

Franz Födermayr / Werner A. Deutsch (Wien)

Zum stimmlichen Ausdrucksrepertoire von Hank Williams

Wie andernorts (Födermayr 1985) bereits ausgeführt, dürften die Gründe für den Erfolg des Country-Sängers Hank Williams (1923-53) über die Gestaltung eines Gesanges hinsichtlich Melodieverlauf, Rhythmus, Form und Text hinaus wesentlich in seiner Präsentation liegen, also in der Vortragskunst des Sängers. Findet man Wege, diese analytisch zu erfassen, könnte das Resultat der Untersuchung so etwas wie das stimmliche Ausdrucksrepertoire des Künstlers sein als wesentliche Komponente seines musikalischen Ausdrucksrepertoires - den Großteil der von Hank Williams vorgetragenen Gesänge bilden ja sowohl textlich wie musikalisch eigene Kompositionen (s. die Diskographie von B. Pinson, in: Williams 1981: 263-303, und G.W. Koon 1983: 131-170).

In der eingangs erwähnten Studie wurde die sonographische Technik als möglicher Weg getestet, um spezielle Vortragsmanieren zu untersuchen. Inzwischen wurde an der Kommission für Schallforschung der Österreichischen Akademie der Wissenschaften eine digitale Arbeitsstation entwickelt (s. Deutsch et al. 1988), die - sehr wesentlich durch die mitlaufende auditive Kontrolle - nicht nur die Deskription verlässlicher macht, sondern auch Möglichkeiten der Resynthese erlaubt. Mit dieser Technik wird im folgenden ein Aspekt untersucht, der in der früheren Studie zwar angesprochen, aber noch nicht bearbeitet wurde: die allgemeine Stimmfarbe als musikalisches Ausdrucksmittel.

Unter allgemeiner Stimmfarbe ist ein Aspekt des Stimmklanges gemeint, der sich in den stationären Klangabschnitten äußert, also vor allem in gesungenen Vokalen, und der eine zur phonetischen Klangfarbe zusätzliche Qualität bringt. Diese Qualität kann auf das Einzelphonem beschränkt sein, sie kann sich aber auch über eine phonetische Segmentkette erstrecken. In der Phonetik wird hierfür der Terminus vocal oder auditory setting verwendet (s. Laver 1980). Ein aus dem Bereich des Gesanges bekanntes Beispiel einer solchen

Über die phonetische Segmentkette gelegten zusätzlichen Qualität bildet der Singformant (Sängerformant) im Bel Canto, also jener zwischen 2800 und 3000 Hz liegende Verstärkungsbereich im Klangspektrum der Stimme, der dieser den charakteristischen metallischen Glanz verleiht (siehe vor allem: Sundberg 1974; 1974a; 1987: 118-24; Hollien 1983).

Neben anderen Mitteln des stimmlichen Ausdrucks hat der Sänger auch die Möglichkeit der vocal settings, und die Untersuchung wird sinnvollerweise an Gesängen ansetzen, die hinsichtlich ihres Stimmungsgehalts gegensätzlich sind. Für die vorliegende Untersuchung wurden zwei solcher Gesänge gewählt: der eine, Hey, good lookin' (im weiteren Gesang 1 genannt), ist ein heiterer Werbegesang mit dem Motto "Wie wär's mit uns beiden?" der andere, I'm so lonesome I could cry (im weiteren Gesang 2 genannt), eine Liebesklage.

Die Hypothese, die getestet werden soll, lautet also: Hank Williams verwendet neben anderen Mitteln des gesanglichen Ausdrucks auch das der vocal settings, wobei der Charakter dieser zusätzlich zum Textklang tretenden Qualitäten vom Stimmungsgehalt des jeweiligen Gesanges oder auch Gesangsteiles abhängt.

Will man solche zusätzliche Qualitäten identifizieren und beschreiben, muß man von der den einzelnen Phonemen der phonetischen Segmentkette eigenen Klangfarbe abstrahieren. Wir versuchten dies durch Erstellung von geglätteten Langzeitspektren über entsprechende Segmente, indem wir uns auf Ergebnisse der Stimmforschung stützten, die besagen, daß die unmittelbare Textabhängigkeit eines Langzeitspektrums ab einer Textlänge von etwa 10 sec. verlorengeht (s. Bunge 1977: 22 f.). Es bleibt dann nur das Stimmspezifische übrig, das z.B. gestattet, einen Sprecher als solchen zu erkennen. Auf unsere spezielle Fragestellung übertragen, heißt dies, daß durch Langzeitspektren bestimmter Gesangssegmente (Abb. 1 und 2) jene zusätzliche Qualität faßbar wird, um die es im gegebenen Zusammenhang geht.

Segment 1 umfaßt die ersten acht Takte von Gesang 1, Segment 2 die melodisch gleichen, textlich unterschiedlichen zweiten acht Takte,

Segment 3 die melodisch und textlich unterschiedlichen dritten acht Takte. Segment 4 ist der mit Segment 1 identische Refrain, und auch das entsprechende Langzeitspektrum (Abb. 3) ist mit dem des Segments 1 praktisch gleich, so daß Segment 4 im folgenden unberücksichtigt bleibt.

Abb. 4 zeigt die Langzeitspektren der Segmente 1, 2 und 3. Trotz geringer Unterschiede zwischen den Kurven 1 und 2 einerseits und der Kurve 3 andererseits, ergibt sich die gleiche Charakteristik: ein um 350 Hz liegender erster Formant (F1), ein zweiter, schwächerer, zwischen 1900 und 2500 Hz (F2) und ein kontinuierlicher Abfall der Amplitude mit zwei kleineren Gipfeln bei 4000 und 4600 Hz. Da als Schlüsselphrasen des Stimmungsgehalts dieses Gesanges die Segmente 1 und 2 anzusprechen sind (s. Abb. 1) - hier erfolgt die Aufforderung "Wie wär's mit uns beiden?", während in Segment 3 dann bloß aufgezählt wird, was die beiden machen könnten -, werden diese der weiteren Untersuchung zugrundegelegt.

Um ungefähr gleiche Segmentlängen miteinander zu vergleichen, wurde beim zweiten Gesang (Abb. 2) die 1. Strophe in zwei Teile geteilt: Segment 1 (in der fortlaufenden Zählung 5) vom Strophenbeginn bis fly, Segment 2 (in der fortlaufenden Zählung 6) anschließend bis zum Strophenende. Daß diese Vorsicht überflüssig war und in Gesang 2 die zusätzliche Qualität stropfenübergreifend ist, zeigt eine Gegenüberstellung der Langzeitspektren für Strophe 1 (Kurve 5/6) und Strophe 2 (Kurve 7), s. Abb. 5. Außerdem ist hierin ein Beweis für die Mittelungstheorie der Langzeitspektren zu sehen.

Abb. 6 bringt die Gegenüberstellung der Langzeitspektren der beiden Gesänge. Während die Kurven 5 und 6 (Gesang 2) trotz melodischer und textlicher Ungleichheit gut zusammenpassen, unterscheiden sie sich deutlich von der Kurve 1 (Gesang 1): F1 liegt - wohl infolge der etwas tieferen Tonlage der Haltetöne - etwas niedriger als in Gesang 1; nach dem steilen Amplitudenabfall nach F1 folgt kein neuerlicher Anstieg wie in Kurve 1, sondern der Amplitudenabfall setzt sich kontinuierlich und leicht wellenförmig fort.

In diese geglätteten Spektren, in denen die von der phonetischen Segmentkette abgehobene Qualität allgemeine Stimmfarbe gesehen

wird, ist nun allerdings die Instrumentalbegleitung mit eingegangen. Wiewohl diese in den Aufnahmen stark zurückgedrängt ist, darf dieses Faktum nicht einfach übergangen werden. Es wurden daher im weiteren Verlauf der Untersuchung von den Haltetönen der durch die geglätteten Spektren 1, 2 und 5 repräsentierten Gesangssegmente Sonagramme hergestellt und daraus unter auditiver Kontrolle je eine Stelle mit besonders geringem Anteil an Instrumentalbegleitung ausgewählt und über die ausgewählten Stellen je eines Segments wieder ein geglättetes Spektrum gerechnet. Als Beispiel zeigt Abb. 7 das Sonagramm des ersten Haltetones in Gesang 1 und eingezeichnet die Stelle geringster Instrumentalbegleitung, Abb. 8 das Spektrum dazu. In Abb. 9 ist das geglättete Spektrum über das ganze 1. Segment des Gesanges 1 jenem über die drei in diesem Segment aufscheinenden Haltetöne gerechneten, gegenübergestellt. Abb. 10 zeigt das gleiche für Segment 2 des Gesanges 1, und schließlich wird in Abb. 11 das Langzeitspektrum des Segments 5 (Gesang 2) dem aus neun Einzeltönen gebildeten desselben Segments gegenübergestellt. In allen drei Fällen zeigt sich grundlegende Übereinstimmung der über die Haltetöne gerechneten Langzeitspektren mit jenen der ganzen Segmente.

Die geglätteten Spektren wurden mit der sog. Cepstrum-Methode (s. Noll 1967) hergestellt, ein Verfahren, das kein psychoakustisches Modell einbezieht, sondern lediglich die physikalischen Daten im Hinblick auf eine optimale Schätzung der Umhüllenden transformiert. Dieses Verfahren beruht auch nicht auf dem akustischen Modell der Sprachproduktion. Bei richtiger Wahl der Parameter erhält man jedoch geglättete Spektren, welche sich durchaus sinnvoll als komplexe Resonanzfunktion eines Vokaltraktsystems interpretieren lassen. Es hat sich gezeigt, daß im Falle von stimmhaften, nicht nasalierten Lauten bei normalen Sprachklängen das über das Cepstrum geglättete Amplitudenspektrum durchaus mit spektralen Daten vergleichbar ist, die über ein Modell der akustischen Sprachproduktion (Linear Prediction Coding) berechnet wurden, weshalb dieses Verfahren auch tatsächlich zur Sprachsynthese verwendet wird (Cepstrum Vocoder). Dieser Weg wurde auch für die vorliegende Studie beschritten, um die von den Segmenten der beiden Gesänge gerechneten geglätteten Spektren zum Erklingen zu bringen. Als Quellenspektren wurden hierbei verwendet: a) ein Klang mit $f_0 = 261$ Hz entsprechend dem

Durationalton (Ton mit der insgesamt größten Gesamtdauer einer Musikprobe, s. Merriam 1963: 111, Anm. 11) des Gesanges 1, b) ein Klang mit $f_0 = 207$ Hz entsprechend dem Durationalton des Gesanges 2, c) - um den Faktor Tonhöhe auszuschalten - weißes Rauschen.

Diese Quellenspektren wurden entsprechend den Langzeitspektren der beiden Gesänge gefiltert (Abb. 12a-d) und hörbar gemacht, wobei der in den geglätteten Spektren deutlich erkennbare Unterschied (s. Abb. 6) auch auditiv deutlich zum Tragen kommt.

Somit kann in den beiden Langzeitspektren und deren auditivem Korrelat jene zur lautsprachlichen Klangfarbe tretende zusätzliche Qualität gesehen werden, die der Sänger zur Verdeutlichung des Stimmungsgehaltes der beiden Gesänge einsetzt. Ob er dies auch in anderen Fällen tut, und ob dementsprechend von einem Charakteristikum seiner Ausdrucksgestaltung gesprochen werden kann, müssen weitere Untersuchungen zeigen. Ferner wird zu untersuchen sein, ob er - was zu erwarten ist - darüber hinaus auch die Nuancierung der einzelnen Vokalqualitäten in den Dienst der musikalischen Ausdrucksgestaltung stellt.

Literatur

- Deutsch, Werner A. & Anton Noll & Heinrich Mislik: Sound-Tools: Integrierte Arbeitsstation zur digitalen Analyse und Synthese von akustischen Signalen (Sprache, Musik, Psychoakustik, Medizin, Bioakustik). In: Ber. Intern. Symposium Das Instrumentalspiel, 12.-14. April 1988, Wien (im Druck).
- Bunge, Ernst: Vergleichend-systematische Untersuchungen zur automatischen Identifikation und Verifikation kooperativer Sprecher, Darmstadt 1977.
- Födermayr, Franz: Populärmusik als Gegenstand musikwissenschaftlicher Forschung. Ein hermeneutischer Versuch. In: *Musica Austriaca* 5 (1985); 49-84.
- Hollien, Harry: The Puzzle of the Singer's Formant. In: Bless, Diana M. & James H. Abbs (Ed.), *Vocal Fold Physiology*, San Diego 1983, 368-378.
- Koon, George William: Hank Williams. A Bio- Bibliography, Westport, Conn. 1983.

Laver, John: The Phonetic Description of Voice Quality, Cambridge 1980.

Noll, M.: Cepstrum Pitch Detection. In: J. Acoust. Soc. Am. 41, 2(1967), 293 ff.

Sundberg, Johan: Articulatory Interpretation of the "Singing Formant". In: J. Acoust. Soc. Am. 55, 4(1974), 838-844; Production and Function of the "Singing Formant". In: Report Eleventh Congr. Intern. Musicol. Soc. Copenhagen 1972, Copenhagen 1974a, 679-686; The Science of the Singing Voice, Dekalb, Ill. 1987.

Williams, Roger M.: Sing a Sad Song. The Life of Hank Williams, Urbana 1981.

Segmente MM. J. = 142

Abb. 1: Hank Williams, *Hey good Lookin'*, 1. Strophe, Grobtranskription nach MGG Rec. Japan, MM 9099 A4, Gesangssegmente 1 bis 4.

Segmente MM. J. = 113

Abb. 2: Hank Williams, *I'm so lonesome I could cry*, 1. Strophe, Grobtranskription nach MGM Rec. Japan, MM 9099 B7, Gesangssegmente 4 und 5.

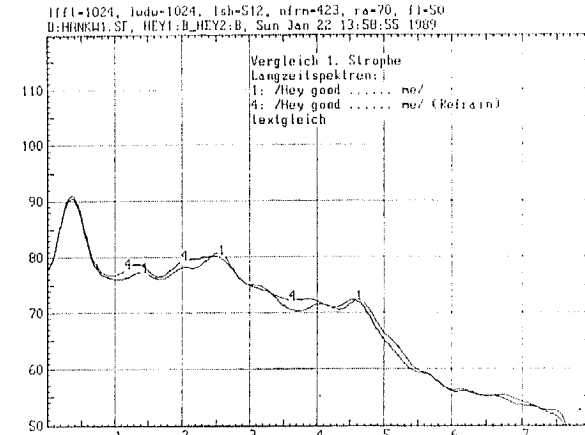


Abb. 3: MM 9099 A4, Langzeitspektren der Segmente 1 und 4. Abszisse: kHz., Ordinate: dB.

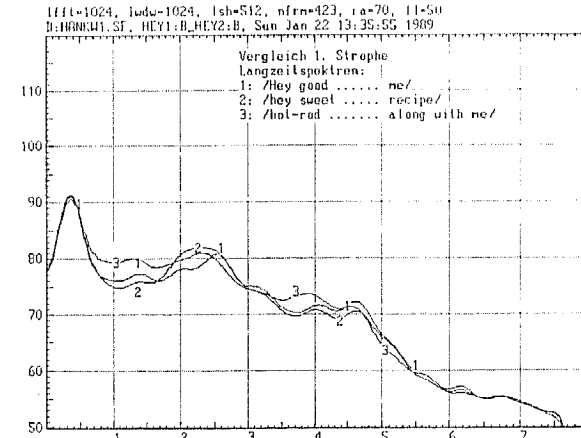


Abb. 4: MM 9099 A4, Langzeitspektren der Segmente 1 bis 3.

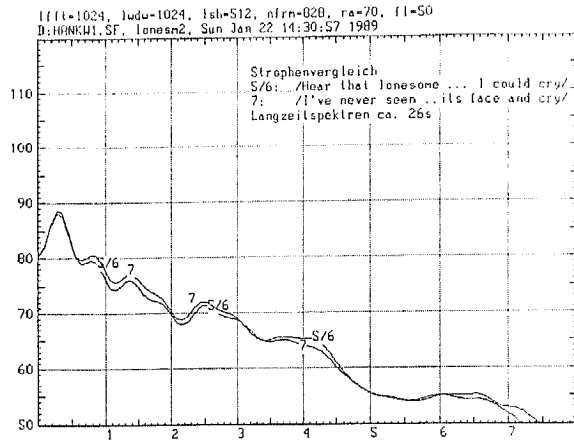


Abb. 5: MM 9099 B7, Langzeitspektren über Strophe 1 (5/6) und 2(7).

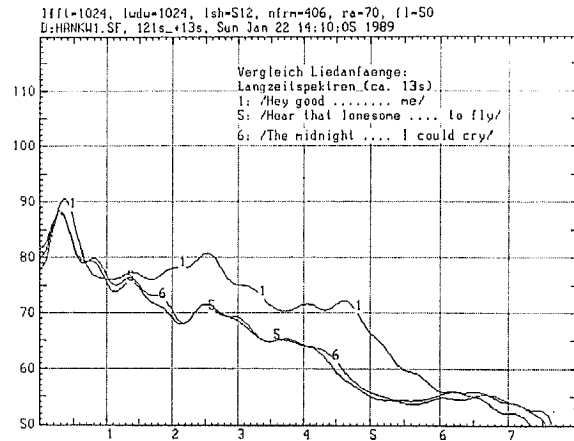


Abb. 6: Gegenüberstellung der Langzeitspektren der Segmente 5 und 6 (MM 9099 B7) mit 1 (MM 9099 A4).

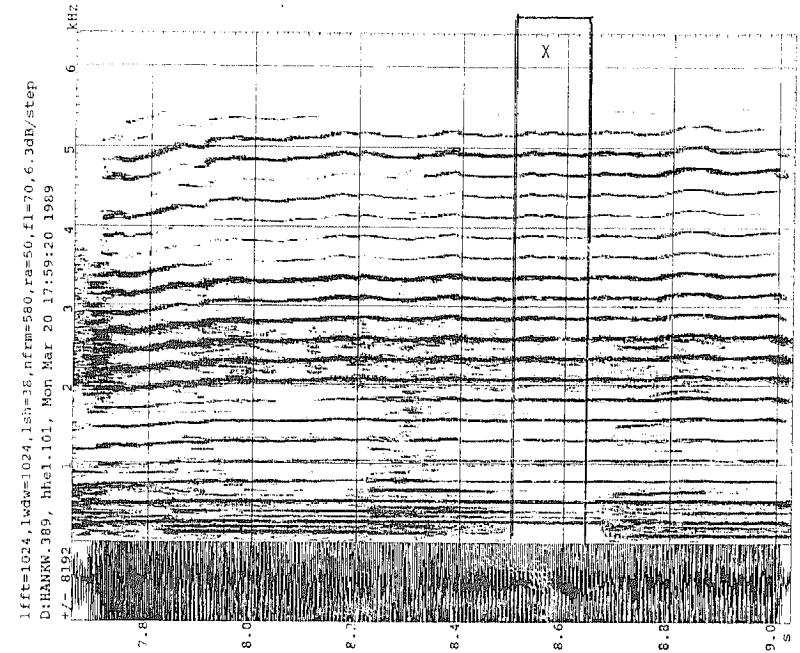


Abb. 7: Sonogramm des 1. Haltetones der 1. Strophe (c' über *hey*).

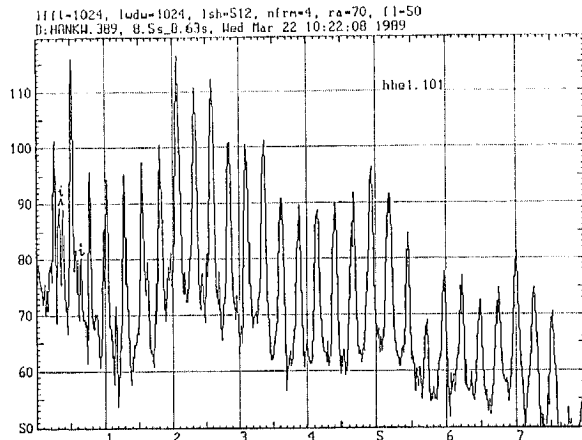


Abb. 8: Spektrogramm der Stelle X von Abb. 7. i = instrumentaler Anteil.

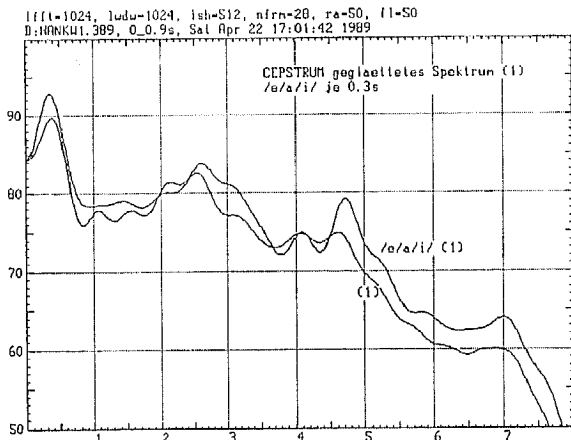


Abb. 9: MM 9099 A4. Gegenüberstellung des Langzeitspektrums von Segment 1 mit dem über die eher instrumentenfreien Abschnitte der Haltetöne c' (hey), c' (what), c' (na) desselben Segments gerechneten.

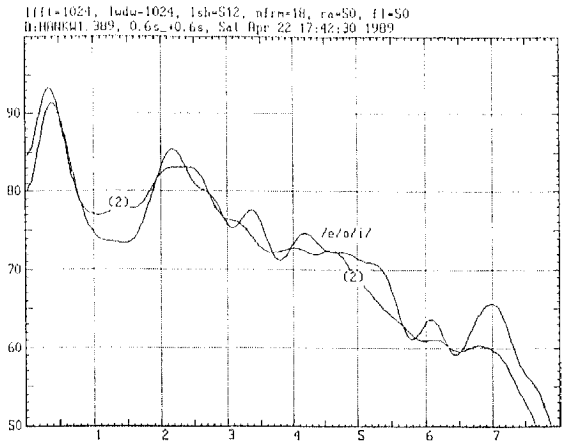


Abb. 10: MM 9099 A4. Gegenüberstellung des Langzeitspektrums von Segment 2 mit dem über die eher instrumentenfreien Abschnitte der Haltetöne c' (hey), c' (don't), c' (pc) desselben Segments gerechneten.

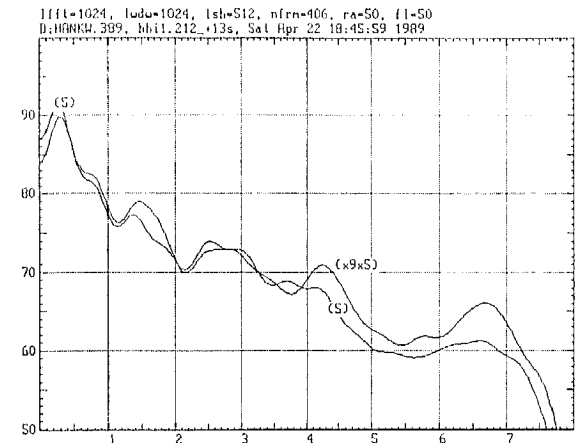


Abb. 11: MM 9099 B7. Gegenüberstellung des Langzeitspektrums von Segment 5 mit dem über die eher instrumentenfreien Abschnitte der 9 Haltetöne des Segments gerechneten.

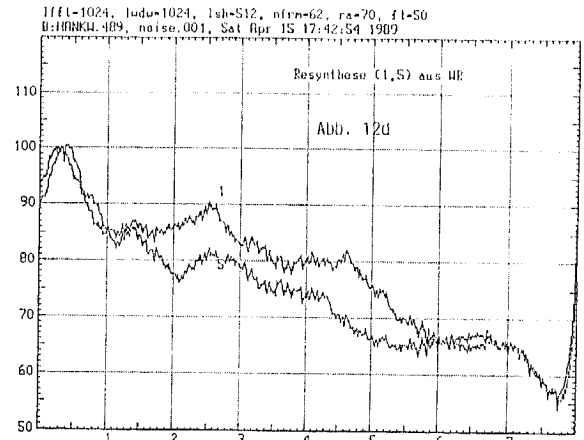
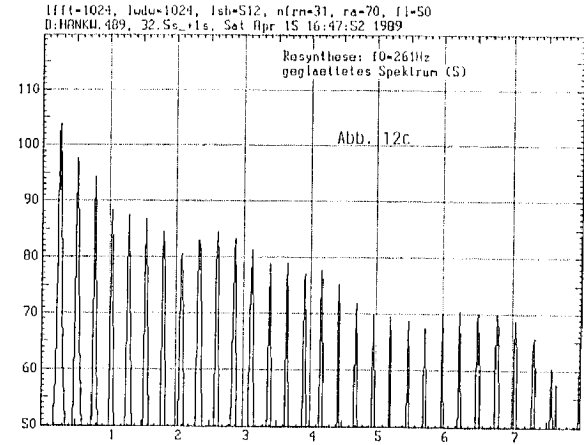
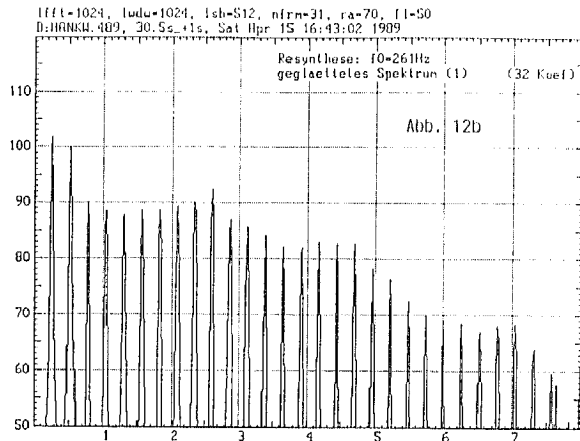
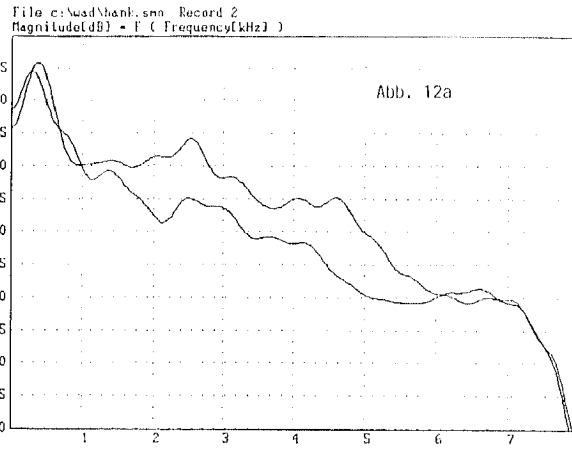


Abb. 12: Resynthese. a) Verwendeter Filterfrequenzgang, b) Quellenspektrum $f_0 = 261$ Hz nach Filterung entspr. Kurve 1, c) Quellenspektrum $f_0 = 161$ Hz nach Filterung entspr. Kurve 5, d) Weißes Rauschen nach Filterung entspr. Kurve 1 und 5.