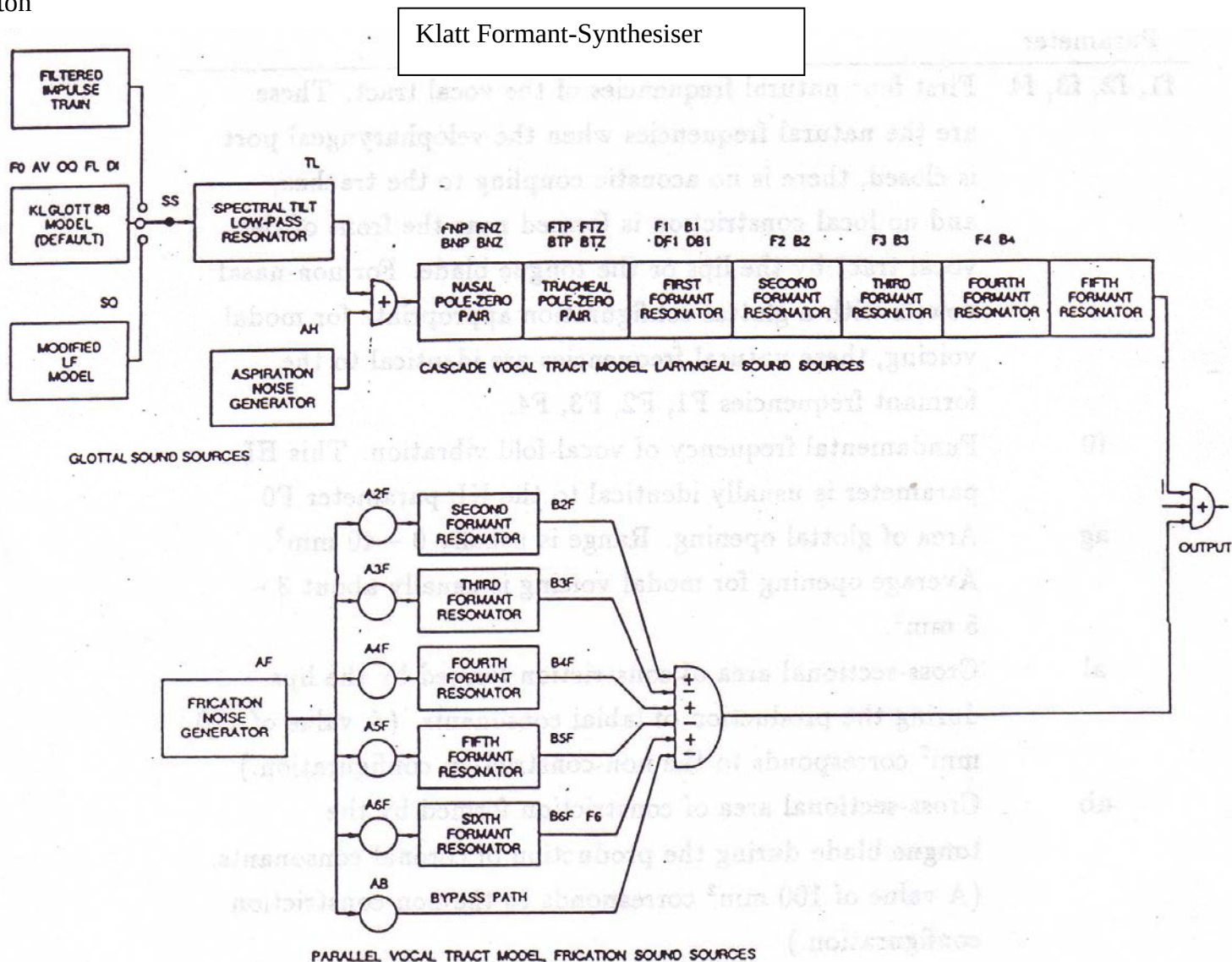


HLSYN: Software für die parametrische Synthese der Sprache

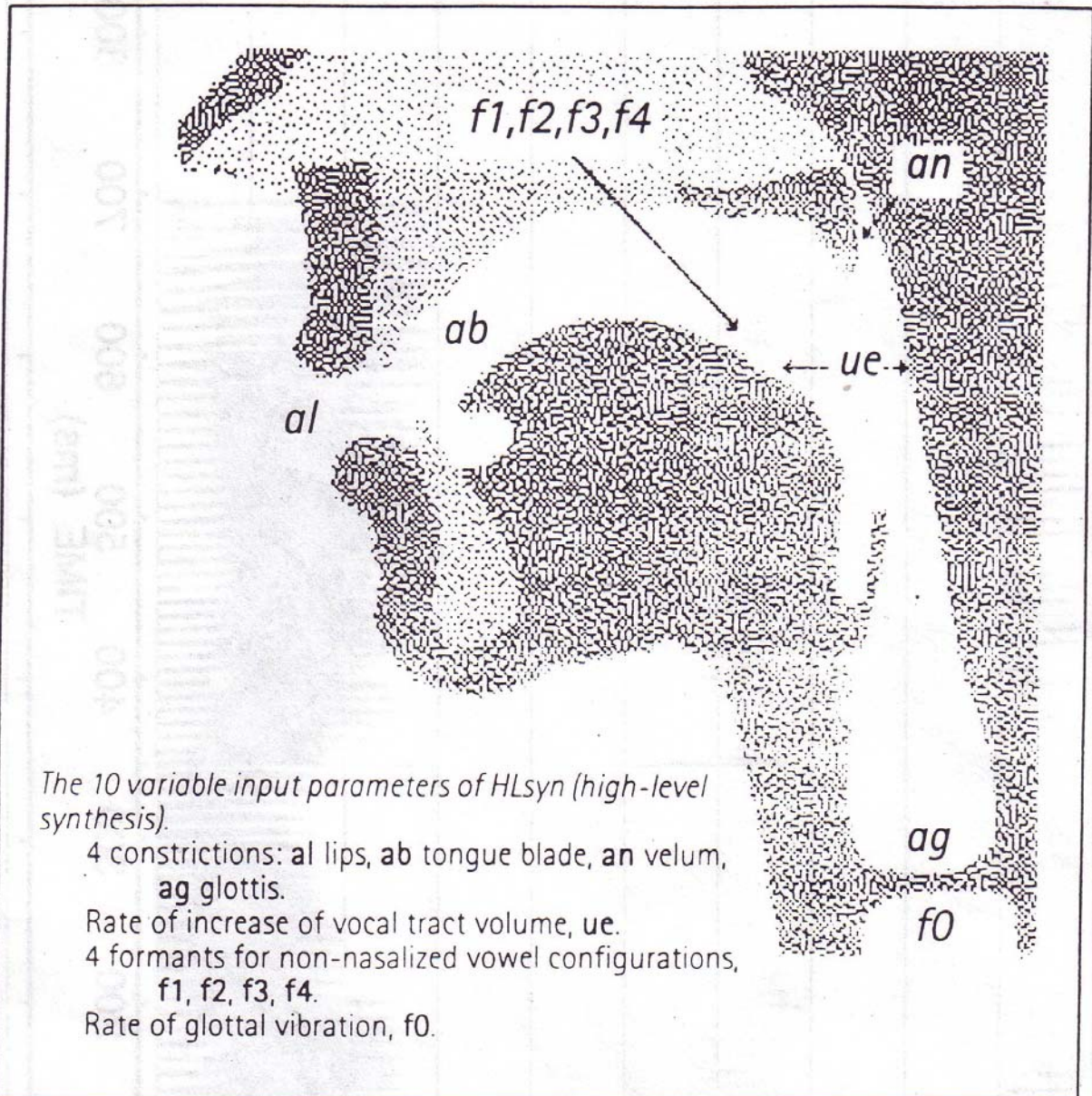
Modul J, parameter.pdf

Die Beispiele (hld und wav Dateien) können von der Modul J Webseite [beispiele.zip](#) heruntergeladen werden.

Jonathan Harrington



Reduzierung der Parameter



Die HL-Parameter

Art. Parameter	Interpretation	Typische Werte
Ag	Glottalverengung	0 (Schließung), 4 mm ² (Normales Sprechen)
Al	Lippenverengung	0 (Schließung) 100 mm ² (Normales Sprechen)
Ab	Zungenblattverengung	0 (Schließung) 100 mm ² (Normales Sprechen)
An	Velumverengung	0 (oral), 30 mm ² (nasal)
ue	Geschwindigkeit der Ausdehnung vom Vokaltrakt hinter der Schließung. Hat nur eine Wirkung, wenn der Vokaltrakt durch Al oder Ab geschlossen ist. Dieses Parameter wird verwendet, wenn die Stimmhaftigkeit während der Schließung erhalten bleiben soll.	0 (keine Vergrößerung); 150 cm ³ /s (Stimmhaftigkeit im Verschluss)

f0, f1-f4

- Folgendes beachten: F1 < 300 Hz bedeutet eine Zungenrückenschließung (wie für [g] – siehe 3.4 unten).
- Alle Werte x 10, daher um eine Grundfrequenz von 100 Hz zu synthetisieren, muss 1000 eingegeben werden.

11 Synthese von Vokalen

[a] Monoton 150 Hz. 11.hld

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps	dc
0.0	4.000	100.0	100.0	0.0	0.0	1500	720	1100	2500	3500	8	0.0
500.0	4.000	100.0	100.0	0.0	0.0	1500	720	1100	2500	3500	8	0.0

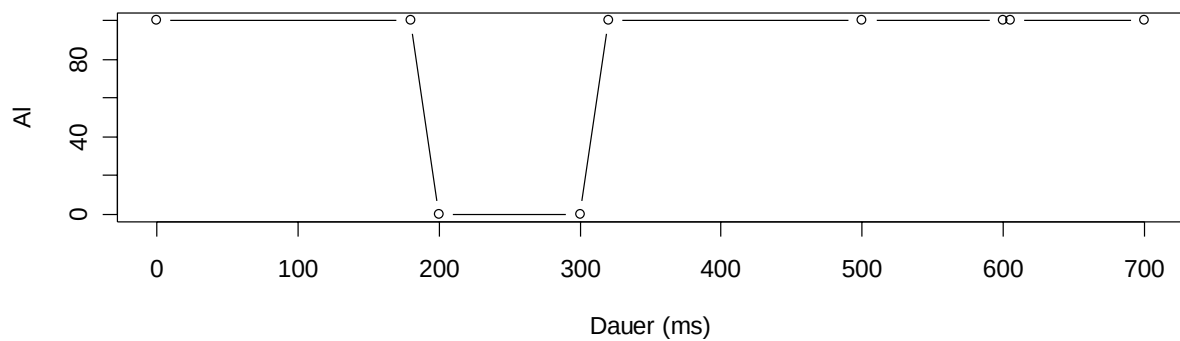
	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps
0.0	0	100.0	100.0	0.0	0.0	0	720	1100	2500	3500	8.000
25	4	100.0	100.0	0.0	0.0	1500	720.0	1100	2500	3500	8.000
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2500	3500	8.000
600	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2500	3500	8.000
605	0	100.0	100.0	0.0	0.0	0	720.0	1100	2500	3500	8.000
700	0	100.0	100.0	0.0	0.0	0	720	1100	2500	3500	8.000

Zwei weitere Beispiele:

[æ] oder [a]: ae.hld

Eine Qualität zwischen [ɛ] und [e]: e.hld

Es gibt immer eine lineare Interpolation zwischen den fettgedruckten Werten, zB hat der eingekreiste Lippenschließung-Parameter diese Interpretation:



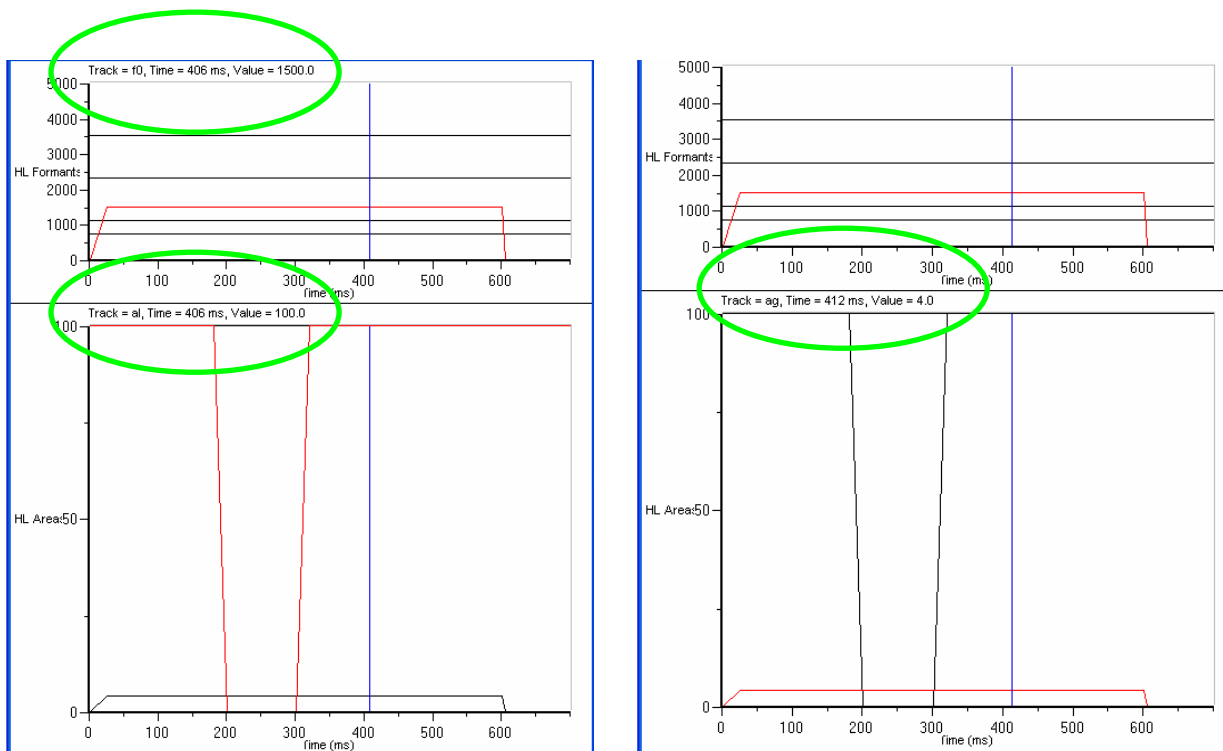
Die Parameter von HLSYN können durch View -> Graphic Panes dargestellt werden:

HL Formants: f0 oder F1-F4

HL Areas: ag, al, ab, an

HL Flow: ue

Der Parameter ist immer derjenige, für den die Werte gezeigt werden – (unten links). Um auf einen anderen Parameter zu wechseln, einen schwarzen Parameter anklicken – der Parameter wird rot und die entsprechenden Werte werden gezeigt (unten rechts).



21 Lippenschließung

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
180	4.000	100	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
200	4.000	0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
300	4.000	0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
320	4.000	100	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
600.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
605.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000
700.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000

22 Wie 21 aber mit einem stimmlosen Ansatz zum zweiten /a/
(dies soll für die Aspiration geeignet sein)

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps	dc	ap
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000	0.0	0.0
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0	0.0
180.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0	0.0
200.0	8.333	0.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0	0.0
300.0	30.00	0.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0	0.0
320.0	23.50	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0	0.0
380.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0	0.0
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0	0.0
600.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0	0.0
605.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000	0.0	0.0
700.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000	0.0	0.0

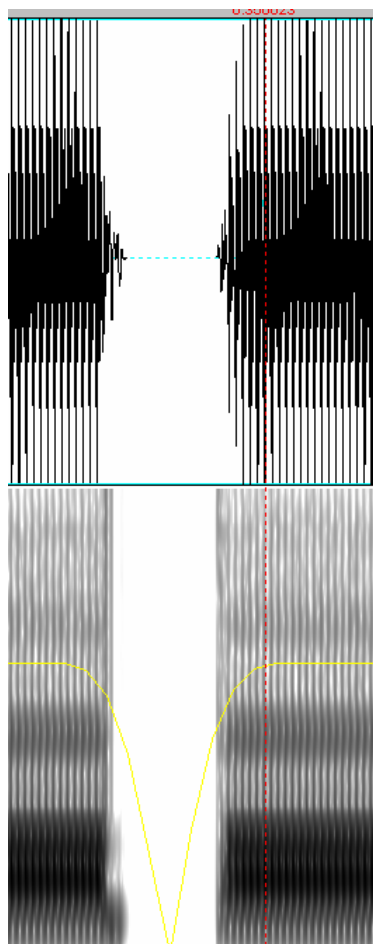
23 Wie 21 und mit Stimmhaftigkeit im Verschluss

Eine Vergrößerung des Vokaltrakts während des Verschlusses um ca. 150 cm³/s
Parameter in /home/jmh/HLSYN/basic/23.hld, 23.wav

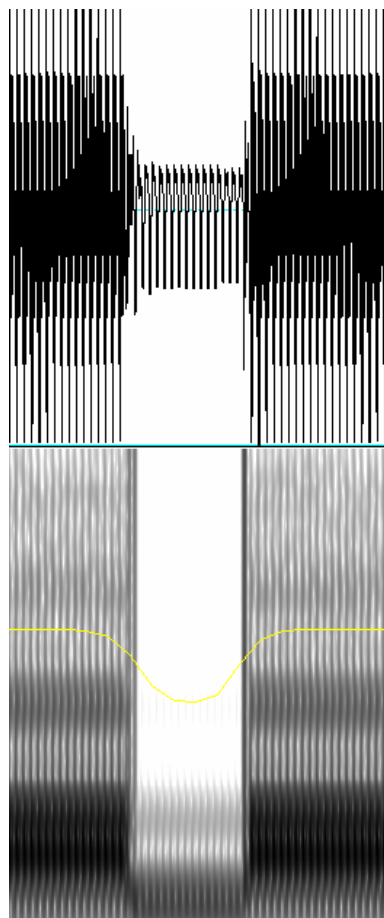
	ag	al	ab	an	ue	f0	f1	f2
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100
180.0	4.000	100.0	100.0	0.0	0	1500	720.0	1100
200.0	4.000	0.0	100.0	0.0	150	1500	720.0	1100
220	4.000	0.0	100.0	0.0	150	1500	720.0	1100
300.0	4.000	0.0	100.0	0.0	150	1500	720.0	1100
305	4.000	25.00	100.0	0.0	0	1500	720.0	1100
320.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100
600.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100

Vergleiche die Spektrogramme (in Praat) von 21 (links) und 23 (rechts) in den Abbildungen unten – gelb ist die Intensität.

21.wav



23.wav



31 Wie 23 und mit Formant-Transitionen für /aba/

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	p
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.00
160	4.000	100.0	100.0	0.0	0.0	1500	720	1100	2300	3500	8.00
180.0	4.000	100.0	100.0	0.0	0	1500	560.0	910.0	2250	3500	8.00
200.0	4.000	0.0	100.0	0.0	150	1500	400	720	2200	3500	8.00
220	4.000	0.0	100.0	0.0	150	1500	400.0	720.0	2200	3500	8.00
300.0	4.000	0.0	100.0	0.0	150	1500	400	720	2200	3500	8.00
305	4.000	25.00	100.0	0.0	0	1500	440.0	767.5	2212	3500	8.00
320.0	4.000	100.0	100.0	0.0	0.0	1500	560.0	910.0	2250	3500	8.00
340	4.000	100.0	100.0	0.0	0.0	1500	720	1100	2300	3500	8.00
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.00

22b wie 22 aber stimmloser /apa/

(also die Formantänderungen in 31 oben werden auf 22 überlagert)

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
160.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
180.0	4.000	100.0	100.0	0.0	0.0	1500	560.0	910.0	2250	3500	8.000
200.0	8.333	0.0	100.0	0.0	0.0	1500	400.0	720.0	2200	3500	8.000
300.0	30.00	0.0	100.0	0.0	0.0	1500	400.0	720.0	2200	3500	8.000
320.0	23.50	100.0	100.0	0.0	0.0	1500	560.0	910.0	2250	3500	8.000
340.0	17.00	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
380.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
600.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
605.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000

33 wie 31 aber /ada/

- ab statt al verwendet – und die Transition ist etwas (5 ms) länger.
- Die Formanttransitionen müssen geändert werden F2=1700 Hz, F3=2650 Hz.
- Die Synthese ist auch ohne ue (ohne Stimmhaftigkeit während des Verschlusses) überzeugender.

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps	dc
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000	0.0
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
160.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
175.0	4.000	100.0	100.0	0.0	0.0	1500	600.0	1325	2431	3500	8.000	0.0
180.0	4.000	100.0	80.00	0.0	0.0	1500	560.0	1400	2475	3500	8.000	0.0
200.0	4.000	100.0	0.0	0.0	0.0	1500	400.0	1700	2650	3500	8.000	0.0
220.0	4.000	100.0	0.0	0.0	0.0	1500	400.0	1700	2650	3500	8.000	0.0
300.0	4.000	100.0	0.0	0.0	0.0	1500	400.0	1700	2650	3500	8.000	0.0
305.0	4.000	100.0	20.00	0.0	0.0	1500	440.0	1625	2606	3500	8.000	0.0
320.0	4.000	100.0	80.00	0.0	0.0	1500	560.0	1400	2475	3500	8.000	0.0
325.0	4.000	100.0	100.0	0.0	0.0	1500	600.0	1325	2431	3500	8.000	0.0
340.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
600.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
605.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000	0.0

24 Zungendorsumschließung und [aga]

Eine Schließung vom Zungendorsum kann durch eine F1 Senkung bewirkt werden: F1=180 Hz ist eine komplette Schließung vom Zungendorsum, wie für ein [g].

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
150	4.000	100.0	100.0	0.0	0.0	1500	720	1100	2300	3500	8.000
180.0	4.000	100.0	100.0	0.0	0.0	1500	396.0	1100	2300	3500	8.000
200.0	4.000	100	100.0	0.0	0.0	1500	180	1100	2300	3500	8.000
300.0	4.000	100	100.0	0.0	0.0	1500	180	1100	2300	3500	8.000
320.0	4.000	100.0	100.0	0.0	0.0	1500	396.0	1100	2300	3500	8.000
350	4.000	100.0	100.0	0.0	0.0	1500	720	1100	2300	3500	8.000
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
600.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000

34 [aga]

Darüber hinaus wirkt (vielleicht) die Synthese von [g] noch überzeugender durch Manipulation von F2 und F3 – zB hier wie 24 aber zusätzlich, F2 = 1600 Hz, F3 = 2200 Hz.

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
150.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
180.0	4.000	100.0	100.0	0.0	0.0	1500	396.0	1400	2240	3500	8.000
200.0	4.000	100.0	100.0	0.0	0.0	1500	180.0	1600	2200	3500	8.000
300.0	4.000	100.0	100.0	0.0	0.0	1500	180.0	1600	2200	3500	8.000
320.0	4.000	100.0	100.0	0.0	0.0	1500	396.0	1400	2240	3500	8.000
350.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000
600.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000

41 /ama/: wie 31 aber mit gesenktem Velum

Parameter in /home/jmh/HLSYN/basic/41.hld, 41.wav

	ag	al	ab	an	ue	f0	f1	f2
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100
160.0	4.000	100.0	100.0	0	0.0	1500	720.0	1100
180.0	4.000	100.0	100.0	15.00	0.0	1500	560.0	910.0
200.0	4.000	0.0	100.0	30	150.0	1500	400.0	720.0
220.0	4.000	0.0	100.0	30.00	150.0	1500	400.0	720.0
300.0	4.000	0.0	100.0	30	150.0	1500	400.0	720.0
305.0	4.000	25.00	100.0	26.25	0.0	1500	440.0	767.5
320.0	4.000	100.0	100.0	15.00	0.0	1500	560.0	910.0
340.0	4.000	100.0	100.0	0	0.0	1500	720.0	1100
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100
600.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100
605.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100
700.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100

14 [aha]

Die ag-Werte im Bezug zu 1.1 müssen modifiziert werden

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps	dc
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000	0.0
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
150	4	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
250	20	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
350	4	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
600.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8.000	0.0
605.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000	0.0
700.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8.000	0.0

25 [afa]

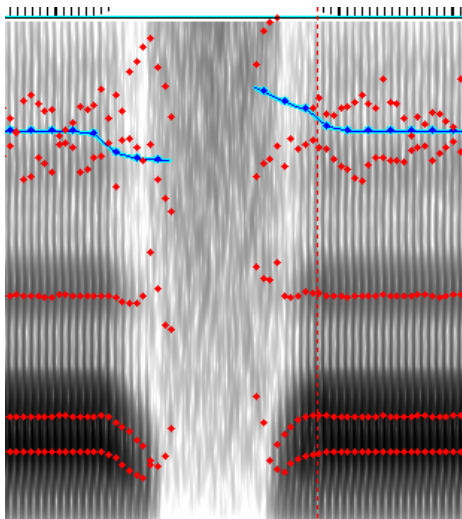
Wie 14 oben, aber jetzt brauchen wir eine Lippenschließung und die Formanten für einen Labial (wie für 31 [aba]) müssen überlagert werden.

	ag	al	ab	an	ue	f0	f1	f2	f3	f4
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500
150.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500
160.0	5.600	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500
200.0	12.00	15	100.0	0.0	0.0	1500	400.0	720.0	2200	3500
250.0	20.00	10.00	100.0	0.0	0.0	1500	400.0	720.0	2200	3500
300.0	12.00	15	100.0	0.0	0.0	1500	400.0	720.0	2200	3500
340.0	5.600	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500
350.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500

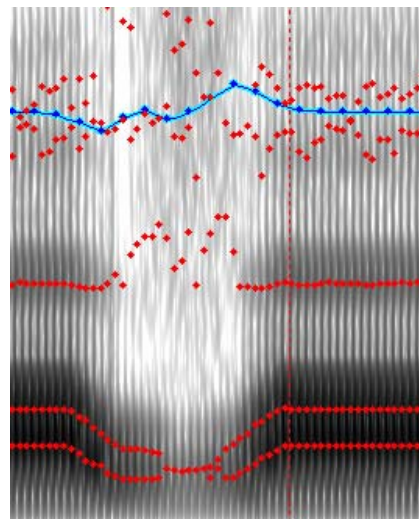
25b [ava]

ava.hld Wie oben, aber ag (die Glottalverengung) zurücksetzen, zB. auf 8 mm² (siehe Spektrogramme unten)

[afa] (25.hld)



[ava] (ava.hld)



35 [asa]

Wir für 25 [afa], jedoch **ab** statt **al** und Formantparameter wir für 33 [ada].

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	
0.0	0.0	100.0	100.0	0.0	0.0	0.0	720.0	1100	2300	3500	8
25.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8
150.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8
160.0	5.600	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8
200.0	12.00	100.0	15	0.0	0.0	1500	400	1700	2650	3500	8
250.0	20.00	100.0	10.00	0.0	0.0	1500	400.0	1700	2650	3500	8
300.0	12.00	100.0	15	0.0	0.0	1500	400	1700	2650	3500	8
340.0	5.600	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8
350.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8
500.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1100	2300	3500	8

[aʃa] kann aus 35 [asa] synthetisiert werden, aber mit F2=1800 Hz, F3=2300 Hz.

5 [ala], [ara]

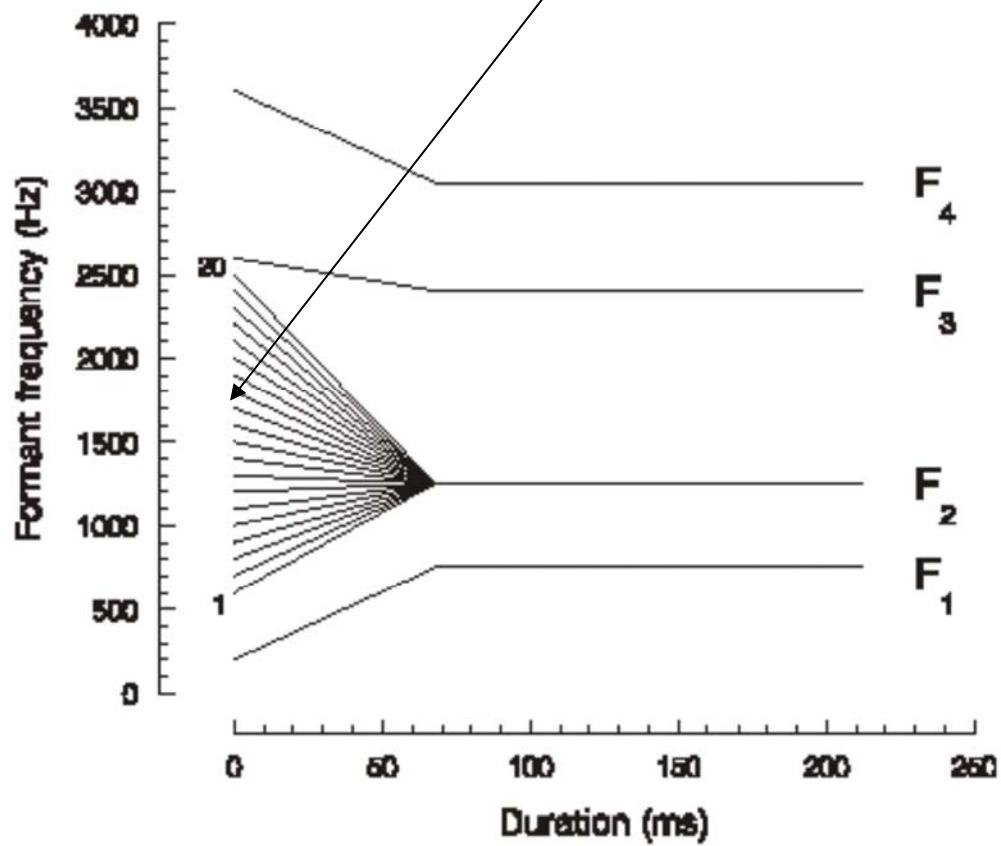
Meine Versuche, [ala] und [ara] (ein'englisches' /r/) zu synthetisieren, sind in ala.hld und ara.hld.

6. Haskins Labs. zwei-Formant Synthese

Prototypische [ba da ga] sind in baf2.hld, daf2.hld, gaf2.hld.

daf2.hld

	ag	al	ab	an	ue	f0	f1	f2	f3	f4	ps
0.0	4.000	100.0	100.0	0.0	0.0	1500	200.0	1800	2600	3600	8.000
70.00	4.000	100.0	100.0	0.0	0.0	1500	720.0	1250	2500	3000	8.000
220.0	4.000	100.0	100.0	0.0	0.0	1500	720.0	1250	2500	3000	8.000



Übungen

Synthese von:

'Steine', 'schnaufen', 'später', 'Zähne', 'glauben', 'Pfosten'.

Ein männlicher Sprecher soll diese Wörter aufnehmen. Die Synthese versuchen, an hand von einem Spektrogramm mit Formanten und f0.