TR－H－066

THE EFFECTS OF COMPLEX EXTRANEOUS SOUNDS ON A VOWEL CONTINUUM

Andrew P. Lea
Alain de Cheveigné(CNRS)

1994. 3. 28

ATR人間情報通信研究所

〒619-02 京都府相楽郡精華町光台2-2　☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan
Telephone: +81-7749-5-1011
Facsimile: +81-7749-5-1008

THE EFFECTS OF COMPLEX EXTRANEOUS SOUNDS ON A VOWEL CONTINUUM

Andrew Lea[1] and Alain de Cheveigné[1&2]

[1]ATR Human Information Processing
Research Laboratories,
Hikaridai,
Seika-cho,
Kyoto 619-02,
Japan.

[2]Laboratoire de Linguistic Formelle,
CNRS/Université Paris 7,
tc 806, 2 place Jussieu,
Paris 75251, France.

Version I - March 28, 1994

ABSTRACT

The experiments described below investigate how listeners segregate two harmonic series when the two harmonic series have a harmonic which is common in frequency - a shared harmonic. Two types of sounds were used: vowel sounds which were members of an /i/ to /e/ vowel continuum and complex extraneous sounds. The extraneous sounds were either the first nine harmonics of a fundamental frequency, or the first nine harmonics, but with the fifth harmonic removed. The extraneous sounds were synthesized so that the fifth harmonic of the extraneous sounds coincided with either the second, third or fourth harmonic of the members of the vowel continuum. In Experiment 1 both sounds were presented to listeners with the extraneous sounds being 0, 3% and 8% mistuned from the shared harmonic. The results suggest that listeners try to disambiguate the energy of the shared harmonic and determine the amount of energy contributed to the shared harmonic from each harmonic series. Also, the effects of mistuning were stronger with the complex extraneous sounds of Experiment 1 than those found with the mistuning of a single harmonic of the vowel continuum in Experiment 2. Thus, suggesting that the complex extraneous sounds were capturing the shared harmonic from the vowels.

INTRODUCTION

In the every day world it is rare for any sound to be heard in isolation. Therefore listeners are continuously having to deal with the segregation of simultaneous sounds. If important features of vowels are masked by extraneous sounds, or if the extraneous sounds are included as part of vowels then errors in vowel perception will occur. It is generally accepted that vowel identification relies on the frequencies of the first and second formants (F1 and F2) (Joos, 1948; Delattre *et al*, 1952). Therefore extraneous sounds near the F1 or F2 frequencies ought to produce noticeable effects on the vowel quality that is perceived by listeners. To ensure that the components of the extraneous sounds are resolved by the auditory system the experiments reported below examine voiced vowel perception where the extraneous sounds are the F1 region of the vowels.

Many previous experiments have examined the perception of vowels in the presence of extraneous sounds in the F1 region. These experiments have used either single pure tones (e.g. Darwin, 1984; Darwin and Gardner, 1986), two pure tones (Roberts and Moore, 1990), or narrow band noise (Roberts and Moore, 1990) as the extraneous sounds. Neither of these are particularly natural sounds, nor do they have any strong harmonic structure like the vowel itself. For that reason in Experiment 1 described below a complex extraneous sound is used in which the many components are harmonically related.

THE HARMONIC SIEVE

Bregman (1990) has suggested that listeners use "primitive grouping cues" which can link together components that have originated from one source, using simple properties of these components to group them together. For example, components that are harmonically related tend to result from a common sound source more often than components that are not harmonically related and therefore harmonicity is a strong primitive grouping principle.

When components of a harmonic series are grouped together the auditory system must ensure that the components do actually belong to the harmonic series and not belong to another sound source. Models of pitch perception have suggested a way in which this may be achieved. A "harmonic sieve" (Duifhuis *et al*, 1982; Scheffers, 1983) can be constructed to filter out components which are too remote and allowing only components within some tolerance to be included in the harmonic series and hence contribute to the pitch of the harmonic series. The tolerance of the sieve for pitch perception has been estimated by Moore *et al* (1985a). They mistuned a single harmonic of a 12-harmonic series with a fundamental frequency (f0) of 100, 200 or 400 Hz and measured the resulting pitch shift. Up to 3% mistuning, the mistuned harmonic was completely included in the harmonic series as it produced a linear shift in pitch. From 4% to 8% mistuning, the mistuned harmonic was gradually rejected from the harmonic series as it produced a non-linear pitch shift at 4% mistuning and reduced amounts of pitch shift at 6% and 8%

mistuning. The pattern of results is in agreement with the harmonic sieve concept and also Terhardt's virtual pitch model of pitch perception (Terhardt *et al*, 1982).

Moore *et al*, (1985b) measured thresholds for the detection of inharmonicity in complex tones using the same listeners as Moore *et al*, (1985a). Listeners were required to distinguish a 12 component, perfectly harmonic complex tone from a similar tone in which one component was slightly mistuned. They found that listeners could detect the inharmonicity of a single component with only 1.1% mistuning with comparative conditions in which Moore *et al*, (1985a) found that the harmonic sieve only begins to reject components when the mistuning is greater than 3%. Thus, it appears that a component can be heard as mistuned, but still contribute to the pitch of a complex tone. This may be a form of "duplex perception" as suggested by Bregman (1990).

VOWEL QUALITY AND THE HARMONIC SIEVE

Darwin and Gardner (1986) examined the perception of vowel quality as a harmonic of the vowel in the F1 region was mistuned. Their hypothesis was that with large amounts of mistuning of the harmonic it would make no contribution to the quality of the vowel and it would be as if the harmonic was not present at all as it had been rejected by the harmonic sieve. They synthesized a vowel continuum between /I/ and /e/ with a 125-Hz f0. They manipulated the harmonics which were just above and below the formant boundary between the vowels, at 375 Hz and 500 Hz respectively. They found that with large mistunings of the two harmonics (10.7% for the 375 Hz harmonic and 8.0% for the 550 Hz harmonic) the F1 boundary shifted in the direction predicted by the total removal of the harmonic, but did not shift as far as total removal of the harmonic would have predicted. Their results with small mistunings are confounded by phase effects as demonstrated by Palmer *et al* (1988).

Roberts and Moore (1990) conducted a number of experiments which examined the effect of extraneous tones in the F1 region on perceived vowel quality on a vowel continuum ranging from /I/ to /ε/. They used as extraneous sounds either two pure tones of various frequencies, or narrow-band noise centered at various frequencies. When the tones or narrow-band noise were added on the low-frequency side of the F1 boundary measured in quiet, the F1 boundary was lowered in frequency. When the tones or narrow-band noise were added on the high-frequency side of the F1 boundary measured in quiet, the F1 boundary was shifted higher in frequency. The results were interpreted by assuming that the additional energy added to the perceived amplitude of the harmonics of the vowels and therefore shifted the phoneme boundary as shown by the modeling of Roberts and Moore (1991b). When the additional energy was mistuned the F1 boundary moved less than when the additional energy was not mistuned, suggesting that some of the energy was rejected by the harmonic sieve.

THE CURRENT STUDY

In previous studies which examine how the harmonic sieve excludes extraneous sounds the extraneous sounds have either been pure tones, two pure tones or narrow-band noise. Thus, previously the extraneous sounds have not had a strong pitch nor a large number of harmonics. The extraneous sounds used in Experiment 1 consist of either eight or nine low-frequency harmonics of an f0 and thus have a strong pitch.

In the natural environment listeners are unlikely to hear a vowel in the presence of an extraneous sound which just consists of one or two tones or a narrow-band noise. A more "ecological valid" stimulus might be one where the extraneous sound has a large number of harmonics itself. In this case listeners might have to construct two harmonic sieves one for the vowel and one for the extraneous sound (Scheffers, 1983). When two harmonic series are present simultaneously it is often the case that at least one harmonic from each series is close to the frequency of a harmonic in the other series. Such instances have been called "over-lapped harmonics" by Assmann and Summerfield (1990) and "shared harmonics" by de Cheveigné (1993) which is the term we will use here.

A shared harmonic can be handled in one of four ways by the auditory system (de Cheveigné, 1993): (a) assign the harmonic to both vowel and extraneous sound, (b) assign the harmonic to neither the vowel nor the extraneous sound, (c) assign the shared harmonic to either the vowel or the extraneous sound, or (d) share the harmonic between the vowel and extraneous sound.

EXPERIMENT 1

The experiment uses a simplified vowel which is a member of a vowel continuum ranging from /i/ to /e/. Two types of complex extraneous sounds were used in Experiment 1. One consisted of the first 9 harmonics of an f0 with equal amplitudes; this will be called the full (F) extraneous sound. The second extraneous sound was the same as the first, except that the 5th harmonic was missing; this will be called the notch (N) extraneous sound. The f0s of the vowels and the extraneous sounds were chosen so that the 5th harmonic of the masker coincided with either the second, third or fourth harmonics of the vowel continuum. Mistunings were also introduced so that the 5th harmonic of the extraneous sounds was mistuned from the harmonics of the vowel continuum.

The listeners were first presented with the vowel continuum in isolation (the parent continuum). Then listeners were presented the vowel continuum in the presence of the extraneous sounds. The task of the listeners was always to identify the vowels. The psychometric curves produced when the continuum was presented in isolation can be compared to when the extraneous sounds were present to infer the change from the psychometric function of the parent continuum caused by the extraneous sounds.

If listeners always assign the energy of the shared harmonic to the extraneous sound, we would expect a large amount of change from the parent continuum with 0%

mistunings and no difference between the two extraneous sounds. This is because the harmonic of the vowel would be completely removed and a large shift in the formant boundary between /i/ and /e/ would be expected. As mistuning increases we would expect less of the energy of the harmonic of the vowel to be assigned to the extraneous sound and therefore the change from the parent continuum would decrease. However, we would still expect the F and N extraneous sounds to have the same effect on the vowel continuum.

If listeners always assign the energy of the shared harmonic to the vowel we would expect no change from the parent continuum for the N extraneous sound, as the N extraneous sound is not adding energy to the harmonic of the vowel. A small change from the parent continuum might be expected due to the other harmonics of the N extraneous sound. However, for the F extraneous sound we would expect a large change from the parent continuum with 0% mistunings which would decrease as mistuning increased. This is because all of the additional energy from the extraneous sound would be included with the vowel at 0% mistuning. As mistuning increased the shared harmonic will decrease in energy due to the frequency separation increasing.

If listeners try to assign the correct amount of energy of the shared harmonic to vowel and to the extraneous sounds then we might expect the N extraneous sound to produce more change from the parent continuum than the F extraneous sound for 0% mistuning. If listeners are performing a kind of profile analysis to work out how much energy of the shared harmonic belongs to the vowel and how much belongs to the extraneous, listeners should be fooled by the N sound. The reason for this is that it unnaturally contains a notch at the frequency of the shared harmonic and therefore listeners might mistakenly assign some of the energy of the harmonic of the vowel to the extraneous sound. However, for the F extraneous sound using the amplitude of the harmonics of the extraneous sound will provide listeners with reliable information to separate the energy of the shared harmonic correctly. The differences between the F and N extraneous sounds should decrease as mistuning increases, as the shared harmonic becomes less well shared between the two harmonic series. The only way listeners could try and disambiguate the ambiguity of the shared harmonic is to use the amplitude of the other harmonics of the two harmonic series to work out how much energy of the shared harmonic came from each of the harmonic series.

The strategy of totally removing the shared harmonic from both the vowel and the extraneous sound is not considered a viable option. The goal of the auditory system is to provide a complete a description as possible of the auditory environment (Bregman, 1990) and ignoring the energy of a shared harmonic would probably have disastrous repercussions.

STIMULI

A vowel continuum from /i/ to /e/ was synthesized using additive harmonic synthesis (16-bits, 44.1 kHz sampling rate). The f0 of the vowel continuum was 150 Hz. The first formant (F1) of the continuum varied from 250 to 650 Hz in 25 Hz steps making 17 vowel sounds in the continuum. Two adjacent harmonics out of the first five were raised above a background of equal amplitude harmonics (from 150 to 1800 Hz) to define the F1 frequency as shown in Figure 1. The two harmonics were raised by a total of 12 dB above the background. The distribution of the 12 dB between the two raised harmonics defined the F1 according to the weighted mean of the additional energy (Carlson *et al*, 1975; Assmann and Nearey, 1987). Both the F2 and F3 of all the members of the continuum were kept constant at 2100 and 2850 Hz respectively. Figure 2 shows the amplitude spectrum of the 450-Hz F1 member of the vowel continuum.

The full extraneous sound (F) consisted of the first 9 harmonics of an f0 and the notched extraneous sound (N) was the same as the full extraneous sound except the 5th harmonic was missing. The extraneous sounds were synthesized such that the 5th harmonic of the extraneous sounds coincided with the second, third and fourth harmonics of the vowels of the continuum. Thus, the extraneous sounds had initial f0s of 60, 90 and 120 Hz so that the 5th harmonic of the extraneous sound was at 300, 450 or 600 Hz. Two further sets of extraneous sounds were generated by mistuning the 60, 90 and 120-Hz f0 extraneous sounds by 3% and 8%. Negative mistunings were not used to reduce the number of conditions. Hence, extraneous sounds with f0s of 60.0, 61.8, 64.8, 90.0, 92.7, 97.2, 120.0, 123.6 and 129.6 Hz were used.

The duration of the vowels and the extraneous sounds was 120 ms including 10 ms cosine ramps at sound onsets and offsets. All the components of the vowels and extraneous sounds were synthesized in cosine phase, so that where components had the same frequency their amplitude was additive.

PROCEDURE

The members of the vowel continuum and the extraneous sounds were summed such that the harmonics of the extraneous sounds were 6 dB more intense than the harmonics of the vowel continuum between 150 and 1800 Hz. This figure excludes the two harmonics of the vowel continuum that were raised to define the F1. Vowels without maskers were presented to listeners at 70.0 dB(A) SPL.

Five listeners with normal hearing, including the first author were used. Each listener was tested individually in a sound attenuated booth. The task for the listeners was a 2 alternative forced choice between /i/ and /e/. At no point in the experiment was feedback given to listeners.

First, listeners were presented with a randomization of the continuum vowels presented in isolation (the parent continuum). Each member of the continuum was presented 20 times in this randomization, making 340 trials in total. Second, a series of randomizations of the continuum vowels added to both the full and notched extraneous sounds was presented to the listeners. The extraneous sounds which shared, or nearly shared the second harmonic with the vowel continuum, i.e. maskers with f0s of 60.0, 61.8 and 64.8 Hz, were grouped together in two experimental sessions of 10 repetitions of each stimulus, making *a total of 1020 trials per session. In the same way, the extraneous sounds* which shared, or nearly shared the third harmonic with the vowel continuum, i.e. maskers with f0s of 90.0, 92.7 and 97.2 Hz, were grouped together. In the same way, the extraneous sounds which shared, or nearly shared the fourth harmonic with the vowel continuum, i.e. maskers with f0s of 120.0, 123.6 and 129.6 Hz, were grouped together. The order of presentation of the experimental sessions was randomized independently for each listener. The entire data set was gathered over seven sessions of listening of an hour each.

RESULTS

Figure 5 shows the results for the five listeners for the members of the vowel continuum presented in isolation. The psychometric curve that results from the vowel continuum presented in isolation will be termed the parent continuum. The letters refer to the individual listeners and the filled circles plots the mean of the listeners. Probit analysis (Finney, 1971) was used to find the perceptual boundaries between /i/ and /e/ of the listeners. The listeners' boundaries were at 507, 365, 427, 481 and 466 Hz for listeners A to E respectively and at 449 Hz for the mean of the listeners.

Figures 6, 7 and 8 shows the results for the five listeners for the members of the vowel continuum presented with the extraneous sounds that shared the second, third and fourth harmonics of the vowel continuum. Normally, with this kind of experiment Probit analysis is used to determine how much an extraneous sound has caused the perceptual boundary between /i/ and /e/ to shift. However, if Figures 6, 7 and 8 are examined it can be seen that many of the listeners psychometric functions are non-monotic and therefore Probit analysis cannot be used. Both Darwin (1984) and Roberts and Moore (1990, 1991a) rejected data from listeners whose psychometric functions were non-monotonic. Rather than reject listeners data we will use two different methods to analyze the data. Another problem with the results is the large individual differences between listeners which means that the results cannot be averaged over the listeners without normalization. Therefore each of the listeners' psychometric functions with the extraneous sounds was normalized by using the individual listener's psychometric function with the vowel continuum presented in isolation.

The normalization was performed in two ways. Firstly, to generate a statistic that highlights the absolute differences between the psychometric curves and secondly a statistic which highlights the relative differences between the psychometric curves. The first statistic computed the root of the mean squared (RMS) difference between each point on the psychometric curve of the parent continuum and each point on the psychometric curve of the other continuums. The average was then taken over all listeners. The results of this analysis for each vowel on the continuum are shown in Figure 9 for both full and notch extraneous sounds and all three harmonics.

Figure 10 shows the same results as Figure 9 except that the average over the different members of the vowel continuum is plotted. Figure 11 shows the same data as Figure 10 except averaged over the five listeners. The results shown in Figure 10 were subjected to a repeated measures analysis of variance (ANOVA) the results of which are shown in Table 1. The factors of extraneous sound, amount of mistuning, shared harmonic and listeners were used in the ANOVA, with listeners as a random effect. All possible interactions were used in the ANOVA with the exception of the fourth order interaction of extraneous sound*mistuning*shared harmonic*listeners. The degrees of freedom were adjusted for the repeated measures design with random effects using Satterthwaite's method (Satterthwaite, 1941). None of the single factors were significant except for listeners, but the interactions of extraneous sound*shared harmonic, mistuning*shared harmonic and extraneous sound*mistuning*shared harmonic were all significant (see Table 1 for details).

Contrast tests were performed to determine the significant effects. For the interaction of extraneous sound*shared harmonic the N extraneous sound produced more RMS error than the F extraneous sound with the 4th shared harmonic ($p<0.027$), other differences were not significant. For the interaction mistuning*shared harmonic the contrast tests are shown in Table 2. The table shows that over the three shared harmonic conditions the difference between 0% and 3% mistuning is larger than then either of the differences between 0% and 8% mistuning, or between 3% and 8% mistuning. For the interaction extraneous sound*mistuning*shared harmonic the contrast tests are shown in Table 3. The table shows that with the 3rd shared harmonic and all three mistunings the N extraneous sound produces more RMS error than the F extraneous sound. Also, for the second harmonic and 0% mistuning the F extraneous produced more RMS error than the N extraneous sound.

The second statistic that was used to analyze the results was the Pearson's correlation statistic. Each of the continua with extraneous sounds was correlated with the parent continuum to measure the amount of deviation caused by the extraneous sounds. Since, the correlation coefficient is not a linear scale it is not possible to directly average the correlation coefficients over listeners. Therefore the correlation coefficients were converted to Fisher z' scores[1] (Minium, 1978), which is a linear scale between plus/minus 3 for correlation scores and is also sometimes called the Fisher r' score. A Pearson correlation coefficient of 1 gives a Fisher z' score of 3, a correlation coefficient of 0 gives a Fisher z' score of 0 and a correlation coefficient of -1 gives a Fisher z' score of -3. Fisher z' scores averaged over the five listeners are shown in Figure 12. The abscissas of Figure 12 have been inverted so that the amount of deviation from the parent continuum increases as we ascend the abscissa for direct comparison with Figure 11.

The Fisher z′ results were subjected to a repeated measures ANOVA in the same way as the RMS error results were analyzed. The results of this analysis are shown in Table 4. The pattern of significant results is identical to the RMS error results except that the factor listeners is not significant. Since the pattern of results with the Fisher z′ score is virtually identical to that found with the RMS score, both statistically and also the pattern of results shown in Figures 11 and 12, the contrast effects of the Fisher z′ score will not be examined.

DISCUSSION

Firstly we will discuss the source of the non-monotonicities in the data, secondly which strategy listeners are using to disambiguate the energy of the shared harmonic and thirdly we will discuss the effects of mistuning.

NON-MONOTONICITIES

The members of the vowel continuum were synthesized so that a constant amount of energy of 12 dB was added to define the F1 of the vowels. This mean that the amplitude difference between the most intense harmonic defining the F1 and the constant level background harmonics varies as a function of frequency. When the F1 coincided exactly with the frequency of harmonic, one harmonic was raised by 12 dB to define the F1. When the F1 was half between harmonic frequencies, two harmonics were raised by 6 dB to define the F1. Other F1 frequencies varied between these two extremes. The extraneous sounds were mixed with the vowels such that the extraneous sounds were 6 dB more intense than the background harmonics of the vowels in the F1 region. Therefore, if we remove 6 dB from the amount the harmonics of the F1 were raised by, we know by how much the harmonics of the F1 rose above the harmonics of the extraneous sounds. This amount we will define as the contrast of the F1. The contrast of the F1 across frequency is plotted in Figure 13.

If we compare the level of contrast of the F1 shown in Figure 13 with the psychometric curves of listeners shown in Figures 6, 7 and 8 that where the non-monotonicities do occur that they are correlated with the contrast of the F1. It is not the case that the simplified synthetic vowels used here are unusual as similar effects of the position of formants relative to the position of harmonics have been found previously using more synthetic vowels which are more natural (Javkin *et al*, 1987).

Another source of the non-monotonicities maybe due to the segregation which is occurring. If energy is removed from one harmonic of the vowel and assigned to the extraneous sound, this will produce different effects depending upon how crucial the harmonic is to locating the position of the F1 of the vowel. The different strategies of disambiguating the shared harmonic are discussed below.

SEGREGATION OF THE SHARED HARMONIC

It is clear from the results that the full and notch extraneous sounds affect the vowel continuum in different ways when the shared harmonic is the 3rd or the 4th harmonic of the vowel continuum. Therefore we can reject the strategy which assigns the shared harmonic entirely to the extraneous sound.

It is also clear from the results that we can also reject the strategy in which the entire energy of the shared harmonic to the vowel. For the 4th shared harmonic the N extraneous sound produces more RMS error than the F extraneous sound whereas this strategy would predict that the N extraneous sound would always produce less RMS error than the F extraneous sound. Also an effect of mistuning for the N extraneous sound would not be predicted, but it can be seen for all three shared harmonics, although it is in the opposite direction than would be predicted for the 3rd shared harmonic.

Having excluded these two strategies the only strategy that can fit the data is the one which tries to disambiguate the energy of the shared harmonic and assign some energy to each of the sources. However, the results vary with the three shared harmonics and the strategy must be able to explain these results before we can conclude that it fits the results.

Let us make three assumptions about what listeners are doing. Firstly, that they try to disambiguate the energy of the shared harmonic and try to assign the correct amount of energy to the vowels and the extraneous sounds. Secondly, that listeners use the amplitudes of the other harmonics of the vowels and the extraneous sounds to work out the source of the energy of the shared harmonic. Thirdly, that even when listeners have disambiguated the energy of the shared harmonic, they are not perfect at doing this and some error results.

With the F extraneous sound and all three of the harmonics, listeners are largely able to disambiguate the energy of the shared harmonic. However, they are not perfect at doing so and hence we have an effect of mistuning. When the fifth harmonic of the extraneous sound is not shared with the second, third or fourth harmonic of the continuum the ambiguity is less and therefore the extraneous sound affects the perception of the vowel less.

With the N extraneous sound and all three harmonics, listeners mistakenly assign some of the energy of the shared harmonic from the vowels to the extraneous sounds. This is because they use the amplitude of the neighboring harmonics of the shared harmonic in each harmonic series to disambiguate the energy of the shared harmonic. The N extraneous sound unnaturally does not contribute any energy to the shared harmonic because of the notch. The removal of the energy of the harmonic of the vowel can shift the perceived F1 frequency in two ways. Firstly, because of the removal of energy the perceived F1 frequency is shifted. Secondly, in pilot experiments of Lea (1993) it was noticed that when the F1 was absent entirely listeners heard the vowel /i/ rather than the vowel /e/, see Appendix A for details. Accordingly, when the F1 contrast is low listeners tend to here the vowel /i/. The three different patterns of results with the N extraneous sound is due to the three harmonics being on different points along the /i/ to /e/ psychometric curve.

With the second harmonic and the N extraneous sound, if the contrast of the vowel is low it makes little difference to what listeners are responding as in the basic continuum they respond with the vowel /i/ in the region. Thus, the N extraneous sound produces little error when the second harmonic is shared.

With the third harmonic and the N extraneous sound, if the contrast of the vowel is low it makes little to what listeners are responding as in the basic continuum listeners respond with approximately 50% /i/ and 50% /e/ in this region. Since the third harmonic is near the average listeners' boundary between /i/ and /e/, a shift in perceived F1 frequency away from the third harmonic will shift stimuli from near the F1 boundary to being further from the F1 boundary and thus a small amount of error is seen with the N extraneous sound and 0% mistuning. As the mistuning increases this effect is reduced and the error increases.

With the fourth harmonic and the N extraneous sound, if the contrast of the vowel is low it makes a large difference to what listeners are responding as in the basic continuum listeners would respond with an /e/in this region. Thus, the N extraneous sound produces a large amount of error when the fourth harmonic is shared.

THE EFFECTS OF MISTUNING

Moore *et al* (1985a) have shown that for the pitch estimation the harmonic sieve begins to exclude mistuned harmonics when the mistuning is greater than 3%. Roberts and Moore (1992b) have shown that for the task of F1 perception the harmonic sieve begins to exclude harmonics when the mistuning is greater than 3%. However, in the results reported here it appears that the harmonic sieve is rejecting energy at mistunings of 3%. This is especially well demonstrated with the N extraneous sound where no energy is added at the frequency of the shared harmonic. In this case the amplitude of the shared harmonic does not change with mistuning and there should be no phase effects of the shared harmonic as the N extraneous sound has a notch at the frequency of the shared harmonic.

One explanation for the apparent narrowing of the harmonic sieve could be that the harmonic series of the extraneous sounds are capturing the shared harmonic and thus reducing the harmonic sieve width. To this test this possible explanation a second experiment was performed which tried to replicate some of the conditions of Moore *et al* (1985a).

EXPERIMENT 2

This experiment was designed to replicate the first experiment of Moore *et al* (1985a) except measuring the effect of mistuning on F1 frequency perception rather than on pitch.

STIMULI

The same vowel continuum as used in Experiment 1 was used here. Nine different variations of the vowel continuum were used. To the second, third or fourth harmonics of the vowel continuum was added a pure tone of the same frequency which was 6 dB more intense than the background harmonics of the vowel continuum in the F1 region. This was done to duplicate as closely as possible the conditions of Experiment 1. The second, third or fourth harmonics were then mistuned by 0%, 3% or 8%. All the components of the vowels and pure tone were synthesized in cosine phase, so that where components had the same frequency their amplitude was additive.

As we are effectively using an extraneous sound which is a sine wave these continua will be abbreviated to S.

PROCEDURE

The procedure was the same as for Experiment 1. Listeners were presented with three randomizations of stimuli, one for each of the second, third and fourth harmonics. Each randomization consisted of 20 repetitions of the 17 vowels in the continuum with the three degrees of mistuning making a total of 1020 trials in each experimental run. The order of presentation of the three randomizations for the second, third and fourth harmonics was randomized for each listener. The entire data set was gathered over three experimental sessions of one hour each.

RESULTS AND DISCUSSION

Figure 14 shows the results for the five listeners for the vowel continuum presented with 6 dB added to the second, third, or fourth harmonic of the vowels. The three mistuning conditions are also shown for each harmonic and listeners' psychometric curves with the parent continuum from Experiment 1 are shown for comparison. Listeners' psychometric functions are not monotonic as were those found for Experiment 1. Therefore the same analysis methods as used for Experiment 1 will be used again here.

Figure 15 shows the RMS errors between each of the S continua and the parent continuum averaged over the five listeners. Figure 16 shows the same results as Figure 15 except averaged over the vowel continua. Figure 17 shows the Fisher z' scores for the Pearson correlations between each of the S continua and the parent continuum averaged over the five listeners.

The results scored by RMS error and Fisher z' were subjected to a repeated measures ANOVA, the details of which are shown in Tables 5 and 6 respectively. None of the RMS error differences are significant, but the Fisher z' scores show a significant effect of mistuning. Contrast tests were performed to locate the source of the difference between the mistuning conditions. Averaged over the three harmonics, 0% produced a lower Fisher z' score than 8% mistuning ($p<0.005$) and no other mistuning comparisons were significant.

Thus, when the 6 dB of energy was added to the second, third and fourth harmonics of the vowels in separate conditions, the 0% mistuned condition deviated further from the parent continuum than did the 8% mistuning condition. This shows that with 8% mistuning not all the energy of the mistuned harmonic was included into the vowel.

GENERAL DISCUSSION

The listeners used in the two experiments reported here produced psychometric curves which could not be analyzed using Profile analysis. Therefore the two statistics of RMS error and Pearson's correlation coefficient were used. The advantage of these two statistics is that listeners psychometric curves did not have to be monotonic. The disadvantage is the insensitivity of the statistics. Both statistics are insensitive to the any shift in the boundary between /i/ and /e/. Fortunately, either the psychometric curves did not have clear boundary shifts, or the boundary shifts were in one direction.

The results of Experiment 1 suggest that listeners try to disambiguate the ambiguity of shared harmonics rather than assign the harmonic exclusively to one harmonic series. To do this listeners must use the amplitude of other harmonics in each harmonic series.

In Experiment 1 there were effects of mistuning at 8% and interestingly at 3% mistuning. Previously the effects of mistuning have only been found at over 3% (Moore *et al*, 1985a; Roberts and Moore, 1991, 1993b). However, mistuning can be detected by listeners at much smaller amounts of mistuning (Moore *et al*, 1985b). In Experiment 2 the effects of mistuning were small and just significant at 8% mistuning. In Experiment 2 only a single harmonic was mistuned, but in Experiment 1 a whole complex sound was mistuned along with the shared harmonic. This difference could be due to the complex sound capturing the shared harmonic away from the vowels. Thus, the effects of mistuning are stronger with the complex extraneous sound than with a single mistuned harmonic, because much of the energy of the shared harmonic from the vowel is been assigned to the extraneous sound.

---

[1] To convert from Pearson's correlation coefficient (r) to Fisher z' scores the following equation is used:

$$z' = \tfrac{1}{2} \times \left[\log_e(1+r) - \log_e(1-r)\right]$$

ACKNOWLEDGMENTS

Thanks to Minoru Tsuzaki for useful discussions.

APPENDIX A

Lea (1993) was interested in the grouping effects of onset and offset asynchronies of a single harmonic of a vowel continuum. During a pilot experiment an attempt was made to avoid using a vowel continuum as there are a number of problems with using vowel continuums. The number of vowels in a continuum can make many hours of listening and also the problem of non-monotonic psychometric functions of listeners. In an attempt to remove these problems two vowel stimuli were synthesized which differed only by the amplitude of one harmonic, one was /i/ and the other was /e/. The stimuli were the same as shown in Figure 1 above, except that all harmonics in the F1 region had the same amplitude for the vowel /i/ and therefore it had no F1. For the /e/ vowel the 600 Hz harmonic was raised by 12 dB to define the F1.

The two vowel stimuli were presented to five listeners including the author with a variety of onset and offset asynchronies of the 600 Hz harmonic. At no point in the experiment were listeners given any feedback. The results are shown in Table A1 for the conditions in which there were no asynchronies. A repeated measures ANOVA was performed and the results show that the listeners could distinguish the vowel-like stimuli ($F_{1,4}$=8.33, $p<0.05$).

The results show that when no F1 is defined for the stimuli shown in Figure 1, then listeners tend to hear the vowel /i/ rather than the vowel /e/.

REFERENCES

Assmann, P. A., and Nearey, T. M. (1987). "Perception of front vowels: The role of harmonics in the first formant region," J. Acoust. Soc. Am. 81, 520-534.

Assmann, P. A., and Summerfield, Q. (1990). "Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies," J. Acoust. Soc. Am., 88, 680-697.

Carlson, R., Fant, G., and Granstrom, B. (1975). "Two-formant models, pitch and vowel perception," in Auditory Analysis and Perception of Speech, edited by G. Fant & M. A. A. Tatham (Academic, London).

Cherry, E. C. (1953). "Some experiments on the recognition of speech with one and two ears," J. Acoust. Soc. Am. 23, 975-979.

Darwin, C. J. (1981). "Perceptual grouping of speech components differing in fundamental frequency and onset-time," Q. J. Exp. Psych., 33A, 185-207.

Darwin, C. J. (1984). "Perceiving vowels in the presence of another sound: Constraints on formant perception," J. Acoust. Soc. Am., 1636-1647.

Darwin, C. J., and Gardner, R. B. (1986). "Mistuning a harmonic of a vowel: Grouping and phase effects on vowel quality," J. Acoust. Soc. Am. 79, 838-845.

Darwin, C. J., and Sutherland, N. S. (1984). "Grouping frequency components of vowels: When is a harmonic not a harmonic," Q. J. Exp. Psych., 36A, 193-208.

de Cheveigné, A. (1993). "Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing," J. Acoust. Soc. Am. 93, 3270-3290.

Delattre, P., Liberman, A. M., Cooper, F. S., and Gertsman, L. J. (1952) "An experimental study of the determinants of vowel color: observations on one- and two- formant vowels synthesized from spectrographic patterns", Word, 8, 195-210.

Duifhuis, H., Willems, L. F., and Sluter, R. J. (1982). "Measurement of pitch in speech: an implementation of Goldstein's theory of pitch perception," J. Acoust. Soc. Am. 71, 1568-1580.

Finney, D. J. (1971). "Probit Analysis" Cambridge U. P., Cambridge, England, 3rd ed.

Javkin, H. R., Hermansky, H., and Wakita, H. (1987). "Interaction between formant and harmonic peaks in vowel perception," Proceedings of the 11th International Congress of Phonetic Sciences, Tallin, Estonia, U. S. S. R.

Joos, M. (1948). "Acoustic phonetics", Lang., 24(Suppl.), 1-36.

Lea, A. P. & Tsuzaki, M. (1993). "Segregation of voiced and whispered vowels in English and Japanese," J. Acoust. Soc. Am. 93 Pt. 2, 2403.

Lea, A. P. (1993). "Grouping in vowel perception: Effects of onset and offset asynchrony of a harmonic," Acoust. Soc. Japan, Technical Report H-93-81.

Minium, W. (1978). "Statistical Reasoning in Psychology and Education," New York, Wiley, 2nd. Edition.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (1985a) "Relative dominance of individual partials in determining the pitch of complex tones," J. Acoust. Soc. Am. 77, 1853-1860.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (1985b) "Thresholds for the detection of inharmonicity in complex tones," J. Acoust. Soc. Am. 77, 1861-1867.

Palmer, A. R., Winter, I. M., Gardner, R. B., and Darwin, C. J. (1987). "Changes in phonemic quality and neural representation of a vowel by alteration of the relative phase of harmonics near F1," in The Psychophysics of Speech Perception, Edited by M. E. H. Schouten (Nijhoff, Dordrecht, The Netherlands), pp371-376.

Roberts, B., and Moore, B. C. J. (1990). "The influence of extraneous sounds on the perceptual estimation of first-formant frequency in vowels," J. Acoust. Soc. Am., 88, 2571-2583.

Roberts, B., and Moore, B. C. J. (1991a). "The influence of extraneous sounds on the perceptual estimation of first-formant frequency in vowels under conditions of asynchrony," J. Acoust. Soc. Am., 89, 2922-2932.

Roberts, B., and Moore, B. C. J. (1991b). "Modeling the effects of extraneous sounds on the perceptual estimation of first formant frequency in vowels," J. Acoust. Soc. Am., 2933-2951.

Satterthwaite, F. E. (1941). "Synthesis of variance," Psychometrika, 6, 309-316.

Scheffers, M. T. M. (1983). "Sifting vowels: Auditory pitch analysis and sound segregation," Doctoral Thesis, University of Groningen, The Netherlands.

Terhardt, E., Stoll, G., and Seewann, M. (1982). "Algorithm for extraction of pitch salience from complex tonal signals," J. Acoust. Soc. Am. 71, 679-688.

Winer, B. J., Brown, D. R., and Michels, K. M. (1991). "Statistical principles in experimental design," McGraw-Hill, New York, 3rd Edition.

LEGENDS TO FIGURES

Figure 1. Simplified spectra of the members of the vowel continuum. The F0 was 150 Hz. The F1 varied from 250 to 650 Hz in 25 Hz steps making 17 vowels in the continuum. The F2 and F3 were fixed at 2100 and 2850 Hz respectively. A total of 12 dB was added to two adjacent harmonics to define the F1.

Figure 2. Amplitude spectrum of the 450-Hz F1 member of the vowel continuum. The third harmonic has been raised by 12 dB to define the F1 frequency.

Figure 3. Amplitude spectrum of a full extraneous sound with a 90-Hz f0.

Figure 4. Amplitude spectrum of a notch extraneous sound with a 90-Hz f0.

Figure 5. Results of the listeners for the continuum vowels presented to listeners in isolation. The abscissa plots the percentage of /i/ responses and the ordinate plots the F1 frequency in Hz. The letters indicate the psychometric curves of the individual listeners and the filled circles the mean of the listeners.

Figure 6. Results of individual listeners for the extraneous sounds which shared the second harmonic of the vowel continuum. The extraneous sounds had f0s of 60.0, 61.8 and 64.8 Hz. F and N refer to the full and notch extraneous sounds respectively. The crosses show the parent continuum for reference. The circles shows the results with the extraneous sounds with no mistuning. The squares shows the results with 3% mistuned extraneous sounds and the stars with 8% mistuned extraneous sounds.

Figure 7. Results of individual listeners for the extraneous sounds which shared the third harmonic of the vowel continuum. The extraneous sounds had f0s of 90.0, 92.7 and 97.2 Hz. F and N refer to the full and notch extraneous sounds respectively. The crosses show the parent continuum for reference. The circles shows the results with the extraneous sounds with no mistuning. The squares shows the results with 3% mistuned extraneous sounds and the stars with 8% mistuned extraneous sounds.

Figure 8. Results of individual listeners for the extraneous sounds which shared the fourth harmonic of the vowel continuum. The extraneous sounds had f0s of 120.0, 123.6 and 129.6 Hz. F and N refer to the full and notch extraneous sounds respectively. The crosses show the parent continuum for reference. The circles shows the results of the extraneous sounds with no mistuning. The squares shows the results with 3% mistuned extraneous sounds and the stars with 8% mistuned extraneous sounds.

Figure 9. RMS error between the parent continuum and the other continua with extraneous sounds. The abscissas plot the RMS error in percent and the ordinates the F1 frequency in Hz. The different symbols indicate the amount of mistuning. The shared harmonic in the vowel continuum is indicated by "H2", "H3" and "H4" for harmonics 2, 3 and 4 respectively. Note that the lower a score in RMS error is, the more like the parent continuum it becomes.

Figure 10. RMS error between the parent continuum and the other continua with extraneous sounds averaged over the vowel continuum. The abscissas plot the RMS error in percent and the ordinates the amount of mistuning of the 5th harmonic of the extraneous sounds in percent. The different listeners and the shared harmonics of the vowel continuum are indicated in the panels of the figure. Note that the lower a score in RMS error is, the more like the parent continuum it becomes.

Figure 11. RMS error between the parent continuum and the other continua with extraneous sounds averaged over the vowel continuum and the five listeners. The abscissas plot the RMS error in percent and the ordinates the amount of mistuning of the 5th harmonic of the extraneous sounds in percent. The error bars show plus/minus one intra-listener standard deviation computed according to the criteria recommended by Winer *et al* (1991) for repeated measures designs. The shared harmonics of the vowel continuum are indicated in the panels of the figure. Note that the lower a score in RMS error is, the more like the parent continuum it becomes.

Figure 12. Fisher z′ scores computed from Pearson correlation coefficients between the parent continuum and the other continua with extraneous sounds averaged over the five listeners. The abscissas plot the Fisher z′ score and the ordinates the amount of mistuning of the 5th harmonic of the extraneous sounds in percent. The error bars show plus/minus one intra-listener standard deviation computed according to the criteria recommended by Winer *et al* (1991) for repeated measures designs. The shared harmonics of the vowel continuum are indicated in the panels of the figure. Note that the higher a Fisher z′ value is, the more like the parent continuum it becomes. This is the opposite of the RMS error plots which is the reason the abscissas have been inverted.

Figure 13. The contrast of the F1 across the different members of the vowel continuum. Contrast is defined by how many dB the harmonics of the F1 of the vowel rose above the level of the harmonics of the extraneous sounds. The abscissa plots the contrast in dB and the ordinate plots the F1 frequency in Hz.

Figure 14. Results of individual listeners for all three shared harmonics for the S continua. The crosses show the parent continua for reference. The circles shows the results with no mistuning. The squares shows the results with 3% mistuning and the stars with 8% mistuning.

Figure 15. RMS error between the parent continuum and the S continua. The abscissas plot the RMS error in percent and the ordinates the F1 frequency in Hz. The different symbols indicate the amount of mistuning. The shared harmonic in the vowel continuum is indicated by "H2", "H3" and "H4" for harmonics 2, 3 and 4 respectively.

Figure 16. RMS error between the parent continuum and the S continua averaged over the vowel continua and the five listeners. The abscissas plot the RMS error in percent and the ordinates the amount of mistuning. The error bars show plus/minus one intra-listener standard deviation computed according to the criteria recommended by Winer *et al* (1991) for repeated measures designs. The shared harmonics of the vowel continuum are indicated in the panels of the figure.

Figure 17. Fisher z′ scores computed from Pearson correlation coefficients between the parent continuum and the S continua averaged over the five listeners. The abscissas plot the Fisher z′ score and the ordinates the amount of mistuning. The error bars show plus/minus one intra-listener standard deviation computed according to the criteria recommended by Winer *et al* (1991) for repeated measures designs. The shared harmonics of the vowel continuum are indicated in the panels of the figure.

Table 1. Repeated measures ANOVA from Experiment 1 using the RMS score. The factors were extraneous sound (full or notched), amount of mistuning (0%, 3% or 8%), which harmonic of the vowel continuum was shared (second, third or fourth) and listeners. Listeners was specified as a random factor. All possible interactions were used in the ANOVA except for Ex_S*Mis*Harm*Listeners. Interactions with listeners are not reported as they are not of interest. The degrees of freedom for the denominator were adjusted using Satterthwaite's method (Satterthwaite, 1941) as are all other repeated measures ANOVA reported.

Source	SS	MS Den	MS Num	DF Den	DF Num	F Ratio	Prob > F
Ex_Sound (Ex_S)	0.27	16.55	0.27	4	1	0.02	0.9043
Mistuning (Mis)	82.65	20.17	41.33	8	2	2.04	0.1913
Ex_S*Mis	19.52	10.66	9.76	8	2	0.91	0.4383
Harmonic (Harm)	1270.61	144.9	635.30	8	2	4.38	0.0518
Ex_S*Harm	159.26	16.78	79.64	8	2	4.74	**0.0437**
Mis*Harm	296.85	7.30	74.21	16	4	10.16	**0.0003**
Ex_S*Mis*Harm	352.70	10.20	88.18	16	4	8.64	**0.0006**
Listeners	2702.53	157.11	675.58	8.8	4	4.30	**0.0333**

Table 2. Contrast tests for the interaction mistunin*shared harmonic. The table shows the significance levels of t-tests between the different mistunings conditions for the three shared harmonic conditions. Note that contrast tests are not *post-hoc* tests and therefore some of the t-tests might be falsely significant due to Type II error.

Mistuning ->	0% vs. 3%	0% vs. 8%	3% vs. 8%
2nd harmonic	**0.0003**	**0.0001**	0.2730
3rd harmonic	**0.0003**	0.6488	0.9461
4th harmonic	0.0861	**0.0489**	**0.0011**

Table 3. Contrast tests for the interaction extraneous sound*mistuning*shared harmonic. The table shows the significance levels of t-tests between the two extraneous sounds for the different mistunings conditions and for the three shared harmonic conditions. Note that contrast tests are not *post-hoc* tests and therefore some of the t-tests might be falsely significant due to Type II error.

Mistuning ->	0%	3%	8%
2nd harmonic	0.3276	0.2602	0.7023
3rd harmonic	**0.0008**	0.9790	0.4604
4th harmonic	**0.0004**	**0.0114**	0.0702

Table 4. Repeated measures ANOVA from Experiment 1 using the Fisher z′ (correlation) score. The factors were extraneous sound (full or notched), amount of mistuning (0%, 3% or 8%), which harmonic of the vowel continuum was shared (second, third or fourth) and listeners. Listeners was specified as a random factor. All possible interactions were used in the ANOVA except for Ex_S*Mis*Harm*Listeners. Interactions with listeners are not reported as they are not of interest.

Source	SS	MS Den	MS Num	DF Den	DF Num	F Ratio	Prob > F
Ex_Sound (Ex_S)	0.0834	0.0707	0.0834	4	1	1.18	0.3387
Mistuning (Mis)	0.4970	0.0691	0.2485	8	2	3.60	0.0768
Ex_S*Mis	0.0877	0.0248	0.0439	8	2	1.77	0.2309
Harmonic (Harm)	0.5188	0.1491	0.2594	8	2	1.74	0.2358
Ex_S*Harm	2.1289	0.2120	1.0619	8	2	5.02	**0.0386**
Mis*Harm	0.6478	0.0387	0.1619	16	4	4.19	**0.0166**
Ex_S*Mis*Harm	0.8483	0.0552	0.2120	16	4	3.84	**0.0266**
Listeners	1.4806	0.0687	0.3701	0.44	4	5.39	0.5114

Table 5. Repeated measures ANOVA from Experiment 12 using the RMS score. The factors were amount of mistuning (0%, 3% or 8%), which harmonic of the vowel continuum was shared (second, third or fourth) and listeners. Listeners was specified as a random factor. All possible interactions were used in the ANOVA except for Mis*Harm*Listeners. Interactions with listeners are not reported as they are not of interest.

Source	SS	MS Den	MS Num	DF Den	DF Num	F Ratio	Prob > F
Mistuning (Mis)	15.96	8.21	7.98	8	2	0.97	0.42
Harmonic (Harm)	324.44	78.49	162.22	8	2	2.07	0.19
Mis*Harm	105.58	11.31	26.40	16	4	2.33	0.10
Listeners	1117.63	75.92	279.41	7.2	4	3.71	0.06

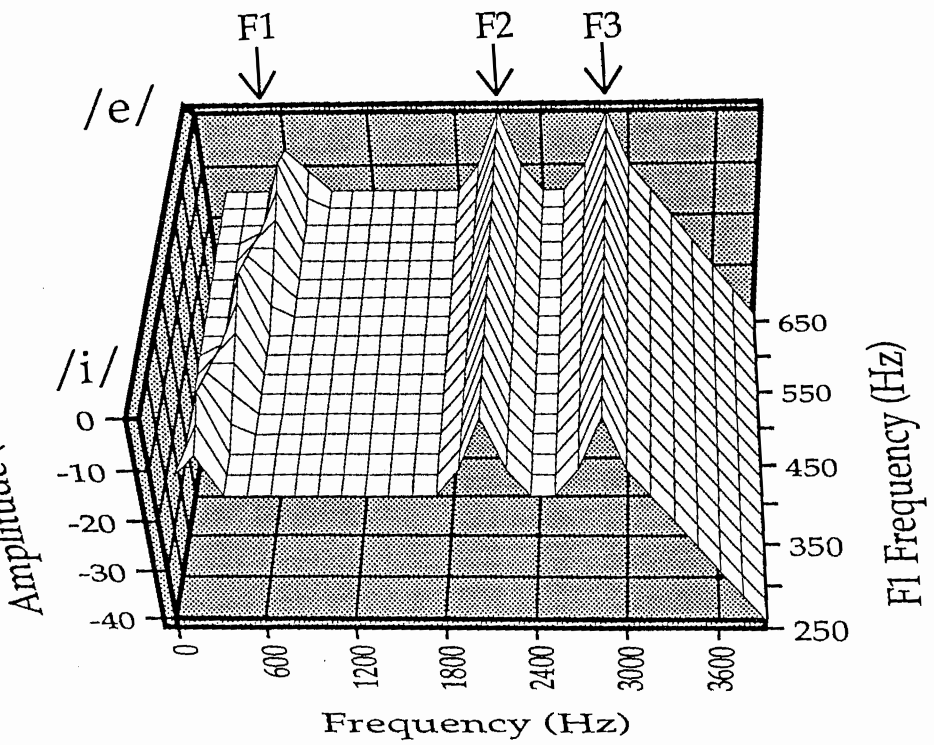
Table 6. Repeated measures ANOVA from Experiment 2 using the Fisher z' (correlation) score. The factors were amount of mistuning (0%, 3% or 8%), which harmonic of the vowel continuum was shared (second, third or fourth) and listeners. Listeners was specified as a random factor. All possible interactions were used in the ANOVA except for Mis*Harm*Listeners. Interactions with listeners are not reported as they are not of interest.

Source	SS	MS Den	MS Num	DF Den	DF Num	F Ratio	Prob > F
Mistuning (Mis)	0.3335	0.0314	0.1668	8	2	5.31	**0.03**
Harmonic (Harm)	1.7204	0.3677	0.8602	8	2	2.34	016
Mis*Harm	0.3823	0.0768	0.0956	16	4	1.24	0.33
Listeners	2.9774	0.3223	0.7444	6	4	2.31	0.17

Table A1. Results of a pilot experiment of Lea (1993). The vowel /i/ had no F1, the F1 of the vowel /e/ was defined by raising the 600 Hz harmonic by 12 dB. The table shows the number of /i/ responses in percent for all five listeners and their mean and standard deviation.

Listener	/e/	/i/
AL	0	100
FP	20	40
IE	20	60
JM	0	10
RW	10	90
Mean	10	60
S. D.	10.0	36.7

Figure 1


F1
F2
F3
/e/
/i/
Amplitude (dB)
0
-10
-20
-30
-40
0
600
1200
1800
2400
3000
3600
Frequency (Hz)
650
550
450
350
250
F1 Frequency (Hz)

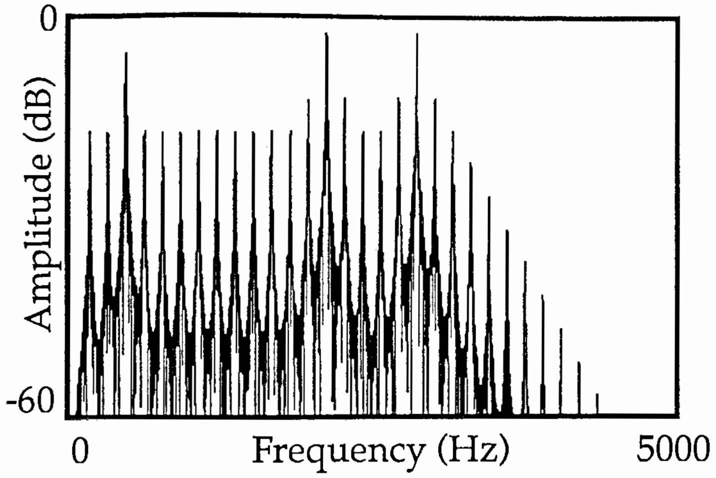


Figure 2

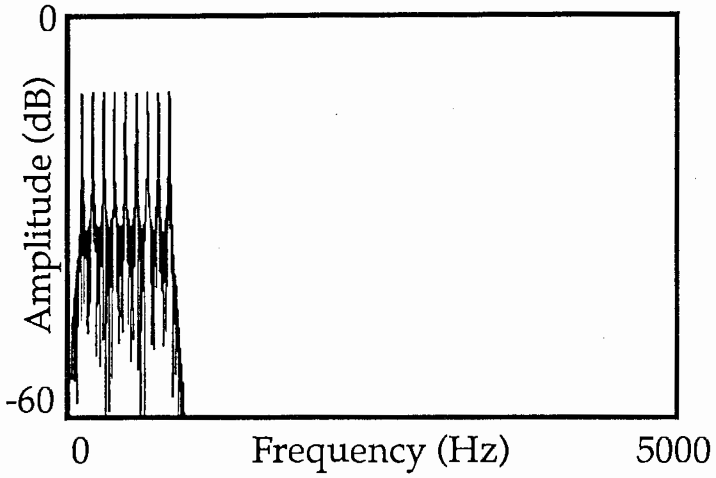


Figure 3

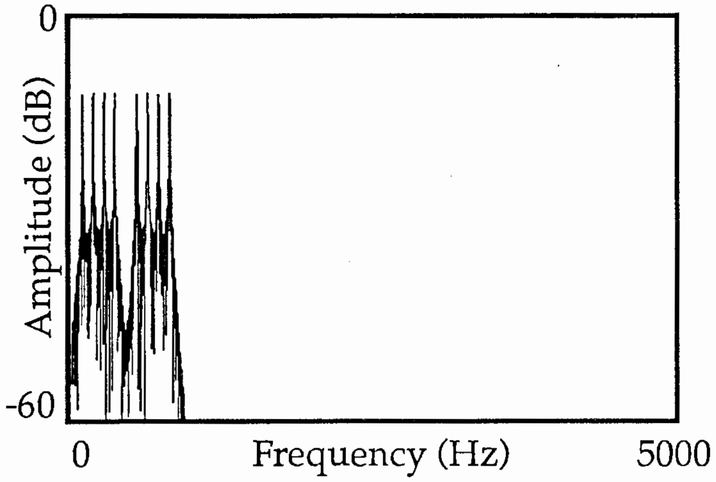


Figure 4

Figure 5

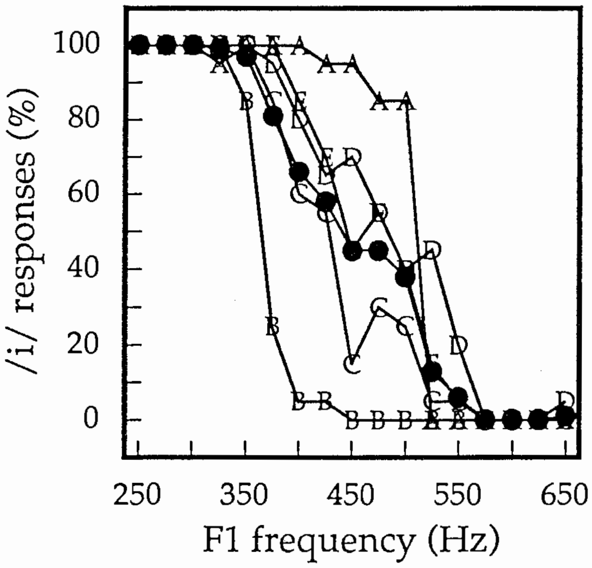

100
80
60
40
20
0
/i/ responses (%)
250
350
450
550
650
F1 frequency (Hz)
A
B
C
D
E

Figure 6

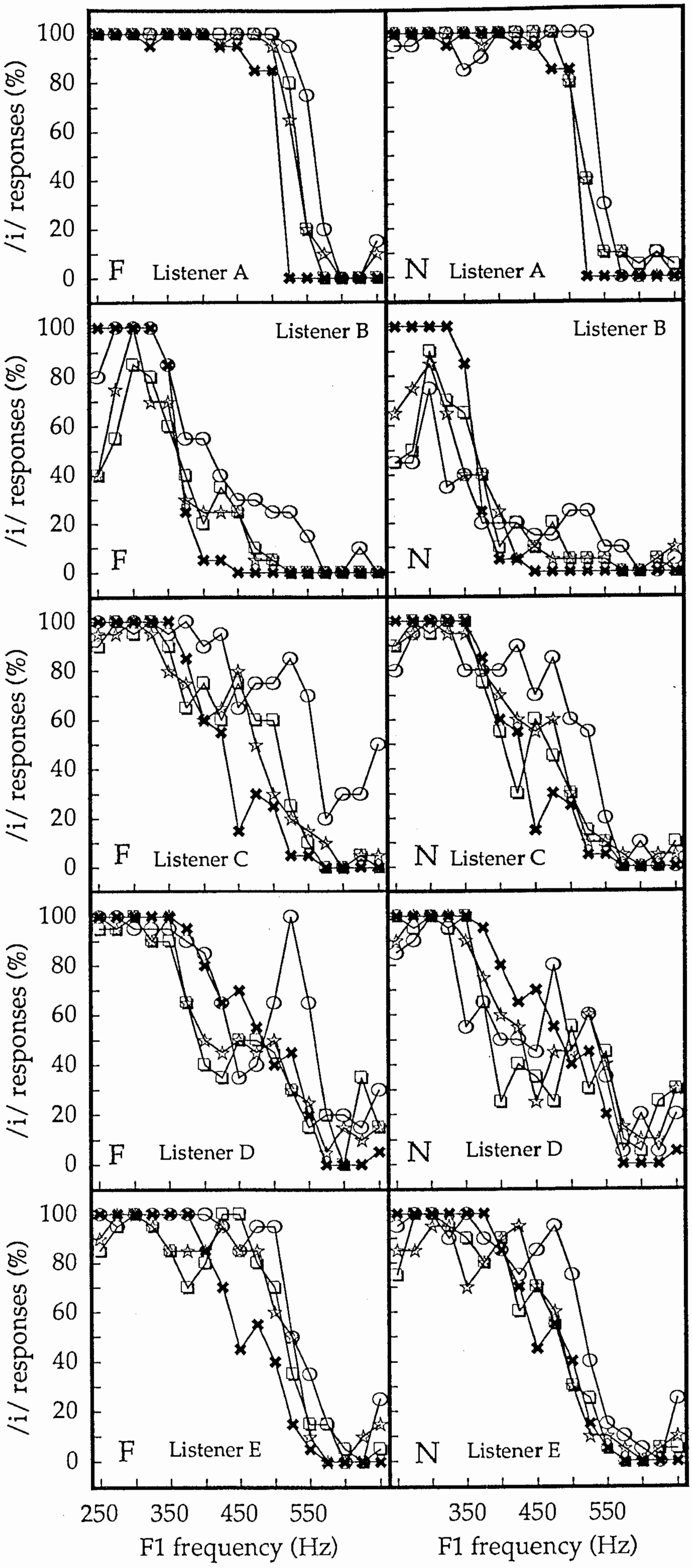

100
80
60
40
20
0
/i/ responses (%)
F
Listener A
N
Listener A
Listener B
F
N
Listener B
F
Listener C
N
Listener C
F
Listener D
N
Listener D
F
Listener E
N
Listener E
250
350
450
550
F1 frequency (Hz)
350
450
550
F1 frequency (Hz)

Figure 7

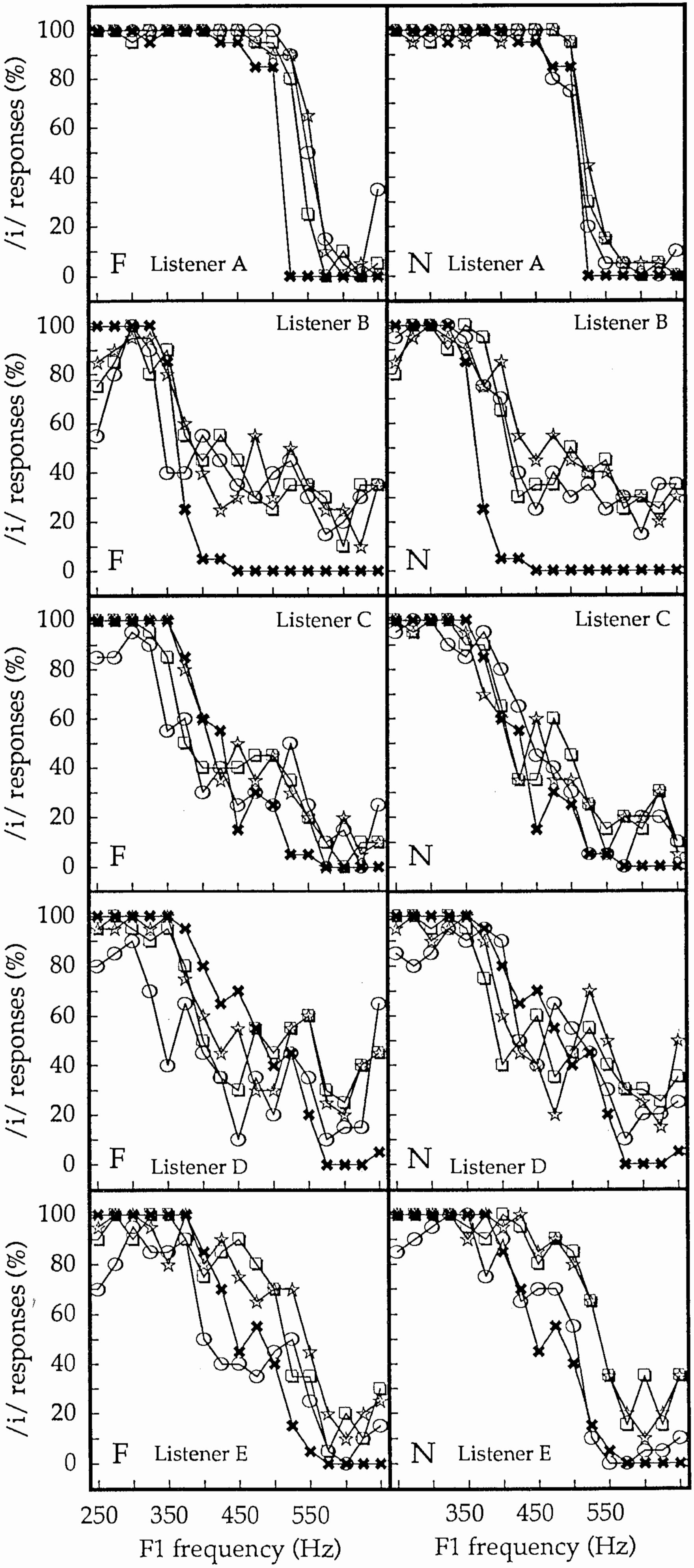

/i/ responses (%)
100
80
60
40
20
0
F
N
Listener A
Listener B
Listener C
Listener D
Listener E
250
350
450
550
F1 frequency (Hz)

Figure 8

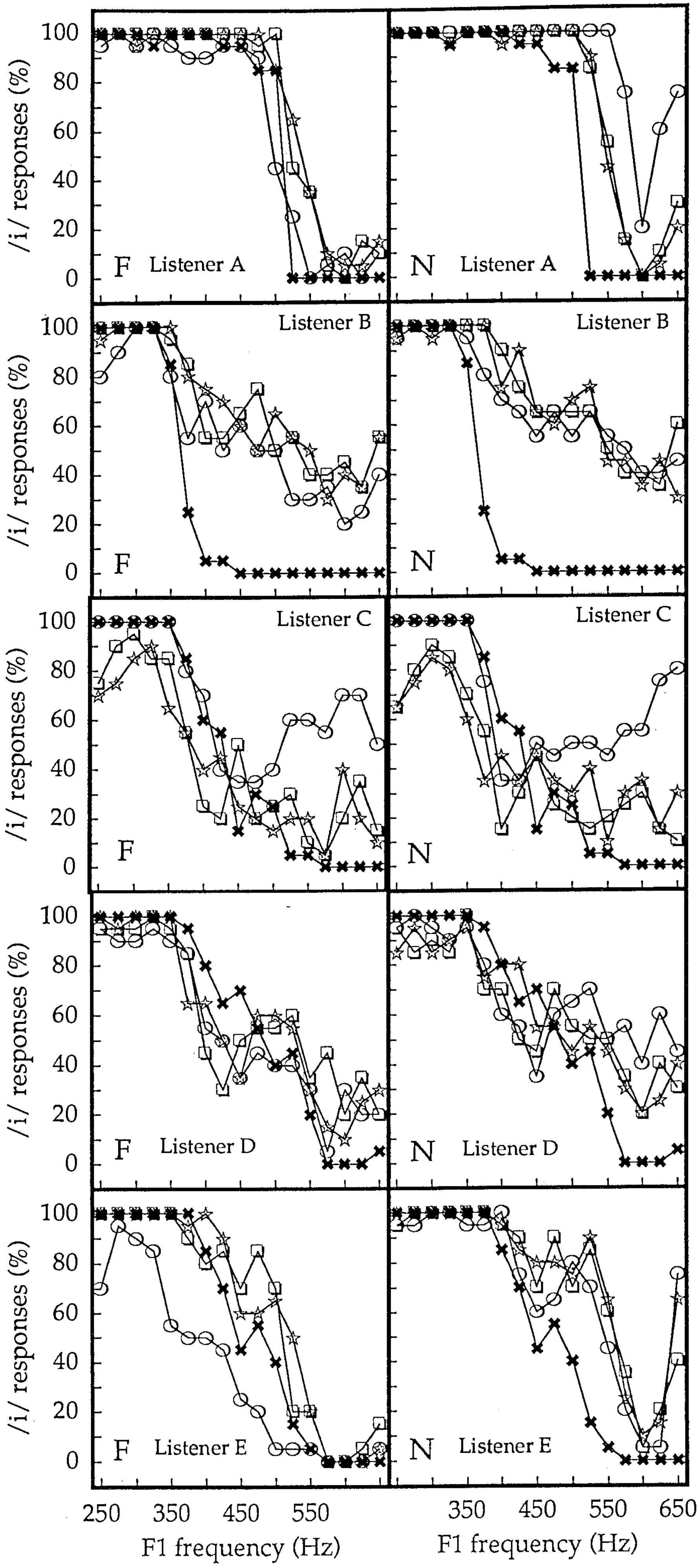

/i/ responses (%)
100
80
60
40
20
0
F
N
Listener A
Listener B
Listener C
Listener D
Listener E
250
350
450
550
650
F1 frequency (Hz)

Figure 9

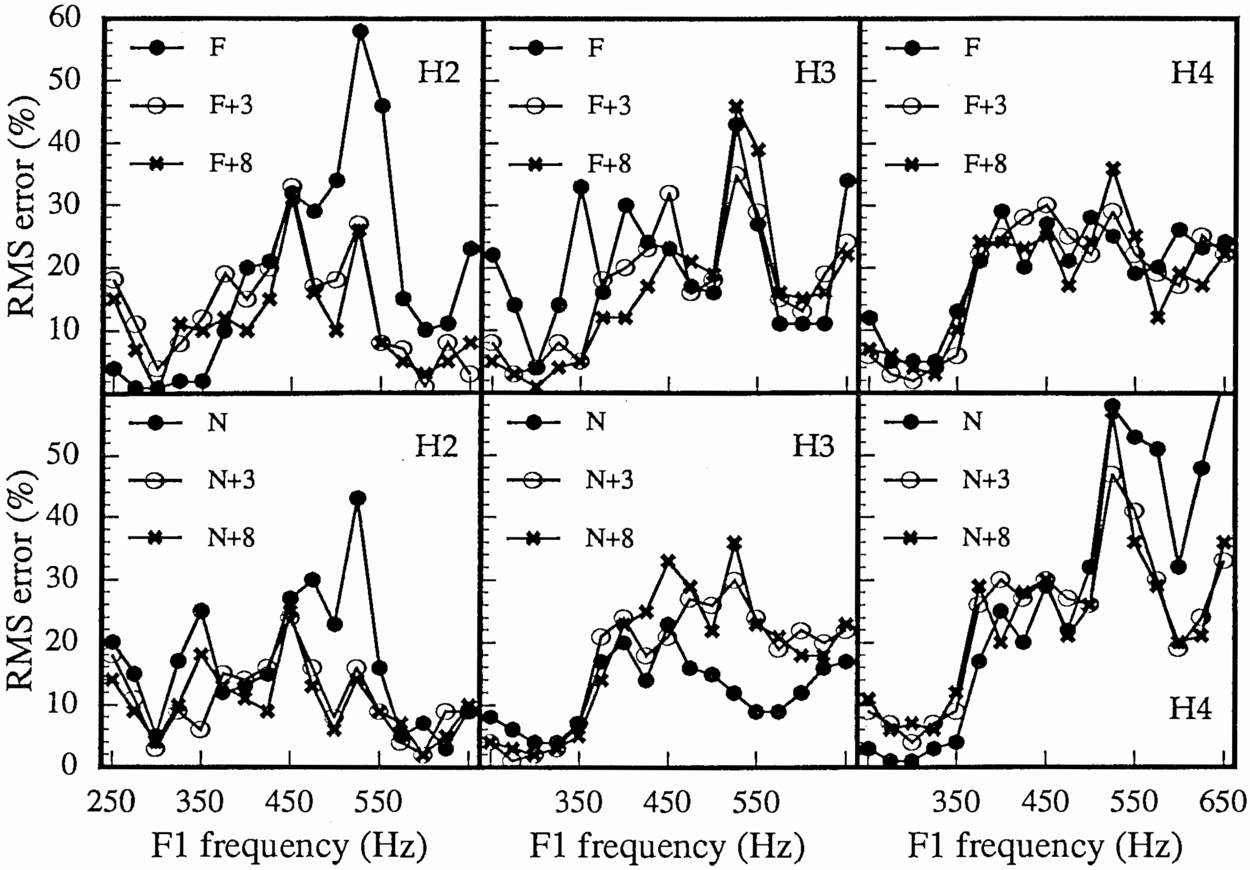

F
F+3
F+8
H2
H3
H4
N
N+3
N+8
RMS error (%)
60
50
40
30
20
10
0
250
350
450
550
650
F1 frequency (Hz)

Figure 10

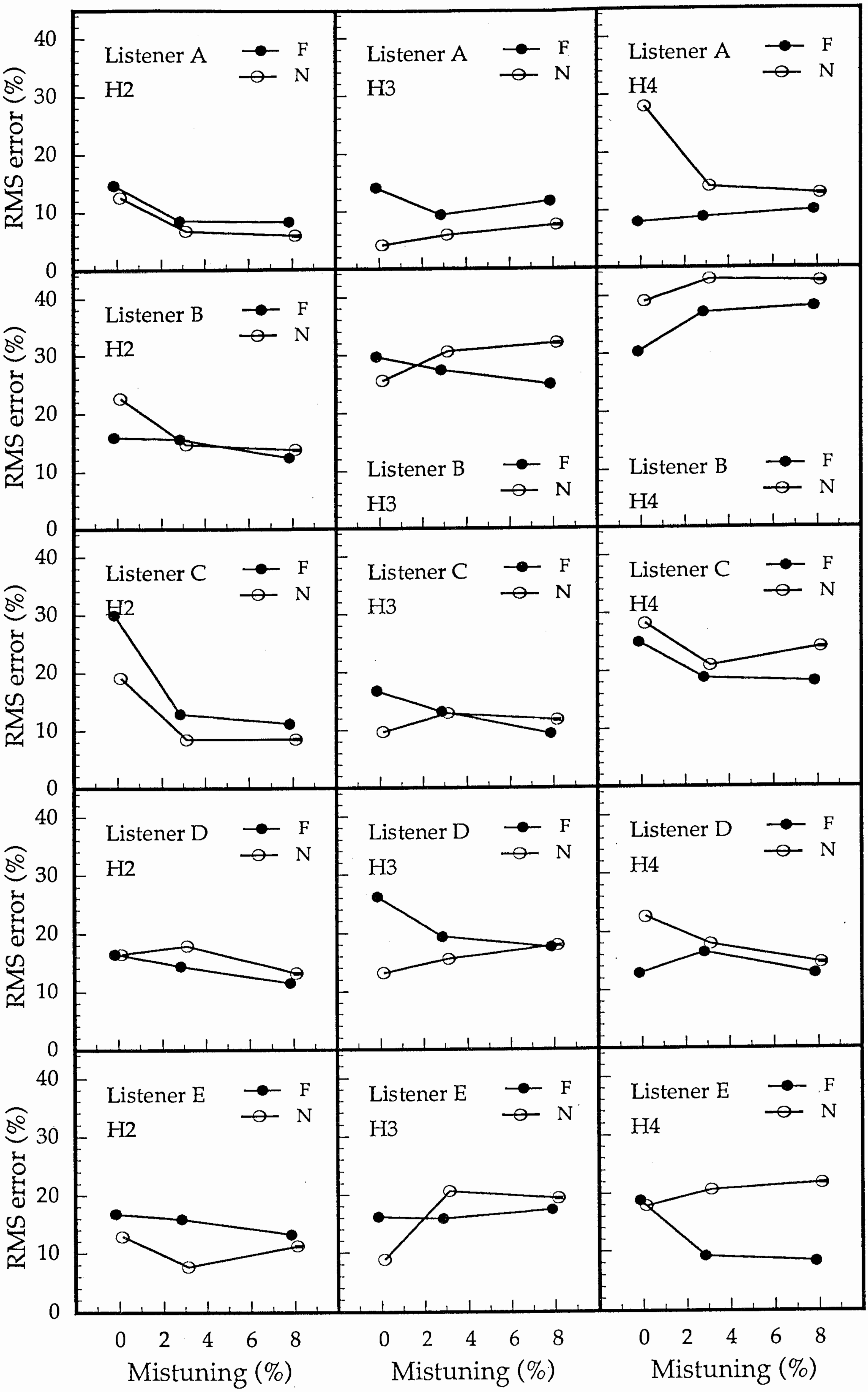

Listener A
H2
Listener A
H3
Listener A
H4
Listener B
H2
Listener B
H3
Listener B
H4
Listener C
H2
Listener C
H3
Listener C
H4
Listener D
H2
Listener D
H3
Listener D
H4
Listener E
H2
Listener E
H3
Listener E
H4
F
N
RMS error (%)
0
10
20
30
40
0
2
4
6
8
Mistuning (%)

Figure 11

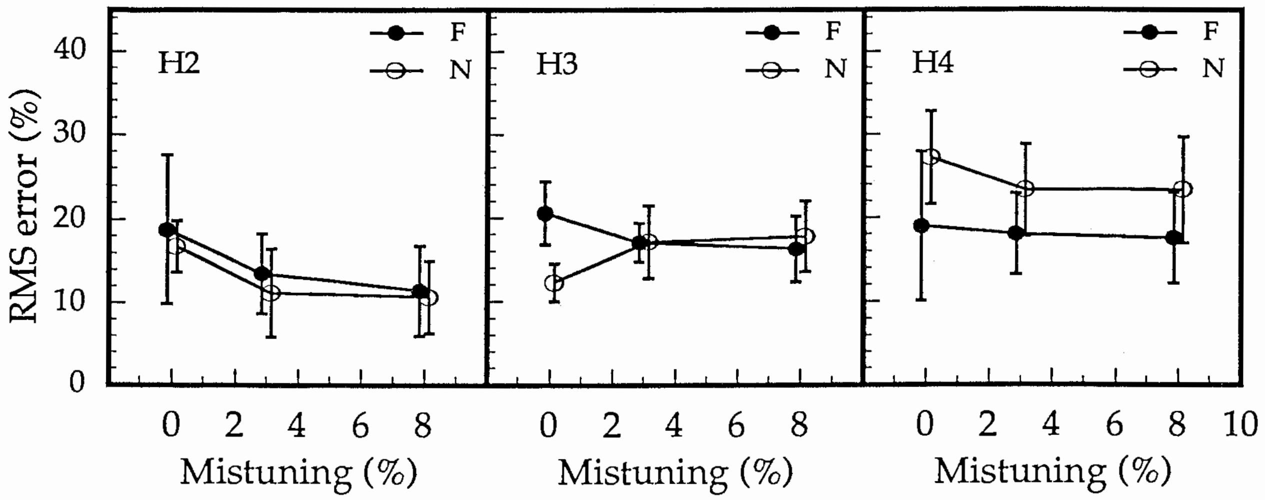
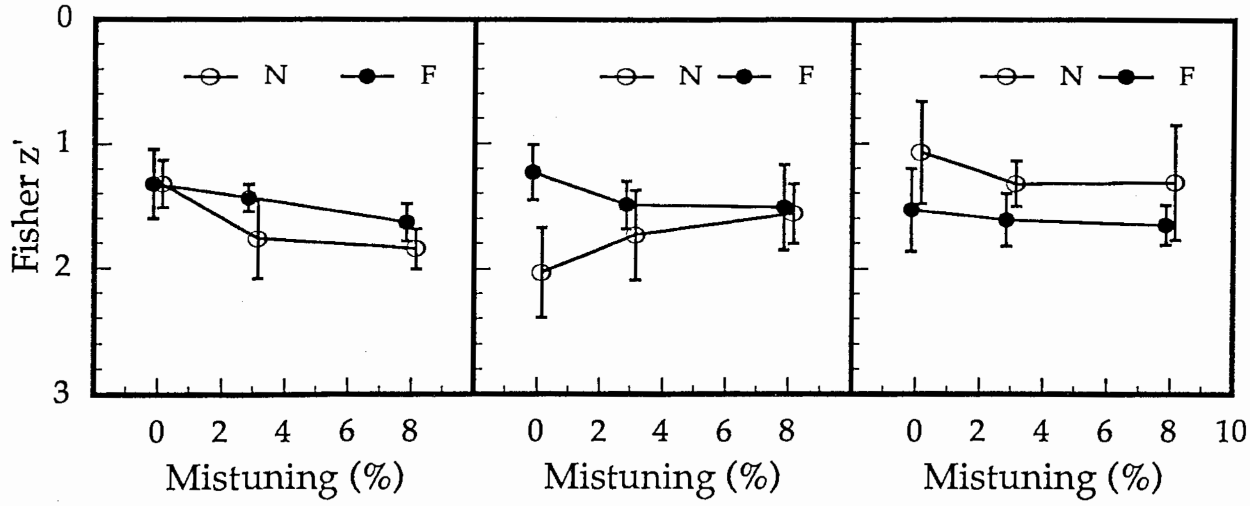


Figure 12



0
1
2
3
Fisher z'
N
F
N
F
N
F
0
2
4
6
8
Mistuning (%)
0
2
4
6
8
Mistuning (%)
0
2
4
6
8
10
Mistuning (%)

Figure 13

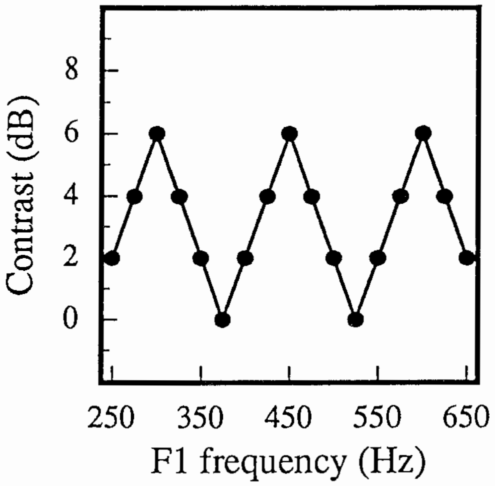
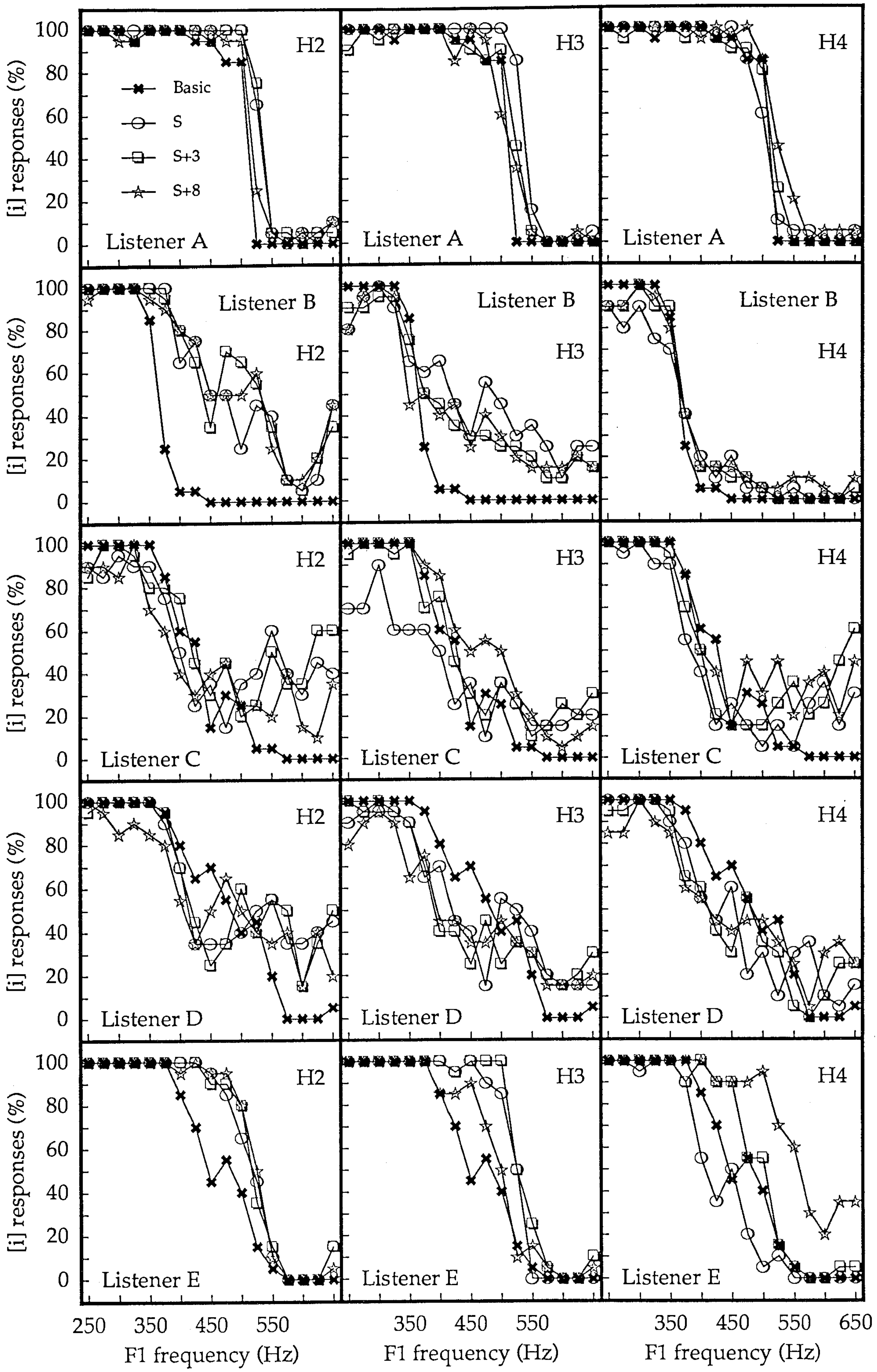


Figure 14



[i] responses (%)
100
80
60
40
20
0
Basic
S
S+3
S+8
H2
H3
H4
Listener A
Listener B
Listener C
Listener D
Listener E
250
350
450
550
650
F1 frequency (Hz)

Figure 15

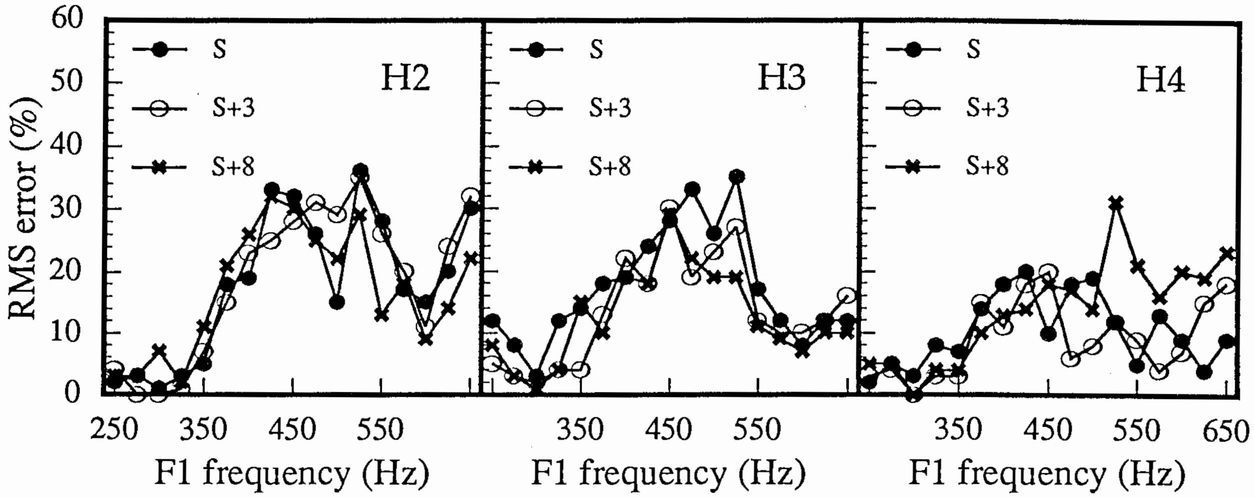

60
50
40
30
20
10
0
RMS error (%)
S
S+3
S+8
H2
H3
H4
250
350
450
550
650
F1 frequency (Hz)
F1 frequency (Hz)
F1 frequency (Hz)

Figure 16

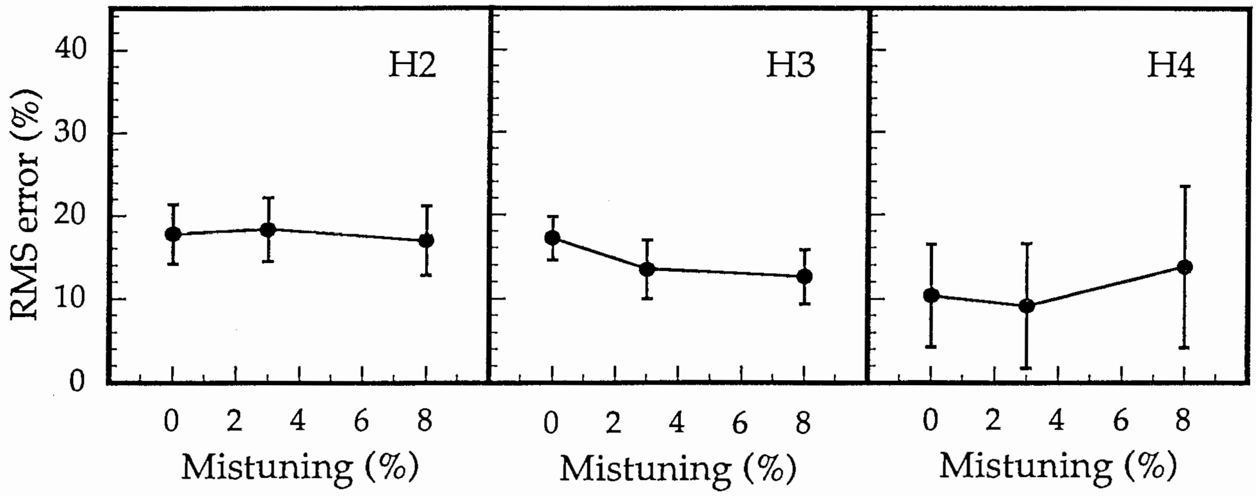
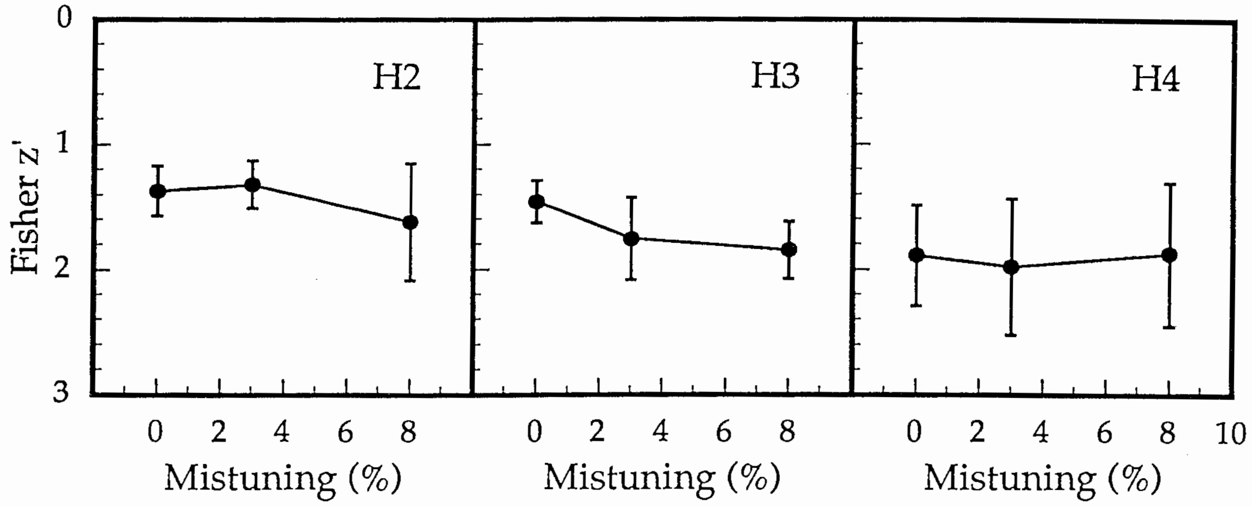

RMS error (%)
0
10
20
30
40
H2
H3
H4
0
2
4
6
8
Mistuning (%)
0
2
4
6
8
Mistuning (%)
0
2
4
6
8
Mistuning (%)

Figure 17



H2
H3
H4
Fisher z'
0
1
2
3
0
2
4
6
8
10
Mistuning (%)
Mistuning (%)
Mistuning (%)