

Human Sound: Formant: what is a formant?

Formant was defined by [Gunnar Fant \(1960\)](#): 'The spectral peaks of the sound spectrum $|P(f)|$ are called *formants*'. This definition is the one broadly used in acoustics research and industry. For instance, [Benade \(1976\)](#) uses a similar definition of formant: 'The peaks that are observed in the spectrum envelope are called **formants**'. In parts of the speech research community, however, 'formant' has come to have other meanings. This page discusses the different usages.

After defining formant, [Fant \(1960\)](#) then defines resonance frequencies of the vocal tract in terms of a gain function $T(f)$ of the vocal tract: 'The frequency location of a maximum' in $|T(f)|$, i.e., the *resonance frequency*, is very close to the corresponding maximum in spectrum $P(f)$ of the complete sound.' He then writes: 'Conceptually these should be held apart but in most instances resonance frequency and formant frequency may be used synonymously.' Hence the problem: resonance and formant are indeed conceptually distinct. Several examples below make this clear. However, some writers about the voice use the terms interchangeably.

There is even a third meaning in voice research. The acoustics of the vocal tract are often modelled using a mathematical model of a filter ([Atal and Hanauer, 1971](#)). The frequencies of the poles of this filter model fall close to those of the formants. As a result, some voice researchers now refer to the frequencies of the poles as formants. So, to some voice researchers, the formant refers to a peak in the spectrum (a property of the sound of the voice), to others it refers to a resonance of the vocal tract (a physical property of the tract), while to a third group it refers to the pole in a mathematical filter model (a property of a model).

In the broader field of acoustics, formant retains its original meaning: a broad peak in the spectral envelope of the sound (of a voice, musical instrument, room etc). When referring to the formant at about 400 Hz in the sound of the French horn, it is obviously a peak in the spectral envelope that is meant, not one of the resonances. Further, in speech and singing research, many writers also use this original meaning, particularly when discussing the singers formant and actors formant, which are broad peaks in the spectral envelope occurring around 3 kHz.

Does it matter? For the voice, a resonance at a frequency R_i gives rise to a spectral maximum at frequency F_i which may produce in a filter model a pole at frequency P_i . Usually, the three frequencies have similar values. However, as Fant observed, they are conceptually distinct. Let's take some examples:

- Consider a vocal tract with a resonance at 500 Hz, which is being excited by

the larynx producing a fundamental frequency of 1 kHz (near C6, the high C for sopranos). There is no spectral maximum at 500 Hz. In this case there is a resonance R1 but no corresponding spectral peak F1. Here of course the difference does matter.

- Consider the singers formant or singing formant, a broad band of enhanced power noticed in the spectral envelope of classically trained male singers (and possible others) in a range. Sundberg (1974) attributes this formant to a clustering of the third, fourth and fifth resonances of the vocal tract. Here, where three resonances are thought to give rise to one formant, the distinction between formant and resonance is important.
- Consider a glottal source with a negative spectral slope, input to a vocal tract that (including radiation impedance) has a resonance at R1. The peak in the spectral envelope of the radiated sound in this case has a frequency less than R1. In this case, if one is estimating the spectral peak from the harmonic spectrum of the output voice, the difference between the two is less than the precision of the estimation, so the distinction is usually not important.
- Consider a musical wind instrument, whose bore radiates weakly below some frequency f , and which is excited by a reed or lip valve whose spectral envelope falls with frequency. Here the output sound has a spectral envelope peak that has nothing at all to do with the resonances of the bore.
- Consider this quote*, from [Stevens and House \(1961\)](#): "When resonant frequencies are sufficiently close, however, they are not necessarily identical with the frequencies of the peaks in the spectrum. For example, when two resonances with bandwidths of about 100 cps are about 100 cps apart, the spectrum envelope may show only one prominence: the frequency of the peak will be somewhere between the two resonant frequencies. In the discussion that follows, the the levels of the resonances will be defined to be the levels of the of the spectral envelope at the frequencies of the resonances (rather than at the spectral peaks)."

In our laboratory, the distinction is important. We routinely measure the resonances independently of the voice ([Epps et al, 1997](#); [Dowd et al, 1997](#); [Joliveau et al, 2004a,b](#)). We are often interested in comparing formants and resonances.

What to do? Our preference would be to retain the original meaning for the word formant. We prefer to say "A resonance at frequency R_i gives rise to a formant at frequency F_i . This may be modelled by a filter with a pole at frequency P_i ". While acousticians will broadly agree with this use, some members of the speech research and modelling community may not. We therefore suggest that, when discussing the voice, the word formant should be defined, to make it clear which meaning is intended. In principle, one could consider abandoning the word. However "broad peak

in the spectral envelope" is a long phrase, so it is useful to retain formant for that reason.

Whatever your choice of definition, you should make it clear. And, in literature and in discussions, prepare for some confusion. For instance, some researchers who use formant to mean resonance will also talk about 'formant level'. When such people then talk of 'formant level', or say that the second formant is 10 dB lower than the first, I suspect that they refer to the amplitude of a peak in the sound spectrum. In a scientific talk, I have heard the sentence: 'Trained sopranos tune the first formant near the note sung, but they usually don't have a strong singer's formant'. When that speaker said 'first formant' he presumably meant 'first resonance' and when he said 'singer's formant' he meant a spectral peak probably due to two or more resonances. So we have the same person using the word in two of its three different meanings in the one sentence.

It's interesting to rewrite the quote from [Stevens and House \(1961\)](#), substituting 'formant' wherever they write 'resonance': "*When formant frequencies are sufficiently close, however, they are not necessarily identical with the frequencies of the formants. For example, when formants with bandwidths of about 100 cps are about 100 cps apart, the spectrum envelope may show only one formant: the formant will be somewhere between the two formants. In the discussion that follows, the levels of the formant will be defined to be the levels of the spectral envelope at the formant frequencies (rather than at the formant frequencies).*"

References

- Atal, B. S. and Hanauer, S. L. (1971) "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave", *J. Acoust. Soc. Am.*, 50, 637-655.
- Benade, A. H. (1976) *Fundamentals of musical acoustics*, Oxford University Press, London.
- Dowd, A., Smith, J.R. and Wolfe, J. (1997) "[Learning to pronounce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time](#)". *Language and Speech*, **41**, 1-20.
- Epps, J., Smith, J.R. and Wolfe, J. (1997) "A novel instrument to measure acoustic resonances of the vocal tract during speech" *Measurement Science and Technology* **8**, 1112-1121.
- Fant, G. (1960). *Acoustic Theory of Speech Production*. Mouton & Co, The Hague, Netherlands.
- Joliveau, E., Smith, J. and Wolfe, J. (2004) "[Tuning of vocal tract resonances by sopranos](#)", *Nature*, **427**, 116.
- Joliveau, E., Smith, J. and Wolfe, J. (2004) "[Vocal tract resonances in singing:](#)

- [the soprano voice](#)", *J. Acoust. Soc. America*, **116**, 2434-2439.
- Stevens, K.N., and House, A.S., (1961). An acoustical theory of vowel production and some of its implications, *J. Speech & Hearing Research*, 4, 303-320.
 - Sundberg, J. (1974) "Articulatory interpretation of the 'singing formant'," *J. Acoust. Soc. America*, **55**, 838-844.

Source: <http://www.phys.unsw.edu.au/jw/formant.html>