**Multimedia Processing – Übung 5**

**LPC-Vocoder/Sprachkompression**

**Melina Hortz**

**5.2. Vorbereitung**

Als erstes lesen wir mit [Read] in PRAAT unsere Sprachdatei ein. Dann wählen wir "Formants&LPC->To LPC (autocorrelation)". Die anderen Verfahren, die PRAAT noch anbietet, unterscheiden sich nur geringfügig in der Methode der Berechnung der Filter-Koeffizienten. Im Einstelldialog sind besonders die Punkte "Prediction Order" (Ordnung des Filters) und "Analysis width" (Framebreite bei der Analyse) von Interesse. Diese wollen wir in dieser Aufgabe variieren. Zusätzlich zu den von PRAAT angebotenen Standardwerten von Order=16/Analysis Width=25ms (kurz: 16/25) werden wir Analysen bei Ordnungen von 10, 6 und 1 durchführen, sowie Framelängen von 50 und 100 ms untersuchen. "Time step" (die "Vorschubbreite" der Analyse) setzen wir auf denselben Wert wie "Analysis width", d.h. wir erlauben kein Überlappen von benachbarten Frames.

Um die verschiedenen Varianten nicht zu verwechseln, geben wir den einzelnen LPC-Objekten mit 'Rename' entsprechende Namen.

Nachdem die verschiedenen LPC-Objekte kreiert wurden, führen wir zunächst eine inverse Filterung des Sprachsignals durch, mit der wir theoretisch die "Quelle" zurückgewinnen (Sound und LPC (16/25) gemeinsam anwählen, "Filter (inverse)"). Das entstehende Signal, das auch als "Residual"-Signal bezeichnet wird, betrachten wir im Edit-Fenster. Wie unterscheiden sich die Spektren vom Original, wie ist die Restverständlichkeit der ursprünglichen Sprache? (Spektrogramme von Original und Residual-Signal ins Protokoll) Nun gewinnen wir das ursprüngliche Signal wieder zurück (Rekonstruktion), indem wir das Residuum mit dem dazugehörigen LPC-Objekt (16/25) filtern (Residuum und LPC gemeinsam anwählen, "Filter", "Use LPC Gain" nicht ankreuzen!).

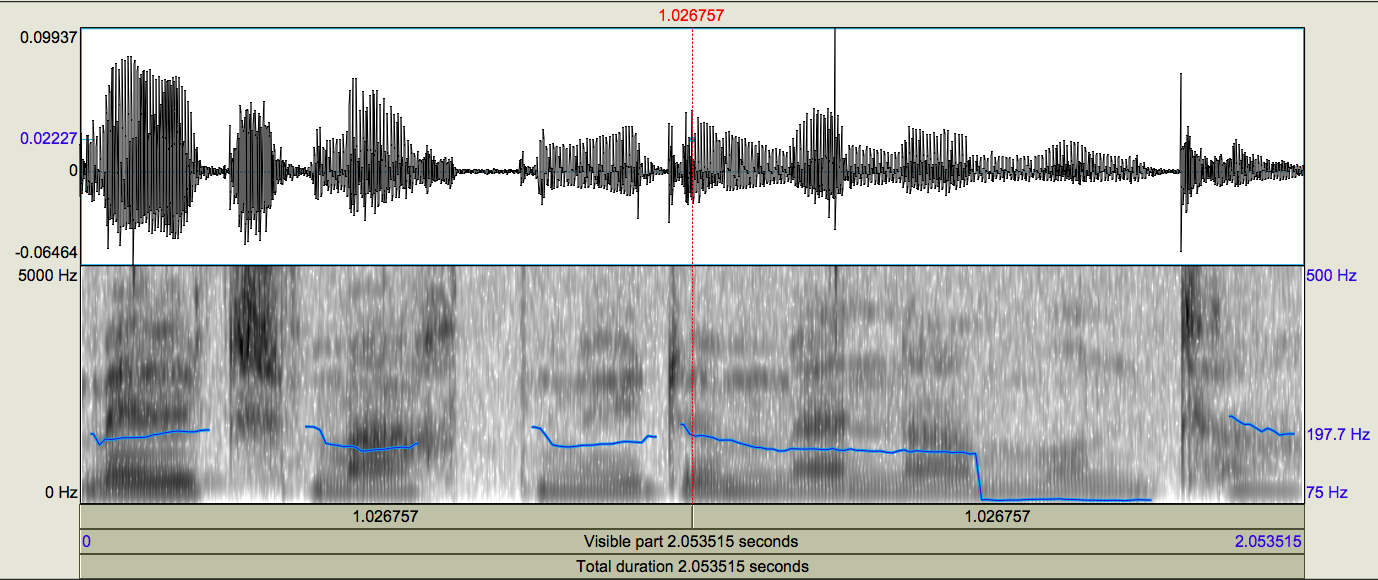


Abb. Originalaufnahme

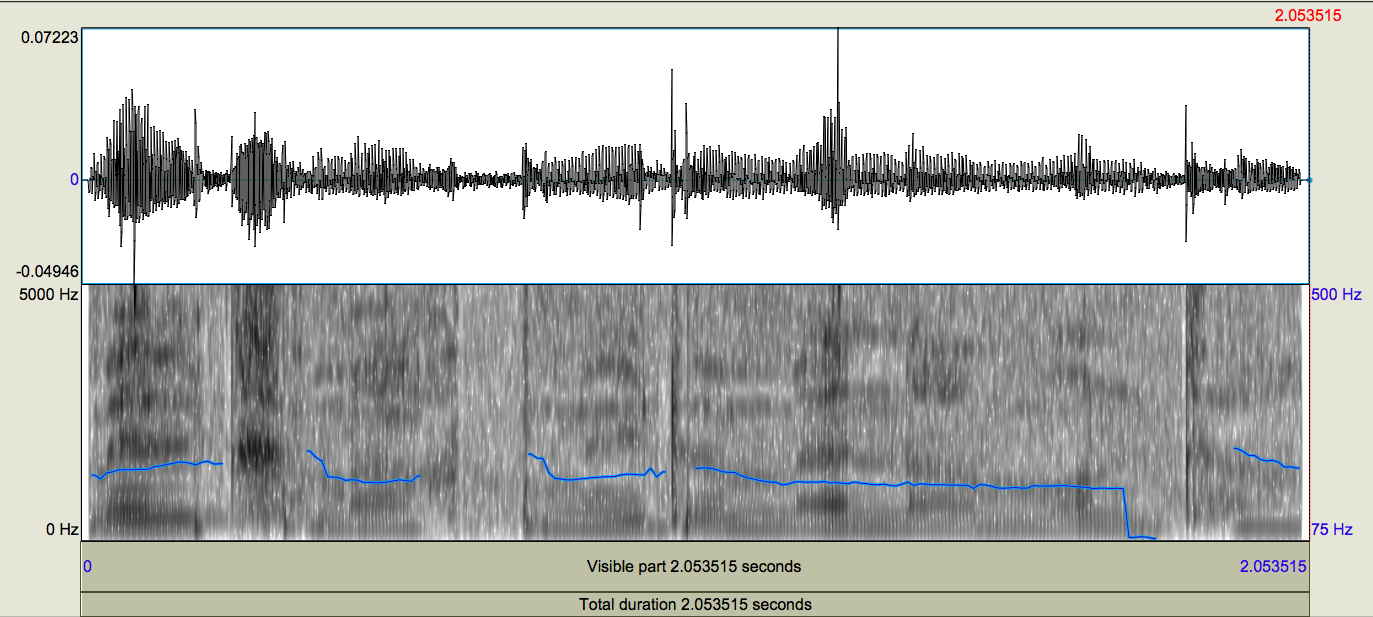


Abb. Residuum

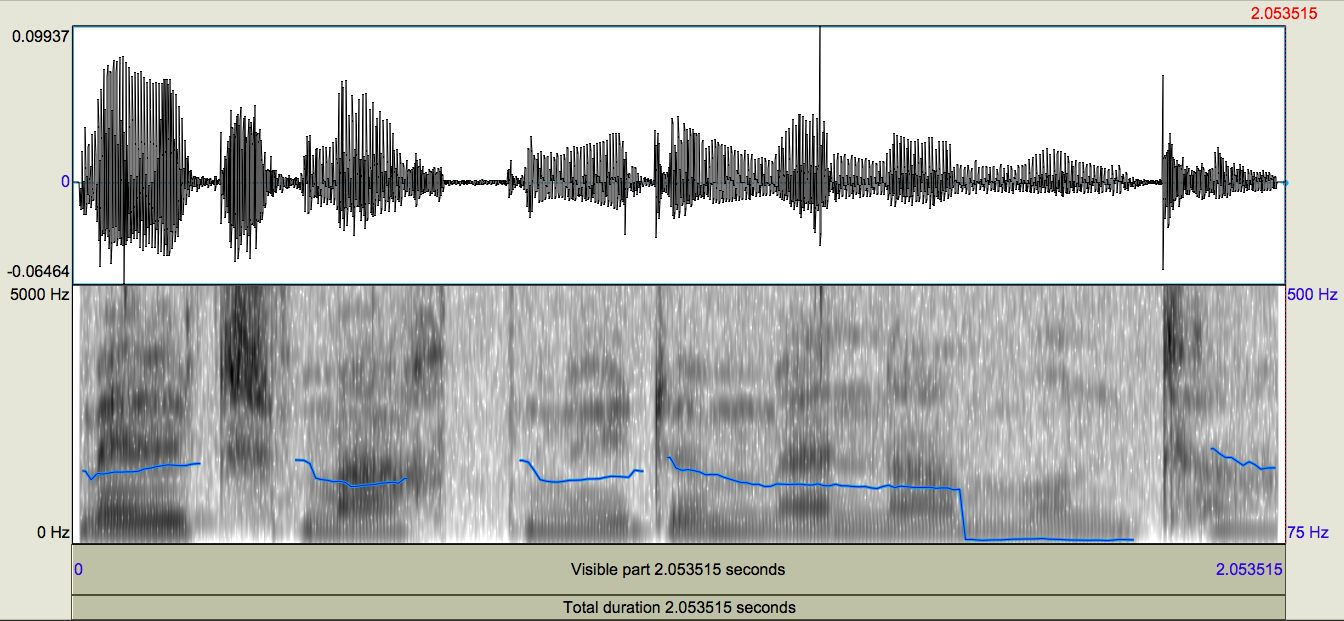


Abb. Rekonstruktion

Originalaufnahme Residuum Rekonstruktion

Das Spektrum des Residual-Signals zeigt insgesamt schwächere Schwingungen als das des Originals. Die Hintergrundgeräusche/Störgeräusche treten bei dem Residual-Signal mehr hervor (sind lauter zu hören) und zeigen daher auch mehr Schwingungen im Spektrum.

Insgesamt ist hört sich das Residual-Signal leiser und nuschelnder/rauschend an und die Hintergrundgeräusche sind lauter.

**5.3 Variation des Filters**

Das Residual-Signal dient uns jetzt zum Testen der anderen LPC-Objekte mit den verschlechterten Parametern. Auch hier wählen wir das Residuum und das gewünschte LPC-Objekt an und filtern wieder. Vergleiche die Qualität der einzelnen Varianten mit dem Original! Wie unterscheiden sich die Beispiele, bei denen die Filter-Ordnung variiert wurde, wie solche, bei denen die Framebreite variiert wurde? Bei welcher Kombination ist die Sprachverständlichkeit nicht mehr akzeptabel?

**Ordnung 1:**

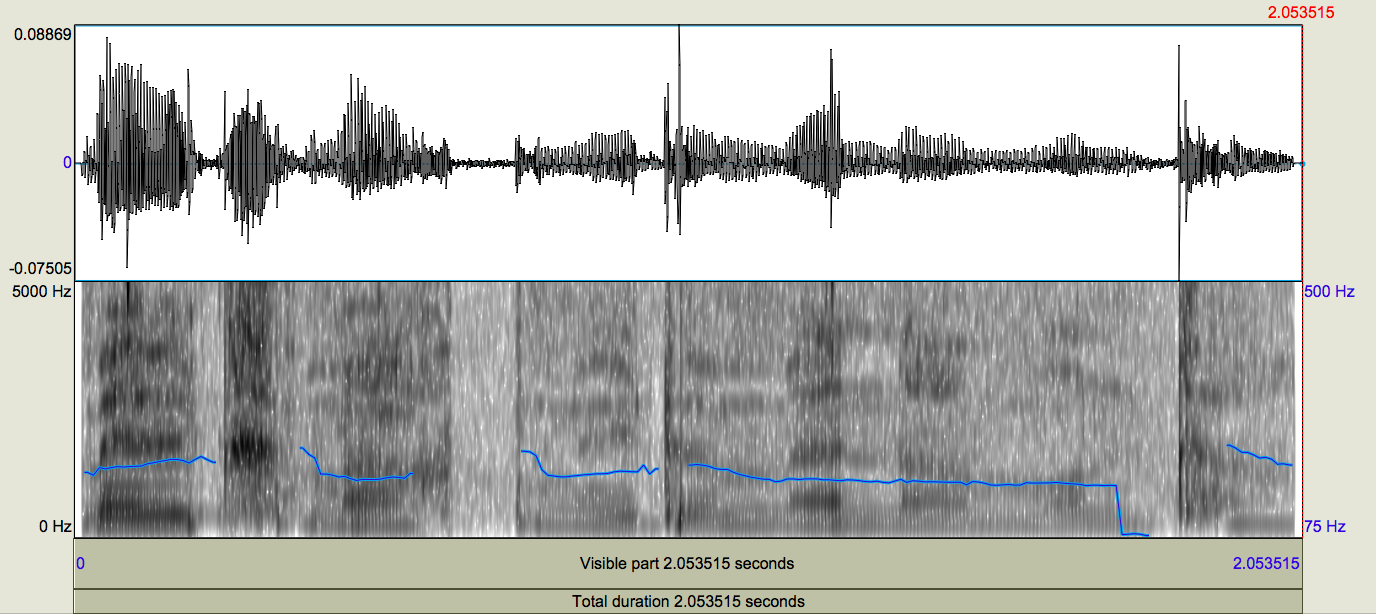
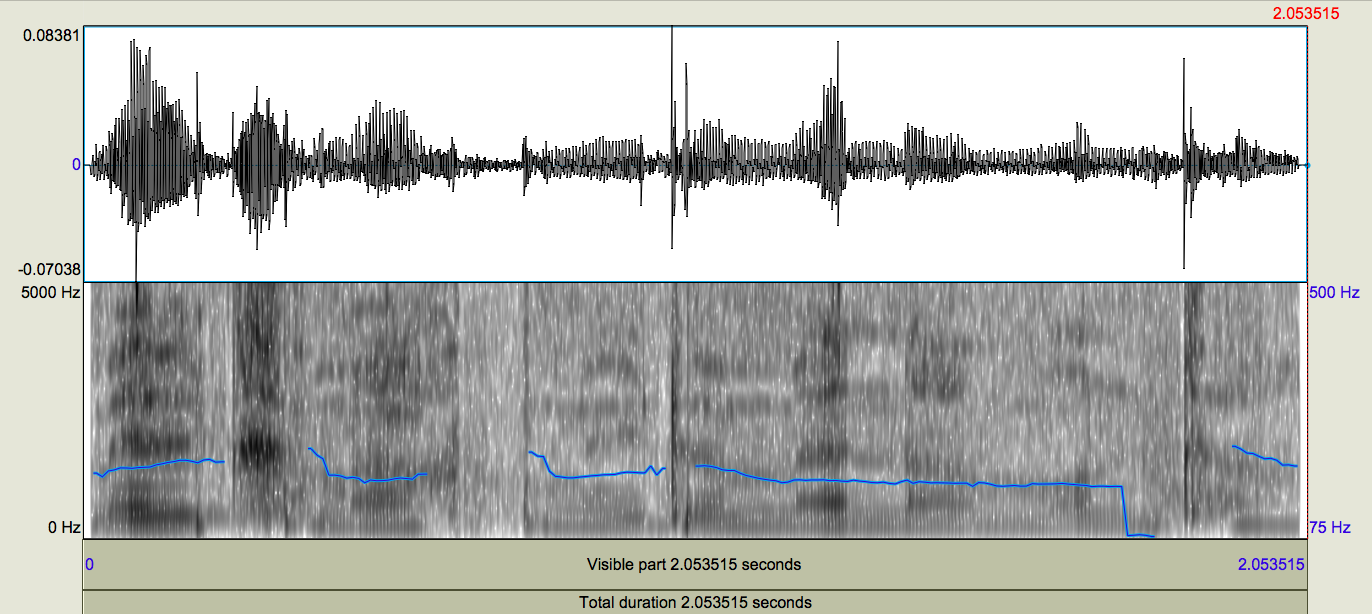


Abb. 1/25

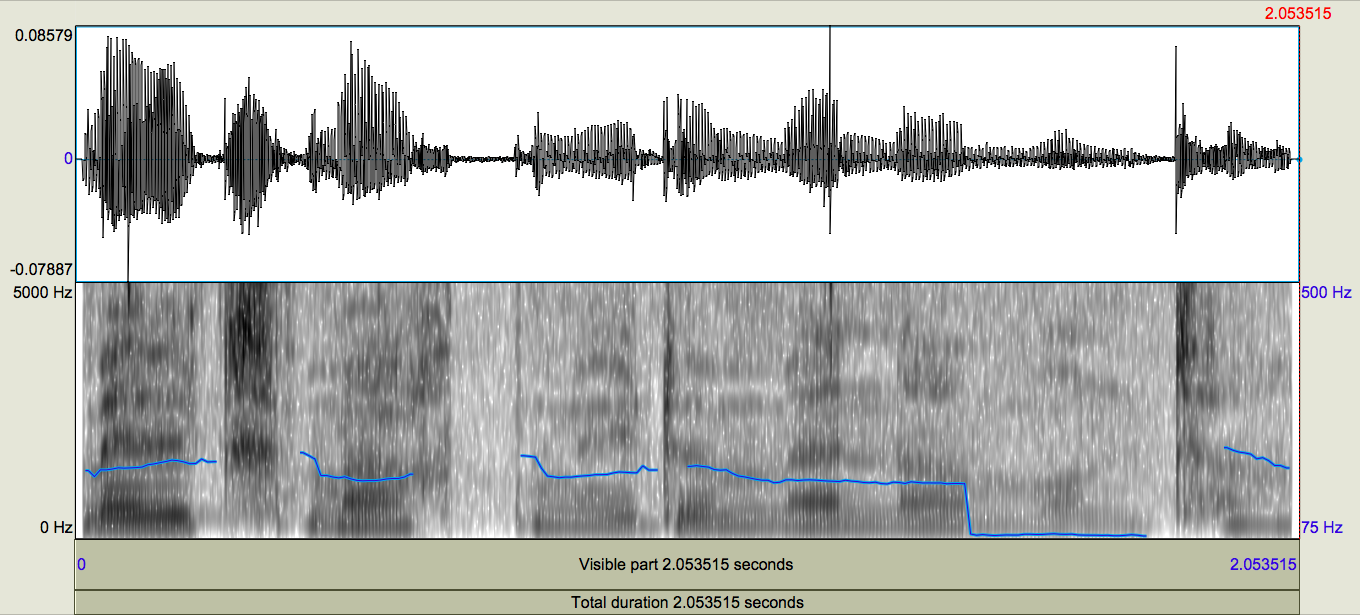
Abb. 1/50

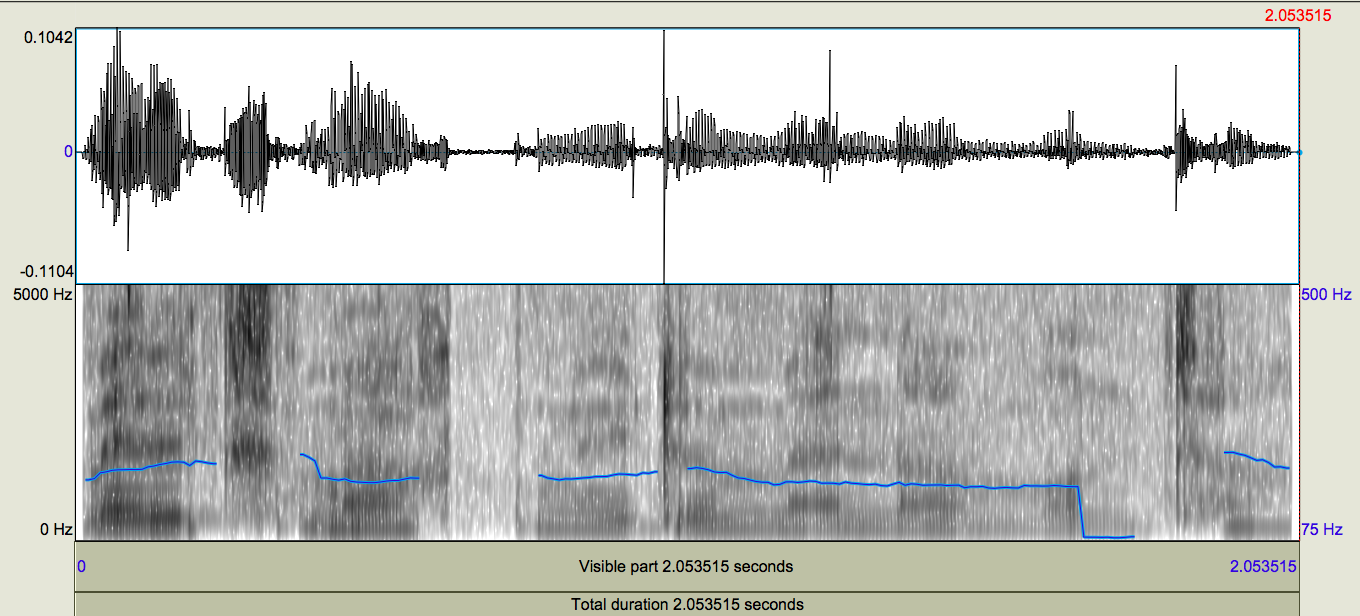
Abb. 1/100

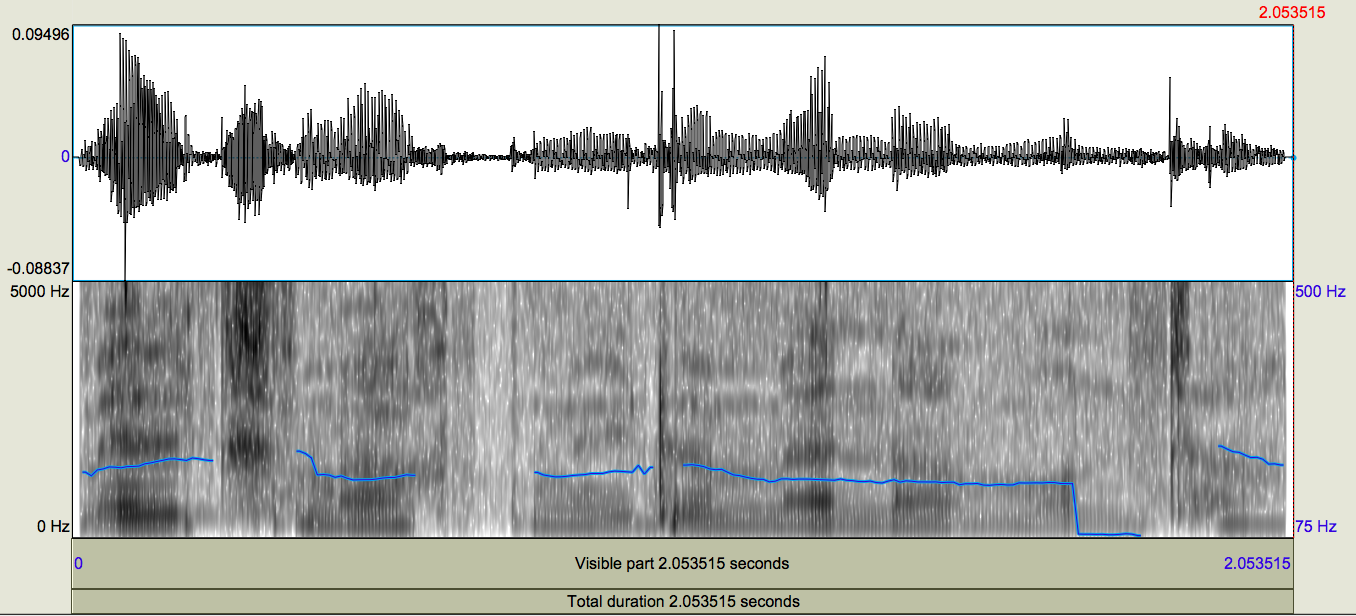
1/25 1/50 1/100

**Ordnung 6:**

Abb. 6/25

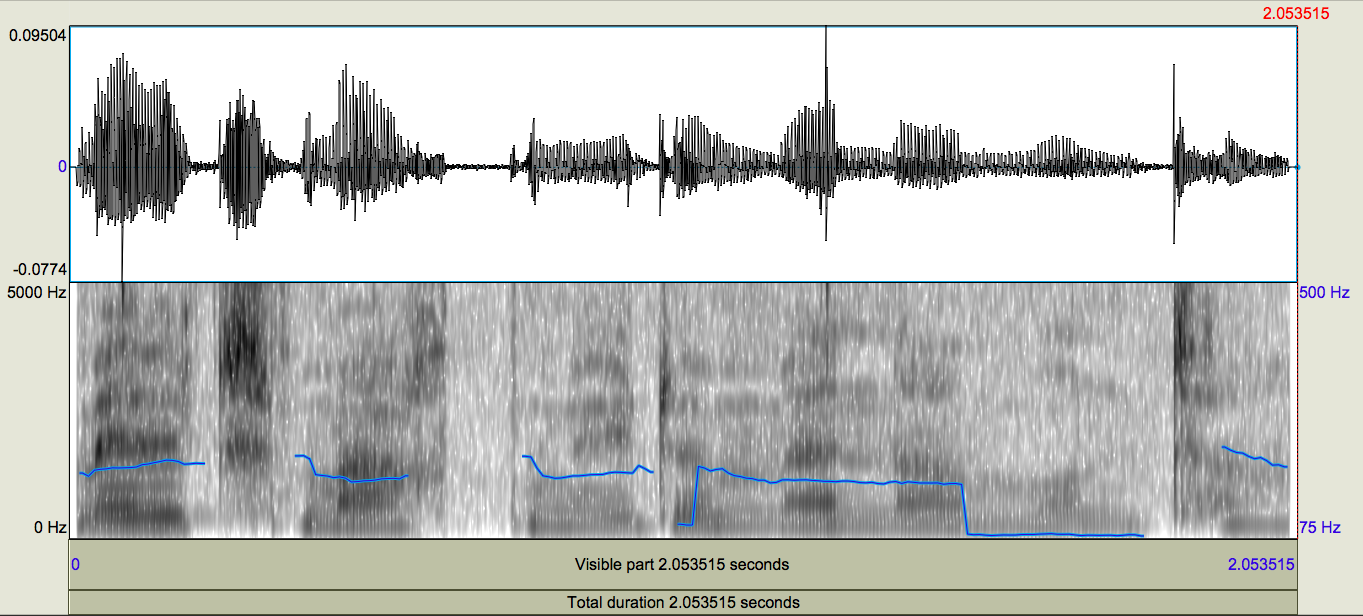
Abb. 6/50

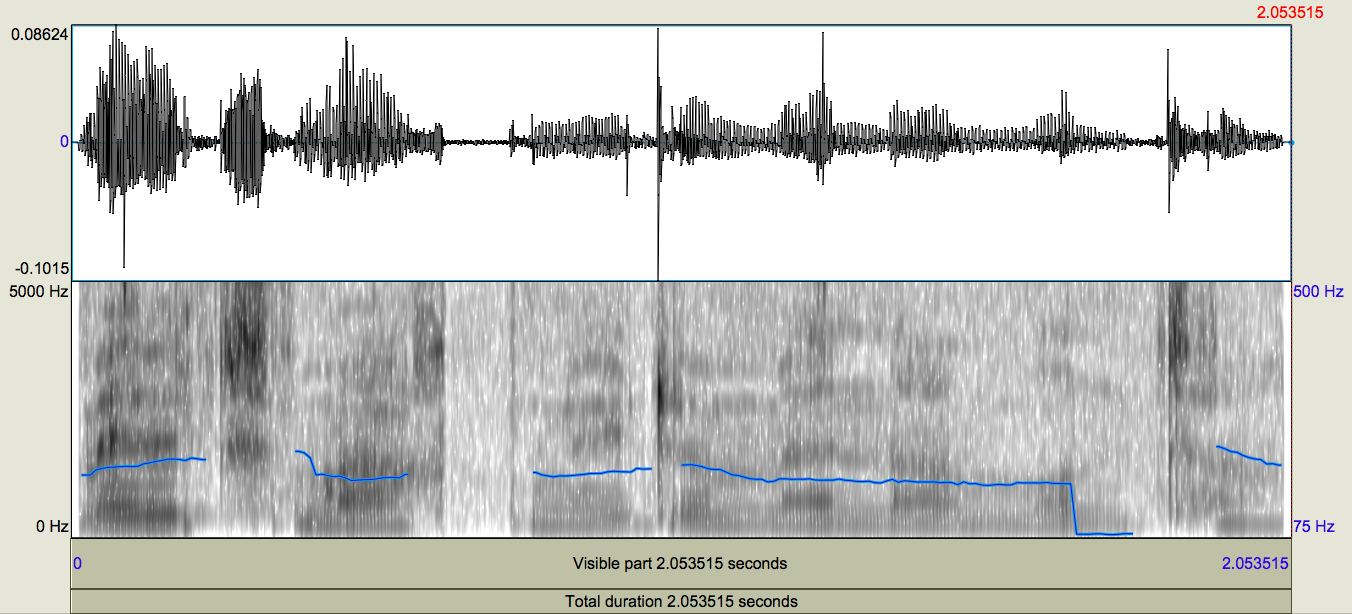
Abb. 6/100

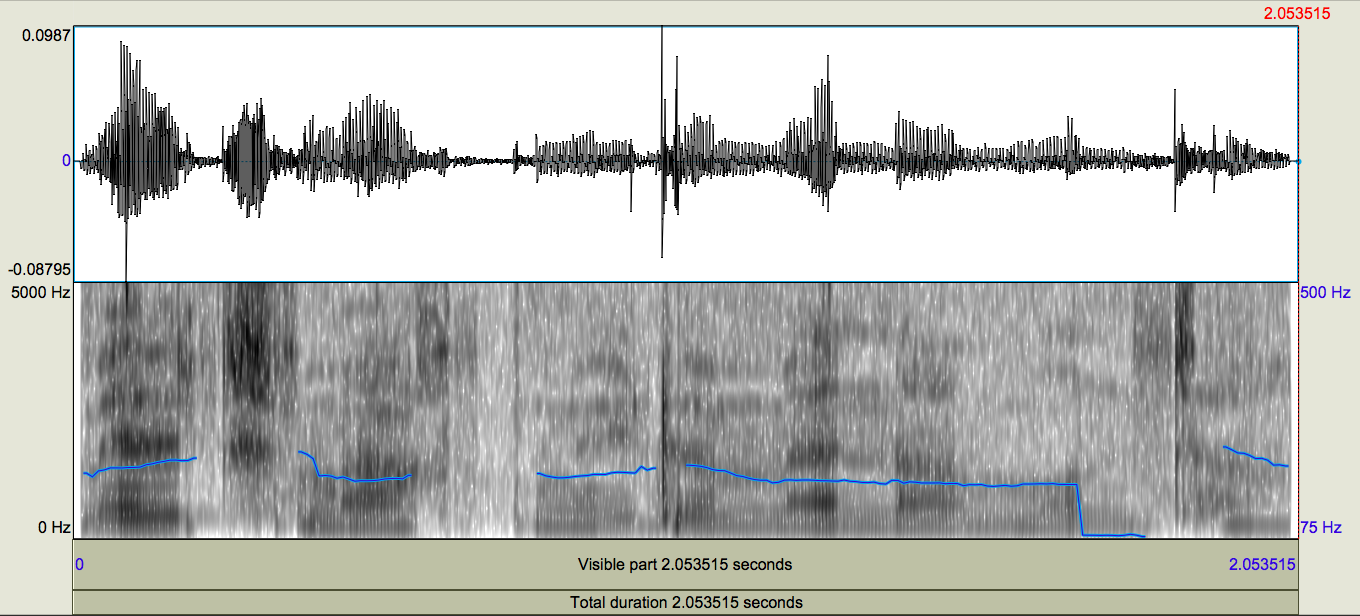
6/25 6/50 6/100

**Ordnung 10:**

Abb. 10/25

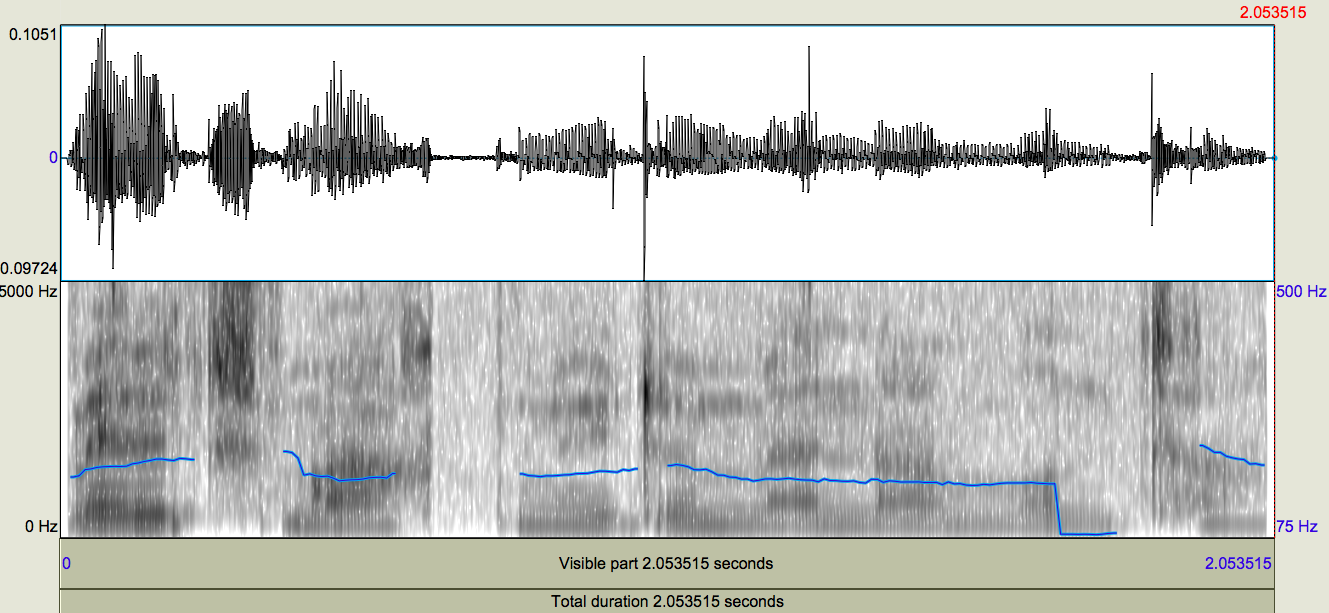
Abb. 10/50

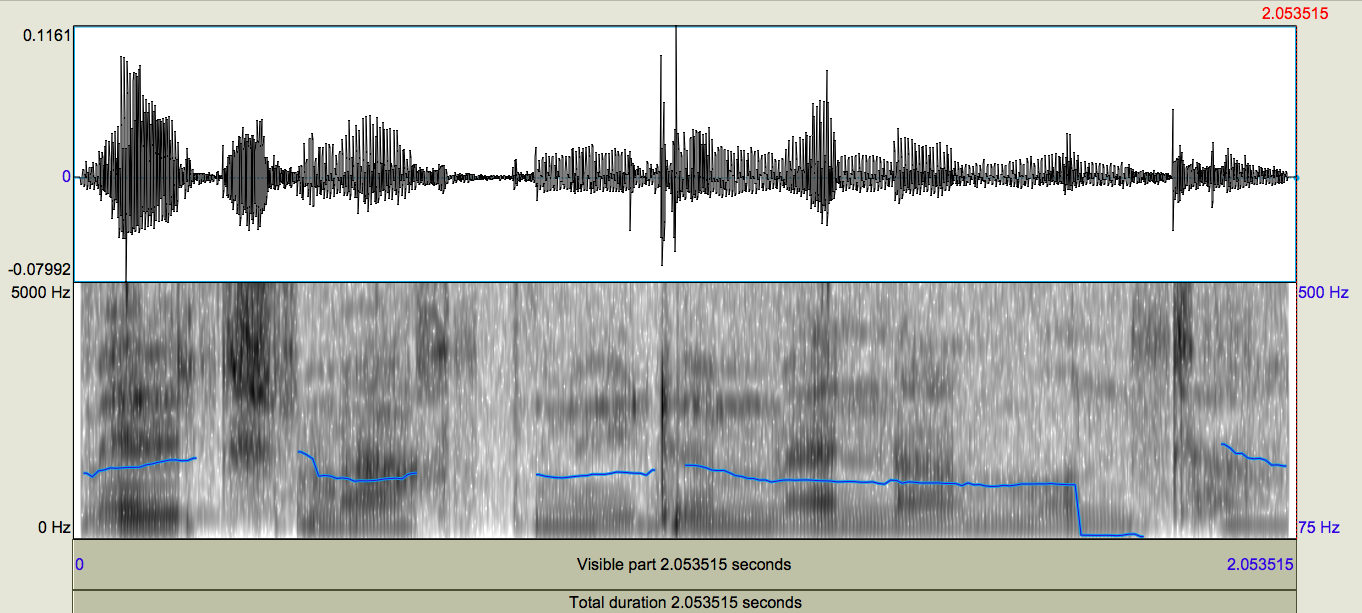
Abb. 10/100

10/25 10/50 10/100

**Ordnung 16:**

Abb. 16/50

Abb. 16/100

16/50 16/100

Das Beispiel mit der 6/25 klingt dem Original am ähnlichsten. Das Beispiel 16/25 auch, jedoch lässt sich hier knacken hören.

Die Verständlichkeit verschlechtert sich von der Variation der Framebreite von 25 zu 100. Es lässt sich mehr rauschen hören und der gesprochenen Satz hört sich nuschelnd an. Des weiteren kommt knacken hinzu, je länger die Framebreite.

Die Beispiele mit der Filter-Ordnung 1 sind allgemein rauschend und nuschelnd.

**5.4 Variation der Quelle**

Nun wollen wir die Anregung (Quelle) verändern. Dafür habe ich zwei Signale bereitgestellt, weißes Rauschen und einen 100Hz-Pulse Train, ein obertonreiches Signal mit Grundfrequenz 100Hz

Weißes Rauschen

Pulsetrain

Wenn wir das Rauschen mit dem LPC-Objekt filtern (jetzt "Use LPC Gain" anwählen), bekommen wir so etwas wie Flüstersprache, beim Pulse Train eine Roboterstimme. Abschließend werden wir auch noch eine Musikdatei als Anregung benutzen, was einen Vocodereffekt bewirken sollte. Am besten eignen sich hier flächige, eher statische Sounds wie Synthesizerakkorde, da der Rhythmus vom Sprachsignal kommen soll. Für die 12 Qualitäten (also Filter-Ordnungen 16, 10, 6, 1 kombiniert mit Framebreiten 25, 50 und 100ms) berechnen wir die Datenraten. Dazu machen wir uns Folgendes klar: Übertragen werden müssen mindestens je Analyseblock (also z.B. je 25ms) die Koeffizienten (Zahl entspricht der Filterordnung) und die Grundfrequenz f0. Setze hier jeweils pro Zahlenwert 16bit an. Vergleiche diese Datenrate mit der ursprünglichen PCM-Datenrate der Sprachdatei (wie in Übung 1 berechnet).

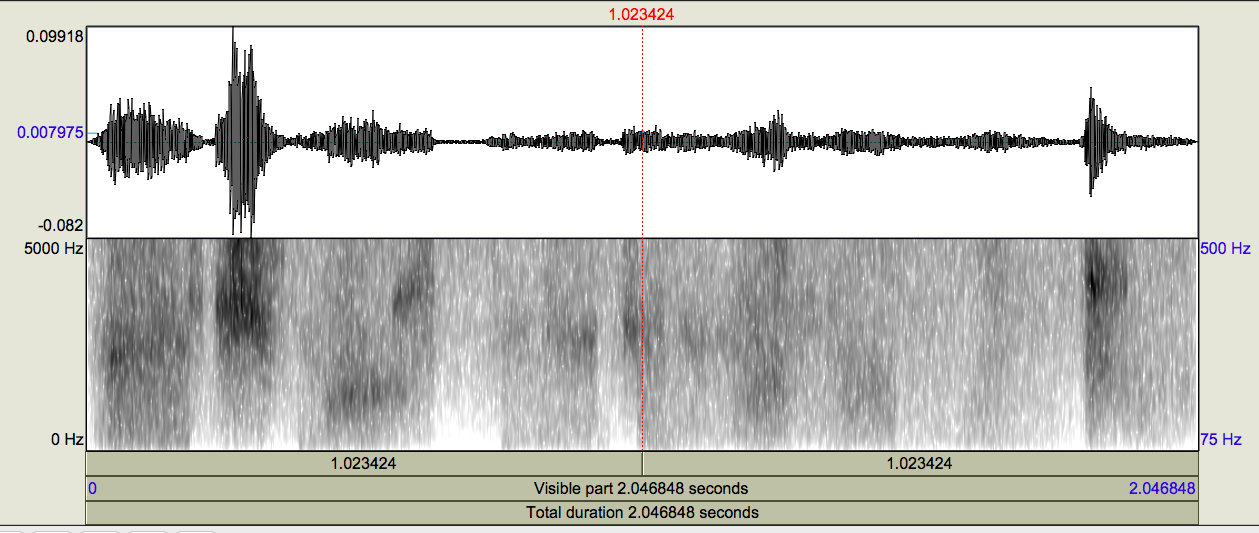


Abb. Weißes Rauschen und LPC-Objekt 16/25

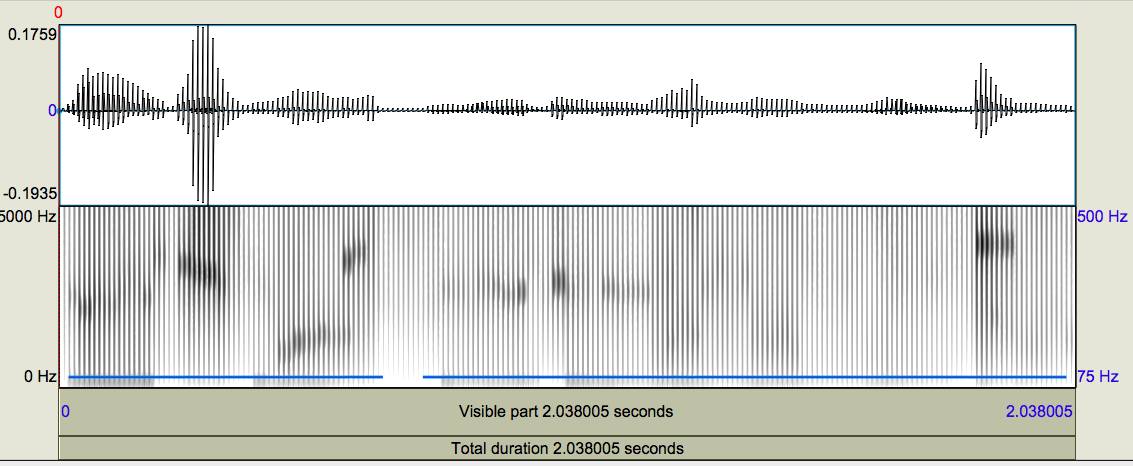


Abb. Pulsetrain und LPC-Objekt 16/25

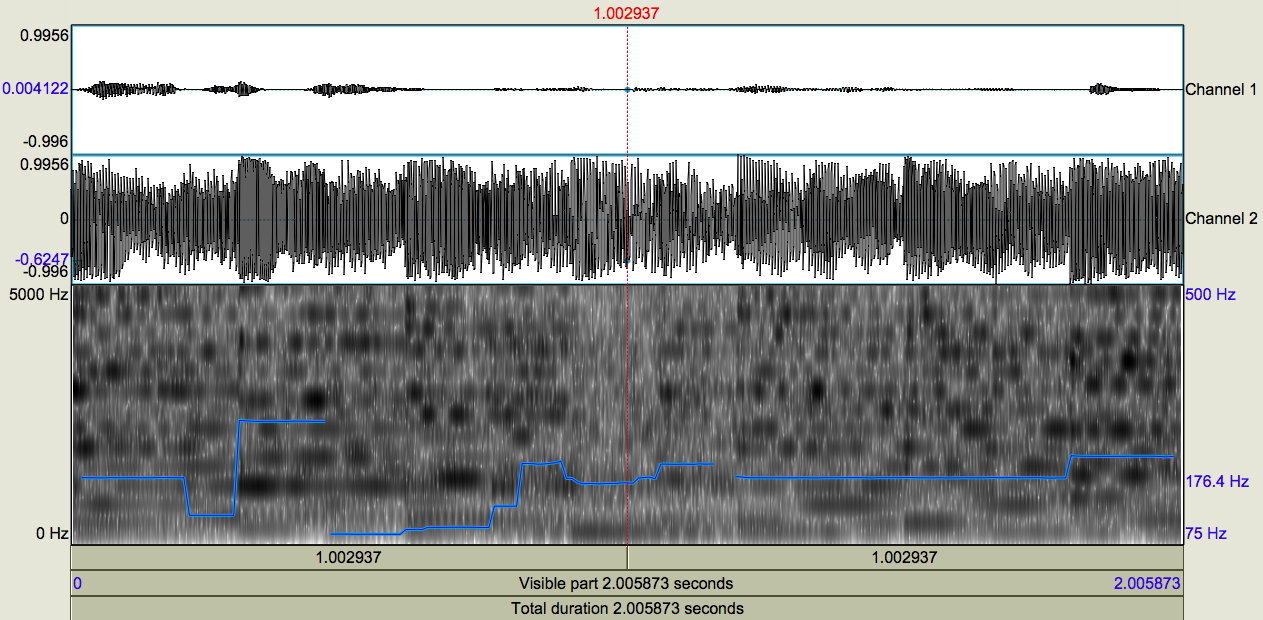


Abb. Vocodereffekt

Whitenoise LPC 16/25 Pulsetrain LPC 16/25 Synthesizer LPC 16/25