The rise and fall of the devil's interval

Stimuli characteristics and hearing threshold may explain the tritone paradox

Christoph Reuter[1], Isabella Czedik-Eysenberg[1,4], Dijana Popovic[2], Esther Rois-Merz[3], Fatima Gerendas Obiols[1,3], Marik Roos[1], Sarah Ambros[1], Jörg Jewanski[1], Felix Klooss[2], Michael Oehler[4], Anja-Xiaoxing Cui[1]
[1] Department of Musicology, University of Vienna, [2] MediaLab, University of Vienna, [3] Audienz, Hearing Examination, Vienna, [4] Institute for Musicology and Music Education, Osnabrück University

Background

If successive **tritone intervals** are made up of **Shepard tones**, the interval can no longer be clearly perceived as ascending or descending (Shepard, 1964, p. 2350). However, when the direction of randomly played Shepard tritone intervals is to be estimated and the results are arranged according to pitch classes, a clear point can be found usually at or near the tritone C-F#, above which subjects consistently recognize a Shepard tritone interval as maximum **ascending** or **descending** (Deutsch, since 1986).

There have been various explanations for this point at which the "peak pitch class" (Deutsch, 1987) is found:

- **highest speech pitch** (Deutsch, North & Ray, 1990)
- **Native language** (Deutsch, 1991, 1994, 1997)
- **Regional origin** (Ragozzine & Deutsch, 1994)

Given the reliability of the C-F# pitch class boundary in almost every study, other factors have been discounted. These include:

- **Number of partials** used (Repp, 1997; Krüger, 2011; Malek, 2018)
- **Center** and **shape** of the envelope (Krüger, 2011)
- **Context**, i.e. preceding or following tones (Repp, 1997; Giangrande et al., 2003; Repp & Thompson, 2009; Krüger, 2011; Englitz et al., 2013; Chambers & Pressnitzer, 2014; Chambers et al., 2017).

Peak Pitch Class	Source
H-Dis	Deutsch, 1987
C-D	Deutsch, Kuyper & Fisher, 1987
C-D	Deutsch, North & Ray 1990
C#-D (vs. G)	Deutsch, 1991
C#-D (vs. G)	Deutsch, 1994
C#-D	Cohen, MacKinnon & Swindale 1994
C-C# (vs. D#)	Ragozzine & Deutsch, 1994
C (vs. F#)	Dawe, Platt & Welsh, 1998
C#-D	Giangrande, 1998
C#-D	Ragozzine, 2002
C#-D (vs. D#-F)	Deutsch, Henthorn & Dolson, 2004

Peak pitch classes observed in studies with the Deutsch stimuli set.

The stimuli used most often for Shepard tritone experiments are each composed of **only six partials** and run under **four envelopes** while their peaks are one tritone apart from each other (**e1** - **e4** with peaks at 300, 450, 600 and 900 Hz).
However, partials of the **highest envelope** (**e4**) are particularly present at the outer ear's resonant frequency at **2-4 kHz**, where we hear especially well. Consequently, the stimuli of the highest envelope may bias the average pitch direction judgments across envelopes.

Hypotheses

H1: Judgements for **e1** and **e3** stimuli **correlate positively** with each other and **negatively** with judgements for **e2** stimuli, given the similarity of the spectral structure for **e1** and **e3** stimuli and differences to the spectral structure of **e2** stimuli.
H2: Judgements for **e4** stimuli **correlate less** with the former and may depend on **hearing thresholds** of the individual ear at 2-4 kHz.

Methods

For n = 23 participants, **hearing thresholds** of both ears were measured before participants completed the **tritone paradox hearing test** (Deutsch 1995, Track 15-18) for **each ear** separately at a uniform sound level (65 dB_{SPL}).
Due to the **uniform level** setting, the levels of the partials in the **spectrally analysed** stimuli could be related equally to the respective measured individual hearing threshold for all participants.
We examined interval judgements for **each envelope** separately, and tested our hypotheses regarding **e4** stimuli by comparing **better-** and **worse-hearing** ears determined using a **median split** of hearing thresholds at 4 kHz.

Results & Discussion

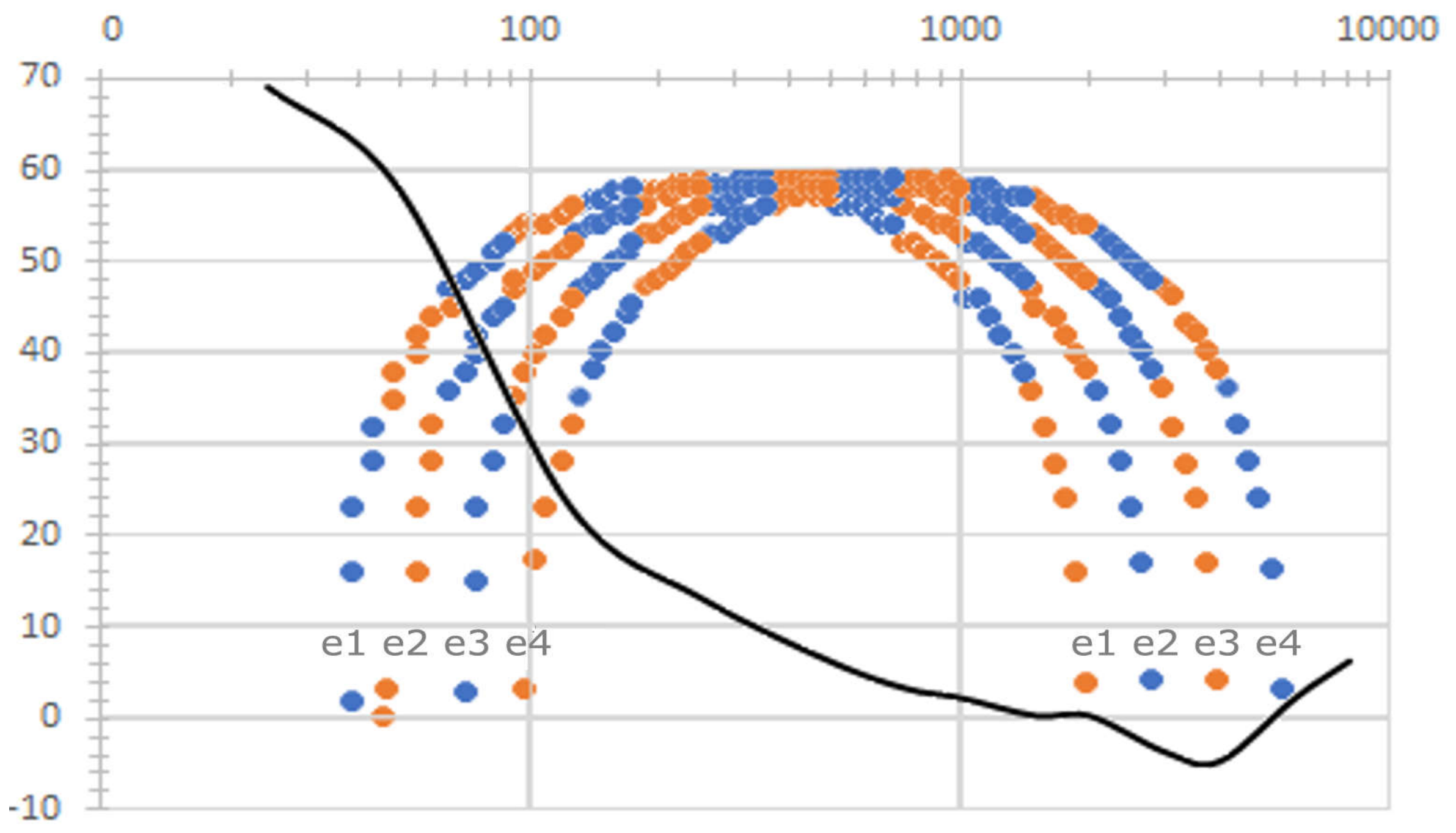


Fig. 1: Shepard envelopes (e1-4) used by Deutsch showing the amplitude of the respective partials of all stimuli related to a healthy hearing threshold (black line).
e1: 30-2000 Hz, **e2**: 40-3000 Hz, **e3**: 60-4000 Hz, **e4**: 90-5500 Hz, ● = partials of the 1st tone ● = partials of the 2nd tone of the respective tritone interval).

Considering the amplitude ratios of the partials under the four different Shepard envelopes (e1-e4), it becomes clear that the **choice of the envelopes** may lead to the **artifact** that the perceived interval direction **changes** from C#/D.

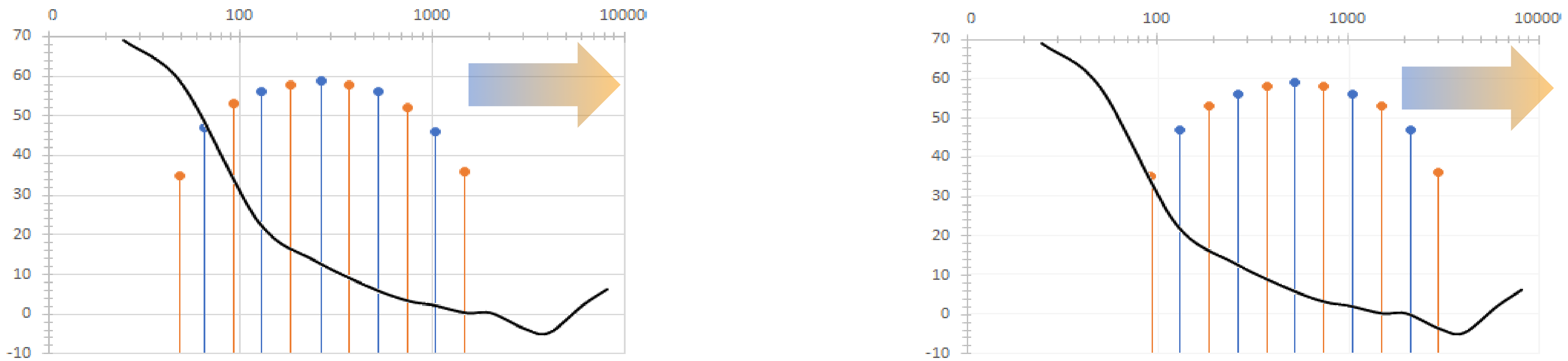
While Repp (1997) already considers this as an explanation for the often observed change of direction of the interval perception with C/C#, Deutsch (2004) argues that the opposite changes of direction caused by the envelopes e1 and e3 vs. e2 and e4 should cancel each other out.
Thus, another explanation for the perceived interval direction is needed, which is why we considered the influence of the **hearing threshold**.

Literature

Chambers, C., Akram, S., Adam, V., Pelofi, C., Sahani, M., Shamma, S., & Pressnitzer, D. (2017). Prior context in audition informs binding and shapes simple features. Nature Communications, 8(1), 15027. <> Chambers, C., & Pressnitzer, D. (2014). Perceptual hysteresis in the judgment of auditory pitch shift. Attention, Perception, & Psychophysics, 76(5 , 1271-1279. <> Cohen, A. J., MacKinnon, K., & Swindale, N. (1994). The tritone paradox revisited: Effects of musical training, envelope peak, and response mode. The Journal of the Acoustical Society of America, 95(5_Supplement), 2937–2937. https://doi.org/10.1121/1.409170 <> Dawe, L. A., Platt, J. R., & Welsh, E. (1998). Spectral-motion aftereffects and the tritone paradox among Canadian subjects. Perception & Psychophysics, 60(2), 209–220. https://doi.org/10.3758/BF03206030 <> Deutsch, D. (1986). A Musical Paradox. Music Perception, 3(3), 275-280. <> Deutsch, D. (1987). The tritone paradox: Effects of spectral variables. Perception & Psychophysics, 41(6), 563-575. <> Deutsch, D. (1991). Pitch Proximity in the Grouping of Simultaneous Tones. Music Perception, 9(2), 185-198. <> Deutsch, D. (1994). The Tritone Paradox and the Pitch Range of the Speaking Voice: Reply to Repp. Music Perception, 12(2), 257-263. <> Deutsch, D. (1997). The Tritone Paradox: A Link Between Music and Speech. Current Directions in Psychological Science, 6(6), 174-180. <> Deutsch, D. (2009). Musical Illusions. In L. R. Squire (Hrsg.), Encyclopedia of Neuroscience (Bd. 5, S. 1159-1167). <> Deutsch, D., Henthorn, T., & Dolson, M. (2004). Speech Patterns Heard Early in Life Influence Later Perception of the Tritone Paradox. Music Perception, 21(3), 357–372. https://doi.org/10.1525/mp.2004.21.3.357 <> Deutsch, D., Kuyper, W. L., & Fisher, Y. (1987). The Tritone Paradox: Its Presence and Form of Distribution in a General Population. Music Perception, 5(1), 79–92. https://doi.org/10.2307/40285386 <> Deutsch, D., North, T., & Ray, L. (1990). The Tritone Paradox: Correlate with the Listener's Vocal Range for Speech. Music Perception, 7(4), 371-384. <> Englitz, B., Akram, S., David, S. V., Chambers, C., Pressnitzer, D., Depireux, D., Fritz, J. B., & Shamma, S. A. (2013). Putting the Tritone Paradox into Context: Insights from Neural Population Decoding and Human Psychophysics. In B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon, & H. E. Gockel (Hrsg.), Basic Aspects of Hearing (Bd. 787, S. 157-164). Springer New York. <> Giangrande, J. (1998). The Tritone Paradox: Effects of Pitch Class and Position of the Spectral Envelope. Music Perception, 15(3), 253–264. https://doi.org/10.2307/40285767 <> Giangrand, J., Tuller, B., & Kelso, J. A. S. (2003). Perceptual Dynamics of Circular Pitch. Music Perception, 20(3), 241-262. <> Krüger, S. (2011). Zur Tonhöhenwahrnehmung von oktav-komplexen Tönen. Psychophysik, psychoakustische Theorie und computationale Modellierung. Martin-Luther-Universität Halle-Wittenberg. <> Malek, S. (2018). Pitch Class and Envelope Effects in the Tritone Paradox Are Mediated by Differently Pronounced Frequency Preference Regions. Frontiers in Psychology, 9, 1590. <> Ragozzine, F. (2001). The Tritone Paradox and Perception of Single Octave Related Complexes. Music Perception, 19(2), 155–168. https://doi.org/10.1525/mp.2001.19.2.155 <> Ragozzine, F., & Deutsch, D. (1994). A Regional Difference in Perception of the Tritone Paradox within the United States. Music Perception, 12(2), 213-225. <> Repp, B. H. (1997). Spectral Envelope and Context Effects in the Tritone Paradox. Perception, 26(5), 645-665. <> Repp, B. H., & Thompson, J. M. (2010). Context sensitivity and invariance in perception of octave-ambiguous tones. Psychological Research, 74(5), 437-456. <> Shepard, R. N. (1964). Circularity in Judgments of Relative Pitch. The Journal of the Acoustical Society of America, 36(12), 2346-2353.

Results & Discussion

The **interval movement** of each partial in the Shepard stimuli may explain the overall tendency of participants to perceive each individual stimulus as **rising** or **falling**. E.g. **partial 6 rises** from 1044 Hz to 1480 Hz from C to F# in envelope **e1** (and equally for each lower partial). Correspondingly, **a rising interval** is perceived by most participants from C to F#.

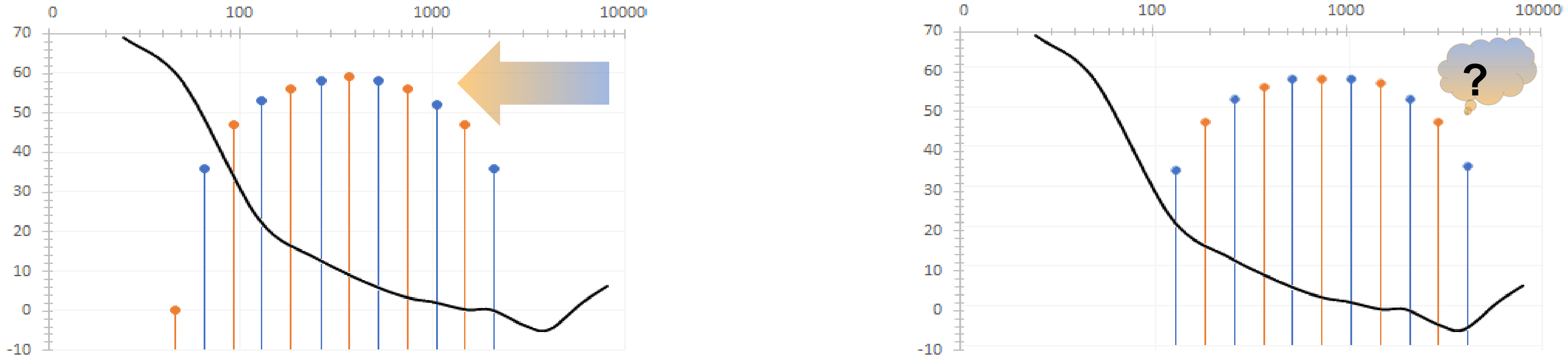


e1	C/F#	F#/C	C#/G	G/C#	D/G#	G#/D
	0.71	0.31	0.78	0.30	0.79	0.22

e3	C/F#	F#/C	C#/G	G/C#	D/G#	G#/D
	0.87	0.23	0.89	0.29	0.88	0.36

Fig. 2: Amplitudes and frequencies of the tritones interval C-F# below the envelopes e1 and e3.
Table: Mean values of the participant's judgments for C-F#, C#-G and D-G# (1 = the interval rises, 0 = the interval falls).
The direction of movement of partials 1-6 (especially the higher ones) determines the perceived interval direction.

This also applies to the envelopes **e2** and **e4**, only here the interval direction of the partials is **reversed** (e.g. 2094 Hz to 1480 Hz for the 6th partial from C to F#). However, the rule found above can only be reproduced for envelope **e2**, but **not** with the highest envelope (**e4**); here the directional perception seems to behave randomly with a slight tendency of all stimuli being perceived as rising.

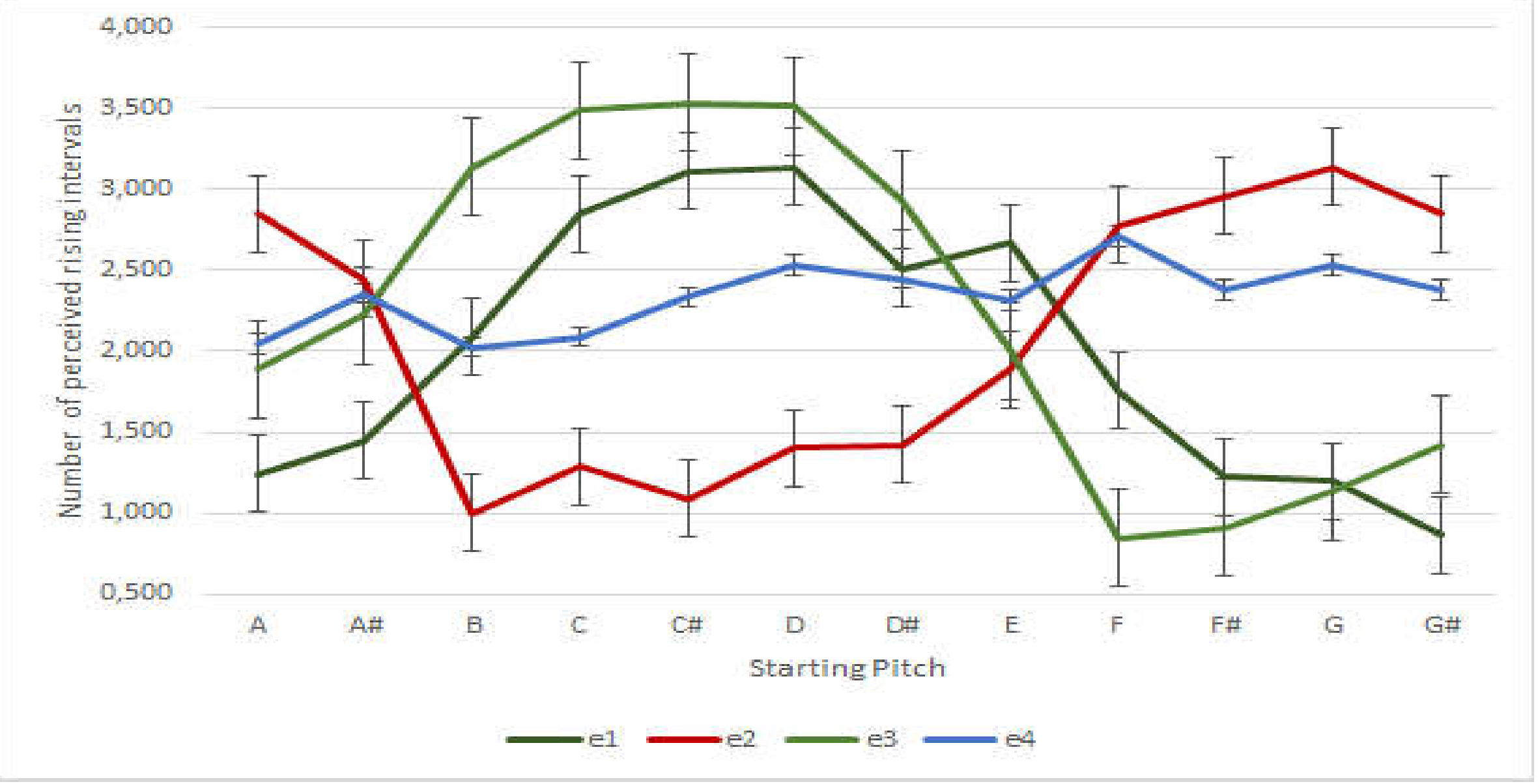


e2	C/F#	F#/C	C#/G	G/C#	D/G#	G#/D
	0.32	0.74	0.32	0.78	0.32	0.71

e4	C/F#	F#/C	C#/G	G/C#	D/G#	G#/D
	0.62	0.60	0.59	0.63	0.64	0.60

Fig. 3: Amplitudes and frequencies of the tritones intervals C-F# below the envelopes e2 and e4.
Table: Mean values of the participant's judgments for C-F#, C#-G and D-G# (1 = the interval rises, 0 = the interval falls). Again, in the case of envelope e2 the direction of movement of partials 1-6 (especially the higher ones) determines the perceived interval direction. However, this does not apply to the stimuli under envelope e4.

The perceived interval directions under the envelopes **e1** and **e3** show a positive correlation ($r_{e1\text{-}e3(10)}$ =.816, p=.001) while the opposite pattern is found between the envelopes **e1** and **e2**, and **e3** and **e2** respectively ($r_{e1\text{-}e2(10)}$=-.871, p<.001; $r_{e2\text{-}e3(10)}$=-.930, p<.001). The perceived interval directions for envelope **e4** do not show significant correlations with any of the other envelopes.



r	e2	e3	e4
e1	-0,871***	0,816**	-0,052
e2		-0,930***	0,350
e3			-0,415
p	**e2**	**e3**	**e4**
e1	0,000	0,001	0,873
e2		0,000	0,265
e3			0,180

Fig. 4: Perceived interval direction for stimuli with **e1**, **e2**, **e3**, and **e4** envelopes
Table: Result of correlation analysis and t-tests on the differences of perceived interval direction of Shepard tones under the envelopes e1-e4.

When comparing **better-** and **worse-hearing ears** determined using a median split of hearing thresholds at 4 kHz, correlations emerged that differed in sign between **better**-hearing ($r_{e1\text{-}e4(10)}$=.64, p=.026, $r_{e2\text{-}e4(10)}$=–.71, p=.010, $r_{e3\text{-}e4(10)}$=.81, p=.001), and **worse**-hearing ears ($r_{e1\text{-}e4(10)}$=–.20, p>.100, $r_{e2\text{-}e4(10)}$=.59, p=.045, $r_{e3\text{-}e4(10)}$=–.68, p=.014):

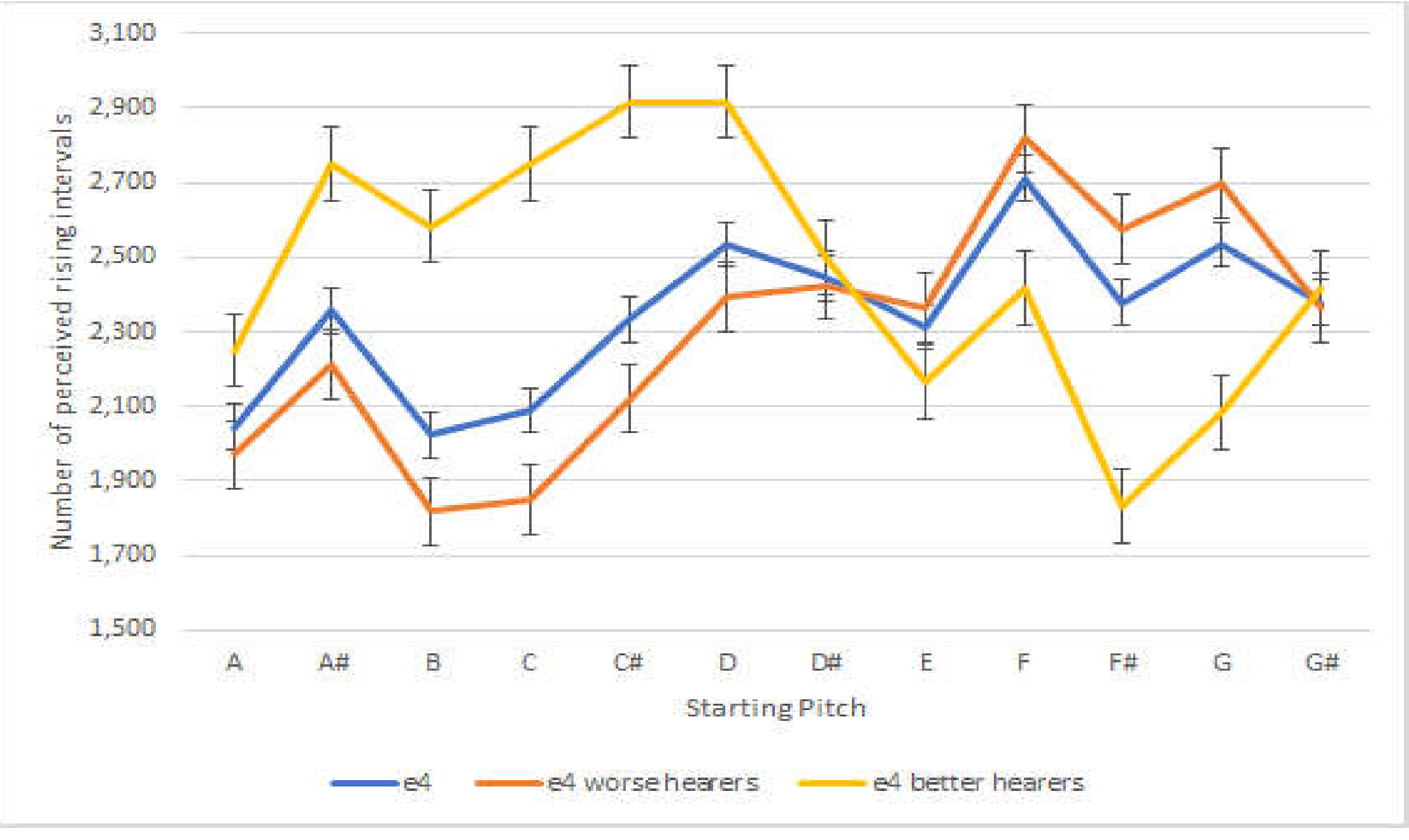


Fig. 5: Interval direction for stimuli with **e4** envelopes perceived by participants with **reduced** and with **healthy** hearing thresholds in the range of 2-4 kHz.

Participants with a reduced hearing threshold in the range of 2-4 kHz show a similar pattern for **e4** stimuli as for **e2** stimuli.

Participants with a healthy hearing threshold in the range of 2-4 kHz show an opposite pattern for the **e4** stimuli, which might be explained by a higher perceptual presence of partials in that frequency range.

Conclusion

While the perceived interval direction in the Shepard tritoni used/provided by Diana Deutsch depends on the **spectral envelope** in the case of **e1**, **e2**, and **e3 stimuli**, the perceived interval direction in the case of **e4 stimuli** may be determined by the **hearing ability in the range of 2-4 kHz**:
For subjects with **reduced hearing** at 2-4 kHz, the perceived interval direction for **e4 stimuli** is akin to that for the **e2** stimuli, whereas for **healthy hearing threshold** it is more similar to that for **e1** and **e3** stimuli, implying the explanatory power of the **hearing threshold** for the tritone paradox.