# FREQUENCY MODULATION OF FORMANT-LIKE SPECTRAL PEAKS

Peter J.Bailey and Kim Bevan

Department of Psychology, University of York,
York YO1 5DD, U.K.

## ABSTRACT

We report the results of two experiments showing that sinusoidal modulation of the centre frequency of one of a pair of formant-like spectral peaks increases its discriminability, as measured by the difference limen for spectral peak frequency. The apparent release from upward spread of masking afforded by modulation occurs for both noise-excited and pulse-excited stimuli and is not closely dependent on stimulus duration, modulation rate or peak centre frequency.

## 1. INTRODUCTION

The specification of formant frequency in vowel perception requires at least two potentially distinct stages: one logically-prior step that isolates a spectral region corresponding to a local energy peak, and another that estimates the peak frequency. Errors are likely in the first step of identifying where formants are when spectral peaks are close in frequency, or when listening in noise, amongst competing sounds or with an impaired auditory system. Such errors will lead in turn to inescapable errors in the second step of formant frequency assignment, and thus to probable inaccuracies in speech recognition performance.

Similar problems attend the visual perception of objects in complex scenes, where errors in locating the contours of an object can lead to conspicuous failures of visual identification. One powerful source of disambiguation in visual scenes is *movement* of the object or observer, which can provide cues to the appropriate parsing of the scene into figure and ground. In essence, the experiments reported here attempted to explore the utility of *auditory* object movement (an auditory object consisting of a single resonance) as a way of specifying for the listener the perceptual coherence of the energy contributing to a spectral peak. We hoped by this means to improve the accuracy of discrimination or recognition tasks that rely on the precision of the representation of peak frequency. We simulated auditory object movement using simple periodic modulation of resonance frequency.

There are demonstrations of the potentially beneficial role of modulation for both auditory detection and segregation tasks. Rasch [2] measured the masked threshold of a harmonic complex tone when it was mixed with a second harmonic complex of lower fundamental frequency. A 5 Hz, 4% vibrato imposed on the fundamental of the higher complex reduced its masked threshold by 17.5dB relative to its threshold when unmodulated. McAdams [1] has shown that modulation of the fundamental frequency of one of a set of three concurrent vowels can increase judgements of its perceived prominence. Our experiments were concerned not with fundamental frequency modulation, but with modulation of spectrum envelope characteristics. In particular they sought to establish whether peak frequency modulation can enhance the discriminability of a spectral peak when presented against the background of an otherwise unmodulated spectrum envelope.

## 2. GENERAL METHOD

Our basic strategy for measuring the perceptual effects of frequency modulation involved four stimuli in each experimental condition. Two of the stimuli had a single spectral peak (the "target" peak). In one case the peak centre frequency was not modulated and in the other it was sinusoidally modulated. The other two stimuli were like these, with the addition of a second lower-frequency spectral peak. In these two-peak stimuli the lower peak was never modulated and was sufficiently close in frequency to the higher-frequency target peak to impair unmodulated target peak discriminability. For each stimulus we measured subjects' difference limen (DL) for an increase in target peak centre frequency.

### 2.1 Stimuli

Stimuli were generated by passing broad-band noise (experiment 1) or a 100 Hz pulse train (experiment 2) through digital second-order resonators. When two spectral peaks were required the outputs of two parallel resonators were summed. Resonator half-power bandwidths were fixed at 100 Hz (target peak) and 80 Hz (lower-frequency peak). Filter coefficients were updated at a rate of 1 kHz (experiment 1) or 200 Hz (experiment 2). The depth of modulation (i.e. the total frequency excursion) for the modulated peak was 16% of the centre frequency. All spectral peaks had approximately equal spectrum level (+/- 2dB). Stimuli were presented at 70 dBA SPL in broad-band background noise at a level set for each subject to give the spectral peaks a presentation level of 10 dB SL.

### 2.2 Procedure

Difference limens were estimated using a two-alternative forced choice trial structure with two pairs of stimuli per trial. In one pair the stimuli were identical and in the other they differed in target peak frequency. The subjects' task was to identify the pair containing the different stimuli. The target peak frequency DL was taken to be the frequency difference corresponding to the 70.7% correct point on the psychometric function, determined by an adaptive staircase. Feedback was given after every response. Subjects were well practised before data collection began.

## 3. EXPERIMENT 1

In addition to the basic question of whether modulation of target peak frequency could improve its discriminability, the first experiment also explored the importance of modulation rate and stimulus duration.

### 3.1 Stimuli and Procedure

Target peak centre frequency was set to 1500 Hz. Target peak frequency DLs were measured for single-peak stimuli and for two-peak stimuli with a lower-frequency peak at 1300 Hz. Since our major concern here was with modulation of spectrum envelope characteristics the resonators were noise-excited, producing whisper-like stimuli with relatively fully-specified spectrum envelopes.

Other stimulus manipulations were as follows. Modulation rate: (i) 0 Hz (unmodulated), (ii) 5 Hz, and (iii) 10 Hz. Stimulus duration: (i) 250 msec. or (ii) 500 msec. Data were collected from seven subjects, including the second author.

### 3.2 Results

Mean DLs for all subjects are shown in Table 1.

**TABLE 1: mean DLs and standard errors (Hz) for experiment 1**

| modul: | none | 5 Hz | 10 Hz |
|---|---|---|---|
| 250 ms | | | |
| 1 peak | 30.29 | 38.27 | 37.64 |
| *sem* | *1.64* | *2.05* | *1.89* |
| 2 peak | 44.59 | 37.55 | 39.34 |
| *sem* | *0.99* | *2.28* | *2.09* |
| 500 ms | | | |
| 1 peak | 25.58 | 33.83 | 33.41 |
| *sem* | *2.26* | *2.20* | *1.28* |
| 2 peak | 41.34 | 32.18 | 35.22 |
| *sem* | *2.19* | *2.57* | *1.48* |

For stimuli with a single spectral peak modulation increased target peak DL. However, the effect of modulation in two-peak stimuli was to *decrease* the target peak DL relative to the unmodulated condition, that is to *increase* discriminability. This was true for both modulation rates and both stimulus durations. DLs were smaller for 500 msec stimuli, but there were no reliable interactions between the effects of modulation rate and duration.

### 3.3 Discussion

The results of this experiment show that sinusoidal modulation of peak centre frequency can lead to reliable improvements in the discriminability of a spectral peak when that peak is presented in an unmodulated spectral context. The absence of any interaction between modulation rate and stimulus duration shows that the effect is not dependent on the number of modulation cycles. Modulation appears to render the target peak perceptually more salient and thus less susceptible to upward spread of masking from the lower peak. This occurs despite the tendency for modulation to spread excitation around the peak frequency in the excitation pattern. The similarity in DL for one-peak and two-peak modulated stimuli suggests that modulation endows the target peak with substantial immunity from the masking effects of the lower peak. In terms of the two-stage sketch of formant perception given in the introduction, it may be that modulation, by providing additional information for perceptual grouping processes, increases the efficiency of the first stage, in which the spectral region corresponding to a spectral peak is identified. The second experiment sought to replicate and extend the generality of these results.

## 4. EXPERIMENT 2

This was concerned with the dependency of the modulation effect on type of resonance excitation and target peak frequency region.

### 4.1 Stimuli and Procedure

Target peak centre frequencies were set to 1500 Hz or 900 Hz, with lower-frequency peaks when present at 1300 Hz and 700 Hz, respectively. All stimuli were pulse-excited with a constant fundamental frequency of 100 Hz. Other stimulus manipulations were as before. DLs were measured in 4 subjects for each target peak frequency. Most of the subjects had also served in the first experiment.

### 4.2 Results

Mean DLs for all subjects are shown in Table 2 for target peak frequency 1500 Hz, and Table 3 for target peak frequency 900 Hz.



**TABLE 2: mean DLs and standard errors (Hz) for experiment 2 (Target Peak frequency 1500 Hz).**

| modul: | none | 5 Hz | 10 Hz |
|---|---|---|---|
| 250 ms | | | |
| 1 peak | 37.26 | 46.12 | 39.60 |
| *sem* | *3.80* | *2.19* | *3.29* |
| 2 peak | 50.88 | 42.29 | 44.24 |
| *sem* | *4.69* | *2.51* | *3.92* |
| 500 ms | | | |
| 1 peak | 30.90 | 36.27 | 34.97 |
| *sem* | *2.86* | *3.62* | *3.17* |
| 2 peak | 45.68 | 37.95 | 35.66 |
| *sem* | *4.19* | *2.74* | *3.60* |

**TABLE 3: mean DLs and standard errors (Hz) for experiment 2 (Target Peak frequency 900 Hz).**

| modul: | none | 5 Hz | 10 Hz |
|---|---|---|---|
| 250 ms | | | |
| 1 peak | 32.52 | 28.60 | 22.68 |
| *sem* | *4.61* | *6.39* | *5.57* |
| 2 peak | 38.45 | 29.95 | 21.45 |
| *sem* | *5.13* | *5.65* | *4.71* |
| 500 ms | | | |
| 1 peak | 31.25 | 21.51 | 21.28 |
| *sem* | *4.00* | *5.54* | *4.41* |
| 2 peak | 37.23 | 25.04 | 23.42 |
| *sem* | *4.24* | *5.14* | *5.49* |

The pattern of results for the two target peak frequencies was somewhat different. For pulse-excited stimuli with target peak frequency at 1500 Hz the results were similar to those obtained in experiment 1 with noise-excited stimuli at the same target peak frequency: as before, modulation apparently gave substantial immunity from the masking effects of the lower-frequency peak. For pulse-excited stimuli with target peak frequency at 900 Hz, modulation had the effect of decreasing the magnitude of the DL for single-peak stimuli as well as two-peak stimuli, relative to the DLs in unmodulated stimuli. As before, longer-duration stimuli tended to have smaller DLs, but there was no interaction between modulation rate and stimulus duration.

### 4.3 Discussion

The similarity between results obtained at the 1500 Hz target peak frequency for pulse-excited stimuli and those from the first experiment for noise-excited stimuli suggests that the enhanced discrimination modulation affords derives from properties of the spectrum envelope itself and not from the acoustic detail underlying it. We have data to suggest that the effect is genuinely attributable to modulation *per se* and not to phasic release from masking as the modulated target peak frequency increases above its mean value. The origin of the differences between the results for 1500 Hz and 900 Hz target peaks is not clear. One speculative suggestion is that modulation of the 900 Hz target peak may lead to detectable modulation of excitation in a larger number of auditory filters.

## 5. GENERAL DISCUSSION

We are aware that our account of the perceptual mechanism by which frequency modulation has its effects is crude and requires refinement. We believe the data are consistent with a role for perceptual grouping processes in the coherence that modulation imposes on the spectral energy contributing to a spectral peak. We are assessing the practical implications of these results by exploring the effect of second formant frequency modulation on vowel recognition.

## 6. ACKNOWLEDGEMENTS

The financial assistance of the U.K.SERC is acknowledged.

## 7. REFERENCES

[1] McAdams S.(1989) J.Acoust. Soc.Amer., 86, 2148-2159

[2] Rasch R.(1978) Acustica, 40,21-33