Beyond existing prosodic dichotomies: Perception of aesthetic prosodic properties of speech and music in a right-hemisphere stroke patient

Ariadne Loutrari[1] & Marjorie Perlman Lorch[2]

[1]*Aston University, Birmingham*
[2]*University of London*
a.loutrari@aston.ac.uk, *m.lorch@bbk.ac.uk*

Abstract
Speech and music processing impairments have been studied in parallel through the investigation of atomistic features such as pitch and duration and gestalt aspects of emotion. The present study explores another holistic dimension of speech and music prosody here termed 'expressiveness.' Novel tasks were designed to investigate whether such hitherto unexplored prosodic aspects of speech and music display processing differences. Five perceptual judgement tasks were employed, two of which involved music and speech stimuli manipulations of 'expressiveness'. Effort was made to maintain more of their natural acoustic complexity, avoiding manipulations which derive music-like stimuli from speech tokens to artificially match items. We examined the performance of IB, an individual who had a right temporo-parietal lesion with frontal extension and compared his performance with 24 neurotypical controls on these prosodic judgements. IB's performance was found to be comparable to that of neurotypical controls on a perceptual discrimination task of 'expressive music prosody', outperforming one-third of them, whereas he displayed severely impaired performance on 'expressive speech prosody'. These results suggest that some prosodic elements may be perceived differently across the domains of language and music. Based on other inter-task comparisons, it is also proposed that the interplay among prosodic features such as loudness and duration might lead to different holistic processing between emotional prosody and 'expressive' prosodic qualities in the speech domain. Inevitably, the current work only provides preliminary evidence and future research with more patients sharing a lesion profile similar to that of IB is warranted.

Keywords: prosody, emotion, music, auditory processing, right hemisphere damage

1 Introduction

The investigation of prosody in non-tonal languages has primarily focused on linguistic and emotional functions. Linguistic prosody communicates lexical, syntactic, and pragmatic information through word stress, sentence focus, and speech segmentation cues (Cutler et al. 1997), while emotional prosody conveys information on the speaker's emotive disposition (Pell 2006). Prosodic processing has recently been investigated with both functional techniques and through the investigation of individuals with acquired focal lesions with regard to features such as pitch, loudness and duration as well as more holistic prosodic qualities using emotional labels.

Functional lateralisation has been systematically investigated in relation to speech prosody processing (Penhune et al. 1996), with a left hemisphere advantage found for processing segmental information and a right hemisphere advantage for tonal processing (Zatorre et al. 2007) that have been attributed to timing differences in the acoustic signal (Poeppel 2003). Left fronto-temporal and right fronto-temporal networks have been proposed to pertain to syntactic-semantic and phrasal prosody processing respectively, with the involvement of the left hemisphere getting more dynamic the more linguistic a stimulus becomes (Friederici & Alter 2004). Although the findings from lesion studies are mixed (Kotz et al. 2003), left hemisphere lesions have been found to be more often associated with linguistic prosody impairments and right hemisphere lesions more often result in emotional prosody impairments (Witteman et al. 2011). Neuroimaging studies have shown some overlap of these prosodic categories in the auditory cortex but different activation of bilateral areas of the inferior frontal gyrus (Belyk & Brown 2014). Strong lateralisation claims are in general hard to make, as damage to the corpus callosum can also compromise prosodic processing (Friederici et al. 2007; Klouda et al. 1988; Sammler et al. 2010). The current picture of prosodic processing pathways has been based on the results of tasks designed in order to explore the prosodic categories already well-established in the literature. We propose that a fuller picture of the perceptual appreciation of the acoustic qualities of speech may require additional aspects of prosody to be included in such accounts.

Although individual differences in the oral renditions of passages have been acknowledged as a crucial aspect of dramatic arts (Jakobson 1958), there has been limited consideration of such aspects in the domain of linguistics. An inexperienced actor will perform differently from an experienced and talented actor that can employ rich prosodic manipulations in the oral production of a given text, and their renditions of the same passage can also vary from performance to performance (Loutrari & Lorch 2018). These prosodic variations differentiate a speech utterance from counterpart utterances that evoke less rich aesthetic appreciation of expressiveness. Acting performance constitutes an example of sound streams containing an additional prosodic layer that complements linguistic well-formedness and emotion.

Consideration of such parallel aspects in music performance may provide a window for considering the relation between different aspects of prosodic processing by comparing music and speech perception. Along with emotional expression and structural well-formedness, artistic individuality is of paramount importance in music performance (see Gingras 2014 for an introduction). If this perceptual dimension of expressiveness is sought contrastively in both music and speech, at one end of the spectrum, a 'deadpan' version of an utterance or melody can be considered to contain a minimal amount of expressive prosodic cues resulting in an aesthetically 'neutral' result. This version would have little prosodic variation in terms of loudness, pause placement, and articulation and no deviation from temporal regularity. Judgements of such a dimension are not dependent on purely deterministic rules, and pertain more to an optional listener preference rather than an absolute judgement of correctness (Palmer & Hutchins 2006). In a previous investigation of an individual with congenital amusia (BZ) (initially studied by Paraskevopoulous et al. 2010) using the same tasks as in the present study, we showed that perception of expressive prosody can be preserved in the presence of compromised perception of widely explored acoustic subcomponents such as pitch contour and pitch intervals (Loutrari & Lorch 2017). BZ's pattern of performance suggests that an individual can manifest major difficulties with standardised music cognition assessment tests, while, at the same time, maintaining the ability to differentiate short music passages based on sensitivity to aesthetic prosodic elements. Further investigation of acoustic perceptual abilities with regard to prosody in a gestalt framework seems warranted in order to obtain a fuller picture of the relationship between speech and music processing.

Approaching the exploration of speech prosody through the consideration of gestalt aspects of stimuli is supported by previous research suggesting that processing of acoustic cues in isolation might not account for the actual processing of prosodic stimuli (Audibert et al. 2005; Menninghaus et al. 2015; Wiethoff et al. 2008).

Moreover, previous research suggests that despite common acoustic correlates, acoustic features present in linguistic and emotional prosody might be encoded through distinct mechanisms, depending on whether they convey linguistic or emotional information (Belyk & Brown 2014). Such findings indicate that it may not be the acoustic cues *per se* that determine how information in the acoustic signal will be processed. For instance, stress (increased loudness and/or duration) can be used to disambiguate a syntactic structure but can be also employed to signal emotional colouring. This, to extend Belyk & Brown's (2014) argument further, does not necessarily entail the engagement of identical processing mechanisms in these two types of instances.

There have been a number of studies which have investigated the perception of features of speech and music in parallel. The acoustic speech stimuli employed by Patel and colleagues (1998) were designed to explore the processing of statement/question discriminations, emphasis, and grouping in parallel with music-like stimuli derived from speech. These tasks, or parts of them, have been used in several studies with individuals manifesting cognitive deficits (e.g., Ayotte et al. 2002; Hutchins et al. 2010; Nicholson et al. 2003; Patel et al. 2005). However, we suggest that using the prosodic structure of speech in order to study music processing separates the prosodic features in question from the context in which they naturally occur.

The main objective of the present study was to explore the perception of aesthetic aspects of speech and music prosody using ecologically valid stimuli, shifting away from music-like speech, in an attempt to examine whether previously unexplored holistic prosodic layers can be perceived in the presence of extensive right hemisphere damage. The investigation of the dimension of speech and music expressiveness aimed at revealing possible patterns of preserved and compromised processing across domains, broadening the debate on shared and distinct processing in language and music. Moreover, the inclusion of emotional prosody tasks aimed at exploring the relationship between previously investigated aspects of emotion and the aesthetic judgement of expressiveness. The inclusion of such tasks addressed the additional objective of comparing similar acoustic cues that, nevertheless, belong to distinct prosodic contexts.

2 Materials and method

2.1 Participants

Right hemisphere brain damaged individual

IB is a 51-year-old Greek monolingual right-handed man with 12 years of education and three years of experience as an amateur drummer in a musical band. He was admitted to Renaissance Rehabilitation Centre, Thessaloniki, Greece, after suffering a stroke resulting in left hemiplegia, having received primary care in another clinic for approximately three weeks. He suffered from an ischemic infarction of the right middle cerebral artery leading to extensive damage in the right temporo-parietal lobe with some frontal extension. IB was reported to have attained all normal developmental milestones in childhood and had no history of mental illness, hearing, or visual difficulties. Upon admission to the Rehabilitation Centre three weeks post-onset, he seemed aware of his condition and did not manifest speech or language difficulties. The rehabilitation therapy he received during his stay only addressed his left-sided sensory-motor impairments.

Several CT scans were obtained between one and nine months post-stroke. An early scan showed an extensive ischemic infarction of the right middle cerebral artery territory and mild pressure in the right lateral ventricle. Subsequent scans indicated a reduced extent of the damaged areas and resolution of the ventricular pressure. The blood flow of the main arterial stems was found to be normal. The scan nine months post-stroke (Figure 1) depicted a relatively large ischemic infarction of the right middle cerebral territory involving cortical structures of the temporo-parietal lobe and subcortical structures with some attraction of the right lateral ventricle. There was some extension rostrally to the frontal lobe. The brain stem and the cerebellum were normally depicted.

On admission to the Rehabilitation Centre, IB’s general neuropsychological functioning was assessed with a mini mental status examination comprised of orientation questions, word registration (repeating words after a short pause), attention and calculation, recall, repetition, word naming, three stage commands, shape copying, and performance of a written command. IB scored 26 out of 30 on the test and when this was repeated a month later (approximately two months post-onset of illness), his score improved to 29 out of 30.

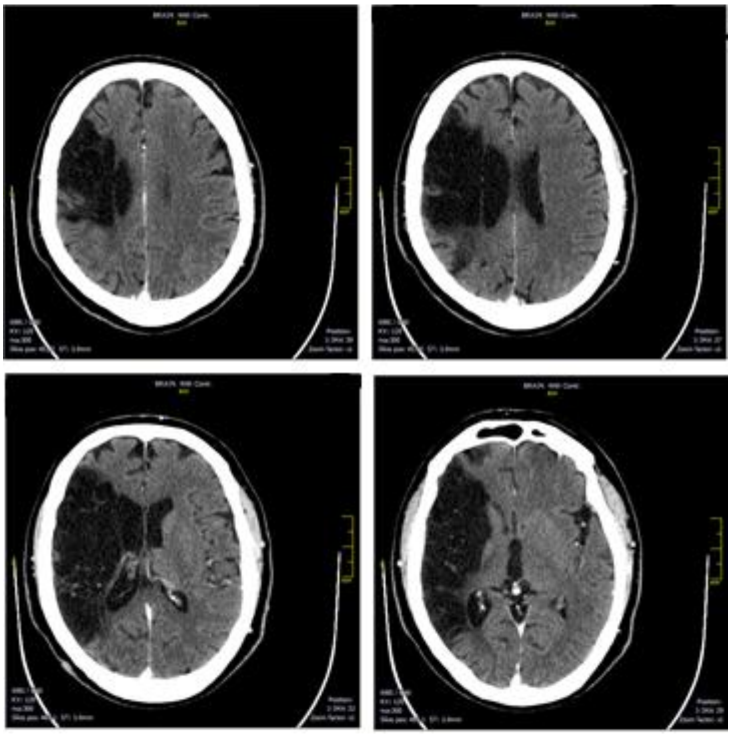

Figure 1. Transverse CT scan obtained nine months post-stroke. The right side of the brain is shown on the left side of the scan. The scans proceed inferiorly, starting from the image appearing in the left upper left-hand corner. The scans show extensive damage in the right temporo-parietal lobe with extension into the frontal lobe in IB.

2.2 Control participants

Twenty-four Greek monolingual individuals served as controls in the study. Their age, education (12 years), and handedness profile was matched with IB’s profile. As previous research has indicated that music experience can have an effect on brain plasticity (Pantev et al. 1998; Zatorre et al. 2007) and can affect processing of speech stimuli (Besson et al. 2007; Marques et al. 2007), music experience has been also taken into account. In light of the fact that IB had approximately three years of music experience, roughly half (n=14) of the neurotypical controls were matched to IB on this variable while the remainder did not have any musical training or performance experience. Although the comparisons of musicians and non-musicians in the literature do not necessarily apply to the limited amount of music experience in the case of IB and controls in this study, music experience was still considered as a variable in order to avoid confounding findings. Summary characteristics of controls are presented in Table 1.

Variables	Control Participants
Total Number (N)	24
No music experience (N)	10
3 years music experience (N)	14
Mean Age in years	50.7
Age SD	5.78
Age range	45-60

Table 1. Summary characteristic of control participants

2.3 Procedure

The presentation of each stimulus or pair of stimuli was preceded by a warning tone. The stimuli of every pair were separated by one second. Practice items were given for each task during testing sessions. The main tasks were administered using Windows Media Player and stimuli were presented binaurally through earphones. An answer sheet was provided for participants to tick the correct answers.

For the present investigation, the majority of testing sessions with IB took place in a quiet room in Renaissance Rehabilitation Centre 11 months post-onset of illness, under the supervision of Dr Harriet Proios, PhD CCC-SLP, Head of the Speech Therapy Department at the Renaissance Rehabilitation Centre and Assistant Professor of Neurocognitive Disorders and Rehabilitation at the Department of Educational and Social Policy, University of Macedonia, Thessaloniki, Greece and the remainder were carried out at his home after discharge. The duration of individual testing sessions was limited to prevent IB from becoming fatigued. Control participants were tested in a quiet room in their house.

2.4 Ethics

Ethics approval was gained from the Ethics Committee of Birkbeck, University of London. The study was considered to entail minimal risk for participants. Consent was obtained from all participants and confidentiality was protected. All participants were treated in accordance with the Declaration of Helsinki. In addition, the Committee of Renaissance Rehabilitation Centre also granted ethical approval for this research project. They granted the researchers' access to IB who was a patient at the Centre. The testing protocol and the consent forms were approved by their research board, the Scientific Director, and the President of the Centre.

2.5 Materials

The stimuli of the study were generated using Digital Performer 8. They were recorded in WAV format, 44.1 kHz, 16 bit in a quiet room. Three of the tasks included 32 pairs of stimuli and two of them required identification of individual stimuli and only contained 32 stimuli in total. Digital Performer 8, Praat software (http://www.fon.hum.uva.nl/praat/), and Audacity software (http://audacity.sourceforge.net/) were used to apply specific manipulations to some stimuli. All speech materials were in Greek and naturalistic manipulations were made by a male Greek native speaker with experience in singing performance. Music stimuli belonged to various styles or European Romantic and late-Romantic music but also native Greek composers of the 20th century. For representative spectrograms for all types of stimuli and more details on their acoustic realisation, please see Loutrari & Lorch (2017) where the methods were originally published.

2.5.1 Baseline auditory judgement tasks

IB was initially assessed on a series of baseline tasks aiming to detect his perceptual sensitivity to basic auditory stimuli and ascertain his ability to be assessed on more complex experimental tasks requiring auditory judgements. An additional part of the rationale behind the use of these tasks was to provide a picture of the relationship between IB's perception of 'low-level' and 'high-level' acoustic stimuli.

These baseline tasks consisted of single tones and vowels and required same/different discrimination judgments. Three conditions varying the pitch, length, and loudness of these stimuli were created using Audacity software. In the music tests, pitch stimuli aiming to elicit a 'different' judgment had a four semitone difference. Length and loudness stimuli displayed a difference devised to be easily perceived by neurotypical listeners. More specifically, the length stimuli were either identical across pairs or had a difference of one second (i.e., the shorter stimuli lasted for one second and the longer for two). The loudness baseline task included stimuli of default loudness used in previous tasks and louder ones, using the command 'effect: amplify' in Audacity software. The analogous speech tests included Greek vowels (a, e, i, o, u) that had undergone the same manipulations.

IB scored well on these tasks, although he did not perform at ceiling; 8/10 on both the speech pitch and the music pitch task; 9/10 on the speech length task; and

8/10 on the music length task. He also performed well on the loudness condition where he scored 8/10 and 7/10 on the speech and music tasks respectively. These scores are also well above chance, although not at ceiling. The participant's basic appreciation of low-level acoustic features of pitch, loudness and length appeared to be sufficient to proceed to more complex judgement tasks.

2.5.2. Judgement tasks of speech and music prosody

i) Speech prosody detection task. This task was used to assess the ability to detect speech in a delexicalised context. It included 64 stimuli that either displayed prosodic inflection or lacked variation in terms of pitch, loudness, pausing, and duration. They had an average duration of 7.93 second and were spoken by a male native speaker of Greek in two conditions: natural prosodic inflection and minimum possible variation of prosodic features.

For the prosodically marked stimuli, the speaker was instructed to employ pitch, loudness, and duration variation and also to pause between tone units. For the prosodically unmarked ones, the speaker had to imitate synthetic speech, limiting, therefore, any variation of pitch, loudness, and duration. This principle aimed to make some stimuli more and other less aesthetically interesting for the listener.

Praat was used to remove lexical information from all stimuli using low-pass filtering. Low-pass filtering removes frequencies above a cutoff frequency in a sound spectrum, resulting in speech that is semantically unintelligible but intact in terms of timing and rhythm (Baer et al. 2002). Stimuli intending to evoke a judgement of being non-speech were the low-pass filtered versions of the stimuli that imitated synthetic speech. Those that intended to evoke a 'speech' judgement were the low-pass filtered versions of the parent stimuli bearing prosodic variation. Participants were blind to the fact that all stimuli were originally derived from speech.

There were four possible responses to the items in this task: the first stimulus is speech, the second stimulus is speech, both stimuli are speech, neither of the stimuli is speech.

ii) Expressive music prosody task. This task aimed to investigate sensitivity to musical prosody and, in combination with the task presented next, to evaluate whether the combination of prosodic features was perceived similarly across domains. It included 32 pairs of stimuli. The melodies of this tasks were derived from a pool of classical, romantic, late romantic, 20th century, and some novel melodies composed by the first author. The stimuli were all monophonic adaptations of the original melodies. The melodies were executed on a digital piano with weighted keys and recorded using Digital Performer 8. Their average duration was 8.48 seconds. Tune and meter was identical across stimuli. There were two groups of stimuli, hereafter, 'expressive' and 'non-expressive'.

The 'expressive' group displayed deviation from temporal regularity, variation in dynamics, connected and smooth transitions between some notes (legato articulation), and grouping of shorter patterns within a melody. The 'non-expressive' group included melodies with temporal regularity, equal loudness across notes, and heavily accented notes (marcato articulation). Using Digital Performer 8, the exact value of every note and rest was assigned and loudness was kept identical across a given melody. The 'marcato' articulation was also technically generated on Digital Performer 8.

Half of the pairs included identical stimuli and half of them consisted of stimuli belonging to different groups. Participants were asked to discriminate between melodies that were played by a human and melodies played by a machine. The response mode of this task required same/different discrimination judgements.

iii) Expressive speech prosody task. This task, designed in parallel with the expressive music prosody task, intended to assess the perception of prosodic inflection in speech in the presence of lexical variation. It included 32 pairs of stimuli. The stimuli of this task were the same as those in the speech prosody detection task before the applying the low-pass filtering manipulation. This task was presented after its low-pass filtered version employed in the speech detection task, as initial piloting of all tasks with other neurotypical participants revealed that when the expressive speech prosody task was presented first, it primed recognition of delexicalised stimuli in the expressive speech prosody task. The task required same-different discrimination judgements. Participants were aware of the fact that utterances across all pairs would be linguistically identical.

iv) Emotional speech prosody. This task was designed to assess emotional prosody perception in semantically neutral utterances and detect possible links between 'expressive' prosody and emotional prosody in speech. It only included 32 stimuli in total. As only two emotions (‘happiness’ and ‘sadness’) were included, this task is not evaluating emotional prosody perception as such but rather seeks connections among prosodic aspects of speech. The stimuli of this task were spoken again by a Greek male native speaker in the above emotional tones. In the ‘happiness’ condition, the speaker produced faster utterances and with higher pitch. His speaking rate in the sad condition sounded slower, his pitch lower, and, in some cases, the voice was softer. For this task, every question of the answer sheet corresponded to a single stimulus. Participants had to identify the intended emotion instead of comparing stimuli. Participants were instructed to only attend to the tone of the speaker and not its semantically neutral content.

v) Emotional music prosody. The purpose of this task was to explore the perception of emotional character in music and provide a parallel assessment of this ability in comparison to emotional prosody in speech and expressive prosody in both domains. It was in total comprised of 32 novel melodies of an average duration of 5.84 seconds. These were composed and played by the first author on a digital piano with weighted keys and recorded on Digital Performer 8. Melodies intending to evoke identification as ‘happy’ were composed in major mode, had a fast tempo, high loudness, and often quick staccato motifs. Those intending to evoke identification as ‘sad’ were comprised in minor mode, had a slower tempo, lower loudness, and the articulation between notes was usually legato. Similar to the previous task, the listener had to identify the intended emotion.

The number of stimuli, the response mode, and the guessing probability for each task are presented in Table 2.

Tasks	Number of stimuli	Response mode	Guessing probability
Speech prosody detection	64	Four answers, forced choice	25%
Expressive music prosody	64	Same-different judgement	50%
Expressive speech prosody	64	Same-different judgement	50%
Emotional speech prosody	32	Identification (two choices)	50%
Emotional music prosody	32	Identification (two choices)	50%

Table 2. Variables for number of stimuli, response mode, and guessing probability of the tasks used in the study.

3 Analysis and results

Results from our five tasks were analysed with independent t-tests (SPSS version 22 for Windows) and Crawford & Garthwaite's (2002) modified *t*-test was used to compare IB's scores with our control sample.

Initial statistical analysis was conducted on the performance of the neurotypical controls in order to determine if the subject variable for musical experience was statistically significant. Independent t-tests were carried out on the performances of the two control groups: Musically naïve group n= 10 and musically experienced group n= 14 on the five tasks. No statistically significant difference was found between the two groups on any of the tasks: 'Speech prosody detection' t(22)= 1.3, p= .18, 'Expressive music prosody' t(22)= -.17, p= .87, 'Expressive speech prosody' t(22)= -.09, p= .92, 'Emotional speech prosody' t(22)= -.87, p= .39, 'Emotional music prosody' t(22)=-.05, p=.95. More detailed results, including the Mean and S.D. of controls for each task are presented in Table 3.

Given that the two groups could not be demonstrated to differ in performance with regard to the variable of musical experience, their results were collapsed. In the results reported below the control participants will be presented as one group n= 24.

a) Speech prosody detection. IB answered correctly on only 11 out of 32 items, scoring only slightly above chance. It is notable that his responses favoured the answer 'both sounds resemble speech' for almost every pair of stimuli. He was not able to distinguish repetitive monotonous sequences from others with pitch, duration, and loudness variation. His performance stands in sharp contrast with the near ceiling performance of the controls (mean score: 28.83/32, min-max: 19-32, SD= 4.12). Crawford & Garthwaite (2002) showed a statistically significant difference in performance between IB and controls (t= -4.240, p= 0.000). IB's performance relative to controls' on this task is displayed in Figure 2.

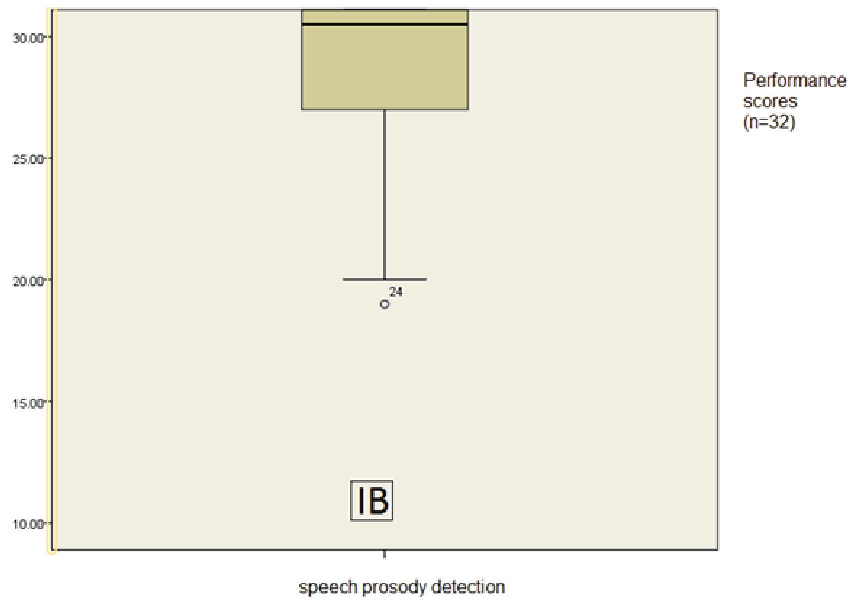

Figure 2. Whisker plot displaying IB's performance relative to controls on the speech prosody detection task. The horizontal line in the box shows the median. There is one outlier represented by an open circle indicating that participant number 24 fell outside the range of the other controls. The chance level for this task is a score of 8/32.

b) Expressive music prosody. On this task, IB performed only slightly below the mean score for the control group. He answered correctly on 23 out of 32 questions, outperforming a number of controls (mean score: 24.16/32, min-max: 17-29, SD= 3.15). Although the performance by the controls were not at ceiling they were all above chance. According to Crawford & Garthwaite (2002) modified *t*-tests, no difference in performance was identified between IB and controls (t= -0.361, p= 0.722).

c) Expressive speech prosody. In contrast to the expressive music prosody task, IB's performance on this task was poor. He scored at chance level, responding correctly on 16 out of 32 questions. In this task, he favoured the answer 'both stimuli are the same', sometimes emphatically uttering 'they sound exactly the same'. Controls' high scores (mean: 30.16/32, min-max: 25-32, SD= 1.49) suggest that the present task was very easy for neurotypical participants. However, it seems that the lexical information of the task (the same stimuli were used in a delexicalised form in task speech prosody detection task) did not have a facilitatory effect on IB's performance. The Crawford & Garthwaite (2002) test showed again statistically significant difference between IB's score and that of controls (t= -9.311, p= 0.000).

Figure 3 compares his performance on the expressive music prosody task and the expressive speech prosody task.

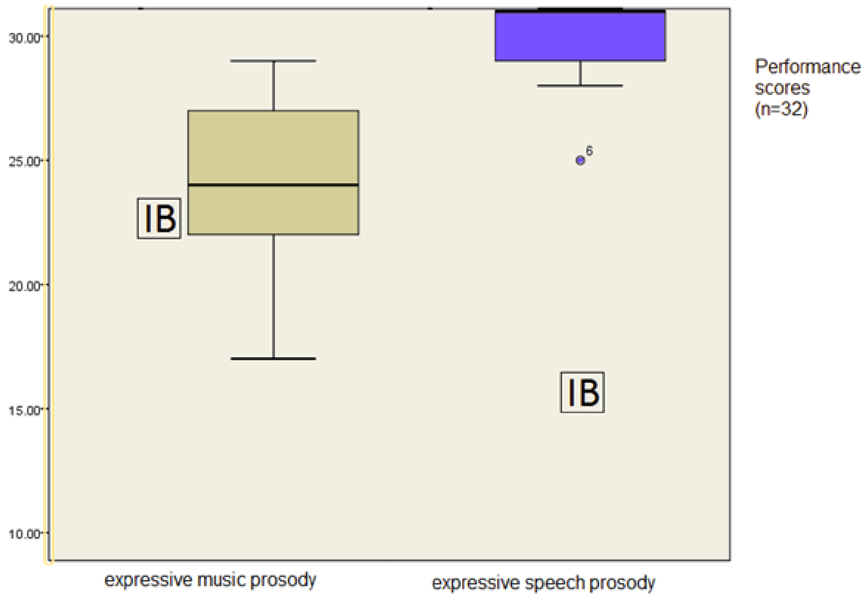

Figure 3. Whisker plot displaying the performance of IB and controls on expressive music prosody and expressive speech prosody. The horizontal lines in the boxes show the median. In the expressive speech prosody box, the median line almost overlaps with the upper line of the box and an outlier (participant number 6) is represented by a small circle. Chance level for both tasks displayed here is a score of 16/32.

*d) Emotional speech prosody***.** IB successfully judged the difference between happy and sad stimuli on 29 out of 32 of the items on this task. The controls' mean score was at ceiling: 31.79/32 (min-max: 30-32, SD= 0.50). It must be noted that, although his performance on this task is quite close to that of the controls, the results from Crawford & Garthwaite's (2002) test suggest that there is a small but statistically significant difference between IB's performance and that of the controls (t= -5.467, p=0.000). This is likely due to the fact that the standard deviation in the controls' scores was very small. Despite the statistically significant difference in comparison to controls, IB's performance remains very high with only three incorrect answers out of 32 questions and thus can still be considered as unimpaired.

e) Emotional music prosody. IB scored 27 out of 32 on this task. The controls' mean was 30.37/32 (min-max: 27-32, SD= 1.88).

IB's performance relative to controls comparing this task and the emotional speech prosody task are displayed in Figure 4. No statistical difference was found between his performance and that of controls (t= -1.756, p= 0.092).

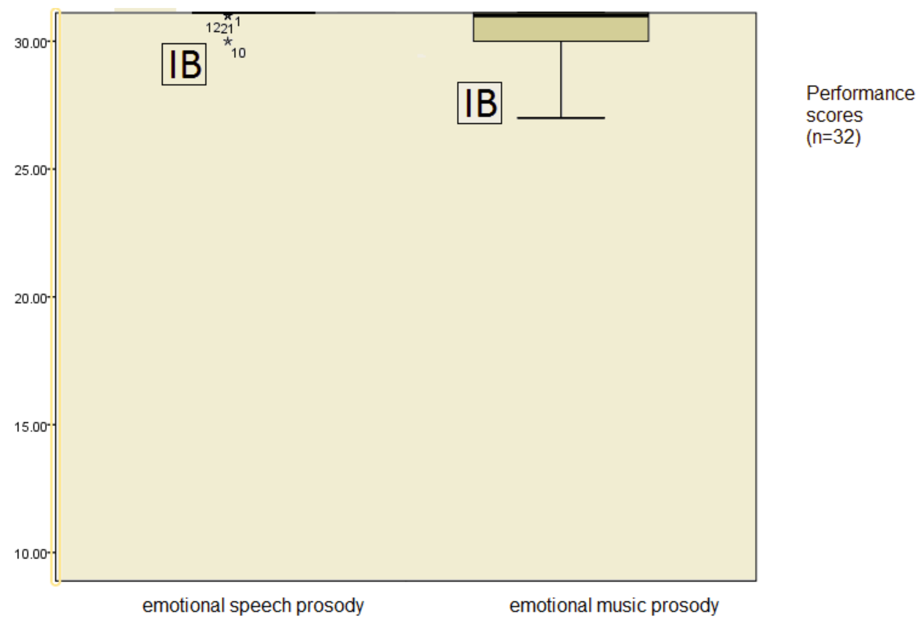

Figure 4. IB's performance relative to controls on the emotional speech prosody task and the emotional music prosody task. Dark lines indicate the median for each task. In the emotional speech prosody task, participants marked with a star (1,12,21) scored 31/32 and participant 10 scored 30/32. These were the only participants that did not perform at ceiling.

Using Crawford & Garthwaite (2002) we ran an additional test for the dissociation in IB's performance on our main tasks, namely the expressive music prosody and the expressive speech prosody. The probability that the absolute difference for a member of the control population would be greater than that of IB (akin to a frequentist two-tailed significance test) is p= 0.000. IB's performance, therefore, fulfilled the criteria for a dissociation on the given tasks.

IB's performance on these tasks indicates preserved abilities in the musical domain and compromised abilities in the speech domain with a strong dissociation in performance on tasks requiring judgements of expressive prosody. His performance on tasks of emotional prosody, however, was preserved across the two domains.

The results on IB's performance on the five tasks of the study are summarised in Table 3.

Tasks Performance items (n=32)	I.B. score	Controls (n=24) Mean score	Controls S.D.	Controls score min-max	*t*-value	2-tailed *p* value	Percentile score
Expressive music prosody	23	24.16	3.15	17-29	-0.361	0.722	36.08%
Expressive speech prosody	16*	30.16	1.49	25-32	-9.311	0.000	0.00%
Speech prosody detection	11*	28.83	4.12	19-32	-4.240	0.000	0.02%
Emotional speech prosody	29*	31.79	0.50	30-32	-5.467	0.000	0.00%
Emotional music prosody	27	30.37	1.88	27-32	-1.756	0.092	4.62%

Table 3. IB's results compared to the normative data. Asterisks denote statistically significant differences from the population using Crawford & Garthwaite (2002) modified *t*-test.

4 Discussion

IB, a 51-year-old Greek monolingual right-handed man, displayed a pattern of preserved and compromised abilities on a set of prosody tasks which explored a number of dimensions employing linguistic and music stimuli. As IB performed well on the baseline tasks and achieved scores comparable to the healthy controls on three of the five tasks, it is argued that his impaired performance on speech prosody was task-specific and did not derive from a general difficulty with the experimental procedure.

His pattern of performance on this battery of tasks offers support for the argument that processing prosodic features with aesthetic value pertains to distinct mechanisms across the two domains. At the same time, the findings of this case study suggest that intact processing of emotional prosodic features in the speech domain does not seem to facilitate processing of aesthetic aspects of prosody despite acoustic similarities across prosodic categories. This evidence derives from results obtained through the use of relatively naturally-occurring stimuli rather than more artificially manipulated sound streams.

Although some dimensions of the tasks included here test novel aspects of prosody, broader comparisons to previous studies involving participants with acquired right hemisphere lesions may be illuminating. For example, Mendez (2001) reported on NS, an individual who had suffered a right temporo-parietal stroke and displayed compromised verbal comprehension and environmental sound perception as well as melody naming. However, his prosody perception and production were spared. More interestingly, NS developed an increased appreciation of music post-stroke. He devoted a great deal of his everyday life to listening to music and also started to attend concerts regularly. (The author does note that the participant was left-handed which may indicate some difference in lateralization.) In this case, right hemisphere impairment led to impairments in speech and music cognition while sparing music appreciation. Although a direct comparison with IB is not possible as they were not tested on the same tasks, both cases suggest that right hemisphere lesions in the temporo-parietal areas did not have a negative impact on the ability to appreciate the aesthetic aspects of music. In contrast, an amateur musician with right temporo-parietal damage was found to have lost his aesthetic appreciation post stroke (Mazzoni et al. 1993). We are not in a position to predict whether Mazzoni et al.'s (1993) patient with his lack of motivation to engage with music would be able to make judgements similar to those IB. The tasks employed in our study do not control for the listener's motivation or their music listening habits but can assess whether one can perceive differences of aesthetic music value. However, we have demonstrated that in the case of a congenital amusic individual, who does not choose to listen to music for aesthetic enjoyment, the ability to make judgements of expressiveness in musical stimuli was preserved (Loutrari & Lorch 2017).

Based on Mazzoni and colleagues' patient's self-reports, they suggested that the right hemisphere is likely to participate in cognitive functions serving global sound appreciation. This hypothesis is not supported by the performance of our case study of IB whose large right lesion including temporo-parietal areas did not affect his ability to discriminate between aesthetically 'expressive' music stimuli and their 'deadpan' versions. In fact, IB was able to judge differences in the prosodic qualities of music streams in a similar fashion to the neurotypical controls. His case suggests that music prosody can be appreciated even when extensive right hemisphere networks are damaged.

Previous studies with individuals suffering acquired brain damage point to the critical involvement of the right parietal cortex in accessing timing information (Griffiths et al. 1997; Nicholson et al. 2003; Patel et al. 1998). In the present study, deviation from temporal regularity in the stimuli of the expressive music prosody task constituted a salient timing difference in the expressive and inexpressive stimuli. Nevertheless, IB was able to successfully judge the difference between the two. It is possible that he may have used other cues (such as variation in dynamics and legato articulation) which were also present in the expressive stimuli to determine his judgements. However, one might have predicted that if he did have difficulties in appreciating timing differences they would have compromised his performance. Instead, he performed no different to neurotypical controls on this task.

The counterpart task, namely the expressive speech prosody task, revealed a different pattern in IB's performance. The presence of lexical information in the expressive speech prosody task did not yield any facilitating effect compared to its delexicalised version, that is, the speech detection task where IB's performance was also poor although his left hemisphere was unimpaired. The lexicalised task was also simpler in terms of its response mode, requiring same/different discrimination judgements but this did not lead to a better performance for IB. These data suggest that extensive right hemisphere damage can compromise global aesthetic prosodic processing in the speech domain sparing the ability to perform analogous judgements in the music domain.

Such performance points to the likelihood that the synergy of prosodic features is encoded differently in speech and music. Despite the fact that variation in pitch, loudness, and duration is present across domains, this case study and others reviewed above indicate that variation patterns may be perceived differently and impairments in one domain may not necessarily entail an analogous deficit in the other domain. This highlights the need of pursuing questions on prosody processing in a more holistic fashion to extend our understanding, rather than continuing to solely examine seemingly similar prosodic qualities in isolation. While the investigation of music perception using speech contours has been carried out in several previous studies (Ayotte et al. 2002; Hutchins et al. 2010; Nicholson et al. 2003; Patel et al. 2005; Patel et al. 1998) the perception of combined acoustic features in a naturally occurring context remains relatively unexplored.

The development of new research instruments such as the battery of tests and stimuli described here may provide a way of furthering this objective.

IB's performance on emotion perception was very high across domains despite his large right-hemisphere lesion. According to fMRI data from various studies, the right mid to anterior superior temporal cortex subserves the processing of affectively charged speech (Kemmerer 2014). This area has been also shown to respond to the combination of several acoustic parameters (duration, intensity, and pitch variation) but not to any of these features taken individually (Wiethoff et al. 2008). However, IB's lesion which included damage to this area did not affect his performance on the 'emotional speech prosody' task which contained affectively charged speech.

Previous research investigating the lateralisation of prosody perception and emotional processing do not yet provide a clear picture. Witteman et al.'s (2011) meta-analysis shows that strong lateralisation claims about emotional prosody cannot be made. Their analysis shows that emotional prosody is processed by bilateral networks, with the right hemisphere being involved to a relatively greater extent in such processing compared to left hemisphere networks. In the light of these findings, a likely interpretation of a patient with an extensive right hemisphere lesion achieving high scores on emotion perception is that it may be through reliance on these (secondary) left hemisphere networks.

It is likely that in the speech domain the appreciation of the significance of different emotions reflected in various prosodic realisations do not necessarily correspond to a common processing mechanism. For example, participants performing a linguistic task in Grimshaw et al. (2009) displayed more significant right hemisphere contribution to linguistic processing when stimuli bore sad prosody but this was not found to be the case for angry and happy prosody. Emotion does not seem to be a monolithic entity in the music domain either, as emotion perception deficits have also been demonstrated to affect selective emotions. A right temporo-parietal patient, JM, in Baird et al. (2014) was shown to be impaired at perceiving sad and peaceful emotions in music but had intact identification ability of happy and scary emotions due to his preserved perception of musical meter, as the authors argue.

It is not known, based on the present data, which cues or what kind of combinations of these cues contributed to IB's successful performance on the emotional prosody task. However, his pattern of performance indicates that the acoustic cues that facilitated his judgements of emotions in speech did not have an equal effect on the expressive speech prosody task. It, therefore, seems that it is not particular features *per se* that can be responsible for intact or impaired processing but their combination depending on a given context. With respect to pitch lateralisation, it has been suggested that hemispheric dominance hinges upon the communicative function that intonation serves (Van Lancker 1980). Additional prosodic elements and their interaction in the acoustic signal can also depend on the function a given stimulus serves rather than the presence of these elements. Emotion, like expressiveness, can be argued to be stimuli that are not processed as a group of distinguishable subcomponents but rather acoustic streams of a holistic character. Expressive prosody, despite relying on similar acoustic cues to those found in emotional prosody, is likely to be processed differentially. This proposal is analogous to Belyk & Brown's (2014) suggestion that while linguistic and emotional prosody share acoustic features, acoustic information might be fed into distinct perceptual processing systems pertaining to syntax, pragmatics or emotions. In order to evaluate this proposal in this new context and also determine whether findings stemming from IB can be generalised to other neurologically compromised individuals with similar lesions, future data collection employing the same materials is required.

5 Conclusions

The dissociation patterns of perceptual judgements of prosody made on speech and music stimuli in this case study of an individual who suffered an extensive right temporo-parietal lesion indicates that aesthetic prosodic streams across domains can be processed independently. Our findings also demonstrate that intact processing of emotional prosodic patterns in speech can dissociate from the perception of its expressiveness. We suggest that the acoustic features contained in the prosodic signal are perceived holistically based on the type of context in which they occur, as features that appear in one context can be perceived without facilitating perception in another context in which they also occur. While our findings await further investigation, this study offers a new approach to the exploration of an overlooked prosodic dimension present in auditory stimuli across the domains of language and music.

Acknowledgements

We thank Harriet Proios for her support during data collection and IB for his patience throughout testing sessions as well as control participants for their time.

Funding

This work was supported by a Wellcome Trust-Birkbeck Institutional Strategic Support Fund [Project # ACR00.ACC01] to the first author.

References

Audibert, N., V. Aubergé & A. Rilliard (2005). The prosodic dimensions of emotion in speech: The relative weights of parameters. In *6th Interspeech 2005 and 9th European Conference on Speech Communication and Technology (Eurospeech)*. Lisbon: International Speech Communication Association (ISCA), 524-528.

Ayotte, J., I. Peretz & K. Hyde (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain* 125: 238-251.

Baer, T., B.C. Moore & K. Kluk (2002). Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *The Journal of the Acoustical Society of America* 112: 1133-1144.

Baird, A.D., D.G. Walker, V. Biggs & G.A. Robinson (2014). Selective preservation of the beat in apperceptive music agnosia: A case study. C*ortex* 53: 27-33.

Belyk, M. & S. Brown (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social cognitive and affective neuroscience* 9: 1395-1403.

Besson, M., D. Schön, S. Moreno, A. Santos & C. Magne (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience* 25: 399-410.

Crawford, J.R. & P.H. Garthwaite (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia* 40: 1196-1208.

Cutler, A., D. Dahan & W. Van Donselaar (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40: 141-201.

Friederici, A.D. & K. Alter (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language* 89: 267-276.

Friederici, A.D., D.Y. von Cramon & S.A. Kotz (2007). Role of the corpus callosum in speech comprehension: Interfacing syntax and prosody. *Neuron* 53: 135-145.

Gingras, B. (2014). Individuality in music performance: Introduction to the research topic. *Frontiers in Psychology* 5: 661.

Griffiths, T.D., A. Rees, C. Witton, P.M. Cross, R.A. Shakir & G.G. Green (1997). Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study. *Brain: A Journal of Neurology* 120: 785-794.

Grimshaw, G.M., J.A. Séguin & H.K. Godfrey (2009). Once more with feeling: The effects of emotional prosody on hemispheric specialisation for linguistic processing. *Journal of Neurolinguistics* 22: 313-326.

Hutchins, S., N. Gosselin & I. Peretz (2010). Identification of changes along a continuum of speech intonation is impaired in congenital amusia. *Frontiers in Psychology* 1: 236.

Jakobson, R. (1958). Closing statement: Linguistics and poetics. In T.A. Sebeok (ed.), *Style and language*. Cambridge MA: MIT Press, 350-377.

Kemmerer, D. (2014). *Cognitive neuroscience of language*. New York: Psychology Press.

Klouda, G., D. Robin, N. Graff-Radford & W. Cooper (1988). The role of callosal connections in speech prosody. *Brain and Language* 35: 154-171.

Kotz, S.A., M. Meyer, K. Alter, M. Besson, D.Y. von Cramon & A.D. Friederici (2003). On the lateralization of emotional prosody: An event-related functional MR investigation. *Brain and Language* 86: 366-376.

Loutrari, A. & M.P. Lorch (2017). Preserved appreciation of aesthetic elements of speech and music prosody in an amusic individual: A holistic approach. *Brain and Cognition* 115: 1-11.

Loutrari, A. & M.P. Lorch (2018). When emotional character does not suffice: The dimension of expressiveness in the cognitive processing of music and language. In P. Gouk, J. Kennaway, J. Prins & W. Thormahlen (eds.), *The Routledge companion to music, mind, and well-being*. New York: Routledge, 249-261.

Marques, C., S. Moreno, S.L. Castro & M. Besson (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience* 19: 1453-1463.

Mazzoni, M., P. Moretti, L. Pardossi, M. Vista, A. Muratorio & M. Puglioli (1993). A case of music imperception. *Journal of Neurology, Neurosurgery, and Psychiatry* 56: 322.

Mendez, M.F. (2001). Generalized auditory agnosia with spared music recognition in a left-hander. Analysis of a case with a right temporal stroke. *Cortex* 37: 139-150.

Menninghaus, W., I.C. Bohrn, C.A. Knoop, S.A. Kotz, W. Schlotz & A.M. Jacobs (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition* 143: 48-60.

Nicholson, K.G., S. Baum, A. Kilgour, C.K. Koh, K.G. Munhall & L.L Cuddy (2003). Impaired processing of prosodic and musical patterns after right hemisphere damage. *Brain and Cognition* 52: 382-389.

Palmer, C. & S. Hutchins (2006). What is musical prosody? *Psychology of Learning and Motivation* 46: 245.

Pantev, C., R. Oostenveld, A. Engelien & B. Ross (1998). Increased auditory cortical representation in musicians. *Nature* 392: 811.

Paraskevopoulos, E., K. Tsapkini & I. Peretz (2010). Cultural aspects of music perception: Validation of a Greek version of the Montreal Battery of Evaluation of Amusias. *Journal of the International Neuropsychological Society* 16: 695.

Patel, A.D., J.M. Foxton & T.D. Griffiths (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition* 59: 310-313.

Patel, A.D., I. Peretz, M. Tramo & R. Labreque (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language* 61: 123-144.

Pell, M.D. (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and Language* 96: 221-234.

Penhune, V.B., R.J. Zatorre, J.D. MacDonald & A.C. Evans (1996). Interhemispheric anatomical differences in human primary auditory cortex: Probabilistic mapping and volume measurement from magnetic resonance scans. *Cerebral Cortex* 6: 661-672.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication* 41: 245-255.

Sammler, D., S.A. Kotz, K. Eckstein, D.V. Ott & A.D. Friederici (2010). Prosody meets syntax: The role of the corpus callosum. *Brain* 133: 2643-2655.

Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Research on Language & Social Interaction* 13: 201-277.

Wiethoff, S., D. Wildgruber, B. Kreifelts, H. Becker, C. Herbert, W. Grodd & T. Ethofer (2008). Cerebral processing of emotional prosody: Influence of acoustic parameters and arousal. *Neuroimage* 39: 885-893.

Witteman, J., M.H. van IJzendoorn, D. van de Velde, V.J. van Heuven & N.O. Schiller (2011). The nature of hemispheric specialization for linguistic and emotional prosodic perception: A meta-analysis of the lesion literature. *Neuropsychologia* 49: 3722-3738.

Zatorre, R.J., J.L. Chen & V.B. Penhune (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews. Neuroscience* 8: 547.