Neuropsychologia 66 (2015) 293–301

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Congenital amusia: A cognitive disorder limited to resolved harmonics and with no peripheral basis

Marion Cousineau [a,*], Andrew J. Oxenham [b], Isabelle Peretz [a]

[a] International Laboratory for Brain, Music and Sound Research (BRAMS) and Center for Research on Brain, Language and Music (CRBLM), Department of Psychology, University of Montreal, C.P. 6128, succursale Centre-ville, Montreal, QC, Canada H3C 3J7
[b] Department of Psychology, University of Minnesota, 75 East River Parkway, Minneapolis, MN 55455, USA

ARTICLE INFO

Article history:
Received 18 July 2014
Received in revised form
1 November 2014
Accepted 25 November 2014
Available online 27 November 2014

Keywords:
Pitch
Auditory perception
Spectro-temporal fine structure
ITD
ILD
Auditory filter bandwith

ABSTRACT

Pitch plays a fundamental role in audition, from speech and music perception to auditory scene analysis. Congenital amusia is a neurogenetic disorder that appears to affect primarily pitch and melody perception. Pitch is normally conveyed by the spectro-temporal fine structure of low harmonics, but some pitch information is available in the temporal envelope produced by the interactions of higher harmonics. Using 10 amusic subjects and 10 matched controls, we tested the hypothesis that amusics suffer exclusively from impaired processing of spectro-temporal fine structure. We also tested whether the inability of amusics to process acoustic temporal fine structure extends beyond pitch by measuring sensitivity to interaural time differences, which also rely on temporal fine structure. Further tests were carried out on basic intensity and spectral resolution. As expected, pitch perception based on spectro-temporal fine structure was impaired in amusics; however, no significant deficits were observed in amusics' ability to perceive the pitch conveyed via temporal-envelope cues. Sensitivity to interaural time differences was also not significantly different between the amusic and control groups, ruling out deficits in the peripheral coding of temporal fine structure. Finally, no significant differences in intensity or spectral resolution were found between the amusic and control groups. The results demonstrate a pitch-specific deficit in fine spectro-temporal information processing in amusia that seems unrelated to temporal or spectral coding in the auditory periphery. These results are consistent with the view that there are distinct mechanisms dedicated to processing resolved and unresolved harmonics in the general population, the former being altered in congenital amusia while the latter is spared.

1. Introduction

Pitch is the perceptual correlate of the periodicity (or repetition rate) of acoustic waveforms. It is a salient characteristic of sounds that plays a fundamental role in music and speech perception, as well as in auditory stream segregation, the mechanism by which we are able to hear out single auditory events within mixtures of sounds. Vocal and instrumental sounds that produce pitch are mostly harmonic complex tones that can be decomposed into a series of sinusoids (or partials) with discrete frequencies at integer multiples of a fundamental frequency (F0). It has long been known that lower-order harmonics (e.g., harmonics 1–5) within a complex can be heard out as separate tones under certain circumstances (e.g., von Helmholtz, 1877; Plomp, 1964; Bernstein and Oxenham, 2003). These lower-order "resolved" harmonics generally provide a strong, or salient, pitch when presented together, and have been shown to dominate the overall pitch of natural harmonic complexes (Plomp, 1967; Moore et al., 1985; Dai, 2000). Listeners can perceive differences in F0 between two successive harmonic complex comprising resolved harmonics that are an order of magnitude smaller than the semitone, the smallest scale step used in Western music (e.g., Micheyl et al., 2006). In contrast, higher-order harmonics (higher than about the 10th) are not readily heard out individually (Bernstein and Oxenham, 2003), and they produce a generally weak pitch sensation, with the minimum detectable change in F0 being much poorer than for resolved harmonics, and often larger than a semitone (Houtsma and Smurzynski, 1990; Bernstein and Oxenham, 2003).

The difference in pitch salience and accuracy produced by resolved and unresolved harmonics has been ascribed to how they are coded in the peripheral auditory system. The cochlea provides a frequency decomposition, resulting in a frequency-to-place mapping of the content of the incoming sounds. Cochlear encoding is thus often likened to a bank of bandpass filters, with bandwidths that are roughly proportional to the center frequency of the filter (Glasberg and Moore, 1990; Shera et al., 2002). The

* Corresponding author. Fax: +1 514 343 2175.
E-mail address: marion.cousineau@umontreal.ca (M. Cousineau).

http://dx.doi.org/10.1016/j.neuropsychologia.2014.11.031
0028-3932/

lower harmonics have been termed "resolved" because they are spaced sufficiently far apart in frequency to fall into distinct auditory filters. Resolved harmonics are thus thought to be represented by the spectro-temporal fine-structure cues associated with the cochlear response to the individual harmonics. In contrast, the pitch of unresolved complexes is thought to be extracted from the repetition rate of the temporal envelope produced by the interactions of multiple harmonics in a given auditory filter. Overall, pitch extraction is a complex process the details of which are not yet fully understood despite many years of research (see Moore and Gockel, 2011; Oxenham, 2012 for recent reviews).

Although pitch plays an important role in speech perception in both tone (Gandour, 1983) and non-tone (Nooteboom, 1997; Binns and Culling, 2007; Miller et al., 2010) languages, it is in music that the demands on pitch processing are highest. In contrast to prosody, where variations as large as an octave can be observed within an utterance (Fitzsimons et al., 2001), differences in pitch as small as one semitone (or a change of 6% in frequency) are both common (Vos and Troost, 1989) and meaningful in music. For instance changes of a semitone define the distinction between the major and minor mode in Western tonal music. Therefore, deficits in pitch processing can lead to a wide variety of deficits in music-related tasks. Here we study pitch perception in congenital amusia, a life-long disorder of music perception that, in its most common form, is thought to reflect impaired pitch processing (sometimes also called tone-deafness).

Impairments related to amusia have been reported for pitch discrimination (Ayotte et al., 2002; Peretz et al., 2003; Foxton et al., 2004; Hyde and Peretz, 2004; Tillmann et al., 2009; Liu et al., 2010), memory for pitch (Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010; but see Jiang et al., 2013), melody recognition (Ayotte et al., 2002), singing in tune (Dalla Bella et al., 2009), and dissonance perception (Ayotte et al., 2002; Cousineau et al., 2012). The condition was originally described as a strictly musical deficit, but this specificity has been challenged by recent evidence for deficits in prosody perception in both non-tonal (Patel et al., 1998, 2008; Thompson et al., 2012) and tonal (Jiang et al., 2010, 2012a) languages, as well as categorical perception of language tones in Mandarin speakers (Jiang et al., 2010, 2012b). The diagnostic tool most widely used for congenital amusia is the Montreal Battery of Evaluation of amusia (MBEA, Peretz et al., 2003), which consists of six tests that assess musical pitch and rhythm discrimination, meter perception and memory. The prevalence of the condition has been estimated to be around 4% of the general population (Kalmus and Fry, 1980; though see Henry and McAuley, 2010).

Here, we use psychoacoustic methods to investigate the origin of the pitch deficit observed in this population. Our main hypothesis, tested in the first experiment, is that the deficits in pitch perception observed in congenital amusia arise from abnormal processing of spectro-temporal fine-structure cues associated with the individual resolved harmonics. The hypothesis is motivated by the observation that the classic deficits found in amusia (pitch discrimination thresholds larger than a semitone and poor melody recognition) can also be observed in normal listeners when the access to these cues is reduced by removing resolved harmonics (Moore and Rosen, 1979; Shackleton and Carlyon, 1994; Kaernbach and Bering, 2001). Accordingly, amusics are expected to perform poorly with single pure tones and low-order resolved harmonics, but normally with high-order unresolved harmonics. Additionally, if amusics suffer from impaired access to the fine spectro-temporal information associated with pure tones and resolved harmonics, then increasing stimulus uncertainty (via pitch transposition) and reducing tone duration will be more deleterious to thresholds for resolved complexes in amusics than in controls. This prediction is based on the hypothesis that amusics use a mechanism to process resolved harmonics that is similar to that used by controls with unresolved complexes and the observation that both these manipulations increase thresholds for unresolved complexes in the normal population (White and Plack, 1998; Cousineau et al., 2009).

Our second experiment tested the hypothesis that the deficits in fine spectro-temporal coding found in amusia reflect a deficit in fine peripheral spectral processing. The processing of low-order resolved harmonics has been shown to relate to frequency selectivity, with poorer frequency selectivity leading to fewer resolved harmonics and poorer pitch discrimination (Bernstein and Oxenham, 2006). If pitch discrimination deficits observed in amusia are related to a peripheral deficit in spectral resolution, then amusics should also show deficits in basic measures of spectral resolution.

Our third experiment tested the hypothesis that deficits in fine spectro-temporal coding found in amusia reflect poor temporal fine-structure coding in the auditory periphery. Spatial hearing in humans relies on the coding of very fine timing differences between the ears in the microsecond range. It has been hypothesized that the same specialized neural mechanisms that permit this exquisite sensitivity to timing may also subserve the processing of periodicity and pitch (Licklider, 1951; Meddis and Hewitt, 1991a, 1991b). If amusia is associated with a deficit in processing temporal fine structure, this deficit may also affect spatial hearing.

The three hypotheses were tested in a series of experiments using the same group of amusic subjects and matched control subjects. In the first experiment, we measured pitch discrimination using either low-order (resolved) or high-order (unresolved) harmonics. In the second experiment, we estimated spectral resolution, or frequency selectivity, using a masking paradigm, known as the "notched-noise method" (Patterson, 1976; Glasberg and Moore, 1990; Oxenham and Shera, 2003). In the third experiment, we measured sensitivity to interaural time differences (ITDs) and interaural level differences (ILDs).

2. Material and methods

2.1. Subjects

Ten amusics and 10 matched controls participated in this study. The two groups did not differ significantly in age, years of education, years of musical training, or audiometric thresholds (see Tables 1 and 2). Each amusic participant scored at least 2SD below the mean of the general population when tested with the MBEA (Peretz et al., 2003; see Henry and McAuley, 2013 for further qualifications).

2.2. Stimuli and procedure

2.2.1. Sound delivery and presentation

All stimuli were delivered via an external soundcard (RME Fireface 800) with a 44.1-kHz sampling rate and 16-bit resolution. Sounds were delivered using closed headphones (Beyerdynamic DT 770 Pro). Participants were tested individually in a double-walled sound-attenuating booth. Responses were collected using a keyboard, and feedback was provided following each trial in all experiments. The overall testing time was approximately three hours per participant.

2.2.2. Experiment 1 – F0 difference limens

Fundamental frequency difference limens (F0DLs) were measured using a dual-pair task with a 2-down, 1-up adaptive method to track the 70.7% correct point on the psychometric function. In this task, subjects are presented with two pairs of sounds, one that contains two identical sounds and one that contains sounds that