Akustische Analyse der Stimmqualität
Ergänzende Information

Ein technisches Problem bei Maßen für ***spectral tilt*** besteht darin, dass sie von der Lage der Formanten beeinflusst werden. Deswegen ist es nicht ohne weiteres möglich, die Stimmqualität verschiedener Vokale (z.B. [a] und [i]) miteinander zu vergleichen. Es sind in der Literatur Korrekturverfahren vorgeschlagen worden, um dieses Problem zu umgehen (s. z.B Garellek & Keating). Die entsprechenden Maße werden oft mit '*' gekennzeichnet (z.B. H1*-H2*, statt H1-H2). Diese Möglichkeit hat das Programm VoiceSauce von Keating et al., Praat aber zur Zeit nicht. Deswegen ist es bei Praat im Moment besser, die Untersuchung auf eine Vokalkategorie (am besten [a]) zu beschränken[1].

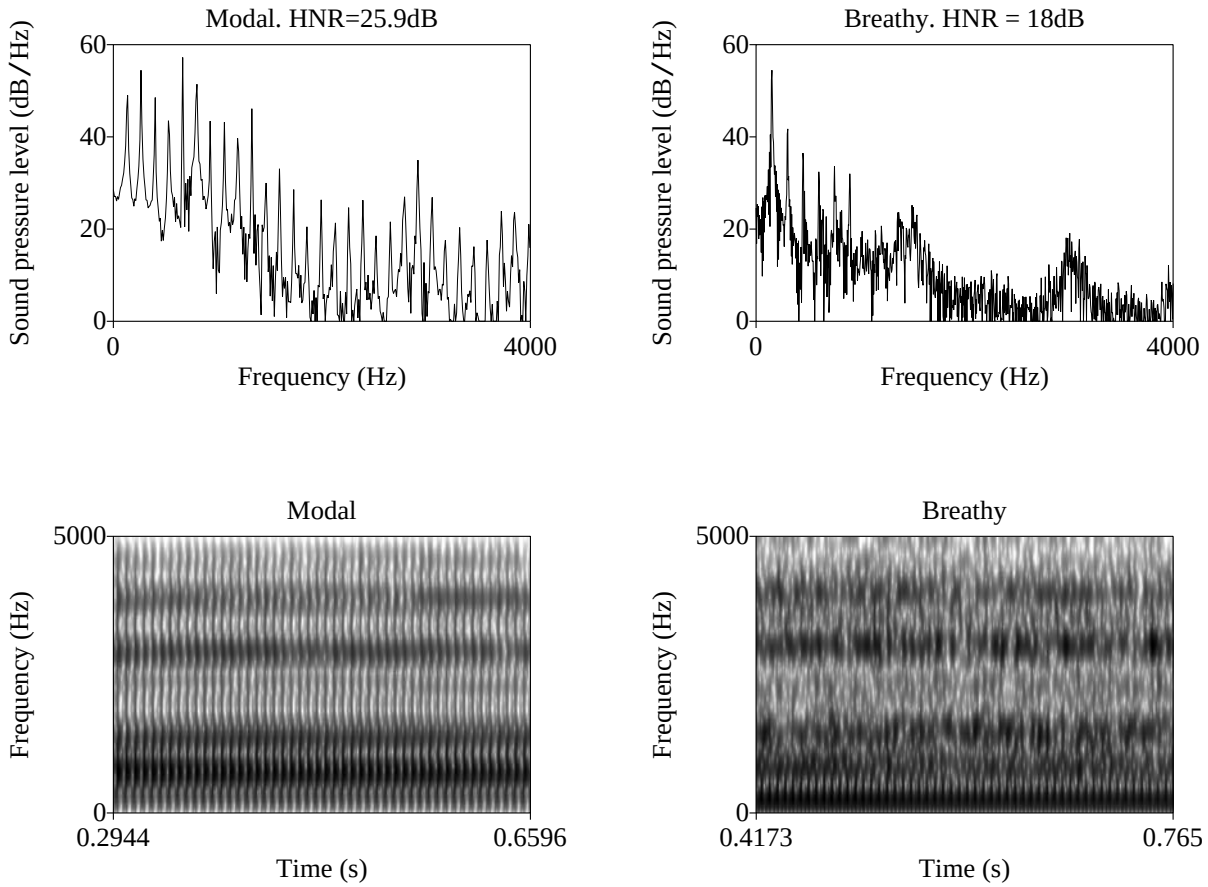
Weiterer Gesichtspunkt bei *spectral tilt*:
Man erwartet bei *breathy* (wegen dem größeren Öffungsquotient, und wegen dem weniger abrupten Schließen der Stimmbänder), dass die Amplitude der höheren Harmonischen schneller abfällt als bei Modal, also größerere Werte für H1-H2 oder H1-A2. Dieser Effekt wird aber u.U teilweise neutralisiert, wenn *breathy* ausgeprägte geräuschhafte Energie aufweist (Aspirationsgeräusch aufgrund des starken Luftstroms durch die offenere Glottis).
Deswegen ist es sinnvoll, die Maße für *spectral tilt*, durch Maße zu ergänzen, die den relativen Anteil von harmonischer und geräuschhafter Energie im Signal bestimmen.
Hier findet man Begriffe wie "***Harmonics to Noise ratio***" oder (in Praat) "*Harmonicity*".
Diese Grundidee wird von Gordon/Ladefoged unter '5.1 Periodicity' erwähnt. Explizit erwähnen Sie lediglich das Maß "Jitter". "Jitter" bezieht sich aber auf die zeitliche Stabilität der Stimmbandschwingungen, erfaßt daher nicht direkt die Stärke der geräuschhaften Energie im Spektrum.
Hier ein Auszug aus der Praat-Hilfe zum Befehl 'Harmonicity':

> *"A Harmonicity object represents the degree of acoustic periodicity, also called Harmonics-to-Noise Ratio (HNR). Harmonicity is expressed in dB: if 99% of the energy of the signal is in the periodic part, and 1% is noise, the HNR is 10*log10(99/1) = 20 dB (approx.). A HNR of 0 dB means that there is equal energy in the harmonics and in the noise.*
> *..... a healthy speaker can produce a sustained [a] or [i] with a harmonicity of around 20 dB ... Hoarse speakers will have an [a] with a harmonicity much lower than 20dB."*

Die folgenden Abbildungen zeigen Spektrum und Sonagramm für den Vokal /a/ jeweils modal und behaucht von der gleichen Sprecherin.
Im Spektrum für Modal, sind die einzelnen Harmonischen über den ganzen Frequenzbereich bis 4kHz leicht zu erkennen. Bei Breathy dagegen ist dies aber nur bis etwa 1kHz möglich. Dementsprechend sind die Harmonicity-Werte für Breathy deutlich niedriger als für Modal (18dB vs. knapp 26dB). Der gleiche Sachverhalt ist in den Sonagrammen daran zu erkennen, dass für Modal die senkrechte Strukturierung durch die Glottisimpulse sich bis zu den höchsten Frequenzen erstreckt, während dies sich bei Breathy auf die niedrigeren Frequenzen beschränkt.

---

[1] In diesem Handout verwende ich die Schreibweise von Garallek/Keating. Für H1-H2 findet man bei Gordon/Ladefoged die Notation h2-f0, also mit umgehrtem Vorzeichen! (und F1-f0 statt H1-A1).



In neueren Arbeiten der UCLA-Gruppe (z.B. Blankenship, 2002; Garallek & Keating, 2011) wird ein weiteres Maß verwendet, um den Grad der Periodizität des Signals zu bestimmen: ***Cepstral Peak Prominence***. Weil dieses Maß zur Zeit in Praat nicht verfügbar ist, gehen wir auf die dem Maß zugrundeliegende Signalverarbeitung hier nicht näher ein (weitere Erläuterungen hierzu in Hillenbrand & Houde, 1996, v.a. Fig. 2)

Literatur

Blankenship, B. (2002). The timing of nonmodal phonation in vowels. J. Phonetics 30, 163-191.

Garallek, M. & Keating, P. (2011). The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. J. Int. Phon. Assoc. 41(2), 185-205

Gordon, M. & Ladefoged (2001). Phonation types: a cross-linguistic overview. J. Phonetics 29, 383-406

Hillenbrand, J. & Houde, R. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. J. Speech Hearing Res. 39, 311-321.

Klatt, D. & Klatt, L. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. J. Acoust. Soc. America 87(2), 820-857.