Investigating computational models of perceptual attack time

Nick Collins
Centre for Music and Science,
Faculty of Music, Cambridge
nc272@cam.ac.uk

ABSTRACT

The perceptual attack time (PAT) is the compensation for differing attack components of sounds, in the case of seeking a perceptually isochronous presentation of sounds. It has applications in scheduling and is related to, but not necessarily the same as, the moment of perceptual onset. This paper describes a computational investigation of PAT over a set of 25 synthesised stimuli, and a larger database of 100 sounds equally divided into synthesised and ecological. Ground truth PATs for modeling were obtained by the alternating presentation paradigm, where subjects adjusted the relative start time of a reference click and the sound to be judged. Whilst fitting experimental data from the 25 sound set was plausible, difficulties with existing models were found in the case of the larger test set. A pragmatic solution was obtained using a neural net architecture. In general, learnt schema of sound classification may be implicated in resolving the multiple detection cues evoked by complex sounds.

Keywords

Perceptual attack time, perceptual onset, computational modeling.

INTRODUCTION

`The notion of onsets is not entirely cut and dried, because the rise times of the notes of different instruments are not the same' (Bregman 1990)

Not all events are impulsive. Slow attack envelopes may shift the perceived onset time later into the physical event. Even with a percussive transient attack, the auditory system imposes small frequency dependent delays in processing the signal in order to trigger event detector units. Perceptual rather than physical onsets give a useful additional feature of signals, and in particular may assist accurate scheduling of a sequence of events, with regard to spacing events within streams, synchronising onset times between streams, and with respect to external time points. In general auditory scenes with multiple streams, there may be interactions in the processing of events in both vertical (simultaneous) and horizontal (temporal) directions due to auditory masking and stream fusion phenomena (Moore at al. 1997, Bregman 1990).

Perceptual onsets were first investigated in prosodic rhythm analysis in the speech processing literature (Marcus 1981, Howell 1988, Pompino-Marschall 1989, Harsin 1997, Scott 1998, Patel at al. 1999, Villing et al. 2003), where they have been termed *p-centers*. It was noted that a sequence of syllables rendered with evenly spaced physical onsets would not sound perceptually isochronous to subjects. Corrections applied to each syllable to produce a perceptually even sequence gave a relative sense of the respective perceptual center of each.

Vos and Rasch (1981) investigated such phenomena for general synthesised tones with variable attack envelope, terming the psychological moment of occurrence the *perceptual onset time* (POT). Following this with experimental work on (analysed and re-synthesised) instrumental tones normalised for energy, duration and pitch, Gordon (1987) differentiated the *perceptual attack time* (PAT) from the POT, arguing that the time of perceptual attack that defined a sound's position within a rhythmic (isochronous) sequence was sometimes separate to the first noticeable sound of a note. Gordon gives the example of a slowly rising bowed string tone. Indeed, the transient whereby the bow first engages with the string may be differentiated from the onset of a stable pitch region, just as consonant may be differentiated from vowel phones in speech work on syllables. A number of noise/tone and modulation detection principles may be operant, and the manner in which auditory objects take on a unified whole, rather than being broken down by a number of different feature cues, is an essential but difficult question underlying research work in this area.

The perceptual attack time will be the preferred term herein, for the applications sought are in managing the scheduling time of sounds within rhythmic sequences. PAT is thus similar to p-centre as commonly presented in speech-specific tasks. A number of onset detection algorithms lay claim to finding an onset close to the perceptual

Proceedings of the 9th International Conference on Music Perception & Cognition (ICMPC9).

moment of occurence (Klapuri 1999, Moelants and Rampazzo 1997), by modeling certain detection principles of the auditory system; these may approach a model of POT, but are inadequate for PAT.

A pre-occupation of the literature is the building of a model that predicts PAT. Some models are founded in properties of the simple amplitude envelope or output of a loudness model of sounds (Vos and Rasch 1981, Howell 1988, Gordon 1987) whilst some take into account more complicated auditory models accepting that a multi-band approach influencing perceptual location is a more likely mechanism psychoacoustically (Pompino-Marschall 1989, Harsin 1997, Villing et al. 2003). A general solution might require an accurate auditory model with differentiated sensitivity to frequency bands, temporal and spectral masking and temporal integration/processing latency effects, change discrimination processes on bandwise energy and modulation rates, and timbral and pitch detection components. Higher-level (learnt schema) mechanisms of active perception and attention may be invoked to explain syllable perception from phones, or analogously, the sense of unified note events belied by stages of initial transient and (pitched) stability present in instrumental tones. For instance, the result from the speech literature that consonant duration in CV pairs is a key factor in p-center location (Harsin 1997, Scott 1998) can perhaps be traced to a noise-tone model, with an interaction between the perception of the initial transient and the onset of pitch for the vowel; (at least) two separate detection mechanisms with different trigger times are implicated in recognising a language specific learnt sound object (the syllable). The quotation from Bregman at the head of this section begins to look like an understatement.

Predicting the PAT allows the early scheduling of the playback of events so as to `sound' at a desired time point. Particularly for slow rising tones, naive scheduling may lead to the perception of the tone occuring after a desired entry point. Knowledge of the attack portion of the perceptual envelope also allows a further parameter for the classification of events in our database. There is a necessary interaction with timbral character, as observed by Wessel (1979): `When we alter the properties of the attack of the tone, we are also likely to influence the temporal location of the perceived onset of the tone'.

Since (especially co-occuring) sounds may interact in the auditory system, the context of a sound's presentation may have an effect upon its PAT. A practical assumption of this work is that if any algorithm is established for PAT determination of isolated events, this PAT will remain valid even in playback situations with multiple streams. A first approximation to the problem of PAT enabled by such study is at least superior to no adjustment whatsoever for slower attacks. It is computationally intensive and unrealistic to render alternative trial outputs of multiple streams to predict combined PAT effects before computer music playback, and of course, it is exactly multi-stream cases that provide the greatest unknowns in current research.

EXPERIMENTS ON PERCEPTUAL ATTACK TIME

Reasons have already been given to suspect that modeling perceptual attack time is a hard problem. In order to further investigate models, ground truth data is required from human subjects.

Collecting such data presents many problems (Soraghan et al 2005, Scott 1998); even when carefully controlled, existing methodologies are indirect and provide relative measures between test sound and reference sound. The essential paradigm is that of an alternating presentation of common reference sound and test sound in a loop, where a subject can adjust the onset time of the test sound until they achieve perceptual isochrony, though simultaneous presentation has also been utilised (Gordon 1987). There are interactions between the need to avoid fusion and masking phenomena through sound overlap, and the need to keep the separation between reference and test sound onset small to improve temporal acuity of subjects in judging isochrony (following Weber's law). Whilst Soraghan and colleagues (2005) have recently suggested the use of Auditory Evoked Potentials as an objective measure of subjective reaction time, this method has not been fully tested, and is beyond the scope of my own investigation.

A preliminary study was carried out by Tom Collins under my supervision as an experimental project for the third year undergraduate Perception and Performance course in the Cambridge music faculty. He prepared a set of male and female sung vocal sounds from recordings of a contemporary composition. These were rated by subjects using a set-up devised by myself following the `seek-isochrony' alternating stimulus paradigm (Vos and Rasch 1981, Gordon 1987). Tom's concern at the time was a statistical comparison of the PAT between male and female voices. His collected data was also useful to myself as ground truth data for prototyping models. It was apparent however that there was great variability between subjects. This could be traced to some flaws in stimulus selection that had only become apparent from running the experiment; namely, that one had to take great care concerning any double-hits, independently attacking formants or strong offsets in the test sounds themselves influencing detections.

To support more extensive PAT modeling work, I prepared a database of 100 short sounds without double hits or strong offset confounds. These were broken down as detailed in Table 1. Recorded sounds were split into two groups of 25, mixing categories evenly between them. The synthesised sounds were self-contained groups of 25. The recorded sounds were selected to provide a cross-section of different sources typically encountered. No attempt to normalise for loudness, pitch or duration was attempted, because of the need for a database of ecologically valid real world examples with a variety of attack envelopes and timbres. The synthesised sounds however had common durations and were normalised by simple attack/decay triangular envelopes; the sines used a linear amplitude scale, the white noise sources a decibel scale.

Sound	**Num**	**Duration (sec)**	**Source/ Construction**
Solo string	6	0.32-0.57	Violin (3), cello (2), double bass
Other solo instrument	10	0.2-0.59	Trumpet (2), sitar (2), clarinet (2), alto sax (2), vibes (1), bell (1)
Voice (sung)	4	0.33-0.56	SATB
Voice (spoken)	4	0.2-0.5	
String orchestra	3	0.57-0.6	
Choir	3	0.56-0.6	
Percussion	6	0.2-0.5	
Orchestral	5	0.28, 0.53-0.6	Beethoven 7th symphony recording
Jazz band	4	0.25-0.6	
Electronic dance music	5	0.16-0.32	Squarepusher recording
Sine at 5 attacks and 5 frequencies	25	0.2	synthesised
Enveloped white noise (25 attacks in steps of 0.01)	25	0.24	synthesised
Reference click	1	0.01	synthesised

Table 1 PAT test sounds.

A few consistency checks were run with experimental subjects based on presentation mode (simultaneous against alternating) and repetition to understand possible problems with ground truth data collection for this problem. All subjects were experienced listeners from the Centre for Music and Science in Cambridge. Data was collected using a program built especially for the task. Subjects assessed sounds in a random order, adjusting the onset time of a sound using a slider by mouse and keyboard shortcuts, so as to seek perceptual isochrony or simultaneity with a reference click (0 msec attack, 10 msec decay impulse). Once happy with a sound, subjects were allowed to proceed to the next by pressing the return key; slider positions were randomised between trials. In order to help reduce fusion effects for simultaneity judgements, binaural presentation of reference and test sound was effected. Of course, crossover of information in the auditory system happens relatively early on in processing, though localisation can be a helpful cue for stream segregation. Correlation scores, and means, standard deviations and ranges of the absolute difference of vectors were calculated to measure the proximity of judgements in different modes.

For group 1 of the recorded sounds, a subject achieved a correlation score of 0.534 between alternating and simultaneous presentation modes for the 25, with absolute difference statistics showing an average discrepancy per sound on the order of 20msec, certainly noticeable as a timing change (mean 0.01908, standard deviation 0.01197, max 0.05625, min 0). In a between subjects test, two further subjects showed a correlation of 0.379 and stats of (mean 0.02742, standard deviation 0.0270, max 0.10425, min 0) between their responses on the second group of 25 recorded sounds. No larger scale study has been carried out to compare the alternating and simultaneous presentation modes on the same test set, and seek statistically significant difference, but this very preliminary report does point to possible inconsistencies in the two collection modes. Because of doubts of the efficacy of data for modeling produced from the more fusion-prone test, It was decided to use the isochrony-seeking paradigm rather than the simultaneous presentation one for further data collection.

To examine the range of responses in the general population under controlled conditions, a small scale study was undertaken using the 25 sinusoid synthesised stimuli. 14 subjects took part, 8 male and 6 female, with ages between 21 and 31, and one subject of age 45 (an electroacoustic composer). Musical experience varied from almost none to professional; no musician/non-musician dichotomy was imposed, for the task is one that can be accomplished by any hearing subject[1]. Each subject rated the 25 sounds twice (in random presentation order within repetition groups), for a total of 50 trials, typically taking around 30 minutes to assess (more if subjects were very particular about the task). A break could be taken at any time; the stimuli were presented over headphones at a common volume across participants. Reference click and test sound alternated within a loop of total time 0.8 seconds, with click at 0.0 seconds and the test sound at a position of 0.4 seconds adjusted by the subject from -0.2 to +0.04 seconds around the centre point of the loop.

Figure 1 gives a breakdown of mean onset correction over sounds, showing both the original run and the repeat, with standard deviation errorbars. The sounds are numbered into five groups of the five different attack times (1-5 = 0, 6-10= 0.01, 11-15= 0.02, 16-20= 0.05, 21-25= 0.1 seconds), where within each group there are the same five frequencies (130.813, 261.626, 523.251, 1046.503 and 2093.005 Hz, corresponding to octaves of a concert pitch middle C) from lowest to highest. Subjects predominantly chose negative corrections, because few sounds can match the reference impulse for suddenness of perceived attack, and the physical onset of the test sound must be pulled forwards to move the perceptual attack to the centre of the loop. As might have been hoped, a trend to larger compensations is evident for slower attack times.

To assess statistical significance, a three factor within-subjects ANOVA was calculated (5*5*2 for five frequencies, five attack times and 2 repetitions) using the SuperANOVA software. The only significant main effect or interaction was that of attack time (F(4,52)= 81.432, p=0.001 (G-G correction), $p<0.01$). Whilst it might have been hypothesised that frequency would have an effect upon results, latencies and time resolution limits due to auditory

[1] Musicians however may have an advantageous familiarity with skills of close listening, temporal acuity and timbral recognition that assist this task; however, for modeling purposes, the best subjects were separated, as described in the main text.

system processing for low against high frequency are relatively negligible compared to the activation envelope; Neely et al (1988) show a 5-10 msec at 250 Hz, and 1-4 msec at 8kHz mechanical (cochlear) delay, and 5 ms constant neural delay.

After taking the experiment, some subjects commented that they could switch between viewing the reference click or the test sound as the head of the loop, and this helped them to assess the isochrony. Such an attentional switch may have a bearing on results if active perception is implicated in the detection.

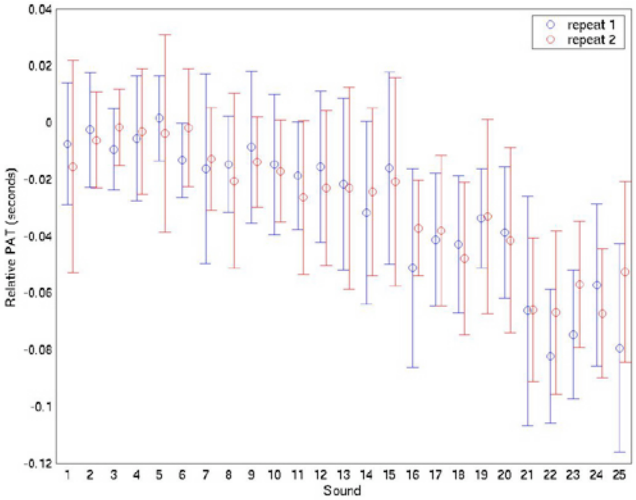

Figure 1 Experimental results showing mean relative PATs (with standard deviation error bars) across sinusoidal sounds.

With results within one standard deviation spanning 20-70 milliseconds of the mean, and total variation from 45 to 155 milliseconds for minimum and maximum values within subjects, the subjectivity of the results makes using ratings from a general population questionable. Expert ground truth is appropriate for modeling.

MODELING GROUND TRUTH DATA

The database of 100 sounds was used to compare the performance of various computational models of perceptual attack time (Figure 2). Those models adapted from the literature had to be constructed from study of the papers; source code implementations by the authors are not available. Where free parameters reside within models, some effort was taken to try to optimise to find the best scores over the test corpus.

Bandwise energy or total signal energy are often used. I have noted below either the use of a loudness like representation within ERB scale bands with equal loudness corrected contours (Collins 2005) or a basic power or windowed maximum representation. Of 42 possible ERB scale bands, those above 4000Hz were combined into a single channel following Gordon (1987), giving 27 bands. This ERB filter bank formed the input to an implementation of the Pompino-Marschall (1989) model.

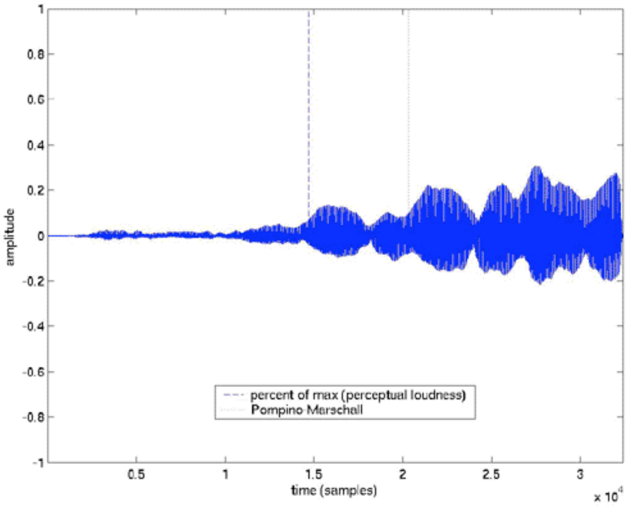

Figure 2 Perceptual attack time: predicted PAT is marked with respect to two models.

Model	**Parameter**	**Sum Squared Error**	**Error per sound**
Constant	0.025 sec	0.0156	0.02498
Time of max (Gordon 1987)		0.0038	0.01233
Percent of max (Vos and Rasch 1981)	97% of max	0.00078543	0.0056
Power in 512 sample windows, integration of normalised power exceeds threshold (Gordon 1987)	Threshold 0.05	0.0016	0.008
Normalized with rise=0.0 (Gordon 1987)	Slope threshold 14dB	0.0274	0.0331
Normalized with rise=0.75 (Gordon 1987)	Slope threshold 12.75dB	0.0015	0.007746
Pompino-Marschall (1989) with 27 ERB scale bands		0.0025	0.01
Pompino-Marschall (1989) with 27 ERB scale bands	Restricted to first 12 frames only	0.0012	0.00693

Table 2 PAT models performance on the 25 enveloped sine tone experimental stimuli.

Models calculated a PAT over a set of test sounds. Each model was further run on the reference click, and the PAT obtained subtracted from those for the test sounds to get a relative measure to be matched to ground truth.

Ground truth was created for the 25 sine sounds by averaging relative PATs from those experimental subjects judged most consistent in their responses. There were six subjects where correlation scores between the first and second repetition were greater than 0.5 and mean absolute difference was less than 20 milliseconds with standard deviation also under 20 milliseconds.

To evaluate the best model over the test database, it was found most straight forward to sum squared absolute error between ground truth relative PAT and model output. A measure of mean error per sound could also be gleaned by dividing the sum by the number of sounds in the database, and taking the square root. Following Gordon (1987), a mean error below 10ms would be preferred as being under the timing discrimination capabilities of human listeners.

Table 2 gives results of fitting models to the stimuli of the experiment alone. A good match is seen for a number of the standard models, the best performing being a simple percent of max model. Unsurprisingly, the particular threshold is different to that found by both (Vos and Rasch 1981) and (Gordon 1987): as the free parameter of the model, it is fitted to the test set.

As a harder test, a larger evaluation was carried out over all 100 sounds in the database. Given the variability of subject data in the general experiment, and some subjectivity perhaps inherent in the task, it was found most consistent for modeling purposes to use ground truth provided by the author, who had spent the most time rating sounds and had provided data across the whole test set over a number of settings.

The range of relative PAT of the data was from a minimum of -0.0100 to a maximum of -0.1813, mean of 0.0375 and standard deviation of 0.0490. The largest relative PATs were scored for some of the enveloped white noise test sounds with very long attacks.

Model	Parameter	Sum Squared Error	Error per sound
Constant	0.038 sec	0.2375	0.0487
Time of max (Gordon 1987)		2.3704	0.1539
Percent of max (Vos and Rasch 1981)	32% of max	0.2129	0.04614
Power in 512 sample windows, integration of normalised power exceeds threshold (Gordon 1987)	Threshold 0.02	0.034	0.01844
Normalized with rise=0.0 (Gordon 1987)	Slope threshold 6dB	0.3121	0.055866
Normalized with rise=0.15 (Gordon 1987)	Slope threshold 5.25dB	0.2288	0.04783
Pompino-Marschall (1989) with 27 ERB scale bands		1.4014	0.11838
Pompino-Marschall (1989) with 27 ERB scale bands	Restricted to first 7 frames only	0.1127	0.03357

Table 3 PAT models performance on the corpus.

Table 3 compares a number of models over the entire test database. Effort has been taken in each case to optimise over free parameters, and the best values found are indicated in the table. Where thresholds are passed by some FFT frame, interpolation since the previous frame was carried out to gain extra time resolution.

Results seem to suggest that the sine stimuli are an insufficient test of models, for none of the standard models predict PAT on the larger test set with greater than 18 millisecond accuracy. Repeating the model fitting process without the synthesised white noise stimuli (so for a database of 75 sounds including the sine tones and `real' sounds) did not improve matters.

In order to prepare a model for practical purposes, neural net models were investigated, which would integrate evidence from a number of signal features during the early portion of each sound.

A number of features were investigated using a simple loudness model without masking. This utilised the energy in ERB scale bands after calculating equal loudness contour correction of ERB band power following the ISO2003 standard; 27 ERB scale bands were so produced, bands 27-40 being combined into a single 27th band following Gordon (1987). These bands were further processed to obtain a number of features for the inputs of a neural net: the time to reach 10% of the total power in a band over a sound's duration, the temporal centroid within 10 FFT frames, combined power summed over combinations of ERB scale bands within the first ten FFT frames (1024 point FFT with hop size 512 samples at 44100 sampling rate, corresponding to 11.6 miliseconds, shorter than any test sound's duration). 3-fold validation was carried out to guard against over-fitting, using a randomised order for the 100 sounds, training on 67 and testing 33 for each fold. Scores in Table 4 give the average error over the folds.

Features	Sum Squared Error	Error per sound
10: ERB scale 1-40 combined, over 10 frames	0.087	0.0295
27: time till 10% of band energy for ERB scale 1-26 separately and 27-40 combined	0.0651	0.0255
20: ERB scale 1-11 and 12-40 combined, over 10 frames	0.0401	0.020025
27: temporal centroids of ERB scale 1-26 separately and 27-40 combined, over first 12 frames	0.0287	0.01694
30: ERB scale 1-11, 12-26 and 27-40 combined, over 10 frames	0.0226	0.015
20: ERB scale 1-11 and 27-40 combined, over 10 frames	0.0159	0.0126
24: ERB scale 1-11 and 27-40 combined, over 12 frames	0.0136	0.0116619

Table 4 Neural net model performance on the corpus.

The best performing features were the combination of power in ERB bands 1-11 and 27-40, over the first 10 frames of each sound. It is not necessarily counter-intuitive that missing out the middle bands helped, for perhaps signal spectral features in this area (667-4417Hz) confound the discrimination. There is a tradeoff between the number of input features for the net and the ability to both fit the training data and show good generalisation. The best model shows a performance around the time resolution of the FFT itself; this is most likely coincidental as it is probable that further investigation of features could reduce the error per sound further. Though the extent to which this identifies plausible physiological mechanisms is very much open to question, it does demonstrate the possibility of preparing relatively accurate predictive models for computer music applications.

This study suggests that work to create a large database of sounds for perceptual attack time modeling is valuable. Future experiments may gather further ground truth data, or perhaps seek to tease out particular signal features of sounds, particularly in terms of spectral envelope, and their contribution to a sense of PAT.

CONCLUSION

This paper has tackled the modeling of perceptual attack time with respect to ground truth data obtained from isochrony judgement experiments. Whilst many existing models were sufficient to match experimental results on a set of simple synthesised stimuli to a high degree of accuracy, a larger scale test database provided more problems. A pragmatic solution for computational purposes utilised a neural net on combined ERB scale band power features. Future work should seek to widen the test database further, and to obtain multiple expert ground truth over this database. Speculatively, the sense of perceptual attack may be related to the interaction of multiple detection principles of the auditory system (bandwise intensity changes, modulation rates, spectral timbral change, transient detection and onset of pitch), and the phonemic construction of sounds which may place into conflict a number of cues. Thus, the PAT may in general be unsolvable on simple principles alone, but reliant on modeling learnt schemata of sound recognition and classification.

ACKNOWLEDGMENTS

I would like to thank my colleagues in the CMS for fruitful discussions, and Tom Collins (no relation) for agreeing to work in this area and run a prototype experiment for an undergraduate project. This work was supported by AHRC grant 2003/104481.

REFERENCES

Bregman, A S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Camb, MA: MIT Press.

Collins, N. (2005). A Comparison of Sound Onset Detection Algorithms with Emphasis on Psychoacoustically Motivated Detection Functions. In *Proceedings of AES118 Convention*.

Desain, P., & Windsor, L. (Eds.) (2000). *Rhythm Perception and Production*. Lisse, the Netherlands: Svets and Zeitlinger.

Gordon, J. W. (1987). The perceptual attack time of musical tones. *J. Acoust. Soc. Am., 82(1)*, 88–105.

Harsin, C. A. (1997). Perceptual-center modeling is affected by including acoustic rate-of-change modulations. *Perception and Psychophysics, 59(2)*, 243–51.

Howell, P. (1988). Prediction of P-center location from the distribution of energy in the amplitude envelope: I. *Perception and Psychophysics, 43*, 90–3.

Klapuri, A. (1999). Sound onset detection by applying psychoacoustic knowledge. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc. (ICASSP)* (pp. 3089–92).

Marcus, S. M. (1981). Acoustic determinants of perceptual center (p-center) location. *Perception and Psychophysics, 30(3)*, 247–56.

Moelants, D., & Rampazzo, C. (1997). A computer system for the automatic detection of perceptual onsets in a musical signal. In Camurri, A. (Ed.), *KANSEI, The Technology of Emotion* (pp140–146).

Moore, B. C. J., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *J. Audio Eng. Soc., 45(4)*, 224–40.

S. T. Neely, S. J. Norton, M. P. Gorga and W. Jesteadt. (1988). Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli. *J. Acoust. Soc. Am., 83(2)*, 652–56.

Patel, A. D., Lofqvist, A. & Naito, W. (1999). The acoustics and kinematics of regularly-timed speech: A database and method for the study of the p-center problem. In *Proceedings of the 14th International Congress of Phonetic Sciences, Volume I* (pp.405–8).

Pompino-Marschall, B. (1989). On the psychoacoustic nature of the p-center phenomenon. *Journal of Phonetics, 17*, 175–92.

Scott, S. K. (1998). The point of p-centres. *Psychological Research, 61*, 4–11.

Soraghan, C., Ward, T., Villing, R., & Timoney, J. (2005). Perceptual centre correlates in evoked potentials. In *3rd European Medical and Biological Engineering Conference (EMBEC 05)*.

Villing, R., Ward, T., & Timoney, J. (2003). P-centre extraction from speech: the need for a more reliable measure. In *Proceedings Irish Signals and Systems Conference (ISSC 2003)*. (pp136–41).

Vos, J., & Rasch, R. (1981). The perceptual onset of musical tones. *Perception and Psychophysics, 29(4)*,323–35.

Wessel, D. (1979). Timbre space as a musical control structure. *Computer Music Journal, 3(2)*, 45–52.