. A pitch accent was imposed on one of the three syllables. This consisted a rise of F0 followed by a reversal to its original value of 200 Hz. Pitch excursions varied between the 200 Hz baseline and a maximum of 408 Hz. Duration and intensity were normalized so that the stimuli varied only in pitch [26, 27]. Figure 1 represents the spectrum of the pseudo word 'manimu' when the first syllable is accented. The WSP test was performed using LPF mode. In this mode, high-frequency cues were removed by low-pass filtering the speech stimuli (MATLAB Filter Function: 300 Hz cutoff frequency, 90 dB attenuation in magnitude over a 50 Hz transition width) and high-pass filtered white noise was added (250 Hz cutoff frequency, 85 dB gain in magnitude over a 50 Hz transition width) [26, 27].

The WSP-LPF test consisted of a four-category identification task. The listener was asked to indicate the syllable (first, second or third) where the accent was perceived, or that there was no detectable pitch accent at all. Test items were presented at 70 dB HL. A training module was used to familiarize the participant with three different stress positions and the procedure. The duration of training never exceeded 10 min. During testing, the software used an adaptive one up-one down procedure to converge to the smallest ΔF0 (Just Noticeable Difference-JND) that was perceived by the listener. The rise in ΔF0 ranged from 0 to 208 Hz. Testing started at ΔF0 of 75 Hz. The ΔF0 changed following the participant's response, i.e., increased for an incorrect response and decreased for a correct one until estimating the 50%-point on a listener's psychometric curve. Presentations at ΔF of 0 Hz were included as internal controls and served to improve the test's reliability by checking if the listener misunderstood the task and to determine the number of reversals required for accurate threshold estimation. Participants received feedback only for their false positive responses where the clinician reminded the participant to indicate only when a pitch accent was reliably detected. If a threshold could not be estimated within 100 trials, the test was ended and JND was set above the maximum ΔF0 value to 220 Hz [26, 27].

The Italian STARR test

The STARR test presented everyday sentences at low, medium, and high levels in a fluctuating speech-shaped noise to estimate an SRT using a roving-level adaptive SNR method [7, 9, 10, 28]. The sentences were selected from the corpus introduced by Cutugno et al. [29]. The test consisted of 10 lists, each containing a total of 15 sentences and allowing 5 presentations at each level (at 50, 65, and 80 dB SPL) within a single list. The task for the participants was

Table 1 Demographic information and audiological findings for CI participants

ID	Age (years)	Gender	CI Ear	Duration of deafness (months)	CI experience (months)	CI model and processor	Sound coding Strategy	Pure tone average (dB HL)	Word recognition score (%)	WSP JND (Hz)	STARR SRT (dB)
P1	77	M	L	36	59	AB 90 K & AURIA	HiRes	34.3	50	220	125
P2	48	M	L	24	27	MEDEL CONCERTO & OPUS2	FS4	33.6	50	220	29.1
P3	52	M	R	3	25	MEDEL CONCERTO & OPUS2	FS4	27.9	90	83	25.6
P4	57	F	L	4	20	MEDEL CONCERTO & OPUS2	FS4	37.9	90	41.5	3.6
P5	60	F	R	120	135	AB 90K & HARMONY	HiRes	32.1	70	67	28.4
P6	58	F	L	6	45	AB 90K & HARMONY	HiRes	38.6	90	6	1.6
P7	61	M	R	12	105	AB 90K & HARMONY	HiRes	36.4	100	24	2.9
P8	40	M	R	6	48	AB 90K & HARMONY	HiRes	25.7	100	24	3.9
P9	59	F	R	360	28	MEDEL CONCERTO & OPUS2	FS4	40	60	220	56.5
P10	54	F	L	9	94	AB 90K & AURIA	HiRes	35.7	90	220	5.4
P11	67	M	L	28	34	AB 90K & HARMONY	HiRes	34.3	80	15	6.6
P12	63	M	R	12	10	MEDEL CONCERTO & OPUS2	FS4	40	80	128.5	15.1
P13	59	M	R	96	15	MEDEL CONCERTO & OPUS2	FS4	40	80	91	13.1
P14	73	M	R	24	15	MEDEL CONCERTO & OPUS2	FS4	40	80	38	24.3
			L	24	15	MEDEL CONCERTO & OPUS2	FS4	38.6	90	55.5	8.3
P15	75	F	L	144	58	AB 90K & HARMONY	HiRes	36.4	80	45	17.3
P16	64	F	L	50	25	AB 90K & HARMONY	HiRes	32.1	80	12	31
P17	81	M	L	12	23	MEDEL CONCERTO & OPUS2	FS4	38.6	80	220	29.9
P18	75	M	L	6	6	MEDEL CONCERTO & OPUS2	FS4	38.6	50	220	94.6

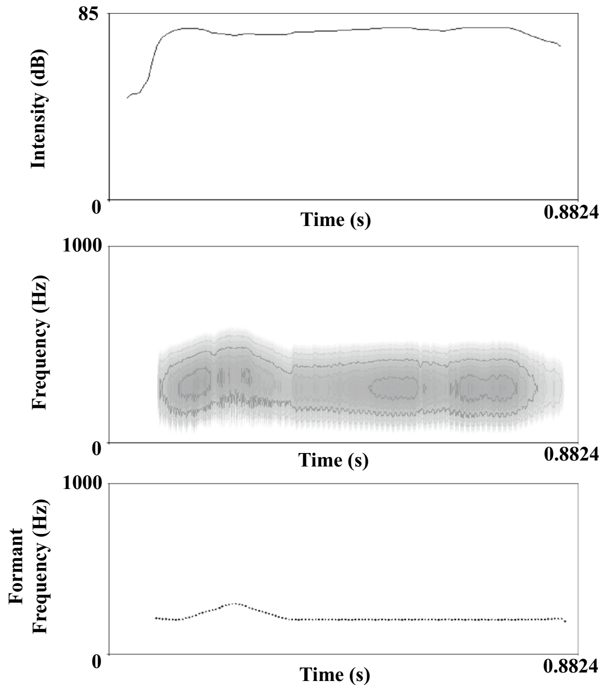

Fig. 1 The spectrum of the pseudo word 'manimu' for WSP-LPF when the first syllable is accented. Duration and intensity are normalized, and the stimuli vary only in pitch. Adapted from the Eargroup, Antwerp-Deurne, Belgium

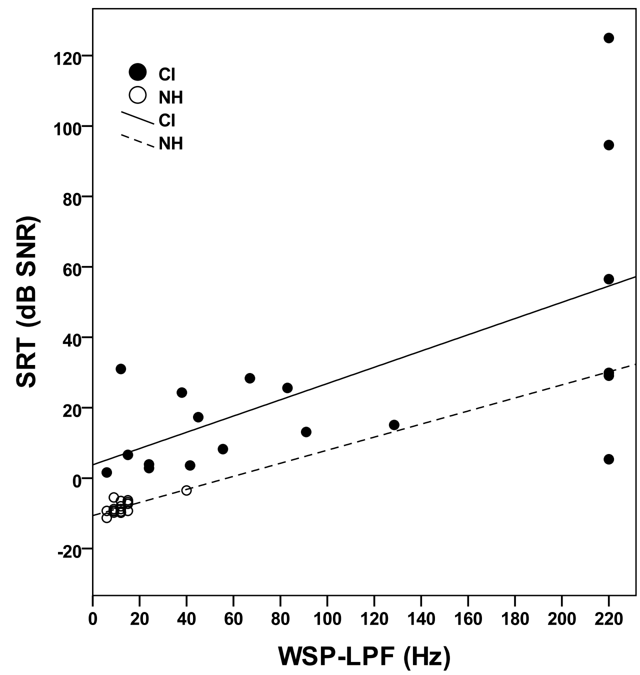

Fig. 2 WSP-LPF outcomes in relation to those obtained from the STARR test for CI and NH listeners

to repeat the sentence as accurately as possible. The participants were administered two test lists following a practice list. The SNR started at + 20 dB and varied adaptively in function of the participant's response. The sentence was considered correct when at least 2 out of 3 key words were correctly recognized, hence, the next SNR became more adverse. If insufficient key words were correctly recognized, then the next SNR was made more favorable. The SNR step size started at 10 dB; dropped to 5 dB after the first reversal of the adaptive track and dropped again to 2.5 dB after a further reversal. The same SNR was used for all three speech presentation levels. The SRT for each test list was computed by averaging the SNRs for the last nine sentences together with the SNR at which a next sentence would have been administered.

Statistical analysis

STARR performance was calculated by averaging scores obtained from two test lists for each participant. Data analysis was performed using Statistical Package for Social Sciences (SPSS) version 19.0 (Chicago, IL, USA). A Shapiro–Wilk test showed that WSP-LPF and STARR outcomes were not normally distributed for both NH and CI groups ($p < 0.001$). Hence, non-parametric statistical tests were used for data analysis. Group differences for WSP-LPF and STARR performance were analyzed using Mann–Whitney *U* test. Relationships between WSP-LPF and STARR scores were tested with Spearman rank-order correlations. The cutoff level for statistical significance was set to 0.05.

Results

NH findings

Median WSP-LPF and STARR scores in NH listeners were 12.0 Hz (range 6.0–40.0 Hz) and -9.1 dB (range − 3.5 to − 11.3 dB) respectively. Figure 2 represents WSP-LPF outcomes in relation to those obtained from the STARR. WSP-LPF and STARR findings showed a significant positive correlation for NH listeners ($r_s = 0.50$, $p = 0.035$).

CI findings

Median WSP-LPF and STARR scores in CI listeners were 67.0 Hz (range 6.0–220.0 Hz) and 17.3 dB (range 1.6–125.0 dB) respectively. CI and NH group performances showed statistically significant differences for both WSP-LPF ($Z = -4.38$, $p < 0.001$) and STARR ($t = -5.27$, $p < 0.001$). As shown in Fig. 2, WSP-LPF and STARR findings showed a significant positive correlation for CI listeners ($r_s = 0.60$, $p = 0.007$).

Discussion

Lexical stress is often expressed by changes in LF pitch or F0 which is the most important acoustic correlate of prosody in a speech signal. CI users usually show great difficulty for understanding speech in noisy listening situations, and this fact has been thought to stem partly from their pitch perception skill, especially in the LF domain [10, 30–32]. Thus, recent years have seen a growing interest in the assessment of LF pitch perception skills in CI listeners using tests mainly based on non-speech tasks such as discrimination of tone complexes [10, 26, 27]. However, the use of linguistically relevant contexts such as lexical stress cued by LF information may better indicate the role of pitch in speech perception. This may happen especially when testing conditions similar to real-world listening where the speech level varies together with the noise, reflecting a listener's capacity to understand speech in a fluctuating background. NH people have the ability to distinguish whether a signal in a fluctuating background is the target speech or the noise. This skill is also known as "listening in the dips" and is thought to derive from phase locking which is a time-based mechanism that locks onto the TFS of the signal [30–32]. CI devices, however, focus at replacing the tonotopic coding or place coding mechanism while they have a limited capacity to convey pitch information and are unable to provide TFS information. Although both place and TFS cues seem to play a role in encoding prosodic cues in speech; temporal coding or more specifically TFS processing and underlying phase locking mechanism are thought to contribute to speech pitch perception especially for resolving the LF information [27, 32, 33].

Non-speech tests based on tone complexes may overestimate listeners' pitch perception performance. This idea is supported by Heeren et al. [26], Schauwers et al. [27] and present findings that show remarkably higher JNDs on the WSP-LPF test in comparison to non-speech discrimination tests based on harmonic or inharmonic pitch glides (Harmonic Intonation-HI and Disharmonic Intonation-DI tests from the ASSE suite) for people with normal hearing and hearing impairment. Like in the WSP-LPF test, the stimuli in the HI/DI tests have an F0 of 200 Hz. In the HI discrimination task, F0 has a sweep together with its three higher harmonics whereas in the DI task only F0 changes. In this sense, the DI test is similar to the WSP-LPF, containing spectral information only below 300 Hz. Hence, the DI and the WSP-LPF tasks seem to address temporal coding of the cochlea and to support the assessment of the availability of TFS cues for speech pitch perception. Abnormally high JNDs on the LPF tasks may indicate the listener's inability to make use of TFS cues or to infer LF pitch information from TFS cues such as people with sensorineural hearing loss or more specifically with cochlear implants [32]. The HI/DI tests have been recently used to investigate the link between TFS sensitivity, LF pitch and speech perception performance in both pediatric and adult CI users, including bimodal listeners who benefit better LF temporal coding provided through the HA. Speech recognition scores in quiet in CI users are shown not to be correlated significantly with HI/DI performance [34], whereas outcome comparisons between electric only and electroacoustic stimulation reveal a significant performance improvement for the DI and word recognition in noise (presented with a fixed SNR of + 10 dB) [35]. A recent study by Dincer D' Alessandro et al. [10] indicates a common performance deterioration for the HI/DI and STARR findings. The use of the STARR test in the study highlights the importance of LF pitch and TFS cues for challenging listening relevant to everyday speech perception where listening in the dips plays an important role. STARR performance shows moderate positive correlations with HI outcomes for CI alone listeners and with those of DI for bimodal listeners, suggesting better ability to encode TFS information with the addition of LF residual hearing and HA use.

Present WSP-LPF and STARR outcomes for both NH and CI listeners are consistent with those reported in the literature [7, 9, 10, 26–28]. Findings show statistically significant group differences, reflecting a big performance gap between the two groups and CI users' great difficulty for LF pitch and challenging speech perception. Statistically significant positive correlations confirm performance changes in parallel for both groups, and the WSP-LPF relationships with the STARR performance for CI listeners tend to be remarkably stronger than those with HI and DI performance reported previously [10] (correlation coefficient 0.60 for the WSP-LPF versus 0.44 for the HI and 0.22 for the DI), supporting the idea that the use of linguistically relevant contexts may better reflect the role of pitch in speech perception, especially when listening to challenging speech with varying levels in a fluctuating background. Such findings might be explained by differences in the processing of speech versus non-speech stimuli as well as by differences in the test difficulty (identification versus discrimination tasks), resulting in higher demands on auditory memory and attention [26, 27].

Although the majority of present CI users show abnormal WSP-LPF and STARR outcomes, surprisingly some seem to be able to perform as well as NH listeners for the WSP-LPF test. However, no CI participant is able to achieve STARR scores within normal range. On the other hand, some CI listeners show good WSP-LPF scores but poor STARR performance or vice versa. The reasons why there is a large variability among CI listeners for both tests and more specifically, why some are able to perform within or near the normal range for the WSP-LPF test remain as important limitations of the present study. One may consider a possible effect of device type, but the present study could not address this issue due to small sample size. Both FS4 and HiRes strategies from the present CI sample are based on channel-specific real-time filters and are designed to provide better temporal cues. The FS4 strategy intends to improve LF TFS perception by adapting the rate of stimulation on four most apical channels to changes in the TFS of the signal in that pulses for these channels are triggered by zero-crossings in the bandpass filter's output. On the other hand, the strategy has a greater extension to the LF range if compared to extended LF filters of HiRes. For FS4 strategy listeners, F0 of 200 Hz is represented by the 2nd apical channel where the frequency band from 198 to 325 Hz would largely keep stimulation within the same channel. Therefore, in Med-El users the stimulation would move to the adjacent electrode and generate a place cue for JNDs higher than 125 Hz. However, in case of HiRes strategy listeners, the frequency band for the most apical electrode varies between 250 to 416 Hz and F0 of 200 Hz is below the pass band of the AB processor. Hence, when F0 sweeps upwards, the processor induces an amplitude change and this fact might have generated an additional cue for achieving lower JNDs in AB users [5].

As mentioned earlier in this paper, Most and Peled [23] have assessed perception of word pattern, intonation and lexical stress in children with bilateral severe to profound sensorineural hearing loss aided either with HAs or CIs. Results indicate perception of stress as the most difficult task for both groups; however, HA users perform significantly better than the CI group in perceiving both stress and intonation. As highlighted by the authors, demographic differences such as late implantation age in their CI sample relative to the HA users who are aided at an early age might have contributed to such findings. On the other hand, HA use is shown to result in better coding of low frequencies and TFS cues [34, 35] while speech perception in noise is known to improve significantly with the addition of an HA in CI users [35–38]. Hence, WSP-LPF and STARR relationships in bimodal listeners may help us to better estimate the role of LF pitch and TFS sensitivity in everyday speech perception as well as to investigate new strategies for bimodal fitting optimization.

Conclusions

NH people have the ability to understand the target speech even in the presence of a remarkable background noise while CI users usually show great difficulty for challenging speech perception. Current cochlear implants have a restricted capacity to convey LF pitch and TFS cues, which are known to contribute to the processing of linguistic information and to be indicative of speech perception, especially when listening in noisy situations. Hence, recent years have seen a growing interest in the evaluation of LF pitch perception skills in CI users. However, the assessment of pitch perception is mainly based on non-speech tasks. The use of lexical stress expressed by changes in LF pitch or F0 may allow a more realistic estimate for the role of pitch in speech perception and may better reflect a listener's capacity to understand speech in a fluctuating background. Present findings confirm this idea and show stronger correlations than those reported previously using non-speech materials. Such findings seem to support that CI listeners' poor speech perception in noise might be strongly associated with their inability for LF pitch perception and TFS processing.

Acknowledgements We would like to thank *Geert DeCeulaer* from *the Eargroup* (*Antwerp-Deurne, Belgium*) for valuable help in providing us the spectrums of the typical stimuli from low-pass-filtered Word Stress Pattern test.

Compliance with ethical standards

Conflict of interest This research was not financially supported. The authors declare that they have no conflict of interest.

Ethical approval Ethical approval was obtained by the Local Ethical Committee. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Written consent was taken from all participants included in the study.

References

1. Geers AE, Sedey AL (2011) Language and verbal reasoning skills in adolescents with 10 or more years of cochlear implant experience. Ear Hear 32:39S–48S. https://doi.org/10.1097/AUD.0b013e3181fa41dc
2. Niparko JK, Tobey EA, Thal DJ, Eisenberg LS, Wang NY, Quittner AL et al (2010) Spoken language development in children following cochlear implantation. JAMA 303:1498–1506. https://doi.org/10.1001/jama.2010.451
3. Nicastri M, Filipo R, Ruoppolo G, Viccaro M, Dincer H, Guerzoni L, Cuda D, Bosco E, Prosperini L, Mancini P (2014) Inferences and metaphoric comprehension in unilaterally implanted children with adequate formal oral language performance. Int J Pediatr Otorhinolaryngol 78:821–827. https://doi.org/10.1016/j.ijporl.2014.02.022
4. Mancini P, Dincer D'Alessandro H, Guerzoni L, Cuda D, Ruoppolo G, Musacchio A, Di Mario A, De Seta E, Bosco E, Nicastri M (2015) Adequate formal language performance in unilateral cochlear implanted children: Is it indicative of complete recovery in all linguistic domains? Insights from referential communication. Int J Pediatr Otorhinolaryngol 79:598–604. https://doi.org/10.1016/j.ijporl.2015.02.003
5. Wilson BS, Dorman MF (2008) Cochlear implants: current designs and future possibilities. J Rehabil Res Dev 45:695–730

6. Lenarz M, Sönmez H, Joseph G, Büchner A, Lenarz T (2012) Long-term performance of cochlear implants in postlingually deafened adults. Otolaryngol Head Neck Surg 147:112–118. https://doi.org/10.1177/0194599812438041
7. Dincer D'Alessandro H, Boyle PJ, Ballantyne D, De Vincentiis M, Mancini P (2018) The role of speech rate for Italian-speaking cochlear implant users: insights for everyday speech perception. Int J Audiol 57:851–857. https://doi.org/10.1080/14992027.2018.1498139.
8. Zeng FG (2002) Temporal pitch in electric hearing. Hear Res 174:101–106. https://doi.org/10.1016/S0378-5955(02)00644-5
9. Boyle PJ, Nunn TB, O'Connor AF, Moore BCJ (2013) STARR: A speech test for evaluation of the effectiveness of auditory prostheses under realistic conditions. Ear Hear 34:203–212. https://doi.org/10.1097/AUD.0b013e31826a8e82
10. Dincer D'Alessandro H, Ballantyne D, Boyle PJ, De Seta E, DeVincentiis M, Mancini P (2018) Temporal fine structure processing, pitch, and speech perception in adult cochlear implant recipients. Ear Hear 39:679–686. https://doi.org/10.1097/AUD.0000000000000525
11. Fry DB (1955) Duration and intensity as physical correlates of linguistic stress. J Acoust Soc Am 27:765–768. https://doi.org/10.1121/1.1908022
12. Chakraborty R, Goffman L (2011) Production of lexical stress in non-native speakers of American English: kinematic correlates of stress and transfer. J Speech Lang Hear Res 54:821–835. https://doi.org/10.1044/1092-4388(2010/09-0018)
13. Colombo L (1992) Lexical stress effect and its interaction with frequency in word pronunciation. J Exp Psychol Hum Percept Perform 18:987–1003. https://doi.org/10.1037/0096-1523.18.4.987
14. Colombo L, Zevin J (2009) Stress priming in reading and the selective modulation of lexical and sub-lexical pathways. PLoS ONE 4:e7219. https://doi.org/10.1371/journal.pone.0007219
15. Thornton AM, Iacobini C, Burani C (1997) BDVBD. Una base di dati sul vocabolario di base della lingua italiana [BDVDB: a database for the Italian basic dictionary]. Bulzoni, Roma
16. Colombo L, Sulpizio S (2015) When orthography is not enough: the effect of lexical stress in lexical decision. Mem Cognit 43:811–824. https://doi.org/10.3758/s13421-015-0506-6
17. Spinelli G, Sulpizio S, Primativo S, Burani C (2016) Stress in context: morpho-syntactic properties affect lexical stress assignment in reading aloud. Front Psychol 7:942. https://doi.org/10.3389/fpsyg.2016.00942
18. Rahmani H, Rietveld T, Gussenhoven C (2015) Stress "Deafness" reveals absence of lexical marking of stress or tone in the adult grammar. PLoS ONE 10:e0143968. https://doi.org/10.1371/journal.pone.0143968
19. Mampe B, Friederici AD, Christophe A, Wermke K (2009) Newborns' cry melody is shaped by their native language. Curr Biol 19:1994–1997. https://doi.org/10.1016/j.cub.2009.09.064
20. Weber C, Hahne A, Friedrich M, Friederici AD (2004) Discrimination of word stress in early infant perception: electrophysiological evidence. Cogn Brain Res 18:149–161. https://doi.org/10.1016/j.cogbrainres.2003.10.001
21. Friederici AD, Friedrich M, Christophe A (2007) Brain responses in 4 month-old infants are already language specific. Curr Biol 17:1208–1211. https://doi.org/10.1016/j.cub.2007.06.011
22. Vavatzanidis NK, Mürbe D, Friederici AD, Hahne A (2016) The Perception of stress pattern in young cochlear implanted children: an EEG study. Front Neurosci 10:68. https://doi.org/10.3389/fnins.2016.00068
23. Most T, Peled M (2007) Perception of suprasegmental features of speech by children with cochlear implants and children with hearing aids. J Deaf Stud Deaf Educ 12:350–361. https://doi.org/10.1093/deafed/enm012
24. Jackson P, Kelly D (1986) The relationship between word and stress pattern recognition ability and hearing level in hearing impaired young adults. Volta Rev 88:279–288
25. Rubin-Spitz J, McGarr N (1986) Perception of stress contrasts by the hearing impaired. J Acoust Soc Am 79:10S. https://doi.org/10.1121/1.2023065
26. Heeren W, Coene M, Vaerenberg B, Avram A, Cardinaletti A, Del Bo L, Pascu A, Volpato F, Govaerts PJ (2012) Development of A§E test battery for assessment of pitch perception in speech. Cochlear Implants Intl 13:206–219. https://doi.org/10.1179/1754762811Y.0000000035
27. Schauwers K, Coene M, Heeren W, Del Bo L, Pascu A, Vaerenberg B, Govaerts PJ (2012) Perception of pitch changes in hearing impaired adults with aided and unaided hearing loss. J Hear Sci 2:OA25–OA34.
28. Dincer D'Alessandro H, Ballantyne D, De Seta E, Musacchio A, Mancini P (2016) Adaptation of the STARR test for adult Italian population: a speech test for a realistic estimate in real-life listening conditions. Int J Audiol 55:262-267. https://doi.org/10.3109/14992027.2015.1124296
29. Cutugno F, Prosser S, Turrini M (2000) Audiometria vocale, vol. IV. Resound Italia, Padova
30. Baer T, Moore BC (1993) Effects of spectral smearing on the intelligibility of sentences in noise. J Acoust Soc Am 94:1229–1241. https://doi.org/10.1121/1.408176
31. Baer T, Moore BC (1994) Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. J Acoust Soc Am 95: 2277-2280. https://doi.org/10.1121/1.408640
32. Moore BJC (2008) The role of temporal fine structure processing in pitch perception, masking and speech perception for normal-hearing and hearing-impaired people. J Assoc Res Otolaryngol 9:399–406. https://doi.org/10.1007/s10162-008-0143-x
33. Rosen S (1992) Temporal information in speech: acoustic, auditory and linguistic aspects. Philos Trans R Soc Lond B Biol Sci 336:367–373. https://doi.org/10.1098/rstb.1992.0070
34. Dincer D'Alessandro H, Filipo R, Ballantyne D, Attanasio G, Bosco E, Nicastri M, Mancini P (2015) Low-frequency pitch perception in children with cochlear implants in comparison to normal hearing peers. Eur Arch Otorhinolaryngol 272:3115–3122. https://doi.org/10.1007/s00405-014-3313-y
35. Vaerenberg B, Péan V, Lesbros G, De Ceulaer G, Schauwers K, Daemers K, Gnansia D, Govaerts PJ (2013) Combined electric and acoustic performance with Zebra® speech processor: speech reception, place and temporal coding evaluation. Cochlear Implants Intl 14:150–157. https://doi.org/10.1179/1754762812Y.0000000008
36. Dorman MF, Gifford RH, Spahr AJ, McKarns SA (2008) The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. Audiol Neurootol 13:105–112. https://doi.org/10.1159/000111782
37. Dincer D'Alessandro H, Sennaroğlu G, Yücel E, Belgin E, Mancini P (2015) Binaural squelch and head shadow effects in children with unilateral cochlear implants and contralateral hearing aids. Acta Otorhinolaryngol Ital 35:343–349. https://doi.org/10.14639/0392-100X-497
38. Mok M, Grayden D, Dowell RC, Lawrence D (2006) Speech Perception for Adults Who Use Hearing Aids in Conjunction with Cochlear Implants in opposite Ears. J Speech Lang Hear Res 49:338–351. https://doi.org/10.1044/1092-4388(2006/027)