Journal of Physiology - Paris 104 (2010) 215–222

Contents lists available at ScienceDirect

Journal of Physiology - Paris

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

Should spikes be treated with equal weightings in the generation of spectro-temporal receptive fields?

T.R. Chang [a], T.W. Chiu [b,e,f], P.C. Chung [c], Paul W.F. Poon [d,*]

[a] Department of Computer Science and Information Engineering, Southern Taiwan University, Tainan, Taiwan
[b] Institute of Basic Medical Science, National Chiao Tung University, Hsin-Chu, Taiwan
[c] Department of Electrical Engineering, National Chiao Tung University, Hsin-Chu, Taiwan
[d] Department of Physiology, National Cheng Kung University, Tainan, Taiwan
[e] Brain Research Center, National Chiao Tung University, Hsin-Chu, Taiwan
[f] Department of Biological Science and Technology, National Chiao Tung University, Hsin-Chu, Taiwan

ARTICLE INFO

Keywords:
STRF
Trigger feature
Response jitter
FM-sensitive cells
Auditory midbrain

ABSTRACT

Knowledge on the trigger features of central auditory neurons is important in the understanding of speech processing. Spectro-temporal receptive fields (STRFs) obtained using random stimuli and spike-triggered averaging allow visualization of trigger features which often appear blurry in the time-versus-frequency plot. For a clearer visualization we have previously developed a dejittering algorithm to sharpen trigger features in the STRF of FM-sensitive cells. Here we extended this algorithm to segregate spikes, based on their dejitter values, into two groups: normal and outlying, and to construct their STRF separately. We found that while the STRF of the normal jitter group resembled full trigger feature in the original STRF, those of the outlying jitter group resembled a different or partial trigger feature. This algorithm allowed the extraction of other weaker trigger features. Due to the presence of different trigger features in a given cell, we proposed that in the generation of STRF, the evoked spikes should not be treated indiscriminately with equal weightings.

1. Introduction

1.1. Neural coding of FM sounds

Since frequency modulation (FM) components are commonly found in communication sounds (Doupe and Kuhl, 1999), the knowledge on neural coding of (FM) sounds is important in understanding speech coding. Psychophysical and functional imaging studies showed that the human brain is sensitive to a variety of FM sounds (Behne et al., 2005; Chen and Zeng, 2004; Hart et al., 2003; Husain et al., 2004; Luo et al., 2006, 2007). Electrophysiological studies of single units in animals also revealed FM-sensitive cells at levels of the central auditory system as low as the midbrain (Felsheim and Ostwald, 1996; Poon et al., 1992; Poon and Yu, 2000; Qi et al., 2007; Rees and Møller, 1983; Suga, 1968). The auditory midbrain has been considered an important center for FM coding because of its large proportion of FM-sensitive cells (~70%; Poon and Chiu, 2000). One characteristic of such FM-sensitive cells is their differential response to the direction of FM sweeps. For a particular cell sensitive to FM sounds, this selectivity to the stimulus is further characterized by FM rate and frequency range (Heil et al., 1992; Rees and Kay, 1985; Whitfield and Evans, 1965). FM selectivity has been described conceptually as the cell's receptive field or receptive space (Poon et al., 1991). More commonly FM receptive fields are depicted as spectro-temporal receptive fields (STRFs; Aertsen and Johannesma, 1981a; deCharms et al., 1998; Depireux et al., 2001; Eggermont et al., 1983; Klein et al., 2000). STRFs display graphically the input characteristics or 'trigger feature' of a given cell. To determine the trigger feature, a variety of random stimuli are used to explore the sound preference of individual neurons. These stimuli include dynamic ripple noise, random chords and random FM tones (Aertsen and Johannesma, 1981b; Shechter and Depireux, 2007; Valentine and Eggermont, 2004; Poon and Yu, 2000). Using random stimuli, trigger features can be determined efficiently (e.g., based on 2 min of spike activity, Chiu and Poon, 2007). The analysis of STRF can hence deepen our understanding of neural coding of speech-like complex sounds (Andoni et al., 2007; Cohen et al., 2007; Kao et al., 1997; Sen et al., 2001; Theunissen et al., 2000).

1.2. Spectro-temporal receptive field (STRF)

STRFs are generated by a procedure known as 'reverse correlation' or spike-triggered averaging (Aertsen et al., 1981; Hermes et al., 1981). Peri-spike segments of the random waveform are first

* Corresponding author. Tel.: +886 6 235 3535x5458; fax: +886 6 236 2780.
E-mail address: ppoon@mail.ncku.edu.tw (P.W.F. Poon).

0928-4257/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jphysparis.2009.11.026

extracted, and then the stimulus power time-locked to the spikes is summed over a peri-spike time-versus-frequency plane to form the STRF. Results are depicted in color-coded pixel maps to reveal areas with high overlaps. For midbrain auditory neurons, these areas typically appear congregated to form a grossly elongated structure (or 'FM band' for simplicity) reflecting the trigger feature of the FM cell (Fig. 1A). In such approach, each spike is given the same weighting, assuming that there is only one trigger feature. If there are two or more trigger features, the method of STRF generation would need modifications. For example, individual spikes may be separated into different groups with each corresponds to a different trigger feature. In this case, spikes should not be given equal weights in spike-trigger averaging.

1.3. Aim of this study

In our previous study (Chang et al., 2005) we had developed an algorithm to sharpen trigger features in the STRF using dejittering (Fig. 1B). Similar but not identical dejittering algorithms have also been developed for sharpening trigger features in the insect mechanoreceptor system (Aldworth et al., 2005; Gollisch, 2006). At the auditory midbrain the synaptic inputs derived from the cochleae are tonotopically structured or stacked into frequency laminae (Clopton and Winfield, 1973; review see Malmierca, 2004). Each lamina receives synaptic inputs confined to a narrow frequency band. It is tempting to speculate that the trigger features of FM-sensitive cells might contain separable components, each corresponding to inputs from individual frequency laminae adjacent to each other. We hypothesized that if spikes from a given cell can be segregated into different groups according to some response property, different trigger features will emerge. In the present study, we developed a novel algorithm to test this hypothesis and the evidence was in support of the presence of two trigger features.

2. Materials and methods

2.1. Animal preparation

Rats (Sprague Dawley, 200–250 gm) were anesthetized with urethane (1.5 g/kg, i.p., maintained at 0.4 g/kg for pain areflexia when necessary) and fixed with a special head holder for recording extracellular single spike activities using conventional electrophysiological procedures. The skull overlying the occipital cortex was first surgically opened and glass micropipette electrode (20–70 MΩ) was advanced into the midbrain (or more precisely inferior colliculus) using a stepping micro-drive (Narishige) to hunt for single units that responded to clicks (0.1 ms pulse, ~90 dB SPL). After unit identification, we recorded its responses to a series of acoustic signals. A computer interface (Tucker Davies Technology System II) with conventional amplification and filtering (Axonprobe 1A, PARC-5113) was used. During experiment, the animal was placed inside a sound-treated chamber (Industrial Acoustic Company) for free-field acoustic stimulations, and its body temperature was maintained with a servo-control heating pad. The procedure was approved by the Animal Ethics Committee, Laboratory Animal Center, NCKU.

2.2. FM stimuli

Two sets of FM stimuli were used. The first set was three sets of random FM tones, generated by digitally low-pass filtering a white noise at 12.5, 25 or 125 Hz. The filtered signal was then fed to the voltage-control–frequency input of a function generator (Tektronix FG280) to control the instantaneous frequency of a continuous tone. The resultant stimulus was a FM signal with frequency randomly varied over time. Such sounds are powerful stimulus for auditory neurons in the rat midbrain (Poon et al., 1991). After determining the unit's most sensitive frequency (or characteristic frequency, CF), the center frequency of the FM was set at this CF, with a modulation depth of ~2 octave. Stimulus intensity was set ~30 dB supra-threshold at CF. We first collected spike responses with all the three sets of FM tones but only the dataset with the maximal spike count was used for subsequent analysis. Each spike dataset represented a continuous recording of 2 min. The second set of stimulus was a family of FM ramps, consisting of six systematically varied triangular linear frequency sweeps. They were delivered intermittently over a period of 2 s (Fig. 2A). The series of FM ramps had both up- and down-sweep phases, and their modulation rates overlapped with the random FM tones (FM rate: 9.8–63.2 octave/s, range: 1.58 octave).

2.3. Data collection

A computer interface (Tucker Davies Technology System II) was programmed to deliver the modulating waveforms while simultaneously collecting spike responses (after pre-conditioning spikes into 0.4 ms pulses). Details of electrophysiological procedures can be found in earlier publications from this laboratory (Poon et al., 1991; Chiu and Poon, 2000).

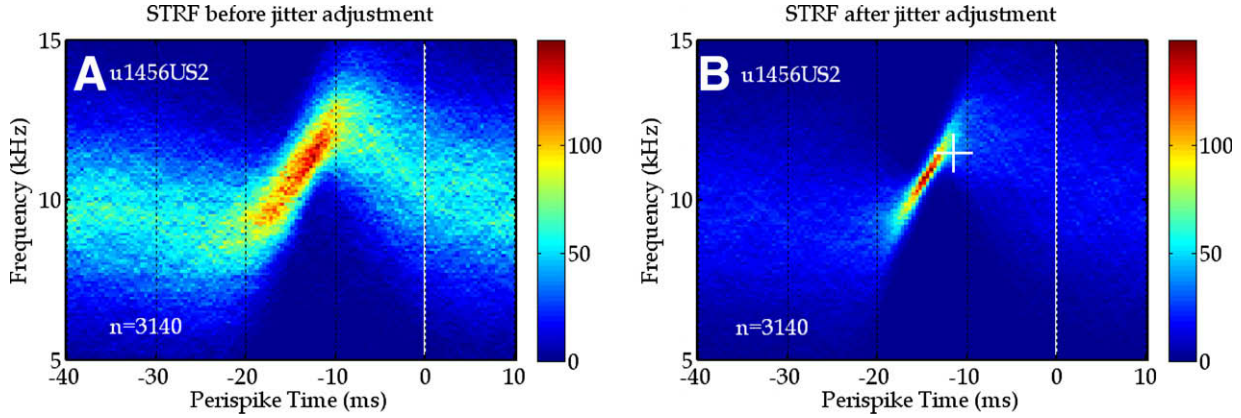

Fig. 1. Spectro-temporal receptive field (STRF) showing the trigger feature of a FM-sensitive cell in response to random FM tones. A: Note the concentration of modulating waveforms (red pixels) indicating a frequency up-sweep occurring from 10 to 20 ms before the spike. B: STRF with a sharpened 'trigger feature' after dejittering. Note the FM trigger point (white '+') is typically located at the end of the 'FM band' (see also Fig. 2D); *n* = number of spikes; color scale: counts modulating waveform overlap at each pixel.

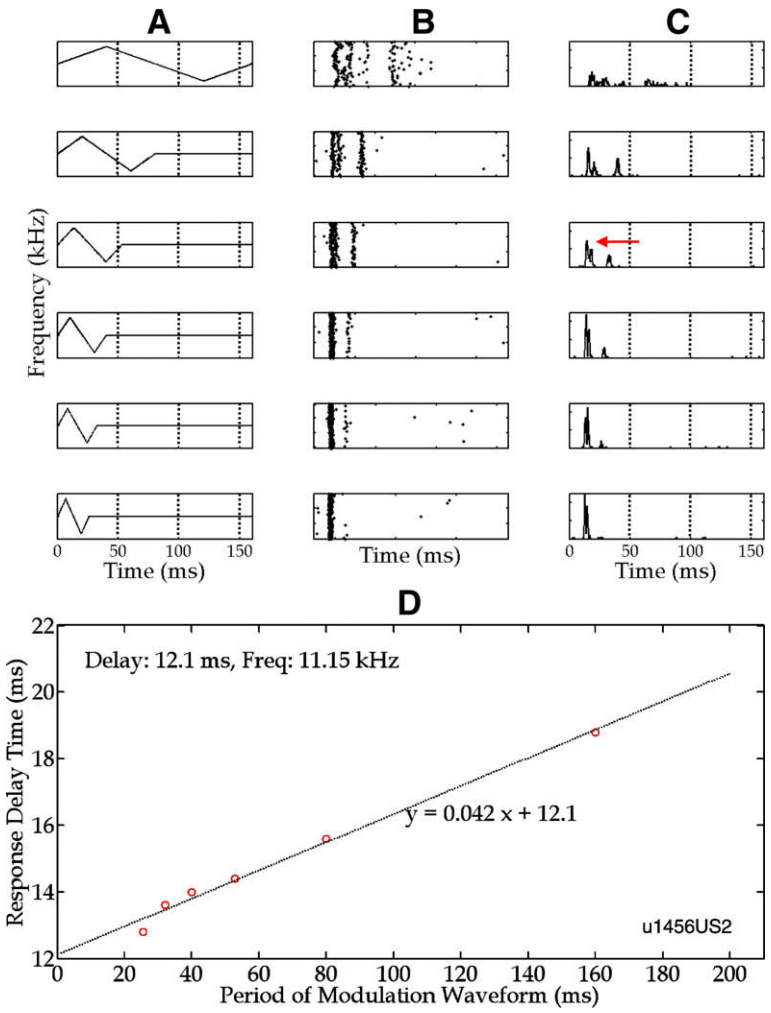

Fig. 2. Determination of the 'trigger point' of a FM-sensitive cell. (A): Frequency profiles of six different FM ramps. (B): Corresponding spike responses shown in dot raster (each dot represents an action potential). (C): Corresponding peri-stimulus time histograms (PSTHs) by summing dot raster in time. An arrow marks the PSTH with the maximal spike count, and this PSTH was low-pass-filtered and overlaid with the jitter adjustment histogram (see later Fig. 4B). (D): PSTH peak delays against the periods of FM ramp overlaid with their linear regression line. The *y*-intercept of the regression line represents the 'central transmission time' from the source to the neuron. Its slope represents frequency point on the FM ramp at which the spike is initiated. This 'trigger point' is also plotted in the STRF (Fig. 1B) to shows its relationship with the trigger feature.

2.4. Sharpening of trigger features in STRF

Individual modulating waveforms were systemically time-shifted to converge on a segment of the STRF which contains the putative trigger feature. The effect on convergence of each time-shifted modulating waveform was optimized further against the time window of the selected modulating waveform (for details see Chang et al., 2005). Its principle is illustrated in Fig. 3 and is briefly explained below.

2.4.1. Mean and variance time profiles

In a typical 2-min dataset, we captured ~1000 peri-spike modulating waveforms of the FM tone (40 ms pre-, 10 ms post-spike). The 'mean frequency time profile' and its corresponding 'variance time profile' were computed.

2.4.2. Time-shift analysis

A modulating waveform was first time-shifted at 0.4 ms steps over a range of −5–20 ms (corresponded to a window spanning from 20 ms pre- to 5 ms post-spike,) at a fixed time window length (chosen systemically at 1 ms steps, from 5 to 15 ms). At each shifted time, a 'similarity index' between the shifted modulating waveform and the 'mean frequency time profile' was computed (with the time window centered at the minimum position of the 'variance time profile'). From the 'similarity index' versus 'shift time' function, the time shift with the maximal similarity was taken as the 'optimal shift time'. This procedure of 'optimal shift time' was computed for each of the ~1000 modulating waveforms, and waveforms were finally shifted according to its individual 'optimal shift time'. A new 'variance time profile' (and new 'mean frequency time profile') was then obtained.

2.4.3. Comparing the new and old 'variance time profile'

We then compared the new and the old 'variance time profiles' according to a 'disparity index'.

2.5. From disparity index to 'sharpened STRF'

Using the new 'mean frequency time profile', the 'disparity index' between two consecutive 'variance time profiles' was again computed (following Sections 2.4.2 and 2.4.3) until the change in 'disparity index' reached a stable level. The STRF was then taken as the dejittered result (or 'sharpened STRF' for simplicity).

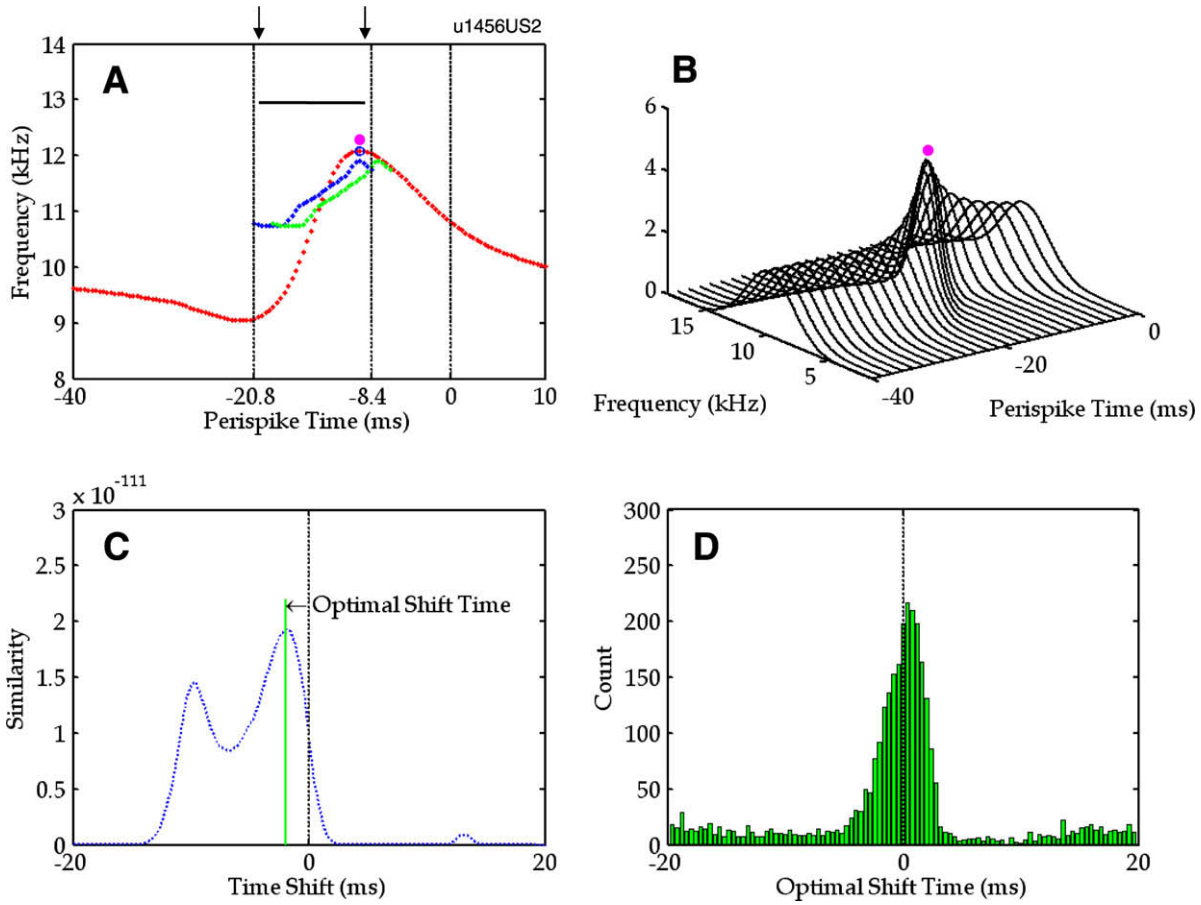

Fig. 3. Procedure of dejittering for a FM cell. (A): A segment of the modulating waveform (blue line between arrows) and its time-shifted version (green line) showed in systematic approximation with the mean (frequency time profile, red line). (B): Distribution of the modulating waveforms around the mean in perspective plot (pink dots indicate the peak, same location in A and B). (C): Similarity indices of the match with the mean frequency time profile over a range of ±20 ms. The peak (green line) represents the best match or the 'optimal shift time'. (D): The distribution of optimal shift times shows a peak riding over a non-zero count profile.

2.6. Match with peri-stimulus time histogram (PSTH)

Using spike responses of the same cell to the family of FM ramps, a family of PSTH was generated (Fig. 2B and C). The peak of PSTH represents the maximal response, and the time spread represents the response jitter. The PSTH that matched best with the slope of the 'FM band' in the STRF was taken as the 'matched PSTH'. Based on the response peak latencies measured from the PSTHs, a plot of delay times (y-axis) against the periods of FM ramp (x-axis) was further generated (Fig. 2D). A linear regression of the data points yielded an equation: $y = ax + b$, where slope 'a' represents the phase and intercept 'b' the central transmission delay (or the time taken from the stimulus source to reach the spike generator of the cell). Phase 'a' gives the frequency position of the 'trigger point'. The position of the 'trigger point' also allows comparison with the 'sharpened STRF' produced in the last session (Fig. 1B). For details of this method see Poon and Yu (2000).

2.7. Grouping spikes based on their dejitter values

The 'matched PSTH' was overlaid on the 'jitter adjustment histogram' (Fig. 4B). Those modulating waveforms with dejitter values within the peak region of the histogram (i.e., overlapping with the temporal extent of spike distribution in the 'matched PSTH') are placed in the 'normal jitter' group (Fig. 4D). Those outliers in the jitter adjustment histogram were placed in the 'outlying jitter' group (Fig. 4E). Stimulus trigger features in the respective STRFs looked different between the two groups, even before any sharpening of trigger features (Fig. 4G and H).

2.8. Sharpening trigger features in the dejittered groups

Using data in each group we computed 'optimal shift time' of the modulating waveforms (Sections 2.4.2, 2.4.3 and 2.5 to produce 'sharpened STRFs', Fig. 5). Here the normal jitter time window (±5 ms) was used for dejittering. Only modulating waveforms with a local optimum inside this time window were used for further analysis.

3. Results

3.1. Trigger features of the 'normal jitter' group

For FM-sensitive cells, about 60% of the spikes belonged to this group. Their STRF showed close resemblance to the original STRF (Fig. 4F and G and Fig. 5A–E).

3.2. Trigger features of the 'outlying jitter' group

Strikingly different trigger features showed up in the STRFs of this group (Figs. 4H, and 5C and F). The 'FM band' was clearly shorter in both frequency and time spans. It usually emerged in the upper frequency range of the 'normal jitter' STRFs, resembling a

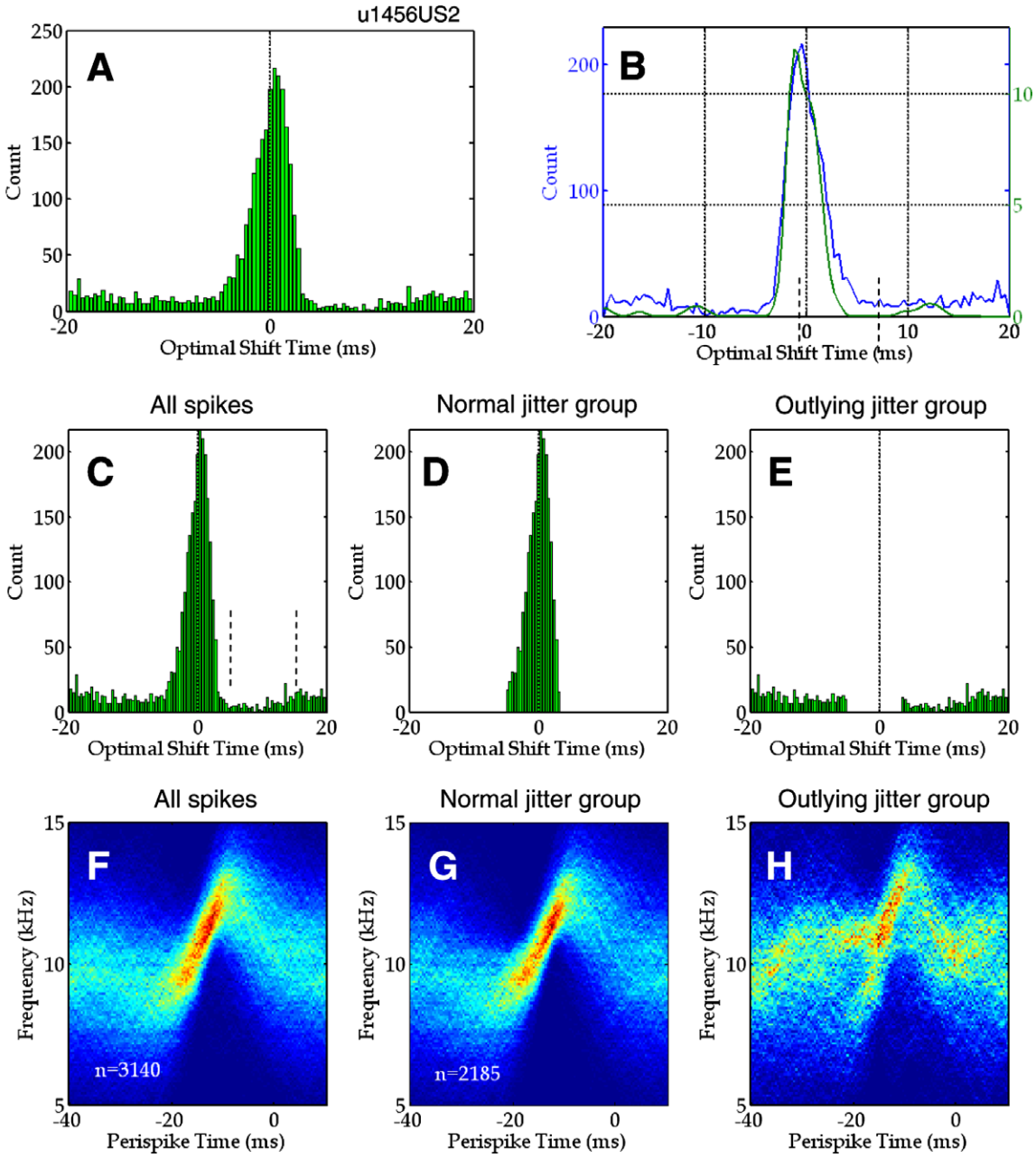

Fig. 4. Segregation of the peri-spike modulating waveforms of a FM cell into two jitter groups: normal and outlying. (A): Distribution of 'optimal shift time' (or jitter adjustment histogram). (B): Overlaying optimal shift time histogram with the PSTH profile as obtained from FM ramps (see also Fig. 2C). Note their similarity at the peak region (vertical dashed-lines mark the range of normal jitters). (C): Splitting the optimal shift times into the normal (D) and outlying (E) jitter groups. F–H: STRF for each group before jitter adjustment. Note the different trigger features in the outlying jitter group (H).

'partial trigger feature'. In addition, the slope of its 'FM band' could be the same or different (Table 1).

3.3. Trigger features of noise-replaced 'outlying jitter' group

As the sharpening procedure always tries to find a structure, even, where there is none, what we found for the 'outlying jitter' group could be an artifact of the method. To rule out this possibility, the sharpening procedure was again applied to a collection of modulating time waveforms ($n > 3000$) containing a mixture of (a) 33% of stimuli from the normal peri-spike dataset from a FM cell and (b) 67% of stimuli selected at random intervals around the spike (i.e., non time-locked to the spikes). We found that the triggering feature was markedly different from the original outlier stimuli (Fig. 6, Table 1). Results ruled out artifacts of the method.

4. Discussion

4.1. Importance of jitter grouping in separating trigger features

For the first time, we were able to reveal different trigger features of a given FM cell, by first segregating spikes into different jitter groups. The 'normal jitter' group showed 'full trigger feature', whereas the 'outlying jitter' group showed a weaker or 'partial trigger feature'. The triggering nature of the 'FM bands' is further supported by its close proximity in the STRF to the 'trigger point' estimated independently using the triangular FM ramps.

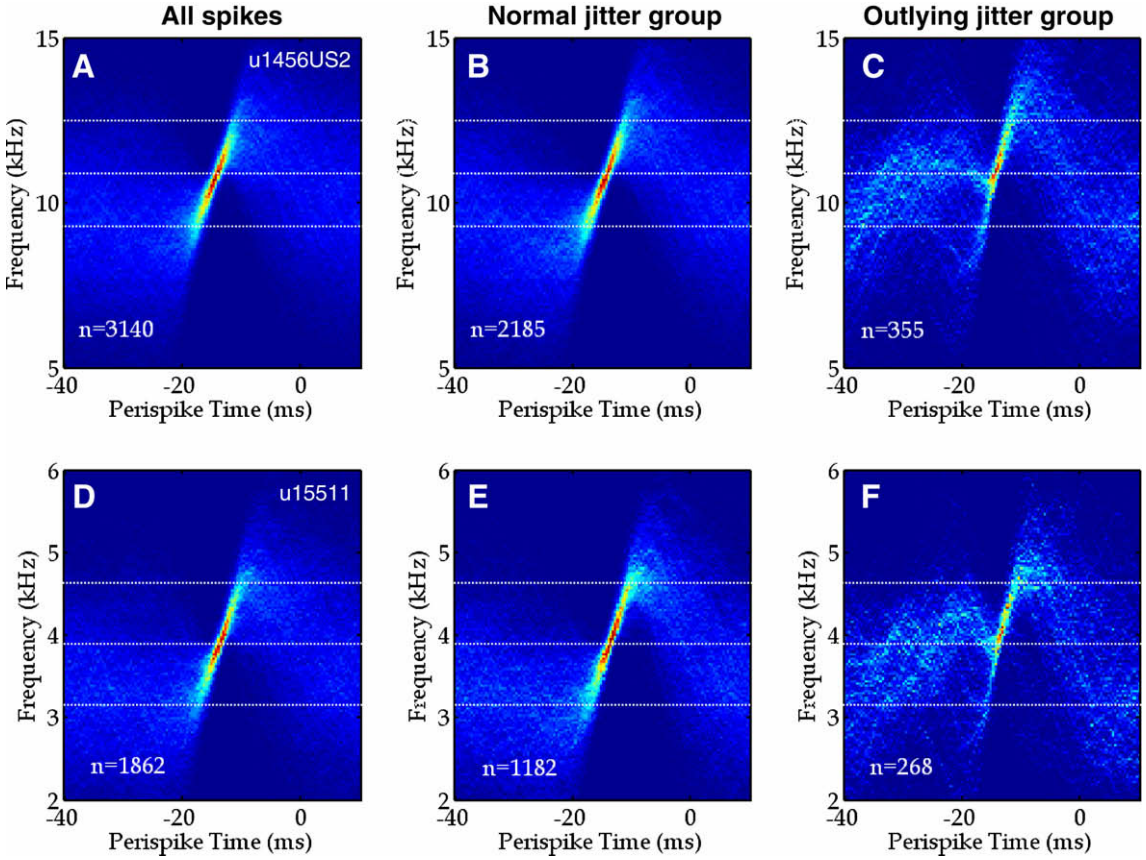

Fig. 5. FM trigger features after sharpening, showing two examples (A–C and D–F). Note in each example, the similarity in trigger feature between the 'all spikes' (original) STRF and the 'normal jitter' STRF, but the greater difference between the 'normal jitter' and 'outlying jitter' (B versus C and E versus F). Horizontal lines are added to facilitate comparison across panels. See Table 1 for a quantitative comparison of the trigger features.

Table 1
Parameters of trigger features extracted from the dejittered STRFs using thresholds set at 20% below the peak values. Note the similarity between the normal jitter group and the original, and its difference against the outlying jitter group. Note also the difference between outlying jitter group and its noise-replaced counterpart.

FM cell	Trigger feature parameter	All spikes (original)	Normal jitter group	Outlying jitter group	Noise-replaced outlying jitter group
U1456US2	FM slope (Hz/m s)	419	454	419	550
	Frequency span (Hz)	1844	2180	1174	1760
	Time span (m s)	4.4	4.8	2.8	3.2
U15511	FM slope (Hz/m s)	165	190	178	224
	Frequency span (Hz)	595	684	357	620
	Time span (m s)	3.6	3.6	2.0	2.8

4.2. Comparison with other dejittering algorithms

Our algorithm of dejittering is similar to those developed by Aldworth et al. (2005) and by Gollisch (2006), in that we share the same goal in trying to sharpen trigger features in the receptive field. However, their algorithms assume a Gaussian distribution of the spike jitter times, whereas ours does not. In the insect mechanical nervous system, where they applied their dejittering algorithms to extract an AM-like trigger feature, the Gaussian assumption seemed to work well. But in the rodent auditory midbrain, the distribution of jitter appeared skewed than Gaussian (Fig. 2C). Our approach also revealed weaker trigger features probably not so evident in the insect nervous system. Each approach may therefore work best for a particular system. Our experimental approach of determining receptive field properties of central auditory neurons using a variety of stimuli (e.g., random FM, triangular FM) allowed additional characterization of the trigger point (see Figs. 2 and 1B) which could be useful in modeling neural response to vocalization sounds (Kao et al., 1997).

4.3. Decomposition of trigger features

The discovery an apparent partial trigger feature of the FM cells is the most interesting finding of this study. This is consistent with our earlier histological findings on FM-sensitive at the auditory midbrain, showing that these cells are typically neurons of large sizes with dendritic fields spanning across adjacent frequency laminae (Poon et al., 1992). One is tempted to speculate that the shorter segment of the 'full trigger feature' is the result of neural integration of synaptic inputs within a single frequency lamina. Our sample of neurons exhibited a preferred FM range of around 0.5 octave. This range corresponds to FM-sensitive cells with dendritic extent of approximately 150 m corresponding to about two frequency laminae in thickness (Poon et al., 1992). Additional evidence is required to test this conjecture, particularly with recordings from FM-sensitive cells of larger frequency range (e.g., 1.0 instead of 0.5 octave); such neurons presumably would have wider dendritic fields. Our recent study in the auditory midbrain (Chiu

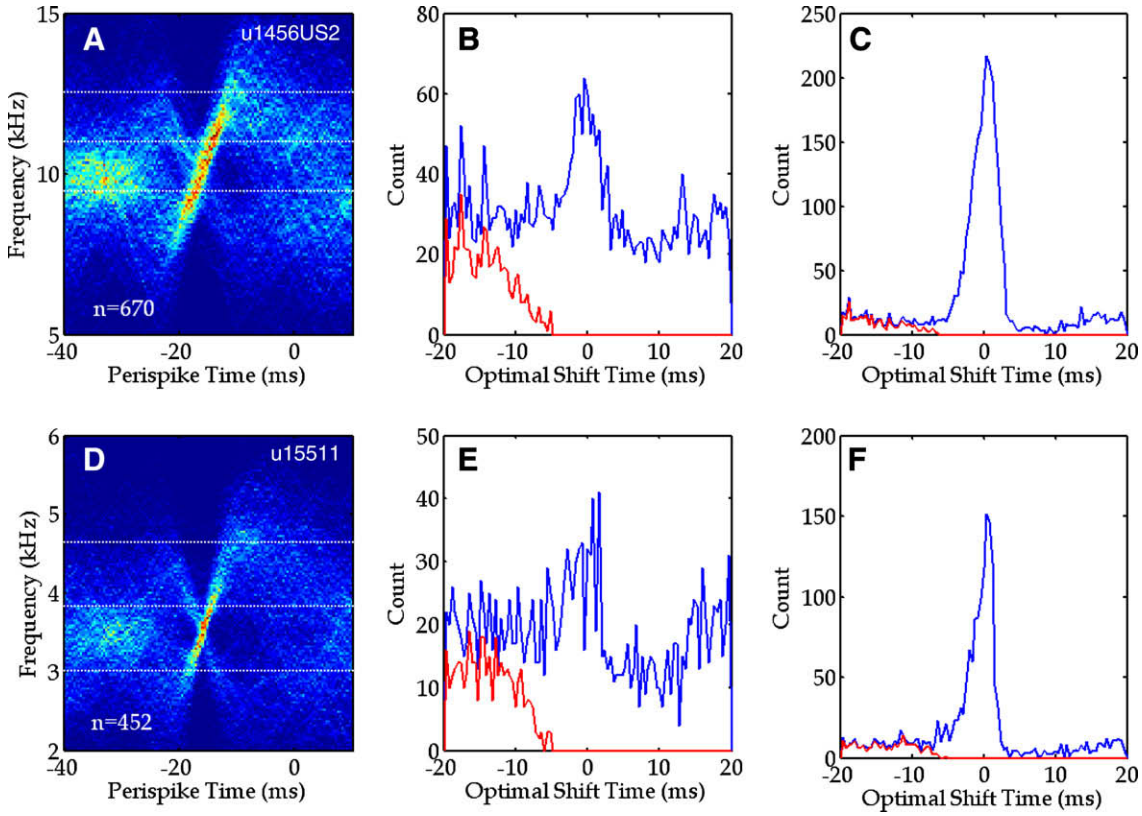

Fig. 6. (A and D): Two examples of sharpened trigger feature of the noise-replaced 'outlying jitter' group (same cells as in Fig. 5). Note the 'FM band' is longer in frequency and time span (see Fig. 5C and F). (B and E): time positions of the outlying jitter group in the noise-replaced datasets (blue line: all spikes including 2/3 noise-replaced signals; red line: outlying spikes that are used to produce A and D). (C and F): Time positions of the outlying spikes in the original datasets (see Fig. 5C and F).

and Poon, 2007) also revealed multiple trigger features in some wide-range STRFs.

4.4. Other non-stimulus related spikes

In some cells, we found small number of spikes that were unrelated to the FM trigger features. This is not surprising since central auditory neurons exhibit spontaneous activity (Bart et al., 2005). These spikes likely formed a noise floor in the jitter adjustment histogram (Fig. 3D). Spikes that occurred as inhibitory rebound might also complicate feature extraction in the STRF. Our study on noise-replaced datasets (Fig. 6) showed that such spontaneous activity (even as high as 67%) did not produce the characteristic partial trigger feature we had observed.

4.5. Complexity of FM coding mechanisms

Our results on different trigger features are consistent with the recent reports on the non-linear properties in FM-sensitive neurons at the cortex (Machens et al., 2004; Christianson et al., 2008; Ahrens et al., 2008). One has to point out that results reported at different neural stations (i.e., midbrain and cortex) should be compared with caution due to the difference in circuit properties. One advantage of studying the auditory midbrain as opposed to the cortex is that its coding mechanism is likely simpler. Furthermore for a simple explanation of results, we used a simple FM stimulus, mono-tone instead of multi-tone. Trigger features that include the simultaneous presence of two or more tones would therefore not be revealed by our present method.

5. Conclusion

Using the information derived from spike jitters, we were able to isolate for the first time different trigger features from STRFs of FM-sensitive cells. Results strongly suggested that in STRF generation, the possibility of more than one trigger features should be considered and the spikes should therefore not be treated indiscriminately with equal weightings.

Acknowledgements

We thank Dr John Brugge for his kind comments. Supported in part by National Science Council, NSC-97-2221-E-218-040, National Health Research Institute, NHRI-EX97-9735EI, Taiwan.

References

Aertsen, A.M., Johannesma, P.I., 1981a. A comparison of the spectro-temporal sensitivity of auditory neurons to tonal and natural stimuli. Biol. Cybernet. 42, 145–156.
Aertsen, A.M., Johannesma, P.I., 1981b. The spectro-temporal receptive field. A functional characteristic of auditory neurons. Biol. Cybernet. 42, 133–143.
Aertsen, A.M., Olders, J.H., Johannesma, P.I., 1981. Spectro-temporal receptive fields of auditory neurons in the grassfrog. III. Analysis of the stimulus-event relation for natural stimuli. Biol. Cybernet. 39, 195–209.
Ahrens, M.B., Linden, J.F., Sahani, M., 2008. Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectrotemporal methods. J. Neurosci. 28, 1929–1942.
Aldworth, Z.N., Miller, J.P., Gedeon, T., Cummins, G.I., Dimitrov, A.G., 2005. Dejittered spike-conditioned stimulus waveforms yield improved estimates of neuronal feature selectivity and spike-timing precision of sensory interneurons. J. Neurosci. 25, 5323–5332.
Andoni, S., Li, N., Pollak, G.D., 2007. Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations. J. Neurosci. 27, 4882–4893.

Bart, E., Bao, S., Holcman, D., 2005. Modeling the spontaneous activity of the auditory cortex. J. Comput. Neurosci. 19, 357–378.
Behne, N., Scheich, H., Brechmann, A., 2005. Contralateral white noise selectively changes right human auditory cortex activity caused by a FM-direction task. J. Neurophysiol. 93, 414–423.
Chang, T.R., Chung, P.C., Chiu, T.W., Poon, P.W., 2005. A new method for adjusting neural response jitter in the STRF obtained by spike-trigger averaging. Biosystems 79, 213–222.
Chen, H., Zeng, F.G., 2004. Frequency modulation detection in cochlear implant subjects. J. Acoust. Soc. Am. 116, 2269–2277.
Chiu, T.W., Poon, P.W., 2007. Multiple-band trigger features of midbrain auditory neurons revealed in composite spectro-temporal receptive fields. Chin. J. Physiol. 50, 105–112.
Christianson, G.B., Sahani, M., Linden, J.F., 2008. The consequences of response nonlinearities for interpretation of spectrotemporal receptive fields. J. Neurosci. 28, 446–455.
Clopton, B.M., Winfield, J.A., 1973. Tonotopic organization in the inferior colliculus of the rat. Brain Res. 56, 355–368.
Cohen, Y.E., Theunissen, F., Russ, B.E., Gill, P., 2007. Acoustic features of rhesus vocalizations and their representation in the ventrolateral prefrontal cortex. J. Neurophysiol. 97, 1470–1484.
deCharms, R.C., Blake, D.T., Merzenich, M.M., 1998. Optimizing sound features for cortical neurons. Science 280, 1439–1443.
Depireux, D.A., Simon, J.Z., Klein, D.J., Shamma, S.A., 2001. Spectrotemporal response field characterization with dynamic ripples in ferret primary auditory cortex. J. Neurophysiol. 85, 1220–1234.
Doupe, A.J., Kuhl, P.K., 1999. Birdsong and human speech: common themes and mechanisms. Annu. Rev. Neurosci. 22, 567–631.
Eggermont, J.J., Aertsen, A.M., Johannesma, P.I., 1983. Prediction of the responses of auditory neurons in the midbrain of the grass frog based on the spectro-temporal receptive field. Hear. Res. 10, 191–202.
Felsheim, C., Ostwald, J., 1996. Responses to exponential frequency modulations in the rat inferior colliculus. Hear Res. 98, 137–151.
Gollisch, T., 2006. Estimating receptive fields in the presence of spike-time jitter. Network 17, 103–129.
Hart, H.C., Palmer, A.R., Hall, D.A., 2003. Amplitude and frequency-modulated stimuli activate common regions of human auditory cortex. Cereb. Cortex 13, 773–781.
Heil, P., Rajan, R., Irvine, D.R., 1992. Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. I: Effects of variation of stimulus parameters. Hear Res. 63, 108–134.
Hermes, D.J., Aertsen, A.M., Johannesma, P.I., Eggermont, J.J., 1981. Spectro-temporal characteristics of single units in the auditory midbrain of the lightly anaesthetised grass frog (*Rana temporaria* L) investigated with noise stimuli. Hear. Res. 5, 147–178.
Husain, F.T., Tagamets, M.A., Fromm, S.J., Braun, A.R., Horwitz, B., 2004. Relating neuronal dynamics for auditory object processing to neuroimaging activity: a computational modeling and a FMRI study. Neuroimage. 21, 1701–1720.
Kao, M.C., Poon, P.W., Sun, X., 1997. Modeling of the response of midbrain auditory neurons in the rat to their vocalization sounds based on FM sensitivities. Biosystems. 40, 103–119.
Klein, D.J., Depireux, D.A., Simon, J.Z., Shamma, S.A., 2000. Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. J. Comput. Neurosci. 9, 85–111.
Luo, H., Wang, Y., Poeppel, D., Simon, J.Z., 2006. Concurrent encoding of frequency and amplitude modulation in human auditory cortex: MEG evidence. J. Neurophysiol 96, 2712–2723.
Luo, H., Boemio, A., Gordon, M., Poeppel, D., 2007. The perception of FM sweeps by Chinese and English listeners. Hear Res. 224, 75–83.
Machens, C.K., Wehr, M.S., Zador, A.M., 2004. Linearity of cortical receptive fields measured with natural sounds. J. Neurosci. 24, 1089–1100.
Malmierca, M.S., 2004. The inferior colliculus: a center for convergence of ascending and descending auditory information. Neuroembryol. Aging 3, 215–229.
Poon, P.W., Chiu, T.W., 2000. Similarities of FM and AM receptive space of single units at the auditory midbrain. Biosystems 58, 229–237.
Poon, P.W., Yu, P.P., 2000. Spectro-temporal receptive fields of midbrain auditory neurons in the rat obtained with frequency modulated stimulation. Neurosci. Lett. 289, 9–12.
Poon, P.W., Chen, X., Hwang, J.C., 1991. Basic determinants for FM responses in the inferior colliculus of rats. Exp Brain Res. 83, 598–606.
Poon, P.W., Chen, X., Cheung, Y.M., 1992. Differences in FM response correlate with morphology of neurons in the rat inferior colliculus. Exp. Brain Res. 91, 94–104.
Qi, Y., Casseday, J.H., Covey, E., 2007. Response properties and location of neurons selective for sinusoidal frequency modulations in the inferior colliculus of the big brown bat. J. Neurophysiol. 98, 1364–1373.
Rees, A., Kay, R.H., 1985. Delineation of FM rate channels in man by detectability of a three-component modulation waveform. Hear Res. 18, 211–221.
Rees, A., Møller, A.R., 1983. Responses of neurons in the inferior colliculus of the rat to AM and FM tones. Hear. Res. 10, 301–330.
Sen, K., Theunissen, F.E., Doupe, A.J., 2001. Feature analysis of natural sounds in the songbird auditory forebrain. J. Neurophysiol. 86, 1445–1458.
Shechter, B., Depireux, D.A., 2007. Stability of spectro-temporal tuning over several seconds in primary auditory cortex. Neuroscience 148, 806–814.
Suga, N., 1968. Analysis of frequency-modulated and complex sounds by single auditory neurons of bats. J. Physiol. 198, 51–80.
Theunissen, F.E., Sen, K., Doupe, A.J., 2000. Spectral–temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J. Neurosci. 20, 2315–2331.
Valentine, P.A., Eggermont, J.J., 2004. Stimulus dependence of spectro-temporal receptive fields in cat primary auditory cortex. Hear. Res. 196, 119–133.
Whitfield, C., Evans, E.F., 1965. Responses of auditory cortical neurons to stimuli of changing frequency. J. Neurophysiol. 28, 655–672.