Measuring the FFR to Speech Syllables

WHAT TO LOOK FOR WHEN USING CABR/FFR

Source/Filter Model of Speech

- Source
  - Glottal pulse: sustained
  - Resonated
  - Vowels and voicing: related to F0 (resilient to noise) D, E, F
- Filter
  - Partial or full obstruction by articulators
  - Tend to have transient portion or aperiodic content
  - Consonants: related to F1 (more susceptible to noise) V,A,O

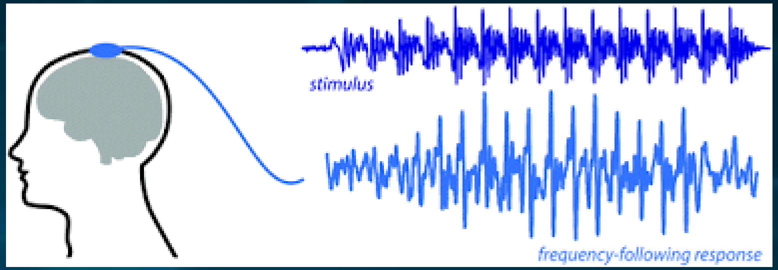
cABR/FFR response to speech syllable

- Has both a transient portion and a sustained portion
  - Transient: onset and offset responses
  - Sustained: FFR, which encodes vowels
- In noise, the transient portion is more vulnerable, and may even be obliterated, while sustained portion is resistant to noise.
  - Consonants are rapid, relatively low in amplitude and transient (rapid change). Comprehending consonants harder in noise.
  - Vowels are periodic, sustained with relatively higher amplitude. Vowels tend to be preserved in noise.

Quality of responses

- Stimulus-to-response correlations and response consistency measurements give insights about the fidelity of the system
- Peak latency, interpeak interval, slope are measurements that provide insight into the accuracy with which brainstem nuclei synchronously response to acoustic stimulation.
- Peak amplitude, RMS amplitude of activation give information about the robustness which the brainstem nuclei respond and the size of spectral components within the response.

Advantage of FFR: Fidelity



- The fidelity of complex signals, such as speech or music, is maintained at the recording electrode.
- The AEP response can be played back through a speaker, and is identifiable as the speech that evoked it.
- The FFR contains more detail regarding speech that evoked it than cortical potentials (which tend to signal detection of stimulus or change in stimulus), and it contains more detail than ABR recorded to simple stimuli like clicks or tones.
- But, IC has an upper limit to frequency coding ability that prevents the FFR from being a perfect representation of the evoking stimulus.

FFR and Phase

- The FFR to a pure-tone is phase-dependent.
  - Timing of response is dependent on phase of the stimulus (this maps onto the fine structure of stimulus).
  - BUT, transient peaks in the response are phase-independent (this maps onto envelope of stimulus).
  - FFR to complex stimulus (syllable) will contain a mixture of phase-dependent and phase-independent components.
- Phase can be used as analysis tool (i.e. cross-phaseogram)

Accuracy, Robustness: Timing, Magnitude

- Timing: latency is measured. Either referenced to stimulus onset or to other waves.
  - Transient peaks: response to stimulus onset and stimulus offset
  - Tonic peaks: response to periodicity of the stimulus during the voiced period of speech (vowel, or voiced formant transition)
- Magnitude: the size of the omnibus response can be measured in the time domain by root-mean-square amplitude over a region of interest.
  - Magnitude of consonant-transition period of a response to consonant-vowel syllable is measured.
  - Magnitude of individual peaks has not been found to be salient, and is usually not measured (or at least, not yet).

Frequency

- Frequency content and magnitude: the frequency component of a response can be obtained through a Fourier transform. Key frequencies in the spectrum are identified and measured.
  - Fundamental frequency (see Pitch Error)
  - Harmonics (not used much at this time)

Inferior Colliculus

- Is excellent at reproducing very fine temporal distinctions
- "Jitter" in a neural system
  - If subsequent responses differ by fractions of milliseconds, when averaged, they could cancel each other out.
  - Poor synchrony (either in short term or over time) results in jitter, meaning fine temporal distinctions are lost or weakly encoded.
  - Jitter will be apparent by measuring Response Consistency.
- "Stable neural responses to sound may be integral to the emergence of sensorimotor synchronization skills."

Intrinsic Factors:

- Stability of response morphology can be assessed by inter-response correlation.
  - AEPs are expected to be highly consistent over recordings, and over time (from session to session)
  - Poor consistency indicates abnormal response.
- Magnitude of background noise level (averaged response to silence) can be measured.
  - AEPs generally have little intrinsic noise.
  - Intrinsic noise can be a warning of abnormality.

Recording Vowels

- Periodicity is imparted by the vocal cord vibration, the rate of which is expressed as pitch.
  - A voice can be categorized as globally high or low on average, with variation above or below average pitch. Absolute pitch does not determine intelligibility.
  - Voicing sets up harmonic series (resonance peaks in vocal tract) including fundamental frequency and integer harmonics due to selective filtering by the articulators.
  - Identify of vowel comes from relationship between the resonance peaks known as formants.
- Fidelity of encoding of vowels by FFR gives useful information.

Vowels, interesting findings

- Fidelity of brain's encoding of fundamental frequency of speech is indicative of the ability to hear in noise.
- Strength of phase-locking to fundamental frequency in FFR to a speech sound of a foreign language relates to the ability to learn that language.
- Poor comprehension of speech in noise can also be a complaint of persons who have had concussions. Their speech-evoked F0 is usually diminished, even after the concussion has been judged resolved.

Harmonic content

- Response to harmonic content of stimulus is phase-dependent: brainstem activity to contrasting vowels can be distinguished by spectral peaks in the FFRs.
  - Individuals with dyslexia or poor readers show reduced harmonic encoding at frequencies corresponding to formant frequencies.
  - Auditory enrichment via musical training boosts the harmonic representation of speech formants in the FFR.
- At this point, only one vowel is used in clinical testing. Research studies have looked at encoding of other vowels.

Consonants

- Consonants: are partial or full obstructions of the breath that produces voicing in speech.
- The response to a consonant in isolation would be expected to be a neural transient (essentially like reaction to click), but since it is very hard to isolate consonants totally from vowels, it is mixed.
- For FFR measurements, consonants may be defined as the *transition to or from a conjoined fully-voiced vowel.*

Stimulus token: 40 ms DA

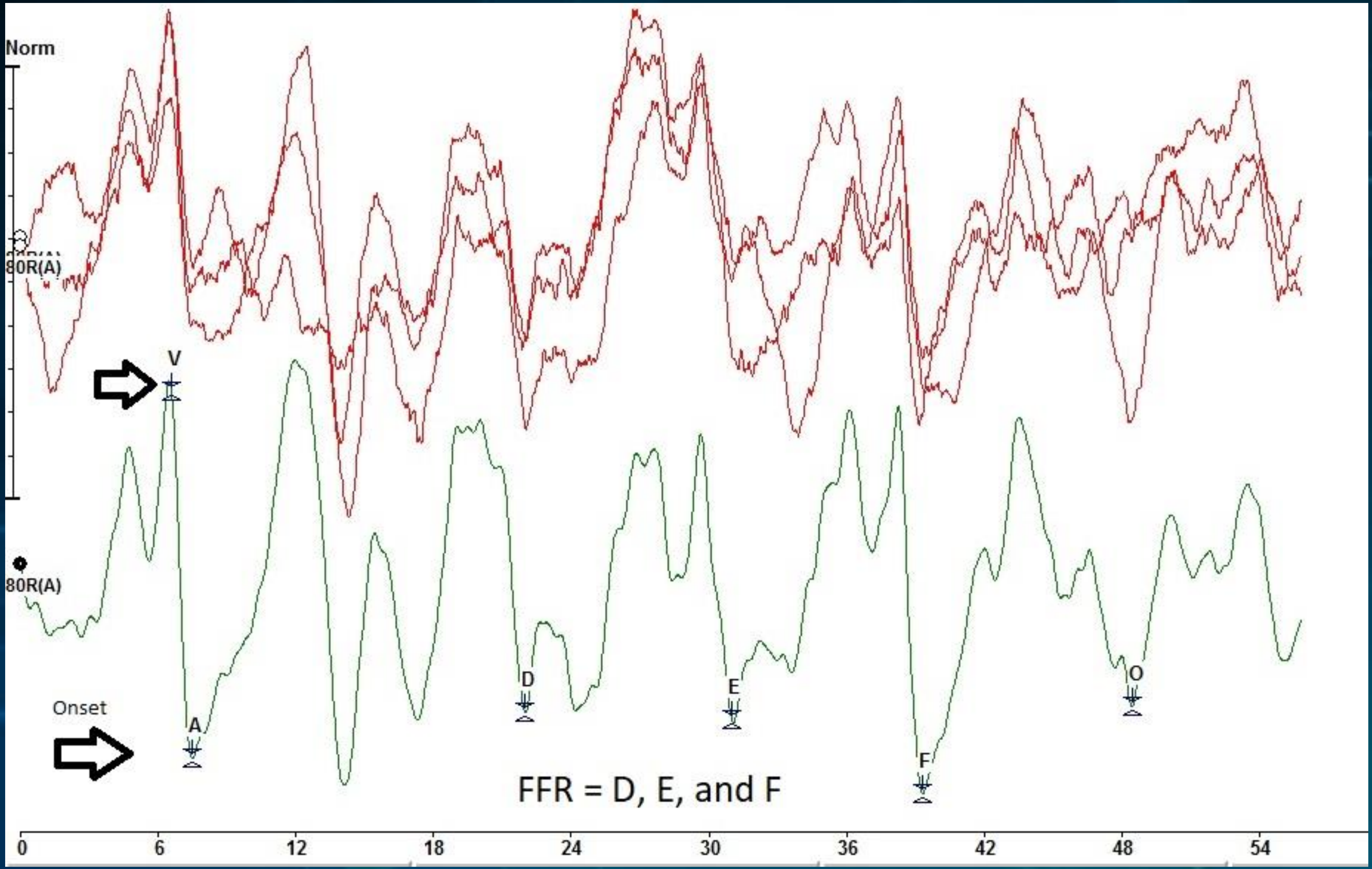
- 40 msec DA was chosen (and synthesized in lab) because
  - 1. it contains key acoustic elements important for speech perception
  - 2. It contains phonemes present in most world languages,
  - 3. Is short enough to be amendable to measurement using evoked response recording.
- It progresses from a relatively broadband release burst to a fully-voiced transition to a vowel, but the vowel is truncated.
- Norms across the age span are available for both the timing of well-characterized discrete peaks and for frequency-domain representation of features of stimulus.

40 ms DA

- Syllable presented monaurally, in alternating polarities, at 80 dB SPL to right ear via insert earphones, with inter-stimulus interval of 51 msec.
- Differentially recorded from Cz (active) to ipsilateral earlobe (reference), with forehead as ground.
- Three blocks of 2000 sweeps [or two blocks of 3000] collected in quiet
    - Some studies have also collected responses in noise.
- Waveforms averaged online with a recording time window spanning 10 msec prior to the onset and 20 msec after the offset of the stimulus.

Waveforms

- Initial waves are similar to those observed to click stimuli. (Assuming normal click ABR):
  - Wave I may be visible (generated at VIIIth Nerve)
  - Wave III is usually visible (generated at low brainstem)
  - Wave V is always visible, but is delayed approximately 1-2 msec from click ABR. (Neural conduction time to brainstem is approximately 7 msec). Represents onset of speech stimulus at upper brainstem).
  - Wave V is positive, and is paired with Wave A which is the trough after Wave V. Wave is the only positive way measured in the speech FFR.
- Wave D, E, F represent phase-locking to the fundamental frequency of the stimulus, arising in response to the periodic information in the vowel at the frequency of the sound source (glottal pulse).
  - Period between Waves D, E, F corresponds to the fundamental frequency of the stimulus (F0), whereas the peaks D, E, F represent phase locking at the frequencies of the first formant (F1).


Norm
80R(A)
80R(A)
V
Onset
A
D
E
F
O
FFR = D, E, and F
0
6
12
18
24
30
36
42
48
54

Peaks of Waveform to 40 ms DA

- Release burst: Wave V and A (Onset)
  - Usually referred to as the “V/A Complex”: 10 ms burst
  - Represents highly synchronized neural response to the onset of the stimulus
  - Wave V is the only positive waveform used; all others are negativities.
  - (Peak C represents response to onset of voicing, 10 ms after stimulus onset.) No longer measured.)
- Periodicity peaks: Wave D, E, F
  - Latency of the peaks can be measured.
  - They are usually considered as a grouping. Period between response peaks corresponds to wavelength of F0 of utterance.
  - Represents glottal pulsing of vocal folds.
- Offset: Wave O
- Encodes the cessation of sound (temporal offset)

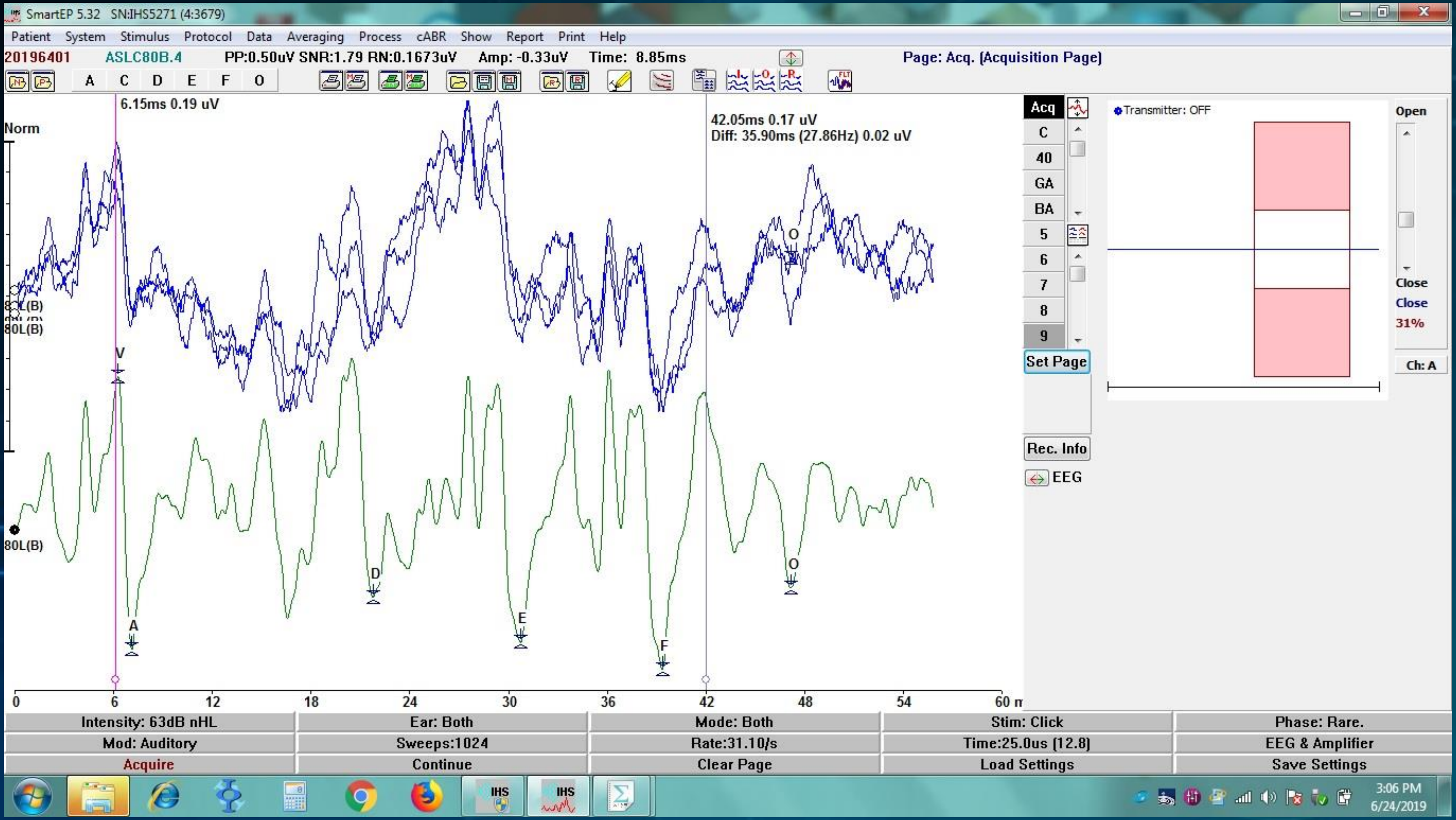
Latency

- Neural response time for impulses to propagate through the nervous system is approximately 6 ms
  - Onset response (Wave V and A) occur at latencies before 10 msec
    - Wave V at approximately 6 ms, Wave A about 1 ms later.
  - FFR waves show up at approximately 22, 30 and 39 msec.
  - The offset response, Wave O, shows up at approximately 48 msec.
- Transient portions (V, A, O) are independent of sustained portions (D, E, F).

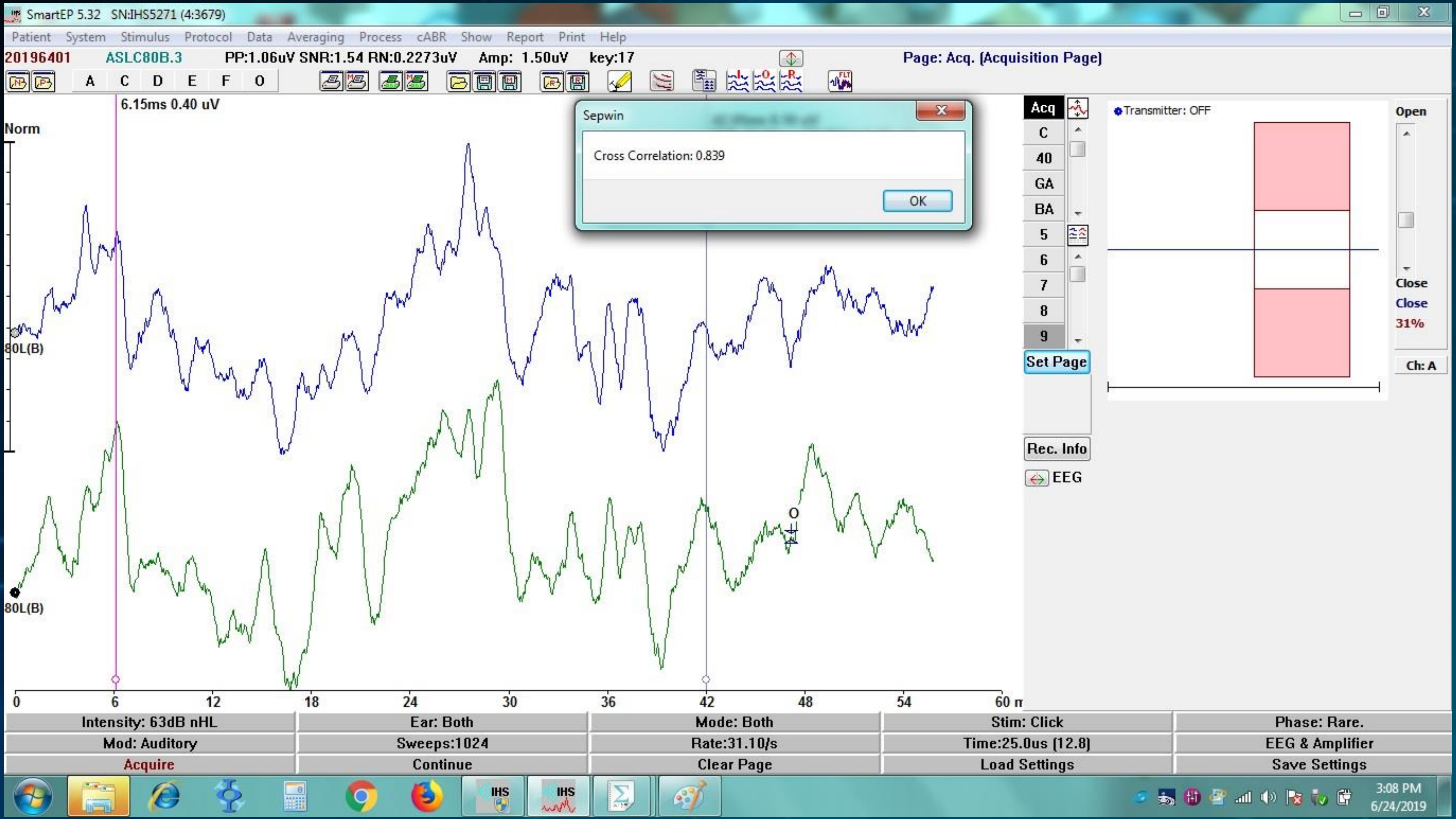
Effect of Noise (generally)

- Addition of background noise interferes with normal brainstem encoding.
- Onset responses are most degraded or even obscured
- If the onset responses are present, they are delayed relative to quiet.
- Peak F tends to be resilient: no change in noise.
- FFR remains present in most subjects, but peak amplitudes are affected.
- Offset response is often degraded or obscured.

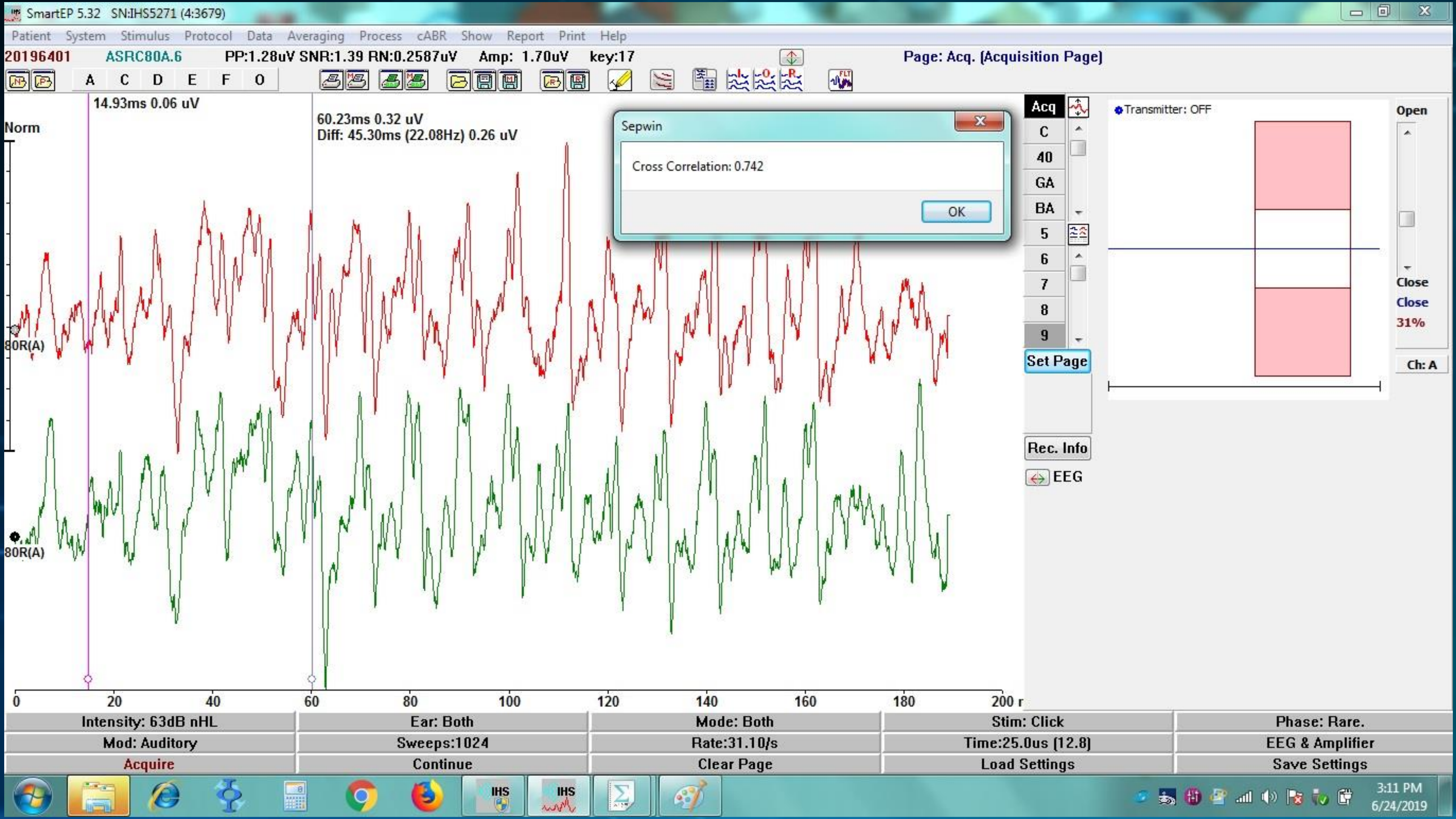
Normal 40 msec DA



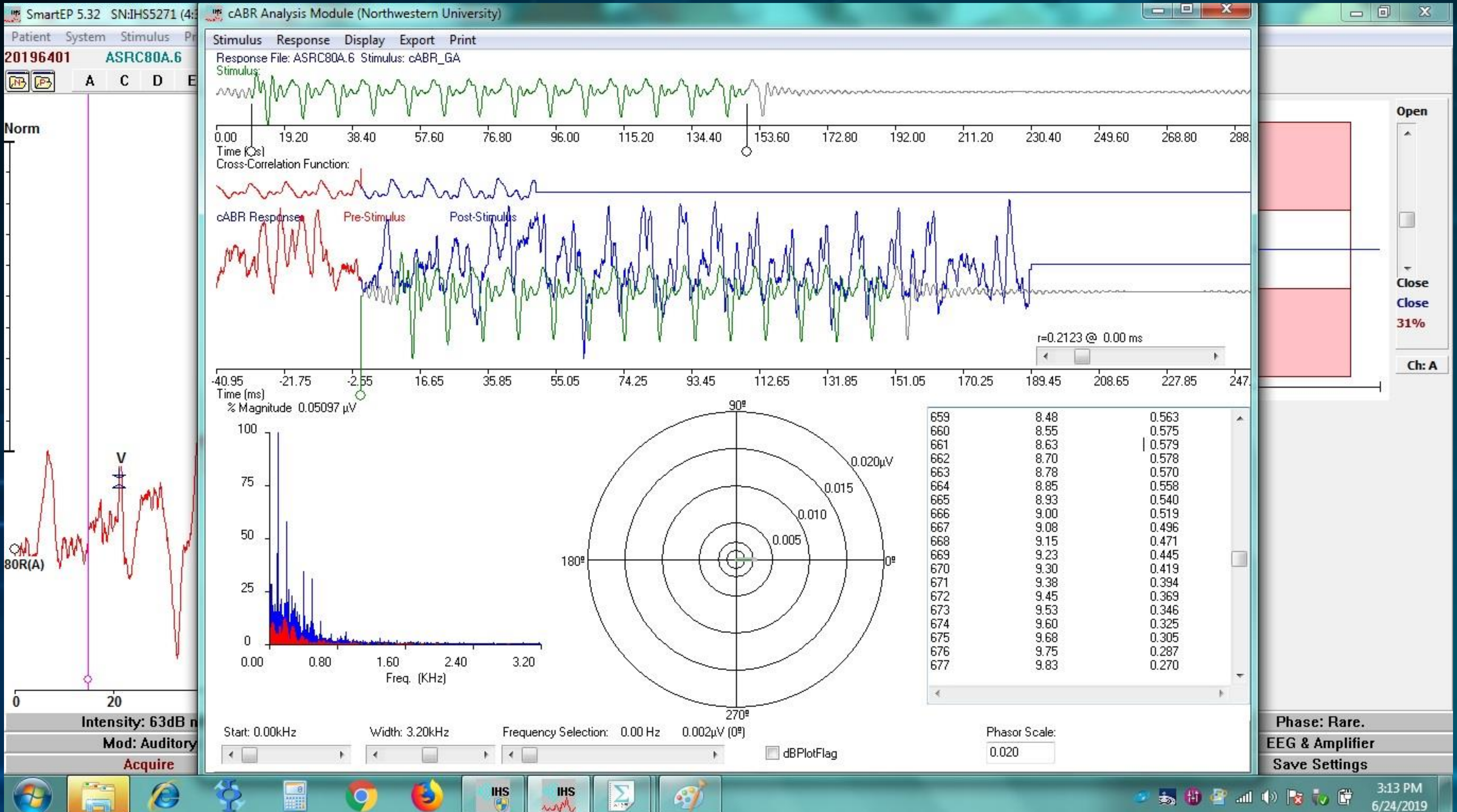
Normal Response Consistency, 40 ms DA



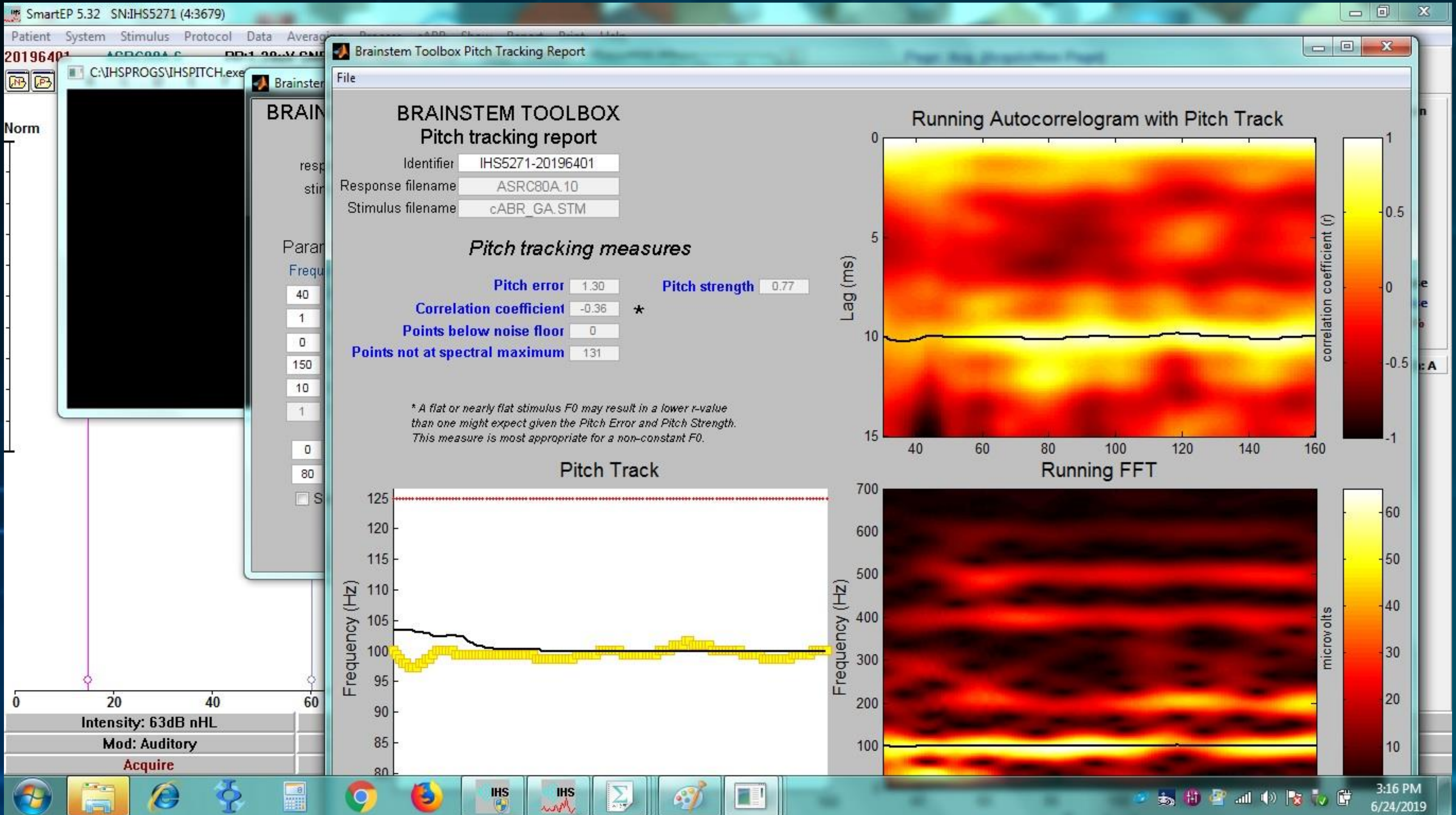
Normal GA CC, 15 to 60 ms



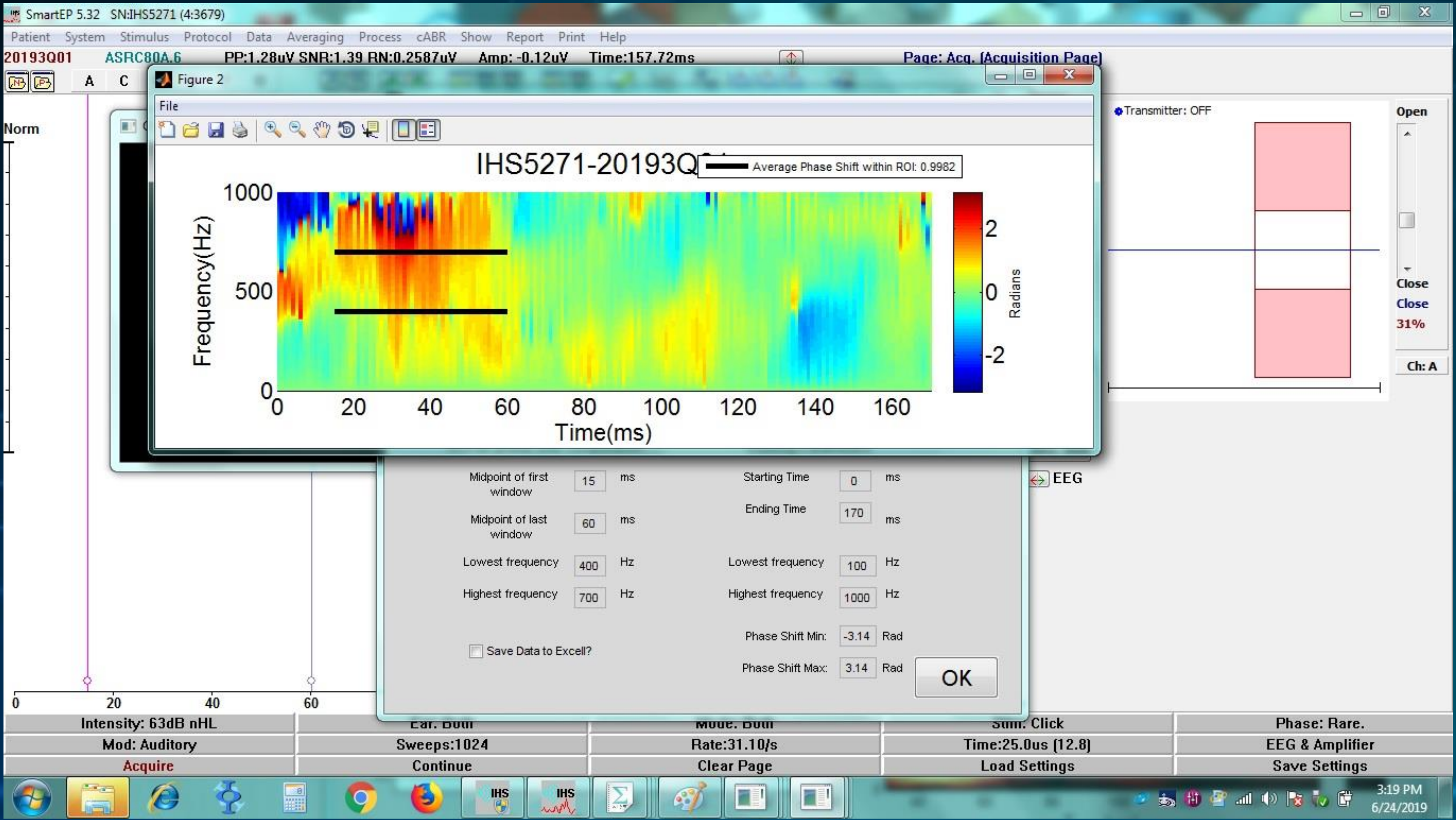
Normal Stimulus-Response Consistency



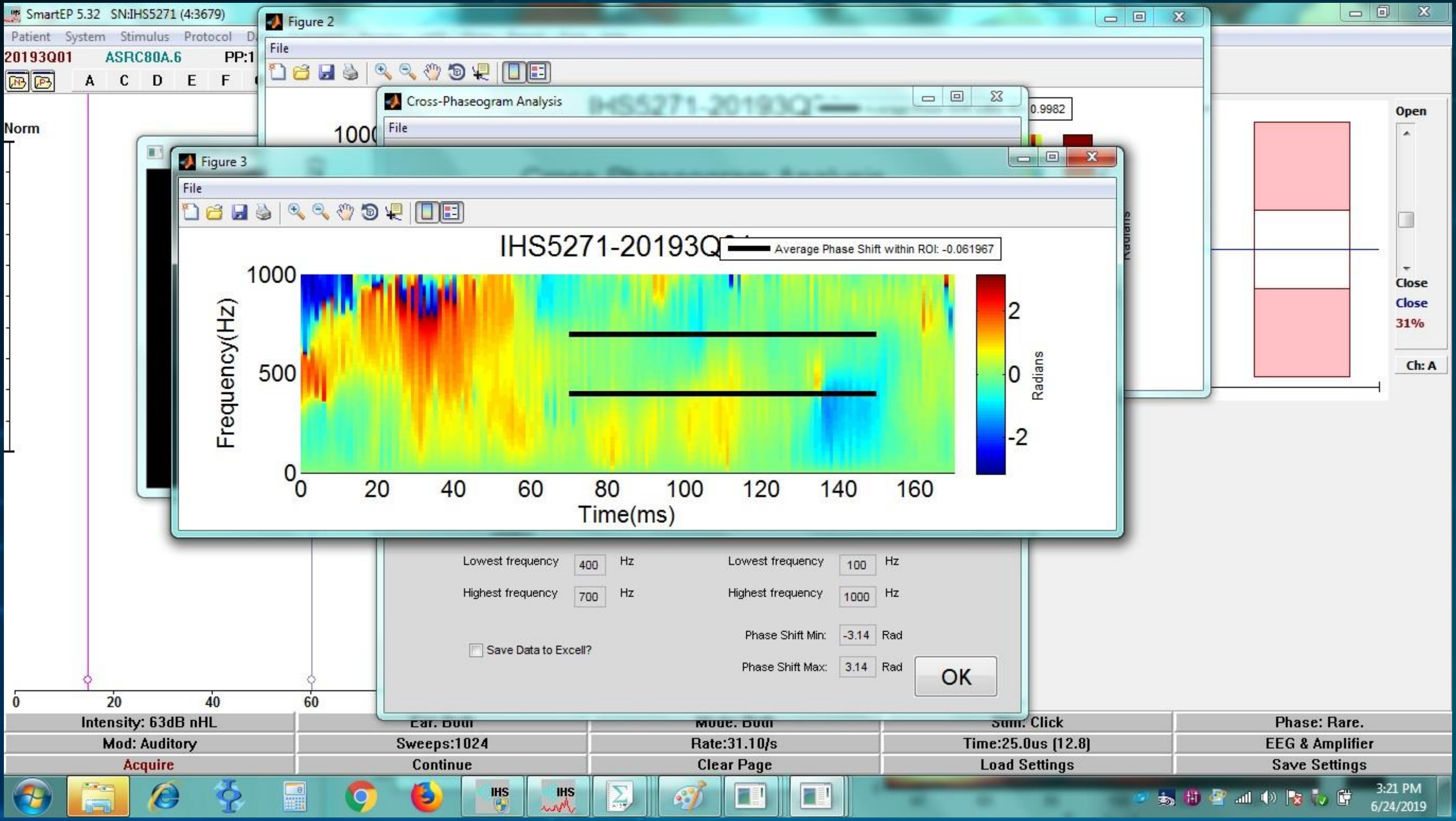
Normal Pitch Tracking, GA



Normal Cross-phaseogram, 15 to 60 ms



Normal Crossphaseogram, 70 to 150 ms



Signs of Abnormal Encoding

- Click ABR latency is normal, although cross-correlation may not be.

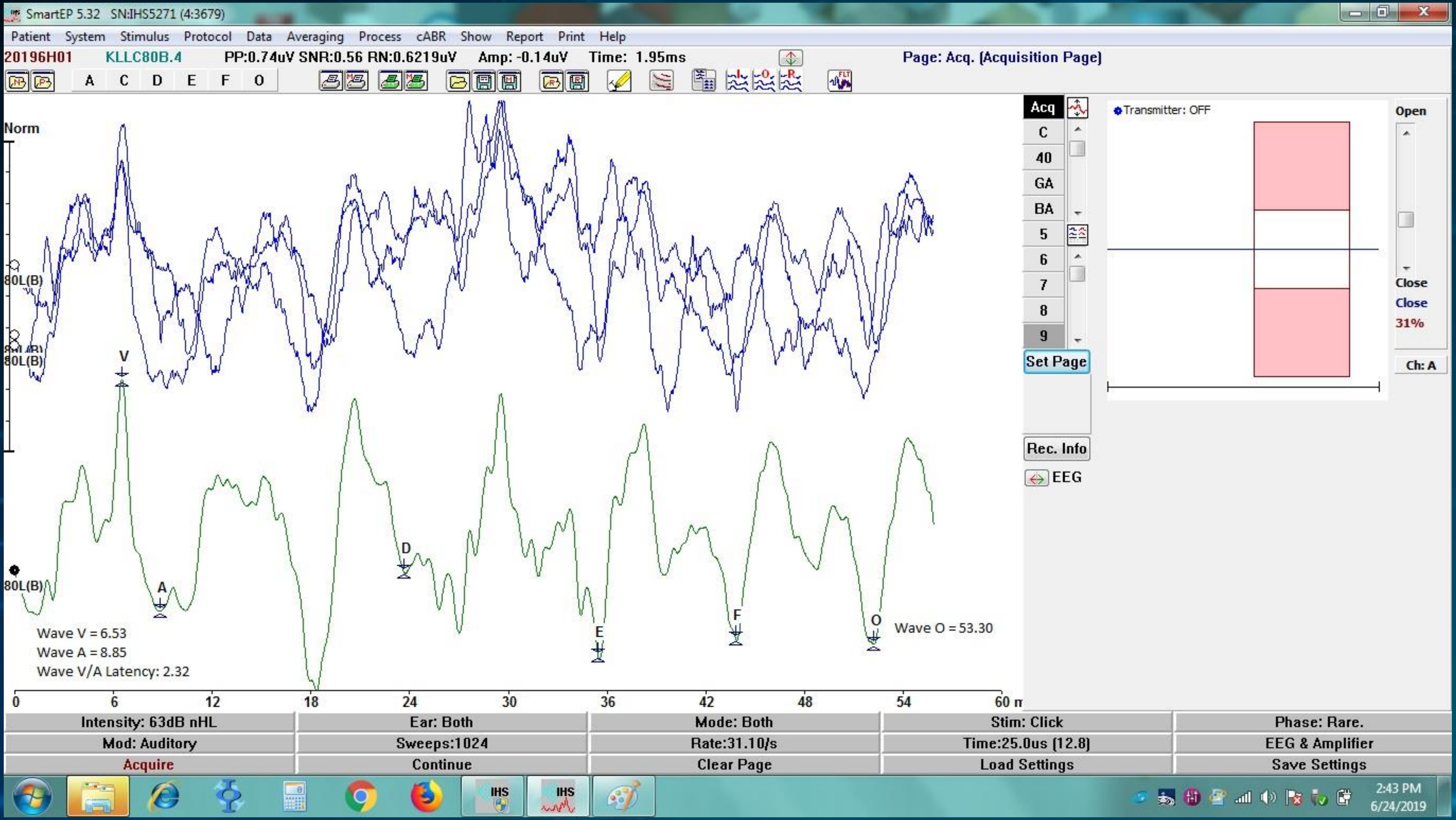
(Click ABR and Speech FFR are not the same. Usually longer conduction time.)

- Wave V may be delayed.
  - Delayed Wave V is found with autism.
  - Some show delay of all waves.
  - Most do NOT have a Wave V delay
- Wave A is often delayed.
  - Very typical with language learning disorders.
  - Normal V with delayed A results in lower slope measurement, which indicates less neural synchrony. (Magnitude/time for V/A complex).

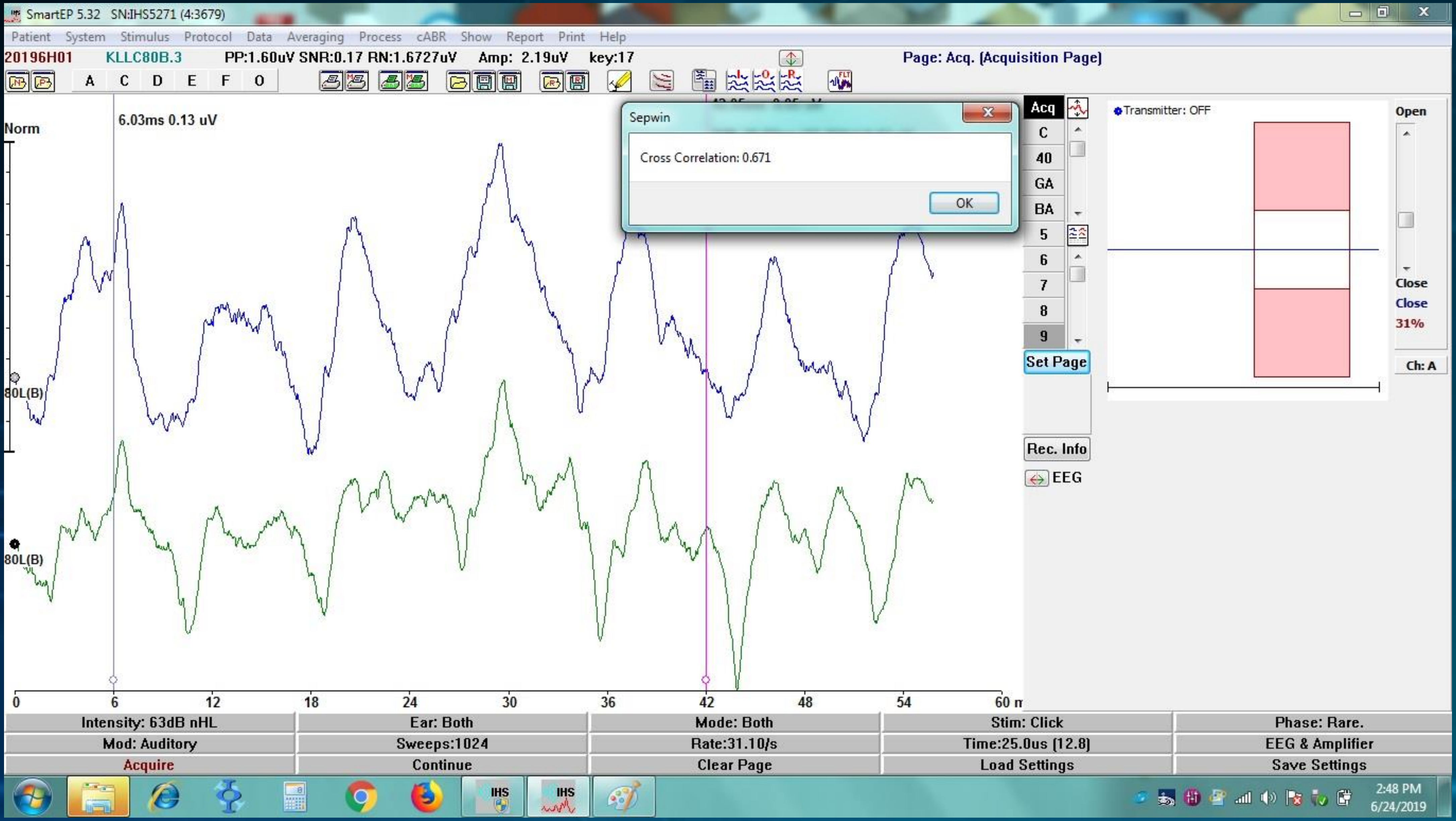
Signs of Abnormal Encoding

- (Wave C was thought to represent onset of voicing. Not used now).
- Waves D, E, F are usually in synch
  - This response to vowel is independent of onset or offset response
  - One of the three can be delayed: we don't know of any significance for this.
  - If all three are delayed, then Pitch Tracking problem.
  - It is common in normals for Wave F to be identical if measurements are made from both ears. Not so for Wave D and E.
- Wave O represents the Offset response.
  - Wave O is often delayed or absent with LD children.
- When deciding whether present or absent, examine individual runs: if have 3, should have concurrence on two out of three

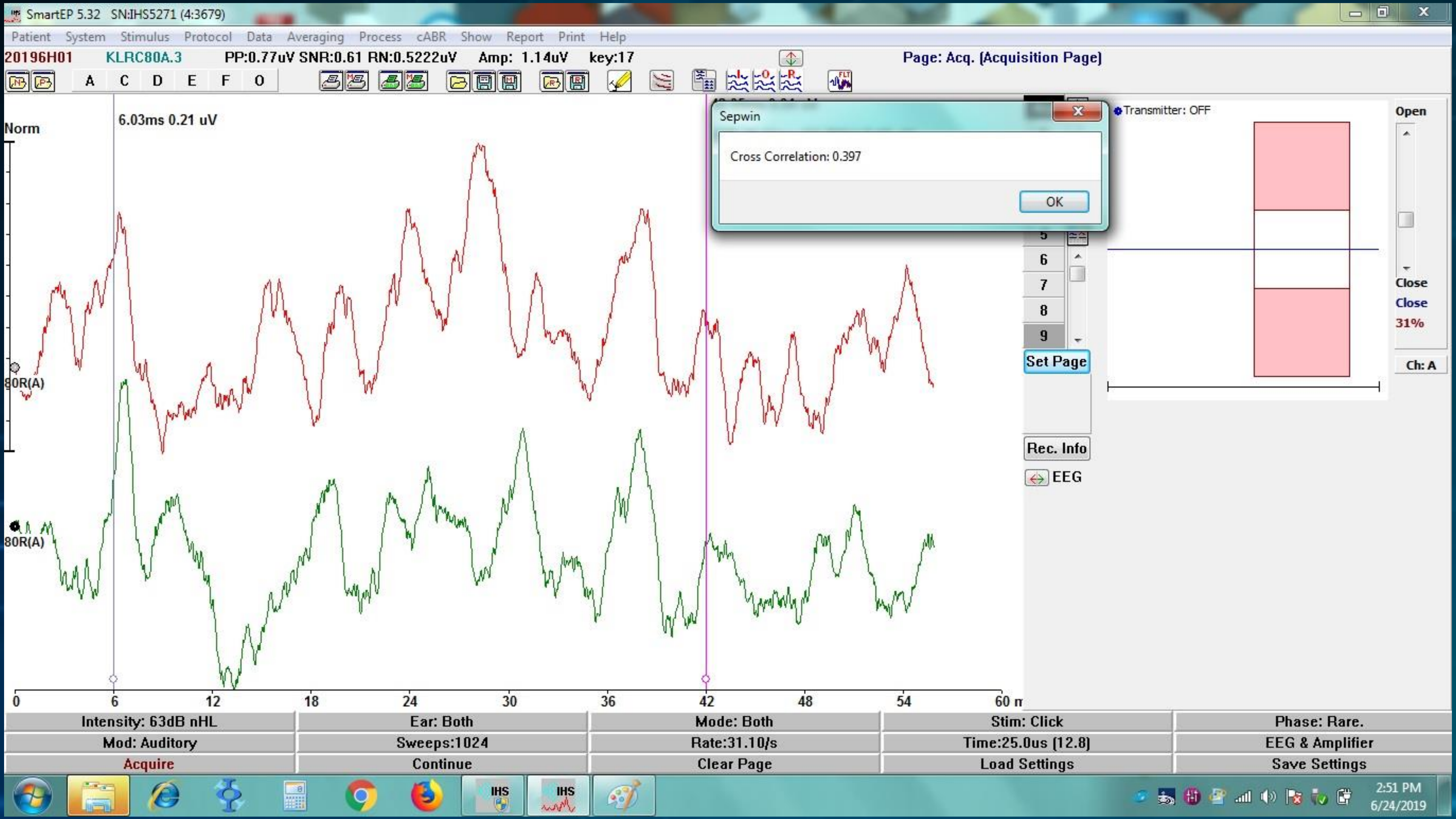
Delayed Onset and Offset



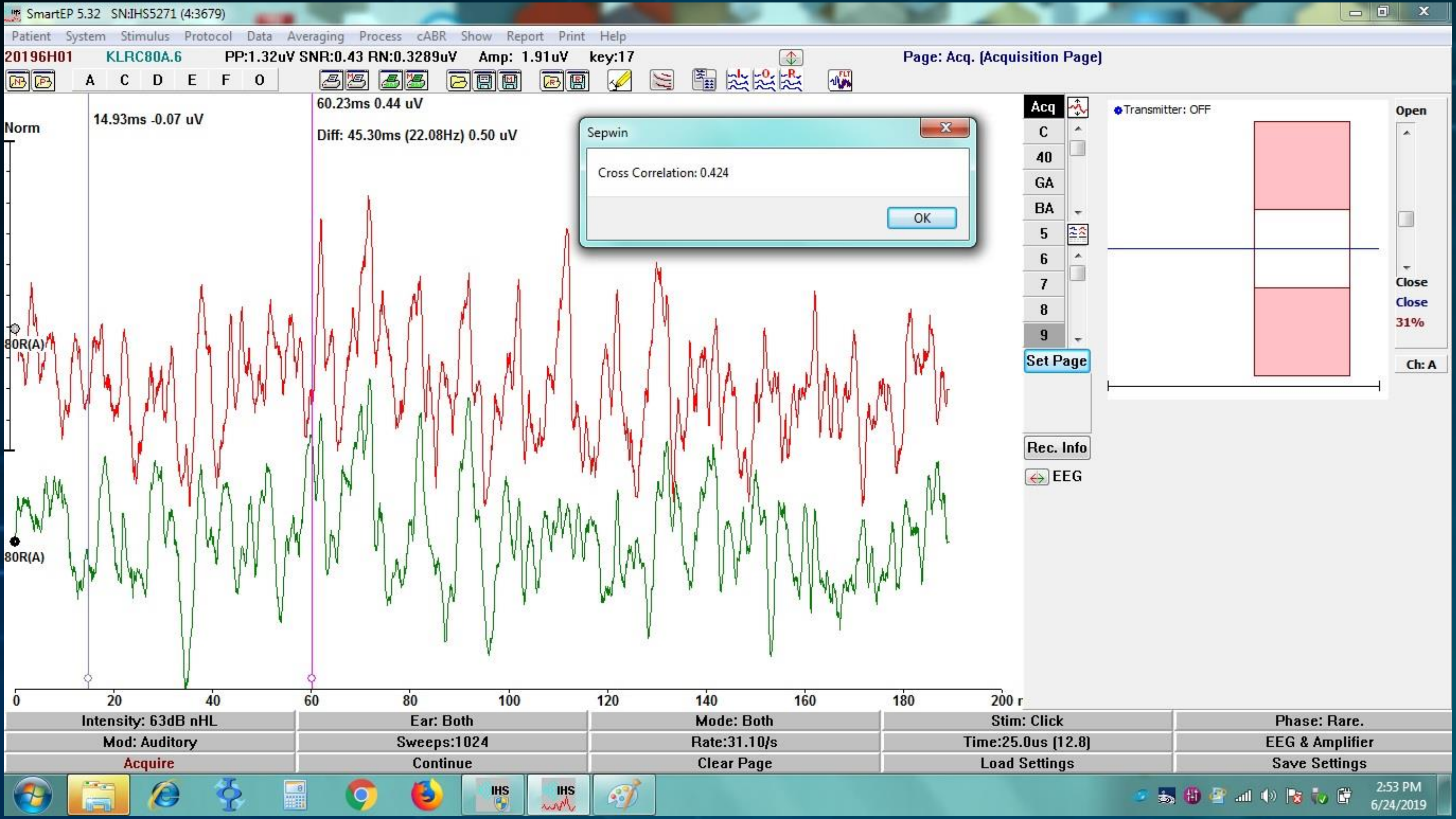
Low CC, 40 ms DA, left



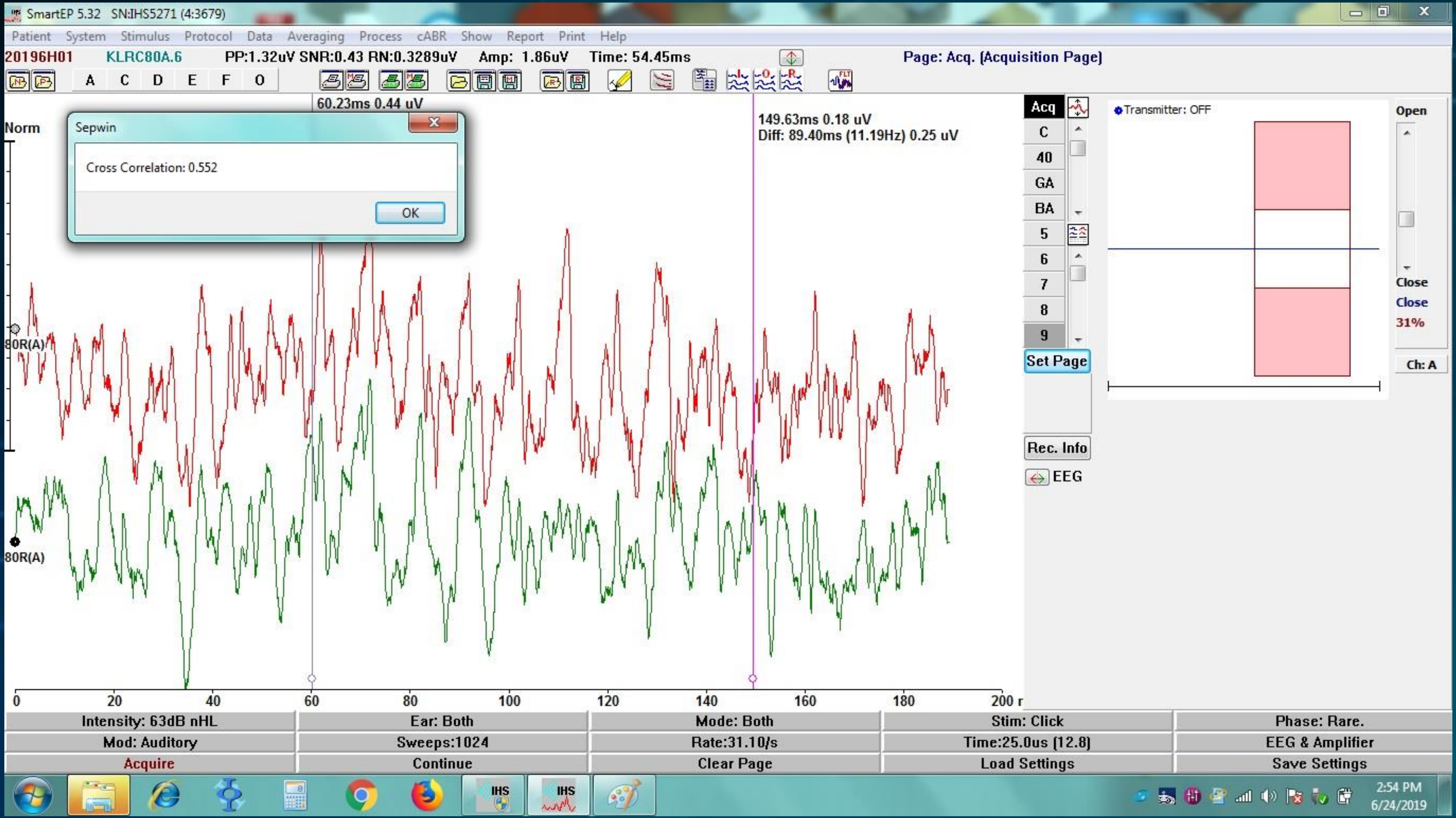
Low CC, 40 ms DA, right



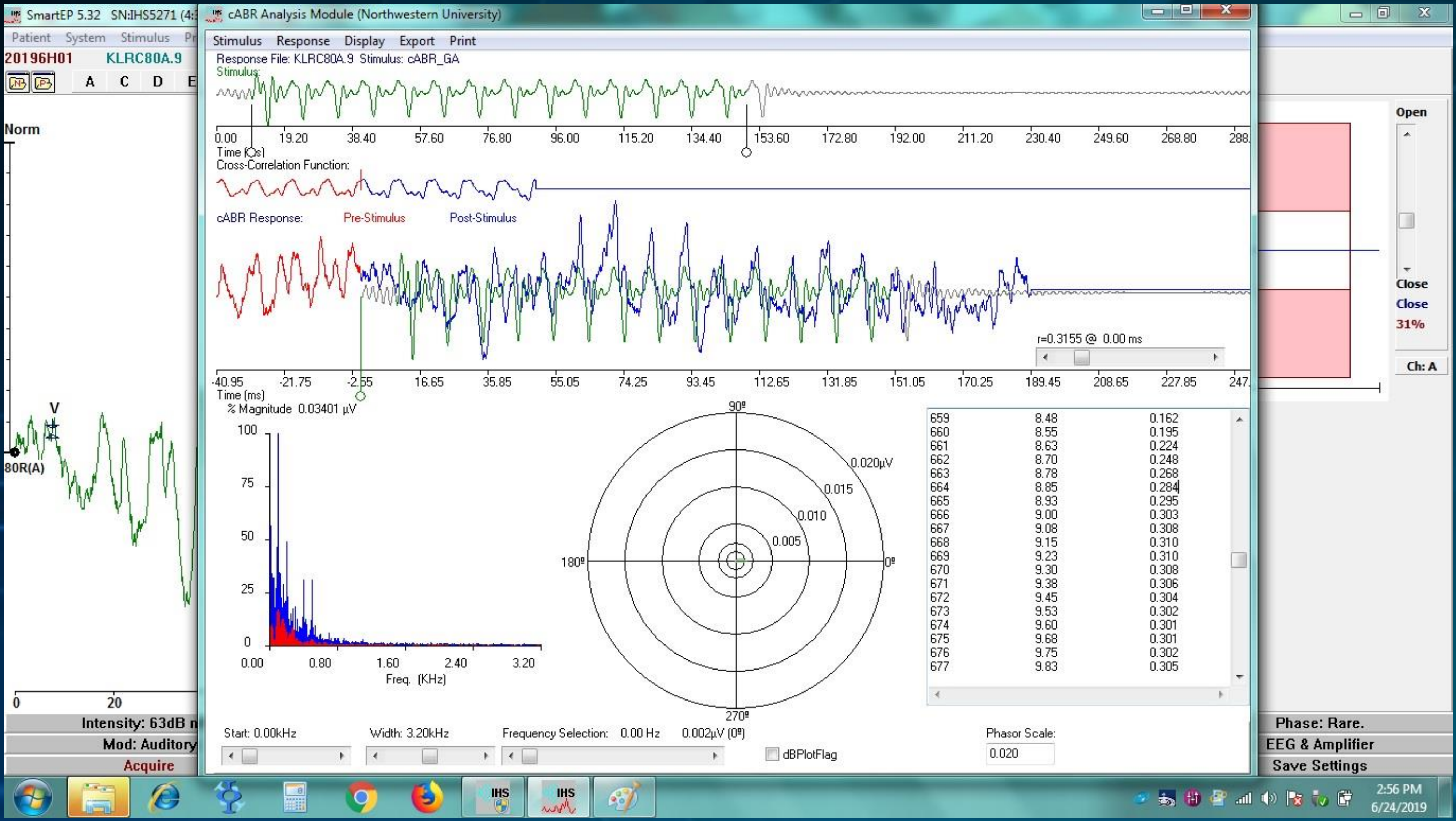
Low CC, GA, 15 to 60 ms



Low CC, GA, 60 to 150 ms



Low Stimulus-Response Correlation, GA



Abnormal Pitch Tracking, GA