24.963 Linguistic Phonetics
Fall 2005

24.963

Linguistic Phonetics

Basic statistics

Reading for week 8:

- Johnson chapters 4 and 9.

Assignments:

- Write up pre-nasal lengthening experiment.
- Come up with a paper topic.

Acoustics assignment - question 2

- The ear canal is a tube open at one end and closed at the other (the eardrum), so the air in it resonates at:

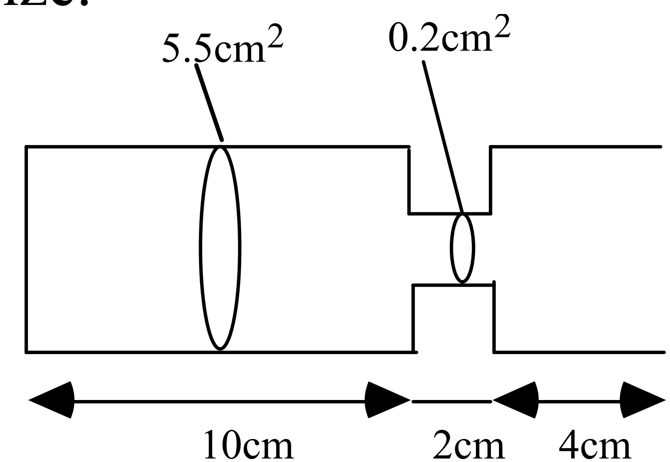
$$F_n = \frac{(2n-1)c}{4L} = \frac{(2n-1)\times 35000}{10}$$

= 3500 Hz, 10500 Hz, etc

- So components of a sound wave which are near these frequencies will be amplified in the signal at the ear-drum, compared to components which are not near a resonant frequency.
- The most relevant boost is to frequencies around 3500 Hz, since there is little energy at 10000 Hz and above in speech signals.

Acoustics assignment - question 2

- The figure below shows a simple tube model for the high front vowel [i]. Ignoring the effects of acoustic coupling, F3 is the first resonance of the front cavity, F2 is the first resonance of the back cavity, and F1 is the helmholtz resonance of the back cavity and constriction.
- The dimensions of this vocal tract are appropriate for an adult male. A typical female vocal tract might be about 90% of this size.



Acoustics assignment - question 2

What are the proportional changes in F2 and F3 if the dimensions of this vocal tract shape are shortened by this amount?

- In the model, F2 is the first resonance of the back cavity, F3 is the first resonance of the front cavity.

$$F2 = \frac{c}{2L_b} \qquad F3 = \frac{c}{4L_f}$$

- So F2 and F3 are inversely proportional to the length of the relevant cavity.
- So a reduction of length by a factor of 0.9 increases formants by 1/0.9 = 1.11

Acoustics assignment - question 2

What is the proportional change in F1 if all dimensions are reduced to 90%?

- F1 is the helmholtz resonance of the back cavity plus constriction.

$$F_1 = \frac{c}{2\pi}\sqrt{\frac{A_c}{Vl_c}}$$

$$femaleF_1 = \frac{c}{2\pi}\sqrt{\frac{0.9^2 A_c}{0.9^3 V \times 0.9 l_c}}$$

$$femaleF_1 = \sqrt{\frac{0.9^2}{0.9^4}} maleF_1 = \frac{1}{0.9} maleF1$$

Acoustics assignment - question 2

- Uniform scaling model predicts that female/male ratios should be equal for all formants (1.11).
- Observed ratios from Peterson and Barney (1952) differ across formants.
- Chiba and Kajiyama's observations suggest that scaling is not uniform: the difference in size of the pharynx is larger than the difference in size of mouth.
- In [i] F2 is a back cavity resonance and F3 is a front cavity resonance, so we expect a larger ratio for F2 than for F3.

		F1	F2	F3
[i]	Males	270	2290	3010
	Females	310	2790	3310
	Ratio	1.15	1.22	1.10

Acoustics assignment - question 2

- Why is F1 ratio also lower than the F2 ratio of F1 is a back cavity (helmholtz) resonance?
- F1 is also affected by constriction area and length, so differences in the scaling of these dimensions could result in the lower ratio for F1.
- E.g. constriction length may be reduced by less than the back cavity, or the constriction area may be reduced by more (F1 is inversely proportional to the square root of constriction length, and proportional to square root of constriction area). (cf. Fant 1972).

		F1	F2	F3
[i]	Males	270	2290	3010
	Females	310	2790	3310
	Ratio	1.15	1.22	1.10

Writing up an experiment

The report on an experiment usually consists of four basic parts:

1. Introduction
2. Procedure
3. Results
4. Discussion

Writing up an experiment

1. Introduction
- Outline of the purpose of the experiment
- state hypotheses tested etc
- provide background information (possibly including descriptions of relevant previous results, theoretical issues etc).

2. Procedure - what was done and how.
- instructions for replication, e.g.
  - Experimental materials
  - Subjects
  - Recording procedure
  - Measurement procedures (especially measurement criteria**).**

Writing up an experiment

3. Results
- Presentation of results, including descriptive statistics (means etc) and statistical tests of hypotheses.

4. Results
- Discuss the interpretation and significance of the results

Some Statistics

Two uses of statistics in experiments:

- Summarize properties of the results (descriptive statistics).
- Test the significance of results (hypothesis testing).

Descriptive statistics

A measure of central tendency:

- Mean:

$$M = \frac{\Sigma x_i}{N}$$

  - M is used for sample mean, μ for population mean.

A measure of dispersion:

- Variance: mean of the squared deviations from the mean

$$\sigma^2 = \frac{\Sigma (x_i - \mu)^2}{N}$$

- Standard deviation: σ (square root of the variance).

Hypothesis Testing

- Mean vowel duration

	__n	__ nd	__d
	176 ms	167 ms	174 ms

- Are these differences in means significant?
- Could the apparent differences have arisen by chance, although the true (population) means of vowel durations in open and closed syllables are the same?

- I.e. given that vowels durations vary, we might happen to sample most of our open syllables from the low end of the distribution, and most of our closed syllables from the high end.
- Statistical tests allow us to assess the probability that this is the case.

Hypothesis Testing: *t*-test

- The *t*-test allows us to test hypotheses concerning means and differences between means.
  1. 'The mean duration of vowels preceding -nd] differs from the mean duration of vowels preceding -n]'.
  2. 'The mean duration of [æ] is 150 ms' (unlikely, but a simpler case).

- We actually evaluate two exhaustive and mutually exclusive hypotheses, a **null hypothesis** that the mean has a particular value, and the alternative hypothesis that the mean does not have that value.
  1. The mean duration of vowels in preceding -nd] is the same as the mean duration of vowels preceding -n] (Null).
  2. The mean duration of [æ] $\neq$ 150 ms (Alternative).

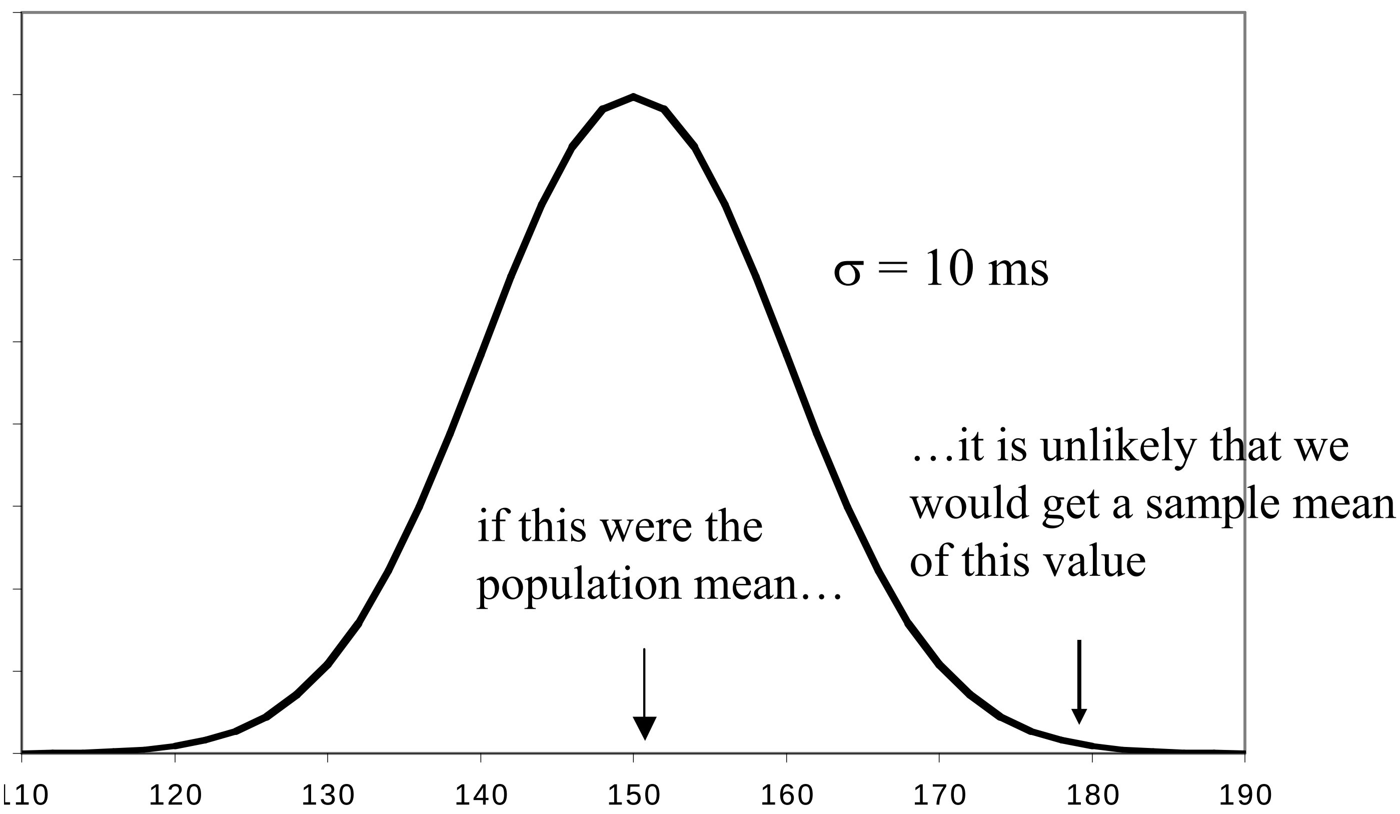
- Statistical tests allow us to assess the probability of obtaining the observed data if the null hypothesis were true.

Hypothesis Testing: *t*-test

- Basic concept: If we know what the distribution of sample means would be if the null hypothesis were true, then we can calculate the probability of obtaining the observed mean, given the null hypothesis.
- We arrive at the parameters of the distribution of sample means through assumptions and estimation.

Hypothesis Testing

Distribution of sample means



Hypothesis Testing

- Basic assumption: The samples are drawn from normal populations.

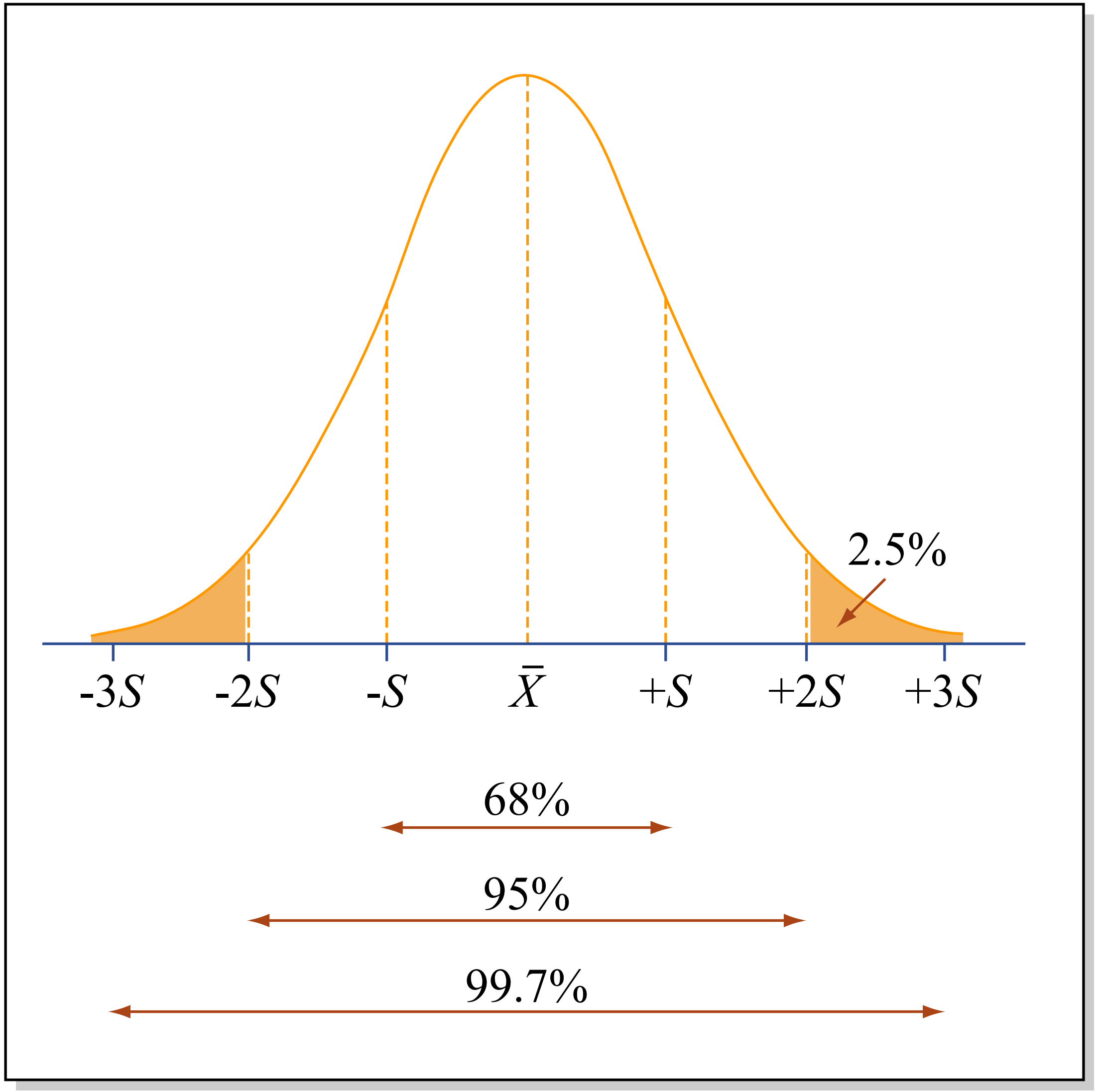


Image by MIT OpenCourseWare. Adapted from Kachigan, S. K. *Multivariate Statistical Analysis*. 2nd ed. New York, NY: Radius, 1991.

Hypothesis Testing

- Basic assumption: The samples are drawn from normal populations.
- Properties of distribution of means of samples of size *N* drawn from a normal population:
  - The sample means are normally distributed.
  - Mean is the same as the population mean.
  - The variance is less than the population variance:

$$\sigma_M^2 = \frac{\sigma^2}{N}$$

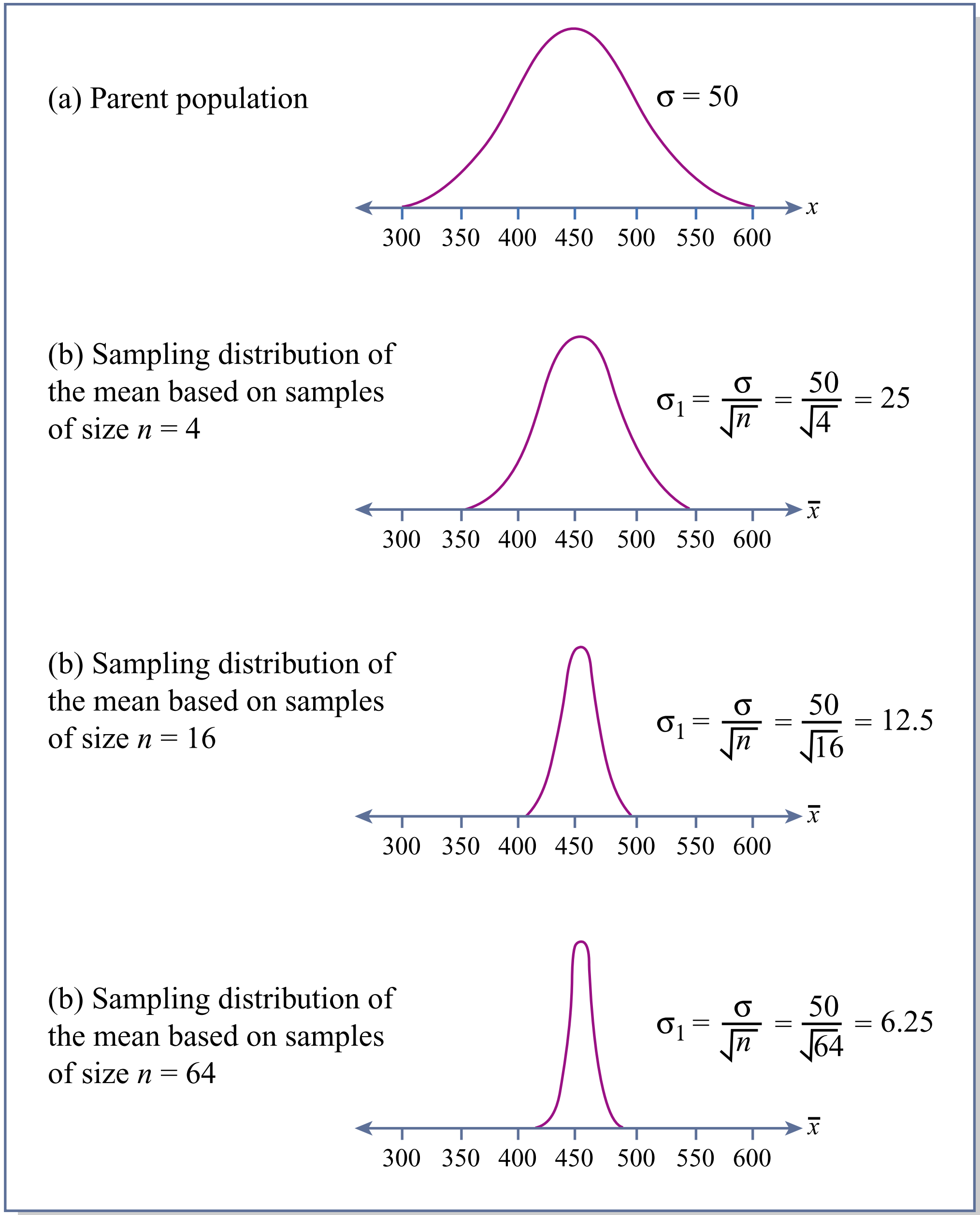


Image by MIT OpenCourseWare. Adapted from Kachigan, S. K. *Multivariate Statistical Analysis*. 2nd ed. New York, NY: Radius, 1991.

Hypothesis Testing

- The mean of the distribution is determined by hypothesis.
  - E.g. mean = 150 ms or mean difference = 0.
- Population variance is estimated from the sample variance. Unbiased estimate of the population variance:

$$S^2 = \frac{\Sigma(x_i - M)^2}{N-1}$$

  - N-1 is the number of degrees of freedom of the sample.
- So estimated variance of distribution of sample means, $S_M^2 = S^2/N$
- *t* score:

$$t = \frac{M - \mu}{S_M}$$

Hypothesis Testing

- *t* scores follow a *t*-distribution - similar to a normal distribution, but with slightly fatter tails (more extreme values) because S may underestimate σ.
- *t*-distribution is actually a family of distributions, one for each number of degrees of freedom.
- Calculate *t*-score then consult relevant *t* distribution to determine the probability of obtaining that *t*-score or greater (more extreme).

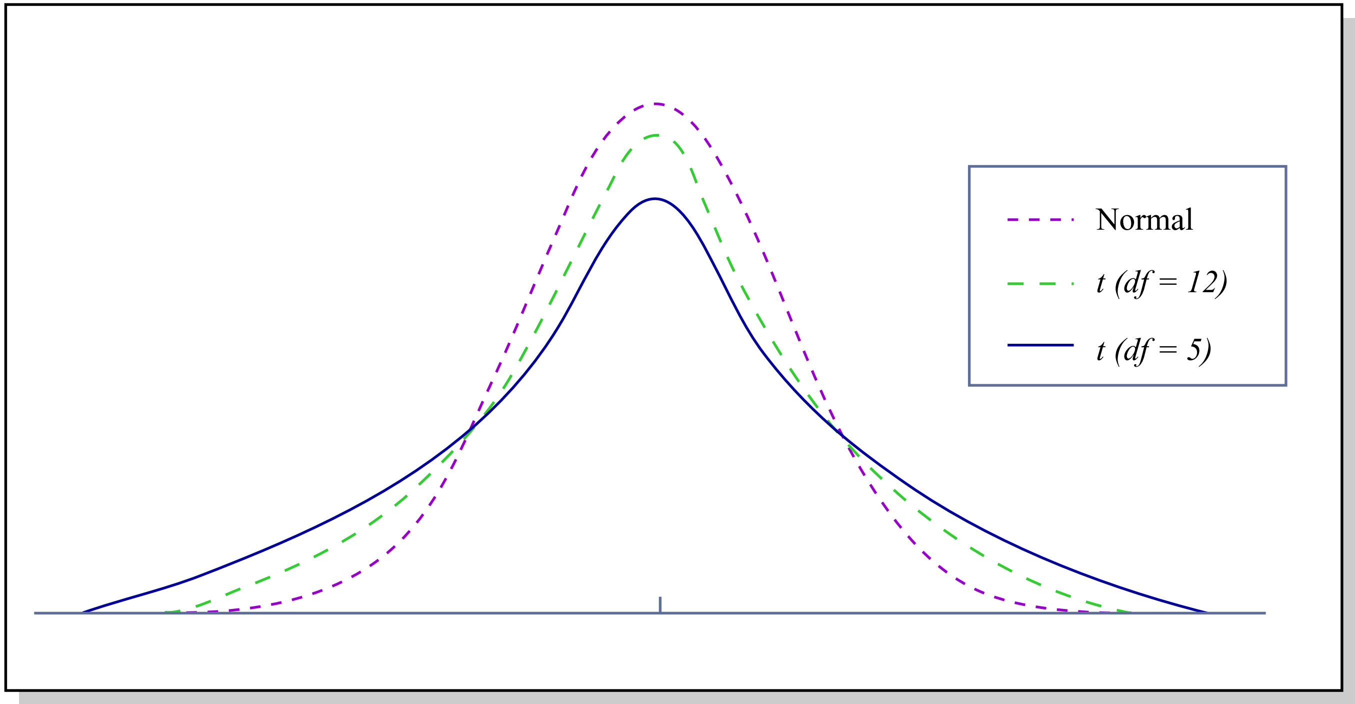


Figure by MIT OpenCourseWare.

t test for independent means

- When we compare means, we are actually sampling a population of differences (e.g. differences in durations of vowels in open and closed syllables).
- If the null hypothesis is correct, then the mean difference is 0.
- Variance of the distribution of mean differences is estimated based on the variances of the two samples.

Hypothesis testing

- Statistical tests like the *t* test give us the probability of obtaining the observed results if the null hypothesis were correct - the ‘p’ value. E.g. $p < 0.01$, $p = 0.334$.
- We reject the null hypothesis if the experimental results would be very unlikely to have arisen if the null hypothesis were true.
- How should we set the threshold for rejecting the null hypothesis?
  - Choosing a lower threshold increases the chance of incorrectly accepting the null hypothesis.
  - Choosing a higher threshold increases the chance of incorrectly rejecting the null hypothesis.
  - A common compromise is to reject the null hypothesis if $p < 0.05$, but there is nothing magical about this number.

Hypothesis testing

- In most experiments we need more complex statistical analyses than the *t* test (e.g. ANOVA), but the logic is the same: Given certain assumptions, the test allows us to determine the probability that our results could have arisen by chance in the absence of the hypothesized effect (i.e. if the null hypothesis were true).

Fitting models

- Statistical analyses generally involve fitting a model to the experimental data.
- The model in a t-test is fairly trivial, e.g.

duration = μ + syllable_type (syllable_type is 'open' or 'closed')

Fitting models

- Statistical analyses generally involve fitting a model to the experimental data.
- The model in a t-test is fairly trivial, e.g.

  duration = μ + syllable_type (syllable_type is 'open' or 'closed')

  $\text{duration}_{ij} = \mu + \text{syllable\_type}_i + \textit{error}_{ij}$

- Analysis of Variance (ANOVA) involves more complex models, e.g.

  $\text{dur}_{ijk} = \mu + \text{vowel}_i + \text{syll\_type}_j + \textit{error}_{ijk}$

  $\text{dur}_{ijk} = \mu + \text{vowel}_i + \text{syll\_type}_j + \text{vowel*syll\_type}_{ij} + \textit{error}_{ijk}$

- Model fitting involves finding values for the model parameters that yield the best fit between model and data (e.g. minimize the squared errors).
- Hypothesis testing generally involves testing whether some term or coefficient in the model is significantly different from zero.

24.963
Linguistic Phonetics
The acoustics of nasals and laterals

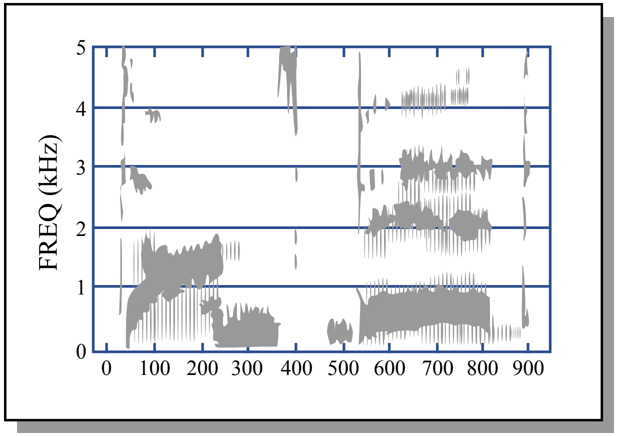


Image by MIT OpenCourseWare. Adapted from Stevens, K. N. *Acoustic Phonetics*. Cambridge, MA: MIT Press, 1999, chapter 6.

Damping and bandwidth

- When an input of energy sets a body vibrating, the resulting vibrations die out as energy is dissipated through friction, transmission of energy to surrounding air, etc.
- A wave of this type is described as **damped**.

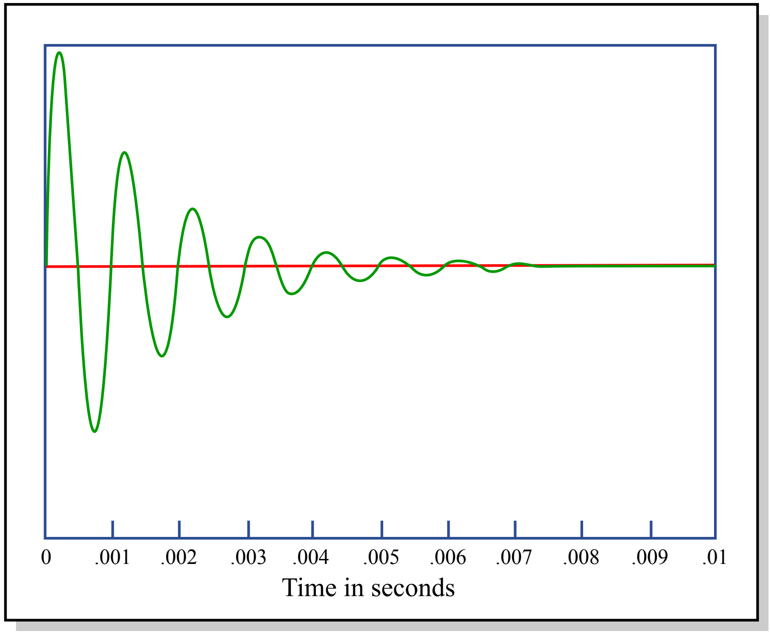
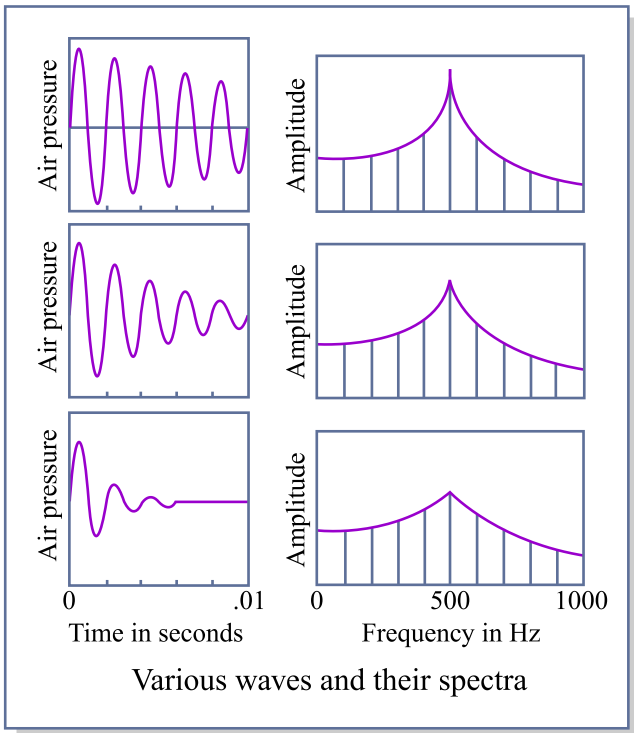


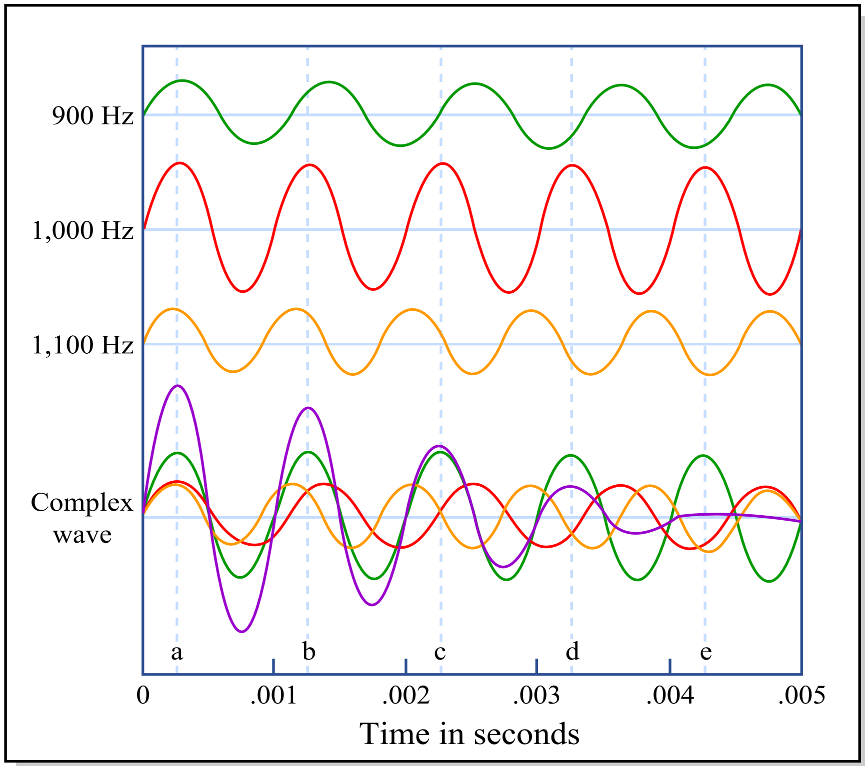
Image by MIT OpenCourseWare. Adapted from Ladefoged, Peter. *Elements of Acoustic Phonetics*. Chicago, IL: Univeristy of Chicago Press, 1962.

Damping and bandwidth

- The more heavily damped a sinusoid is, the wider the bandwidth of its spectrum.
- A wider bandwidth implies a lower spectral peak, other things being equal, since the energy in the wave is distributed over a wider range of frequencies.



Various waves and their spectra



Images by MIT OpenCourseWare. Adapted from Ladefoged, Peter. *Elements of Acoustic Phonetics*. Chicago, IL: Univeristy of Chicago Press, 1962.

Damping and bandwidth

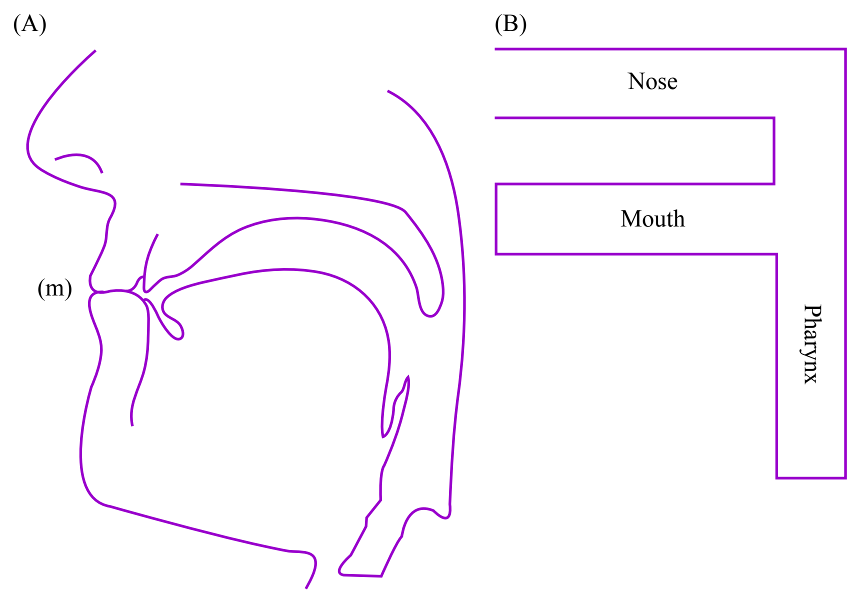
- Standing waves in the vocal tract are damped due to friction, radiation losses to the outside air, absorption of energy by the elastic vocal tract walls etc.
- Greater surface area results in more damping (and broader formant bandwidths).
- Coupling the nasal cavity to the vocal tract in in nasal sounds increases the surface area of the vocal tract.

Nasal stops

- A uvular nasal can be modeled in terms of a tube which is closed at the glottis and open at the nostrils.
- It is not a uniform tube.
- The effective length of the tube is greater than the length from glottis to lips
  - 21.5 cm (Johnson) or 20 cm (Stevens) vs. 16 cm.
- So the resonances of the tube are at lower frequencies (and closer together) than in a vowel.

Nasal stops

- In nasal consonants formed with an oral constriction further forward than the uvula, the oral cavity forms a side branch on the pharyngeal-nasal tube.
- This side branch is longest in labials, shortest in velars.

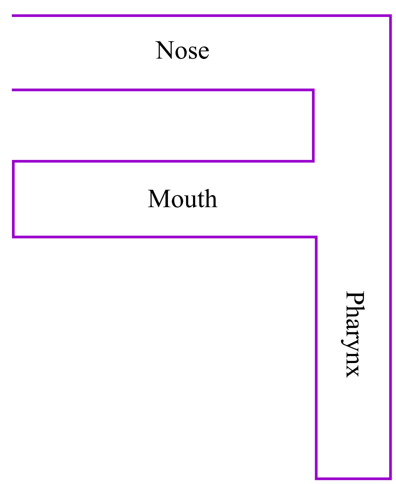


(A) An X-ray tracing of the positions of the articulators during (m).
(B) A tube model of this vocal tract configuration.

Image by MIT OpenCourseWare. Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997.

Nasal stops

- The side branch is a tube open at one end and resonates accordingly.
- The side branch is not coupled to the outside air. Its resonances are zeros in the signal radiated from the nose - i.e. those frequencies are attenuated.
- Zeroes also reduce the amplitude of higher frequencies, so nasals are characterized by less high frequency energy than vowels.
- Frequencies of the zeroes depend on the size of the side branch, e.g. 8 cm for [m], 5.5 cm for [n].



A tube model of this vocal tract configuration.

Image by MIT OpenCourseWare. Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997.

Laterals

- Laterals also have a side branch - the pocket behind the tongue tip is a side branch to the main tube(s) passing around the side(s) of the tongue.
- Laterals are thus also characterized by zeroes - the lowest appears between F2 and F3, often significantly reducing the amplitude of F2.
- The presence of zeroes and the coronal constriction reduce the intensity of laterals compared to most vowels.
- On spectrograms, laterals look similar to nasals, but differ in the location of formants and zeros, and in their effects on neighboring vowels.

Laterals

- Laterals also have a side branch - the pocket behind the tongue tip is a side branch to the main tube(s) passing around the side(s) of the tongue.

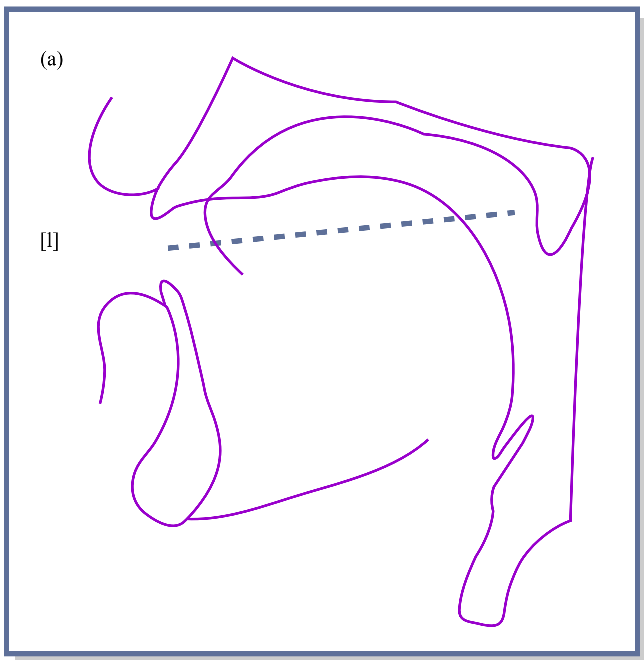


Image by MIT OpenCourseWare. Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997.

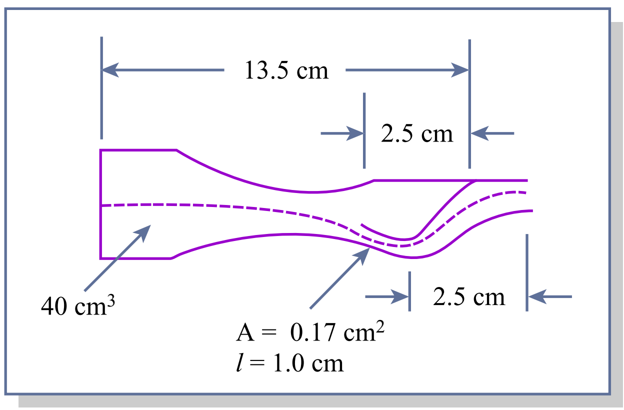


Image by MIT OpenCourseWare. Stevens, K. N. *Acoustic Phonetics*. Cambridge, MA: MIT Press, 1999.

Nasalized vowels

- Nasalized vowels involve a complex vocal tract configuration.

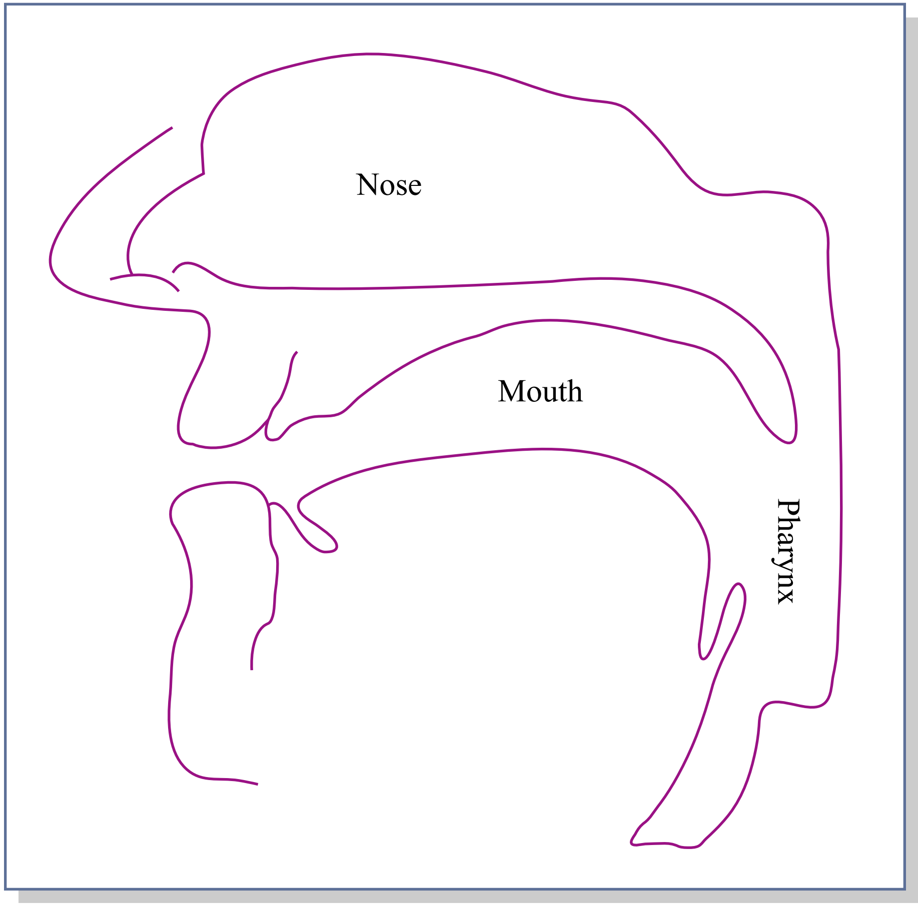


Image by MIT OpenCourseWare. Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997.

Nasalized vowels

- The coupled nasal cavity contributes both formants and anti-formants (poles and zeroes) which combine with the formants of the oral tract.
- The net acoustic effect of velum lowering depends on the locations of the formants in the corresponding oral vowel (Maeda 1993)

Nasalized vowels

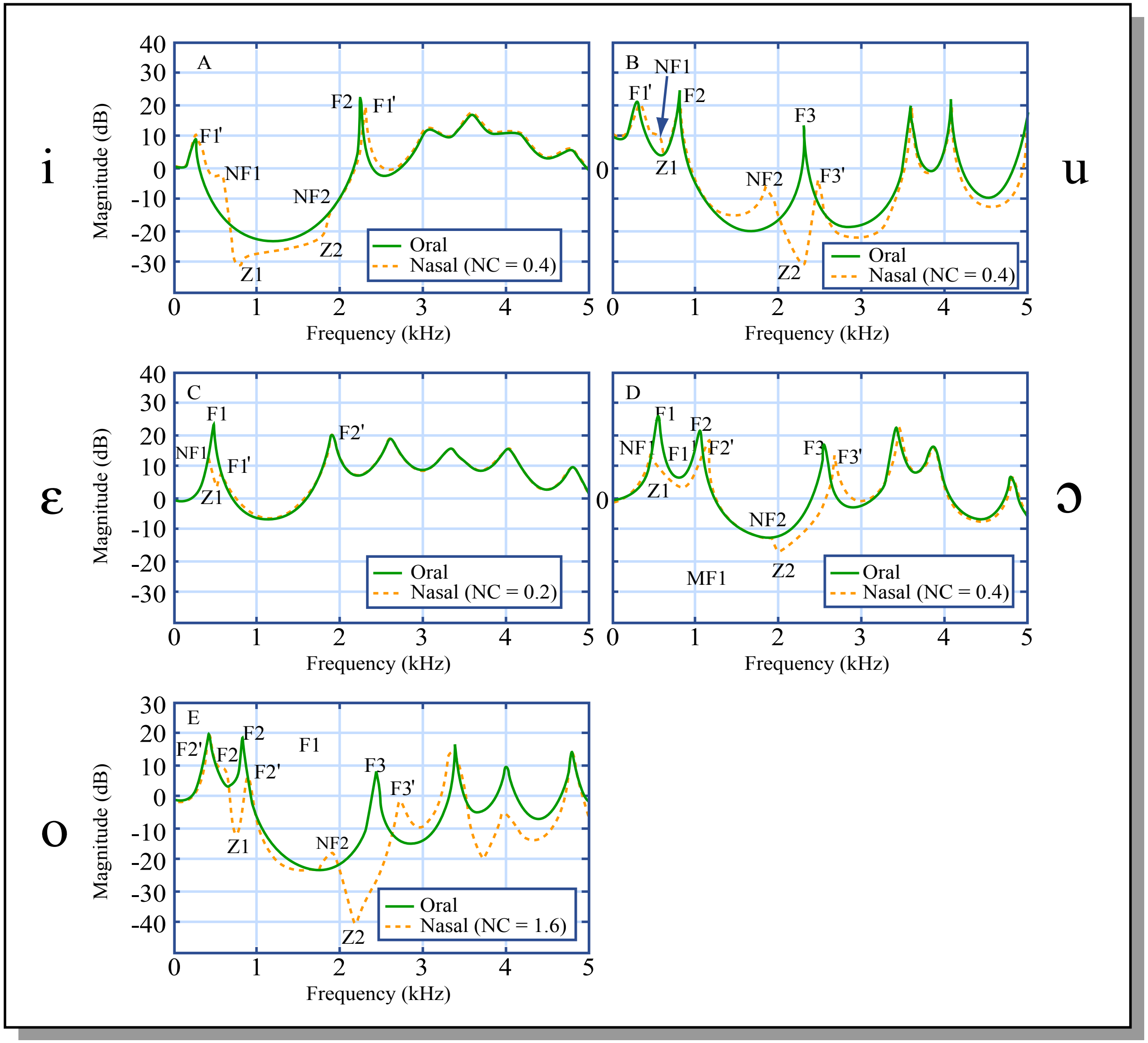


Image by MIT OpenCourseWare. Adapted from Maeda, Shinji. "Acoustics of Vowel Nasalization and Articulatory Shifts in French Nasal Vowels." In *Nasals, Nasalization and the Velum*. Edited by M. Huffman and R. Krakow. New York, NY: Academic Press, 2003, pp. 147-170.

Nasalized vowels

- As a result it is difficult to give a general acoustic characterization of nasalization on vowels.
- Maeda: nasalized vowels generally have a flat, low intensity spectrum in the low frequency region.
  - Addition of the lowest nasal formant in combination with broad bandwidth F1 can create a broad, low frequency prominence.

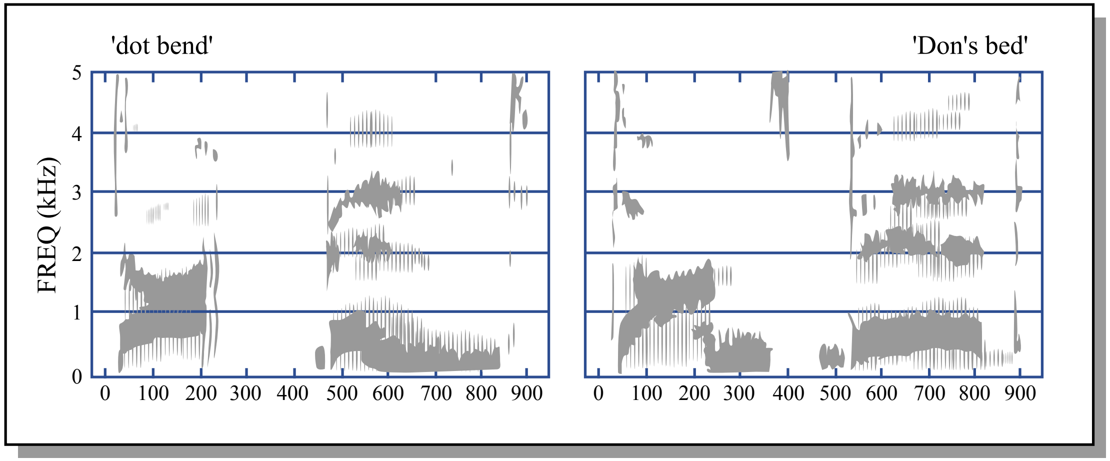


Image by MIT OpenCourseWare. Adapted from Stevens, K. N. *Acoustic Phonetics*. Cambridge, MA: MIT Press, 1999.

Nasalized vowels

- Chen (1997) proposes the measures A1-P0 and A1-P1 to quantify nasalization.
  – A1 is amplitude of F1
  – P0 is amplitude of the low frequency nasal peak.
  – P1 is amplitude of a higher nasal peak, between F1 and F2.

- A1 amplitude of F1
- P0 amplitude of lower frequency nasal peak - intensity of harmonic with greatest amplitude at low frequency.
- P1 amplitude of nasal peak - measured as amplitude of highest harmonic near to 950 Hz.
- These measures are sensitive to vowel quality - Chen proposes corrections to adjust for this.

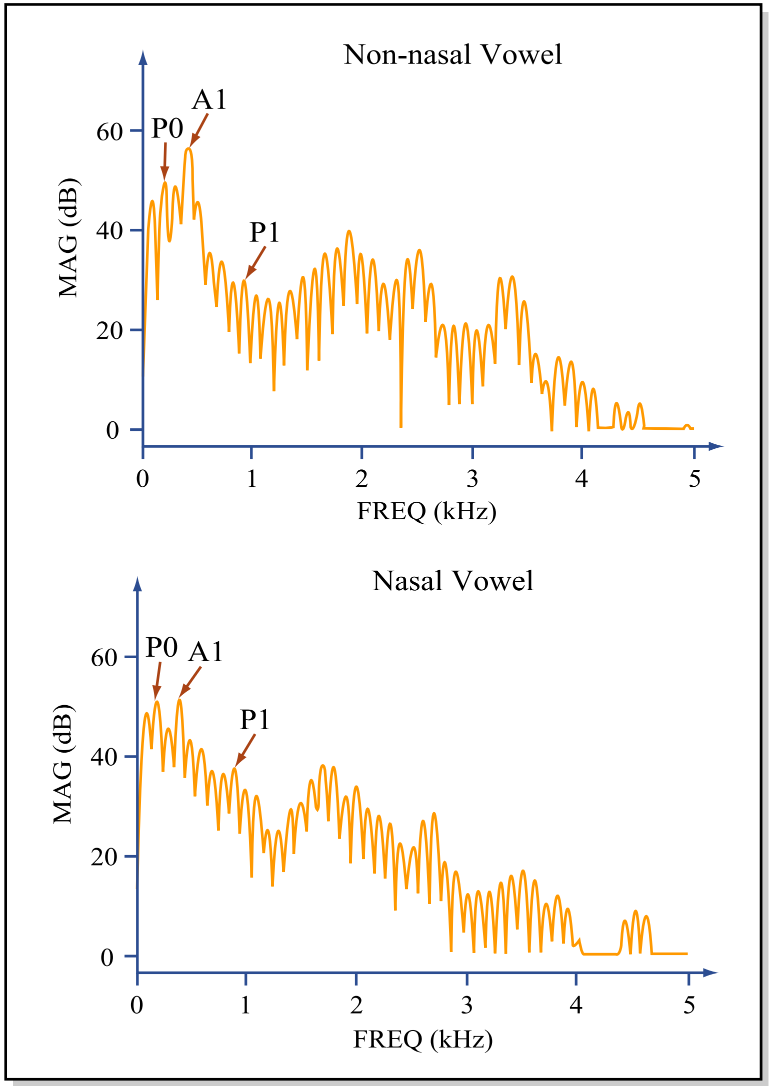
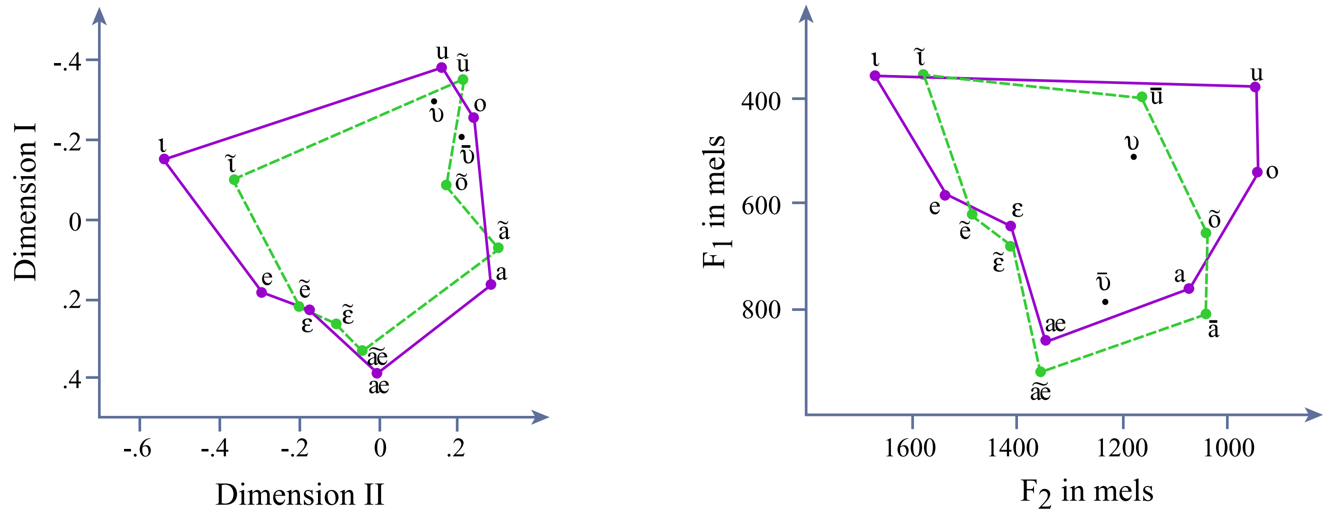


Image by MIT OpenCourseWare. Adapted from Chen, Marilyn Y. "Acoustic Correlates of English and French Nasalized Vowels." *The Journal of the Acoustical Society of America* 102 (1997): 2360.

Nasalized vowels

- Perceptually, nasal vowels are less distinct from each other than oral vowels, both in terms of confusability (Bond 1975) and similarity judgements (Wright 1986).



The first two dimensions in the perceptual space as determined by INDSCAL

F_1 versus F_2 plot of the mel values of the stimuli

Image by MIT OpenCourseWare. Adapted from ÊWright, James T. "The Behavior of Nasalized Vowels in the Perceptual Vowel Space." In *Experimental Phonology*. Edited by J. J. Ohala and J. J. Jaeger. Orlando, FL: Academic Press, 1986, pp. 45-67.

Nasalized vowels

- This greater confusability is reflected in the fact that, in languages with contrastive vowel nasalization, the nasal vowel inventory is always the same as or smaller than the oral vowel inventory, never larger (Ferguson 963, Ruhlen 1973).

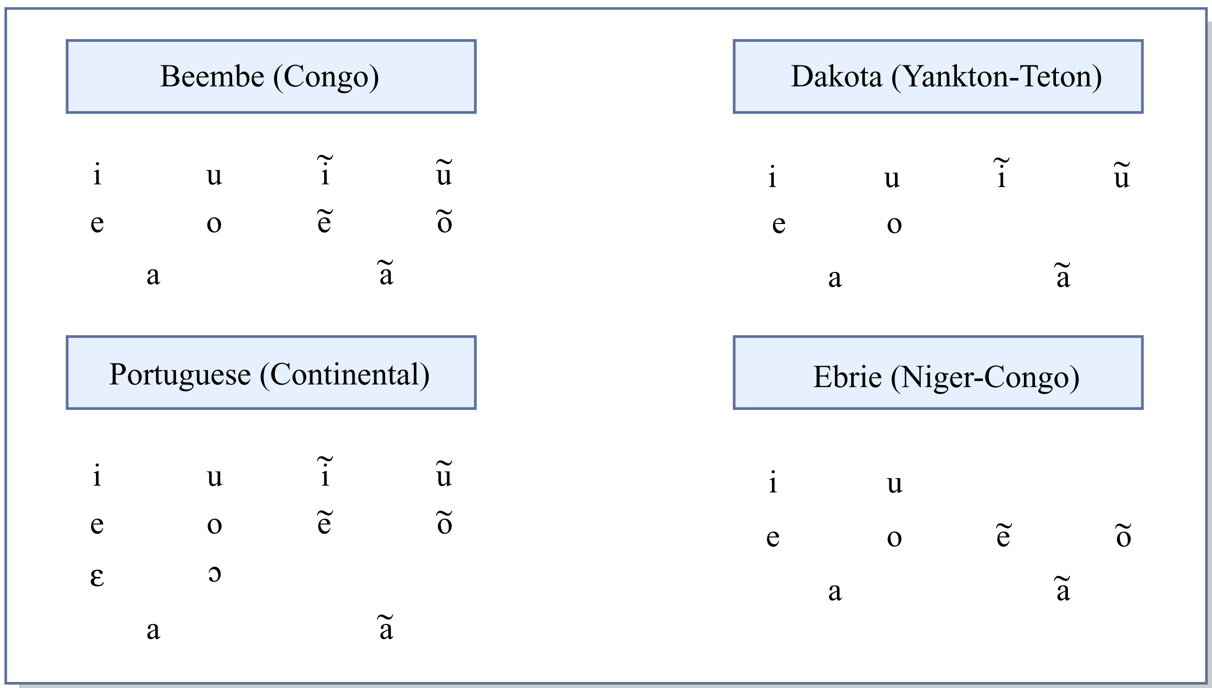


Image by MIT OpenCourseWare.

Nasalized vowels

- More specifically, high nasalized vowels are perceived as lower than their oral counterparts and low nasalized vowels are perceived as higher than their oral counterparts.
- Wright speculates that this may be related to the fact that FN1 is below F1 in low vowels and above it in high and mid vowels in his stimuli.

Nasalized vowels

- Nasalization alternations are sometimes accompanied by height alternations, e.g. French.

i	y	u				iN ⟶ æ̃
e	ø	o			õ	e.g. finiʁ 'finir' (to finish)
ɛ	œ	ɔ		œ̃		fæ̃ 'fin' (end)
	a		æ̃	ɑ̃		

Image by MIT OpenCourseWare.

- Beddor (1982) surveys 75 languages and concludes that, with a few exceptions, nasalized high vowels tend to lower, low vowels tend to raise, and mid vowels can raise or lower depending on context. However, interpretation of the data is often problematic.
- E.g. English dialects ɛn → in
  - unclear whether neutralization is due to raising or vice versa.