

logarithmierte Flächen und LSF-Koeffizienten (line spectral frequencies) betrachtet. Für die Berechnung der LSF-Koeffizienten wird das Nennerpolynom $A(z)$ der Übertragungsfunktion H benötigt. Wie in [1] beschrieben, ist die Darstellung des verlustbehafteten Rohrmodells durch Zähler- und Nennerpolynom numerisch problematisch. Als Näherung für $A(z)$ werden die Reflexionskoeffizienten unter der Annahme eines verlustlosen Rohrmodells in das Polynom $A'(z)$ konvertiert, woraus dann die LSF-Koeffizienten γ ermittelt werden können. Da das verwendete Rohrmodell Verluste aufweist, werden verteilte Verluste in der Polynomdarstellung berücksichtigt. In Abb. 1 ist zu sehen, dass jeder Zustandsspeicher z^{-1} mit einem Verlustsystem zu \mathcal{G} multipliziert wird. Bei Verwendung einer reellen Verlustkonstanten ν kann daher durch die Substitution $z \rightarrow z/\nu$ von der verlustlosen Darstellung in eine verlustbehaftete Darstellung des Polynoms gewechselt werden. Die aus $A'(z/\nu)$ berechneten LSF-Koeffizienten werden mit γ_ν bezeichnet; ν hat hier den Wert 0,965.

Für die Evaluierung der Parameterdarstellungen werden die Fehler-Maße verwendet. Dafür werden für jeden Laut 45 Diphone herangezogen. Die verwendete Diphonatenbank del ist mit 16 kHz aufgenommen und enthält eine weibliche Stimme; jedes Diphon ist nur einmal vorhanden. Für alle sinnvollen Kombinationen der Diphone werden die Fehler-Maße berechnet und die zu einem Laut gehörenden Fehler gemittelt. Anschließend wird nochmals über alle Laute gemittelt; die resultierenden über alle Diphonkombinationen gemittelten Fehler \bar{e}_1 , \bar{e}_2 und \bar{e}_3 erlauben eine Bewertung der Parameterdarstellungen bezüglich der Diphonübergänge. Tabelle 1 zeigt die gemittelten Fehler in Abhängigkeit der Parameterdarstellung. Für jeden gemittelten Fehler wurden 15760 Diphonübergänge herangezogen. Die Anzahl der interpolierten Parametersätze ist mit $M = 5$ gewählt.

Tabelle 1: Fehler-Maße der Parameterdarstellungen

	\bar{e}_1	\bar{e}_2	\bar{e}_3
r	2,78	1,57	2,06
log.AREA	2,97	1,70	2,21
γ	1,96	1,18	1,50
γ_ν	1,89	1,10	1,42

Wie in Tabelle 1 zu erkennen ist, ergeben die Übergänge mit LSF-Koeffizienten die kleinsten Fehler. Die in Abb. 2 gezeigten Übergänge zeigen, welchen Einfluss die Parameterdarstellungen auf die Spektralübergänge haben. Abb. 2 a), b) zeigt, dass insbesondere bei den Übergängen mit Flächen oder Reflexionskoeffizienten fehlerhafte Resonanzverläufe bzw. Zuweisungen stattfinden können im Gegensatz zu den LSF-Koeffizienten. Der Spektralübergang mit Reflexionskoeffizienten ist mit dem von Abb. 2 a) vergleichbar. Die Unterschiede zwischen den LSF Versionen γ und γ_ν sind eher gering ausgeprägt. Wie bei der markierten Resonanz in Abb. 2 c), d) zu sehen ist, werden durch die Verwendung von γ_ν in d) die Bandbreiten der Resonanzen bei größeren Frequenzverschiebungen erhöht; dies verkleinert den Fehler.

Die Untersuchungen favorisieren die Darstellungen mit den LSF-Koeffizienten im Gegensatz zu denen, die unmittelbar aus den Reflexionskoeffizienten oder Flächen gebildet werden. Von den beiden LSF Versionen erweist sich γ_ν als etwas günstiger. Zusätzliche Verbesserungen lassen sich erzielen, wenn die Pole von H als Parameter verwendet werden.

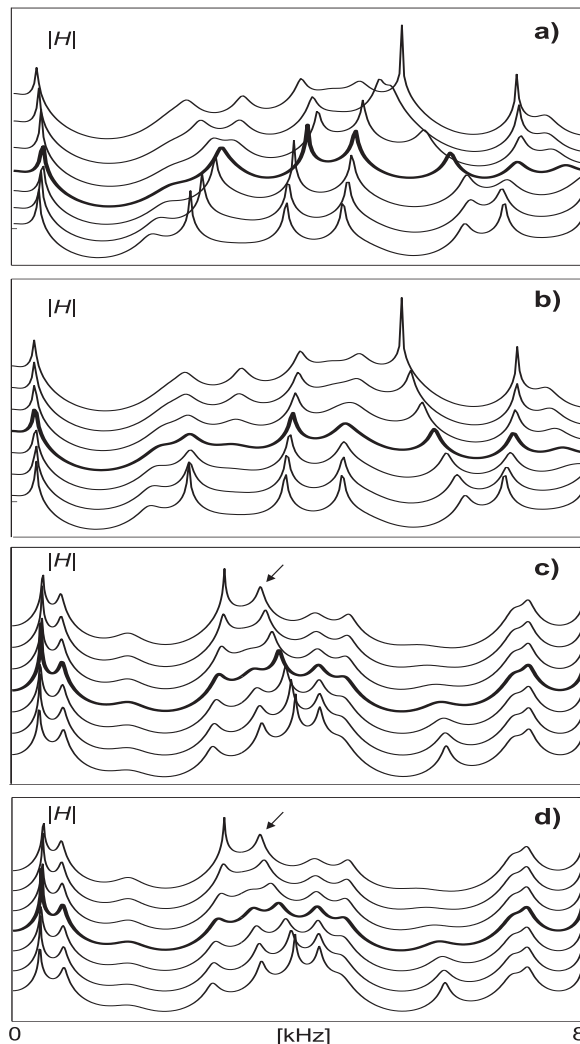


Abbildung 2: Betragsgänge des Übergangs: a) Diphonverketzung im Laut /j/ mit logarithmierten Flächen als Parameter, b) wie Graph a) jedoch mit LSF-Koeffizienten, c) Diphonverketzung im Laut /o:/ mit LSF-Koeffizienten γ als Parameter, d) wie Graph c) jedoch mit LSF-Koeffizienten γ_ν . Der Betragsgang der Übergangsmittle besitzt eine kräftigere Strichdicke.

Literatur

- [1] Schnell K., Lacroix A.: 'Analysis of lossy vocal tract models for speech production', Proc. EUROSPEECH-2003, Geneva Switzerland, pp. 2369-2372, 2003.
- [2] Chappell D.T., Hansen J.H.L.: 'A comparison of spectral smoothing methods for segment concatenation based speech synthesis', Speech Communication (36) pp. 343-374, 2002.
- [3] Schnell K., Lacroix A.: 'Model Based Analysis of a Diphone Database for Improved Unit Concatenation'. Proc. EUROSPEECH-2005, Lisbon, 2005.