

HEINRICH-HERTZ-INSTITUT FÜR SCHWINGUNGSFORSCHUNG
BERLIN-CHARLOTTENBURG

Technischer Bericht Nr. 129

Erfassung und Verarbeitung von Daten
zur Sprachintervallstatistik

von

Dipl.-Ing. Hans Dietrich Höhne

Berlin

1 9 7 0

Technischer Bericht Nr. 129

Erfassung und Verarbeitung von Daten
zur Sprachintervallstatistik

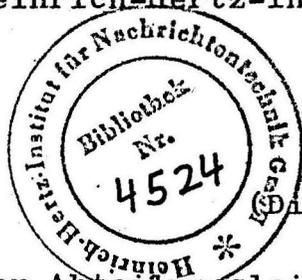
Zusammenfassung

Die Zeiten des Vorhandenseins oder Fehlens von Sprachenergie sind wichtig für die Beurteilung der Kommunikation und für den Entwurf von sprachgesteuerten Schaltern in Echosperren oder TASI-Systemen. Im vorliegenden Bericht wird über die Aufbereitung der auf Tonband gespeicherten Sprache zweier Gesprächsteilnehmer zur Erstellung von Verteilungsfunktionen berichtet. Die für ein automatisches Verfahren entwickelten Geräte werden beschrieben. Die im manuellen und im automatischen Verfahren ermittelten Verteilungsfunktionen werden mitgeteilt.

Es wird ein Algorithmus vorgeführt, der es gestattet, aus den am "nahen" Leitungsende aufgenommenen Kommunikationsereignissen die bei Laufzeit im Übertragungsweg von ihnen verschiedenen am "fernen" Ende zu berechnen.

Heinrich-Hertz-Institut für Schwingungsforschung

Der Bearbeiter:



Hans D. Höhne

(Dipl.-Ing. Hans Dietrich Höhne)

Der Abteilungsleiter

E. R. Berger

(Prof. Dr.-Ing. E. R. Berger)

Der Institutsdirektor

P. Matthieu

(Prof. Dr. phil. P. Matthieu)

Berlin-Charlottenburg, den 1. Dezember 1970

	Seite
1. Einleitung und Inhaltsübersicht =====	2
2. Sprachdetektor =====	3
2.1 Begriffsbestimmungen	3
2.2 Aufgabe und Prinzip	5
2.3 Variable Aussprechschwelle	5
2.4 Blockschaltbild	7
2.4.1 "Hüllkurven"-Erzeugung	7
2.4.2 Störunterdrückung	9
2.4.3 Frequenzabhängigkeit	10
2.4.4 Schwellenkurve	10
2.5 Überprüfung der Wirkungsweise	16
3. Verschlüsselung der Sprech-Ereignisse =====	20
3.1 Kennzahlen für Beginn, Ende und Abstand	20
3.2 Ermittlung der Zeitabschnitte	21
3.3 Manuelle Verschlüsselung	21
3.4 Automatische Verschlüsselung	22
3.4.1 Direkte Verschlüsselung	23
3.4.2 Abtastung und Umrechnung	24
4. Algorithmus zur Bestimmung des Zeit- punktes der Ereignisse am "fernen" Leitungsende =====	28
4.1 Zustandekommen und Wirkung der Laufzeit	28
4.2 Definition der Zeitangaben	29
4.3 Satelliten- und Versuchsstrecke	30
4.4 Rechnerische Bestimmung der Ereignisse am fernen Leitungsende	34
4.4.1 Verschiebealgorithmus	35
4.4.2 Sortieralgorithmus. Verwendetes Ver- fahren, Beschreibung und Einordnung	36
4.4.2.1 Auswechseln in einem Durchgang	37
4.4.2.2 Differenzen als Sortierkriterium	40
4.5 Ergebnisse der manuellen und der automa- tischen Auswertung	46

Anhang: Sortierverfahren

1. Einleitung:

Das Vorhandensein oder Fehlen von Sprachenergie wird von sprachgesteuerten Schaltern, im folgenden kurz Sprachdetektoren genannten, angezeigt. Sprachdetektoren finden Anwendung z.B. in Echosperrern oder TASI-Systemen (Time Assignment Speech Interpolation). Sie wurden im Rahmen eines Forschungsvorhabens eingesetzt, um Zeiten des Vorhandenseins oder Fehlens der Sprachenergie zu registrieren und daraus eine Aussage über die Beeinträchtigung der Kommunikation durch Laufzeit zu gewinnen.

Zu diesem Zweck wurde nach einem Vorschlag des Verfassers ein Sprachdetektor mit adaptiver Ansprechschwelle entwickelt. Da dieser Detektor und eine nachzuschaltende Codiereinrichtung zunächst nicht zur Verfügung standen, wurde das zu untersuchende Material - Tonbandaufnahmen der Testgespräche über eine mit Laufzeit behaftete Vierdrahtverbindung - anfänglich mit einer Lauffilmkamera, des Weiteren mit einem Lichtstrahloszillographen registriert. Der Inhalt der Oszillogramme wurde anhand der Tonbandaufnahmen identifiziert, die Länge der sprachlichen Äußerungen wurde ausgemessen.

Die hieraus erstellten Verteilungsfunktionen verschiedener Sprechereignisse werden beigelegt.

Die gleichen Tonbandaufnahmen wurden anschließend automatisch verschlüsselt, und die zugehörigen Verteilungsfunktionen wurden erstellt. Sie bestätigen im wesentlichen die Brauchbarkeit des Sprachdetektors, der damit für größere Datenmengen einsatzbereit ist.

Im folgenden werden der Detektor und die Sprachcodiereinrichtung beschrieben.

Weiterhin wird die manuelle Codierung anhand von Oszillogrammen erläutert.

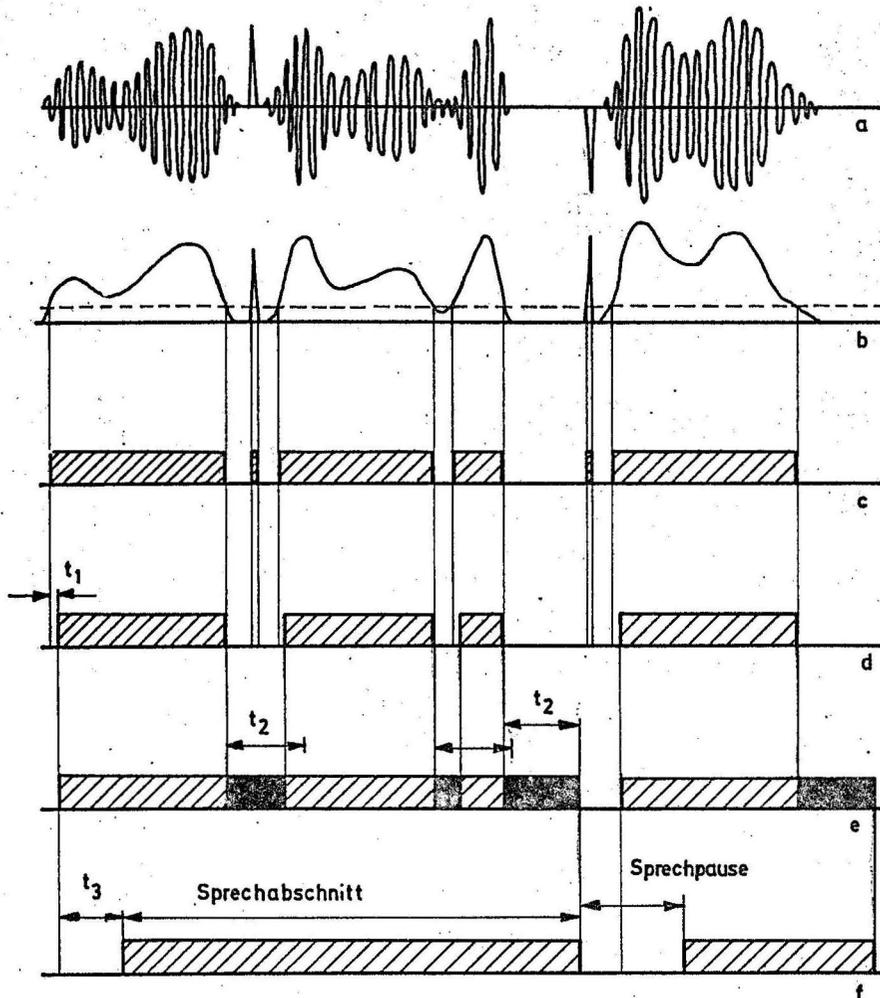
Abschließend wird ein Algorithmus "Reduktion auf fernes Leitungsende" vorgeführt, der es gestattet, aus den am nahen Leitungsende aufgenommenen Sprachereignissen die bei Laufzeit im Übertragungsweg von ihnen verschiedenen am "fernen" Ende zu berechnen. Wesentlich hierbei ist die Wahl eines zweckmäßigen Sortierverfahrens.

2. Sprachdetektor

2.1 Begriffsbestimmungen

Als Sprachdetektor wird eine "Einrichtung zur Erkennung des Vorhandenseins von Sprache" bezeichnet, - so auch der Titel einer Studienarbeit von H.v.Zerboni /14/, die der Verfasser dieses Berichtes betreute.

Bei der Aufgabenstellung im Sept. 67 waren Arbeiten BRADY's (1965) /2/, /3/ bekannt. In "A Technique for Investigating ON-Off Patterns of Speech" beschreibt BRADY ein Verfahren, bei dem ein Schwellwertdetektor alle 5 ms abgefragt wird, ob im letzten 5 ms-Intervall die Schwelle überschritten wurde. Solange die Schwelle überschritten ist, entsteht eine Folge von Pulsen mit einer Folgefrequenz von $1/(5 \cdot 10^{-3}) = 200$ Hz. BRADY definiert nun als Abschnitt (spurt) die Länge eines Pulszuges und als Lücke (gap) die Pause zwischen zwei Pulszügen. Dem wird gegenübergestellt der Sprechabschnitt (talkspurt) als Zeit, die ein Zuhörer als zusammenhängende Äußerung empfindet - während derer er normalerweise nicht unterbrechen würde - und als Sprechpausen (pause) die Zeit, in der ein Zuhörer zu sprechen beginnen könnte, ohne dem Partner ins Wort zu fallen. - Obgleich die beiden letzten Definitionen nicht sehr scharf sind, zeigen sie, in welcher Weise die Abschnitt-Lücken-Folge noch korrigiert werden muß, um daraus eine Sprech-



- a) Sprechwechselspannung
- b) Hüllkurve und Triggerschwelle
- c) Ausgang des Triggers
- d) Störunterdrückung
- e) Silbenpausenunterdrückung
- f) Korrektur der Abschnittslängen

Bild 1
Stufen der Verarbeitung im Sprachdetektor

abschnitt-Sprechpausenfolge zu machen: Pulszüge unterhalb einer Mindestlänge von 2 Pulsen werden entfernt (throwaway-time). Damit werden Knackstörungen der Länge $t < 5$ bis 15 ms (je nach Phasenlage) beseitigt. In natürlicher Sprache kommen so kurze Abschnitte nicht vor.

Lücken unterhalb 40 Pulsen entsprechend ≈ 200 ms Dauer werden rechnerisch aufgefüllt (fill-in time). Damit wird berücksichtigt, daß die Lücken bei Verschlußlauten, kleinen Hemmungen und kurzen Atempausen sicher nicht zu den Sprechpausen (pause) im obigen Sinne gehören.

Unter Berücksichtigung dieser Begriffsbestimmungen wurde die Studienaufgabe gestellt, einen Sprachdetektor mit Bauteilen der Analogrechentechnik zu entwerfen und aufzubauen.

2.2 Aufgabe und Prinzip

Es ergibt sich folgendes Prinzip für die Verarbeitung der Sprechwechselspannung im Sprachdetektor: (Bild 1) Es wird eine "Hüllkurve" (b) der Sprechwechselspannung (a) gebildet. Es ergeben sich (c) die "Abschnitte" und "Lücken" (spurts und gaps), in denen die "Hüllkurve" eine Triggerschwelle über- bzw. unterschreitet. Die Störunterdrückung wird mit einem ansprechverzögerten Trigger ($t_1 < 5$ ms) erreicht (d). Diese kurze Verzögerungszeit ist ausreichend zur Störunterdrückung. Sie empfiehlt sich aus einem später zu erläuternden Grund. Die Silbepausenunterdrückung wird mit einem abfallverzögerten Trigger $t_2 \approx 200$ ms erreicht (e). Ein ansprechverzögerter Trigger mit $t_3 = t_2 - t_1$ gleich 195 ms stellt die ursprüngliche Länge von Abschnitten (spurts), die > 5 ms sind und einen Abstand von mehr als 200 ms von den Nachbarabschnitten haben, wieder her (f).

2.3 Variable Ansprechschwelle

Ein wichtiger Gesichtspunkt für den Entwurf dieses Detektors ist es, die Abhängigkeit der Abschnitt-Lücken-Vertei-

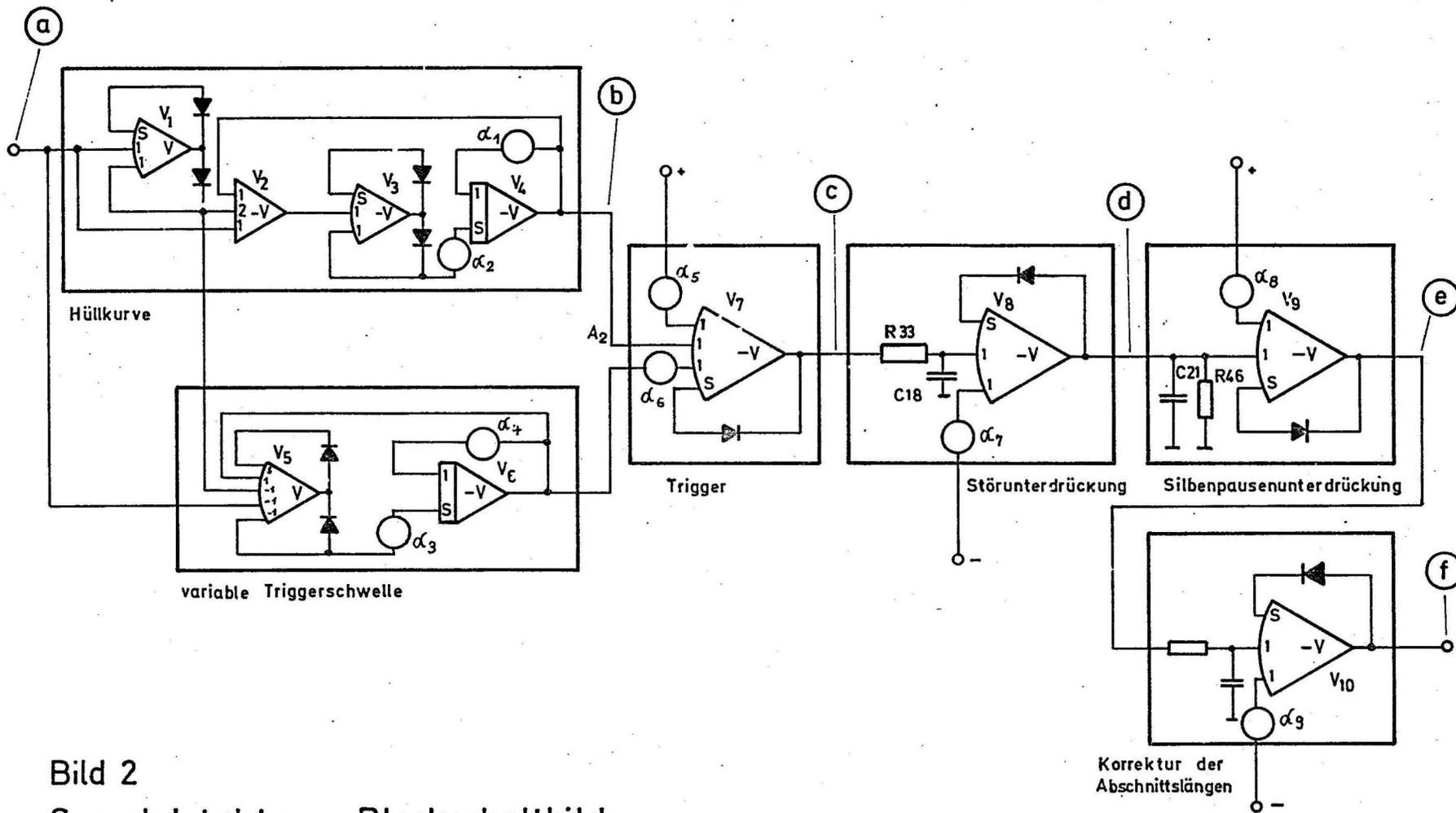


Bild 2
Sprachdetektor Blockschaltbild

lung von der Triggerschwelle zu verringern. Untersuchungen von BRADY an einem Sprachdetektor ergaben eine Verdreifachung der Zahl der Lücken bei Heraufsetzen der Schwelle von -44 dbm auf -32 dbm (/2/, S.10). Der Verfasser schlug deshalb bei der Stellung der Studienaufgabe vor, die Ansprechschwelle der Sprechlautstärke nachzuführen, d.h. aus der "Hüllkurve" durch Integration mit geeigneter Abklingzeitkonstanten als Regelgröße eine variable Ansprechschwelle zu gewinnen. Damit der Sprachdetektor nicht auf Rauschen anspricht, ist eine untere Grenze für die Schwelle vorgesehen.

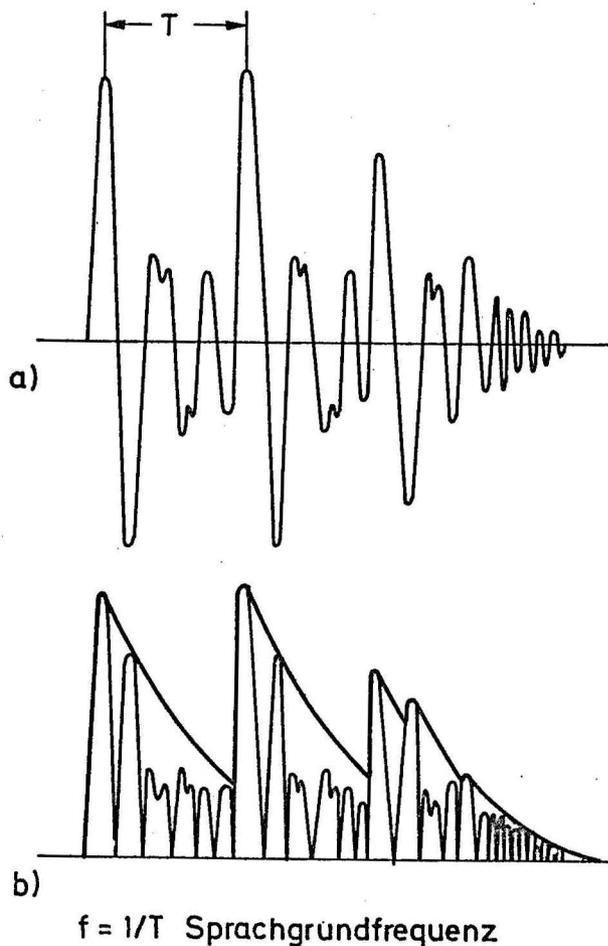
2.4 Blockschaltbild

Im Bild 2 ist das Blockschaltbild des Sprachdetektors angegeben. Die verschiedenen in Bild 1 schematisch angegebenen Spannungen können an den in Bild 2 entsprechend gekennzeichneten Punkten abgegriffen werden.

2.4.1 "Hüllkurven"-Erzeugung

Die Schaltung zur Erzeugung der "Hüllkurve" ist GILOI-LAUBER /5/ S. 59, 80 entnommen. Ihre Arbeitsweise wird jedoch dort nicht beschrieben. Eine ausführliche Beschreibung nimmt v. ZERBONI /14/ S. 30 f. vor. Die Schaltung wirkt im wesentlichen so, wie ein Kondensator, der über einen Gleichrichter von einer Spannungsquelle mit $R_1 = 0$ rasch aufgeladen und über einen Widerstand langsam entladen wird.

Man sollte daher anstatt von einer Schaltung zur Erzeugung der Hüllkurve /5/ S. 59, 80 besser von einem Spitzenwertfolger mit exponentieller Alterung sprechen. Der Ausdruck "Hüllkurve" für den im Bild 3b) gezeigten sägezahnartigen Verlauf ist jedoch bequem. Es sei daher gestattet, ihn weiter zu verwenden.

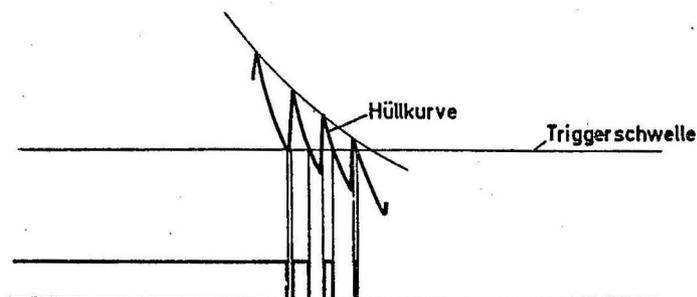


Gleichrichtung und Hüllkurvenbildung
Bild 3

In Bild 3a) ist eine Sprachwechselspannung zu erkennen (nach STEVENS /11/ S. 1059). In b) ist sie gleichgerichtet.

Die überlagerte "Hüllkurve" ist zu erkennen. Die Entladung wird so eingestellt, daß die "Hüllkurve" auch auf Null angekommen ist, wenn der Laut ausgeklungen ist. So erhält man eine Spannung, deren langsam veränderlicher Anteil überlagert ist mit Bögen einer Exponentialfunktion. Der Vorteil der Anwendung einer solchen Schaltung gegenüber derjenigen von Filtern

zur Abtrennung des langsam veränderlichen Anteils besteht im Fortfall einer Einschwingzeit, der Nachteil in der Welligkeit der Ausgangsspannung. Aus dem Folgenden wird klar werden, weshalb die "gedachte" Kurve entlang der oberen Spitzen des Sägezahns zu führen ist.



Abschnitts- und Pausenverteilung infolge welliger Hüllkurve bei schleifendem Schnitt
Bild 4

Die "wellige Hüllkurve" bringt es mit sich, daß bei schleifendem Schnitt mit der Triggerschwelle der Trigger mehrmals anspricht und abfällt (Bild 4). Aus Bild 3 ist bekannt, daß die Grundfrequenz der überlagerten "Sägezahnspannung" gleich der Sprachgrundfrequenz ist.

2.4.2 Störunterdrückung

Rechnet man bei Männern mit ca 100 Hz Sprachgrundfrequenz entsprechend 10 ms Periodendauer, so würden bei einer auf 10 ms eingestellten Störunterdrückung alle von der Welligkeit der "Hüllkurve" herrührenden Rechtecke des Triggers weggeschnitten. Die gedachte Kurve, die zum Schnitt mit der Triggerschwelle gebracht wird, verläuft also entlang den unteren Spitzen des Sägezahns. Umgekehrt würden ohne Störunterdrückung alle Rechtecke erhalten bleiben und durch die nachträgliche Silbenpausenunterdrückung miteinander verbunden werden, so daß die Schaltung so wirken würde, als ob die gedachte Kurve entlang der oberen Spitzen des Sägezahns verlief. Eine Störunterdrückung in der Größenordnung von 5 ms wirkt so, als ob die gedachte Kurve in der Mitte zwischen oberen und unteren Spitzen verlief.

2.4.3 Frequenzabhängigkeit

Die Lage der gedachten "Hüllkurve" ist unerwünschterweise von der Sprachgrundfrequenz abhängig. Da nur männliche Sprecher codiert wurden, fiel der Mangel wenig ins Gewicht. Er könnte jedoch behoben werden, indem vor die Störunterdrückung ein weiterer monostabiler Multivibrator zur Unterdrückung von Pausen ($< 17 \text{ ms}^*$) geschaltet würde. Dadurch werden die aufgrund der welligen "Hüllkurve" abgetrennten Abschnitte wieder angefügt, was für den Schnittpunkt der "Hüllkurve" mit der Schwelle einem gedachten Hüllkurvenverlauf entlang der oberen Spitzen des Sägezahnes (Bild 4) entspricht. Bei höherer Sprachgrundfrequenz bleiben (bei sonst gleicher Kurvenform und Entladungszeitkonstanten) die oberen Spitzen erhalten, während die unteren näher heranrücken.

Das spricht dafür, sich grundsätzlich nur auf die oberen Spitzen der welligen Hüllkurve zu beziehen. Mit der vorgeschlagenen Ergänzung der Schaltung wäre es zu erreichen. In der Wahl der Zeit für die anschließende Störunterdrückung ist man dann frei.

2.4.4 "Schwellenkurve"

Die "Schwellenkurve" wird durch einen weiteren Spitzenwertfolger mit langsamerer Alterung als bei der "Hüllkurve" gewonnen. Sie soll während der Sprechabschnitte dem Pegel - repräsentiert durch die Spitzen^{**}) der "Hüllkurve" - folgen und in Sprechpausen wieder auf den Ausgangswert zurückgehen. Die Schaltung zur Erzeugung der Schwellenspannung ist in /14/ S. 32 f. beschrieben. Sie stellt eine interessante Abwandlung der Hüllkurvenschaltung durch v.ZERBONI dar.

*) Pausen $\geq 17 \text{ ms}$ werden zugelassen, da sie nicht mehr von der Welligkeit der "Hüllkurve" herrühren können, sofern man als niedrigste Sprachgrundfrequenz 60 Hz annimmt.

***) In /3/ weist BRADY nach, daß der a.p.l. (average peak level) geeignet zur Charakterisierung der Lautstärke von der Sprache ist und anstelle der VU (volume unit) verwendet werden sollte.

Die "Schwellenkurve" und die "Hüllkurve" werden dem in Bild 2 als Trigger bezeichneten Komparator zugeführt (im Schaltplan Bild 5 b Karte 5, S. 13). Damit der Trigger nicht beim Störpegel anspricht, wird eine untere Schwelle fest eingestellt (α_5 Bild 2 ; R 34, Karte 5, Bild 5 b).

Die "Schwellenkurve" folgt - wie schon gesagt - während der Silben den Maxima der "Hüllkurve" und fällt während einer Pause im Redefluß exponentiell auf den voreingestellten Wert ab. Damit wird der Einfluß der Sprachlautstärke ausgeglichen. Da die "Schwellenkurve" jedoch träge abfällt, bringt leise Sprache, die auf laute Sprache folgt, den Trigger nicht zum Ansprechen. Beim Ohr ist eine Heraussetzung der Schwelle ebenfalls bekannt, wenn auch nur für einige 10 - 100 ms nach Geräuschbelastung. (STEVENS /11/ S. 101)

Den günstigen Abstand der variablen Schwelle von der "Hüllkurve" wählt man nach Oszillogramm (Bild 6). Er wird am Regler α_6 eingestellt, und zwar bei Sinusspannung am Eingang des Detektors so, daß die Spannung hinter α_6 um 20 db kleiner ist als an A_2 (Bild 5 b Spannung vor R 32 gegenüber Spannung vor R 36).

Die in der Aufgabenstellung geforderte zeitliche Voreinstellung der Triggerschwelle gegenüber der Sprechwechselspannung wird erreicht, indem die Aufladezeitkonstante für den Spitzenwertfolger der "Hüllkurve (in V_4 an α_2) etwas größer gewählt wird als diejenige für den Spitzenwertfolger der Triggerschwelle (in V_6 an α_3). Von ZERBONI /14/ (S.35) führt den Begriff Aufladegeschwindigkeit ein und zeigt, daß bei gleicher Aufladegeschwindigkeit (Flankensteilheit der Integralkurve) die Integrationszeitkonstante von V_6 um so größer gewählt werden muß, je größer bei konstanter Eingangswchselspannung der Abstand - a [dbm] der "Hüllkurve" ist.

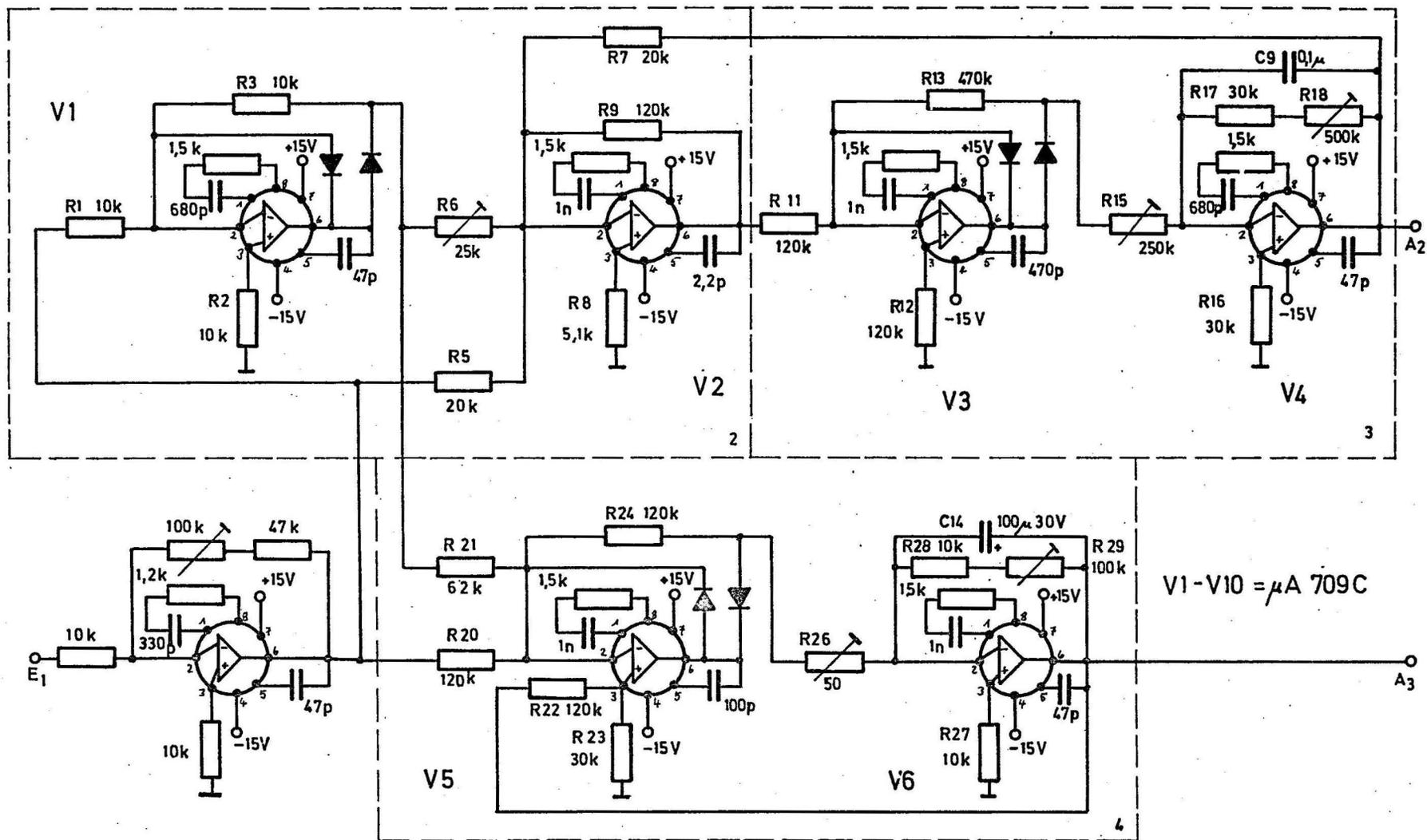


Bild 5a Sprachdetektor

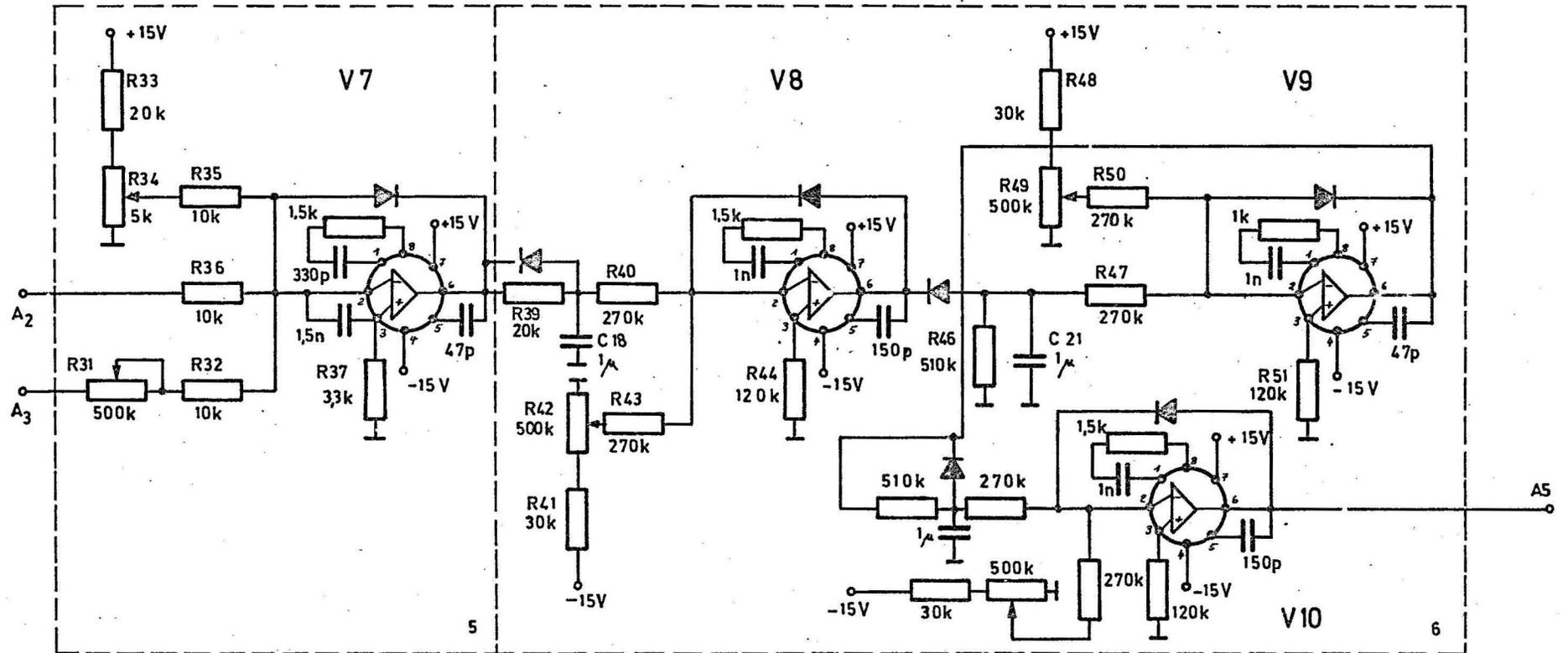
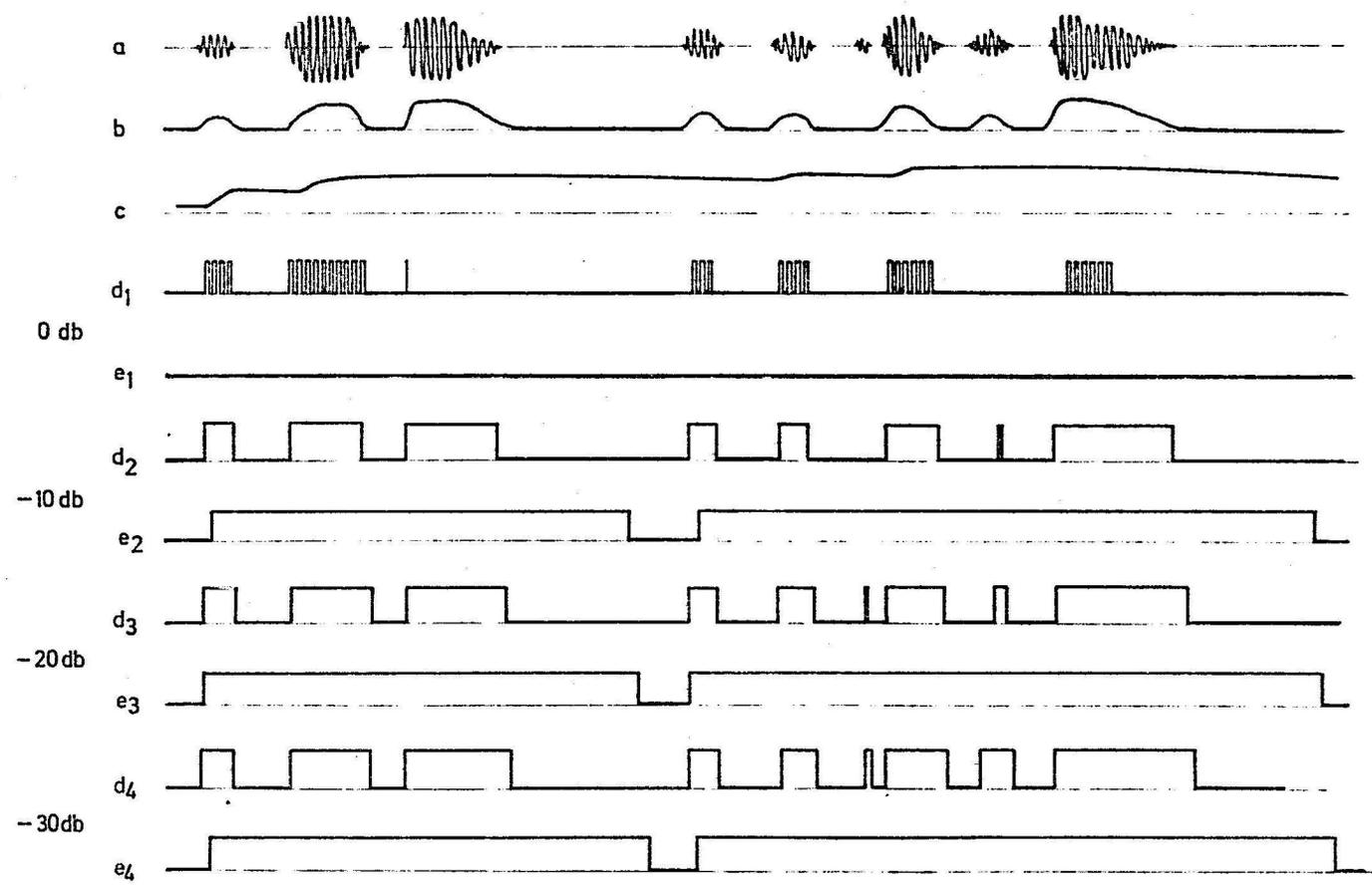
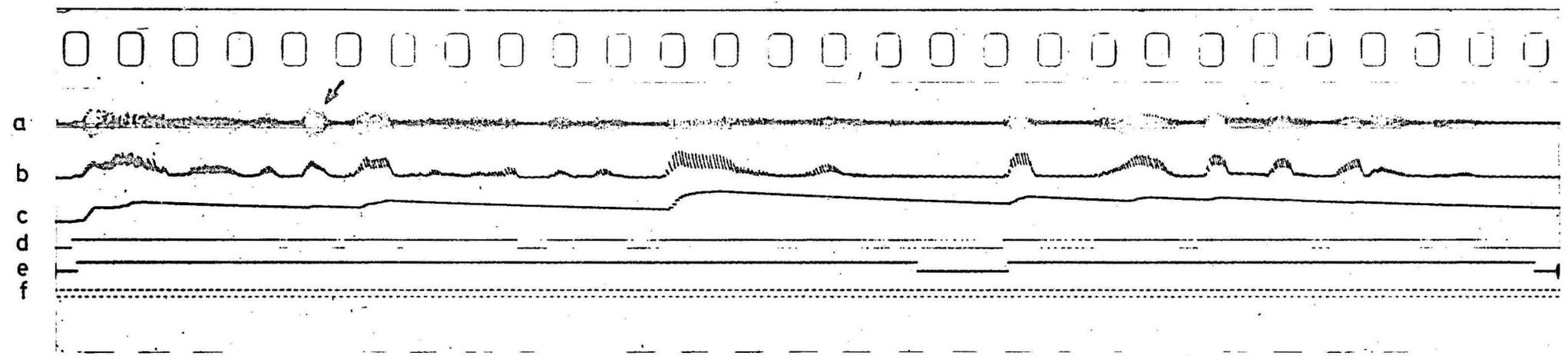
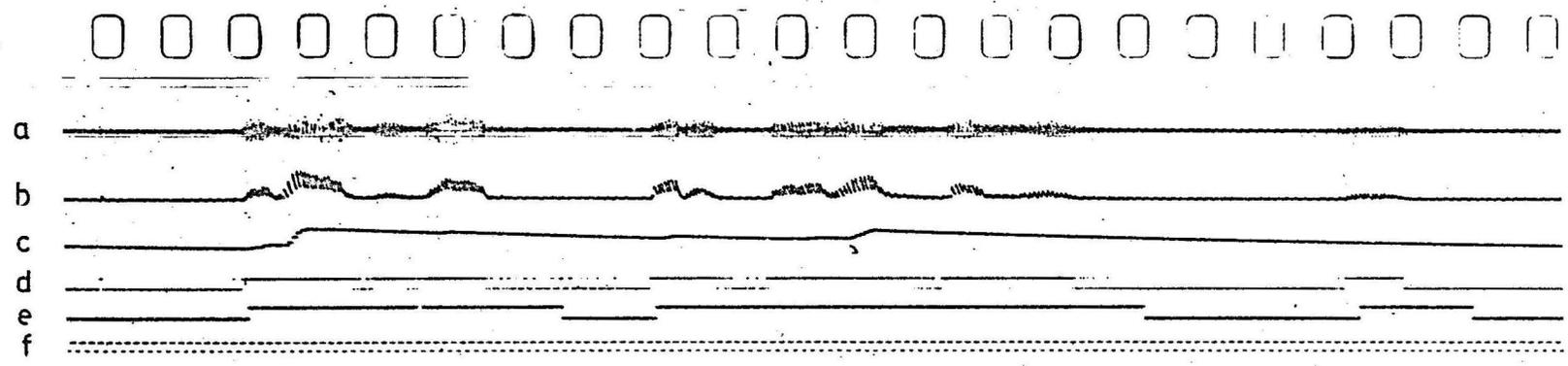


Bild 5b Sprachdetektor



a) Sprechwechselspannung c) Schwellenkurve e) Ausgang des Detektors ohne
 b) Hüllkurve d) Ausgang des Triggers Korrektur der Abschnittslängen

Bild6: Spannungen an verschiedenen Punkten des Sprachdetektors. Gezeichnet nach Lichtstrahloszillogramm auf Blauschriftpapier. Verschiedene Abstände der Triggerschwelle von Hüllkurve, gemessen bei Dauersignal am Eingang.



- a) Sprechwechselspannung
- b) Hüllkurve
- c) Schwellenkurve
- d) Ausgang des Triggers
- e) Ausgang des Detektors ohne Korrektur der Abschnittslängen
- f) 50Hz Rechteck, $6\text{mm} \approx 5T = 100\text{ms}$

Bild7: Spannung an verschiedenen Punkten des Sprachtriggers, aufgenommen mit Lauffilmkamera an Elektronenstrahloszillograph. Laufgeschwindigkeit 3cm/sec, Vergrößerung 2:1.

2.5 Überprüfung der Wirkungsweise

Bild 7 zeigt mit Lauffilmkamera aufgenommene Oszillogramme vom Arbeiten des Triggers. Das Korrekturglied für die Abschnittslängen (Bild 2 S.6 V 10) wurde nachträglich gebaut. Im Oszillogramm ist die Korrektur daher nicht enthalten. Man sieht, soweit die Sprechwechselspannung (a) erkennbar ist, z.B. im unteren Teil v. Bild 6 (Pfeilmarkierung), die Übereinstimmung des oberen Teils der Einhüllenden der Sprechwechselspannung mit den geometrischen Orten für die Spitzen der "Hüllkurve". (b) Von der "Hüllkurve" sind an Stellen höherer Intensität nur die abfallenden Flanken (Bild 3 b S.8) zu erkennen, an den steilen Anstiegsflanken hat der Elektronenstrahl eine so hohe Geschwindigkeit, daß der Film nicht geschwärzt wird. Aus der "Hüllkurve" kann man die Sprachgrundfrequenz des aufgenommenen Sprechers ablesen. Sie beträgt ≈ 120 Hz.

In Kurve (d) (Bild 7) ist das wiederholte Ansprechen des Triggers (wie an Bild 4 erläutert) zu erkennen. In Kurve (e) sieht man die Wirksamkeit der Störunterdrückung (Ansprechverzögerung) und der Silbenpausenunterdrückung (Überhang).

Tabelle I zeigt einen kurzen Vergleich einer Abschnittsbestimmung durch Ausmessen der Sprechabschnitte und -pausen im Oszillogramm mit einer Auswertung durch den Sprachdetektor (Zeiteinheit 10 ms).

Ausführlichere Informationen enthalten die Verteilungsfunktionen Bild 22 bis Bild 35, die einmal unter Verwendung von Detektor und Sprachdodierer, zum anderen unter Verwendung handausgewerteter Oszillogramme erstellt wurden.

Dauer des Sprechabschnittes		Dauer der Sprechpause	
nach Oszill.	nach Detekt.	nach Oszill.	nach Detekt.
54	54	23	40
56	24	65	89
282	272	27	34
65	55	12	28
348	337	48	53
236	210	47	70
51	40	33	45
120	110	30	39
172	166	25	29
216	214	18	23
25	13	25	31
154	132	8	29
26	17	27	32
108	106	14	20
114	90	37	54
154	146	4	25
26	11	156	159
12	5	48	55
88	75	972	985
98	76	39	66
156	146	33	34
284	285		

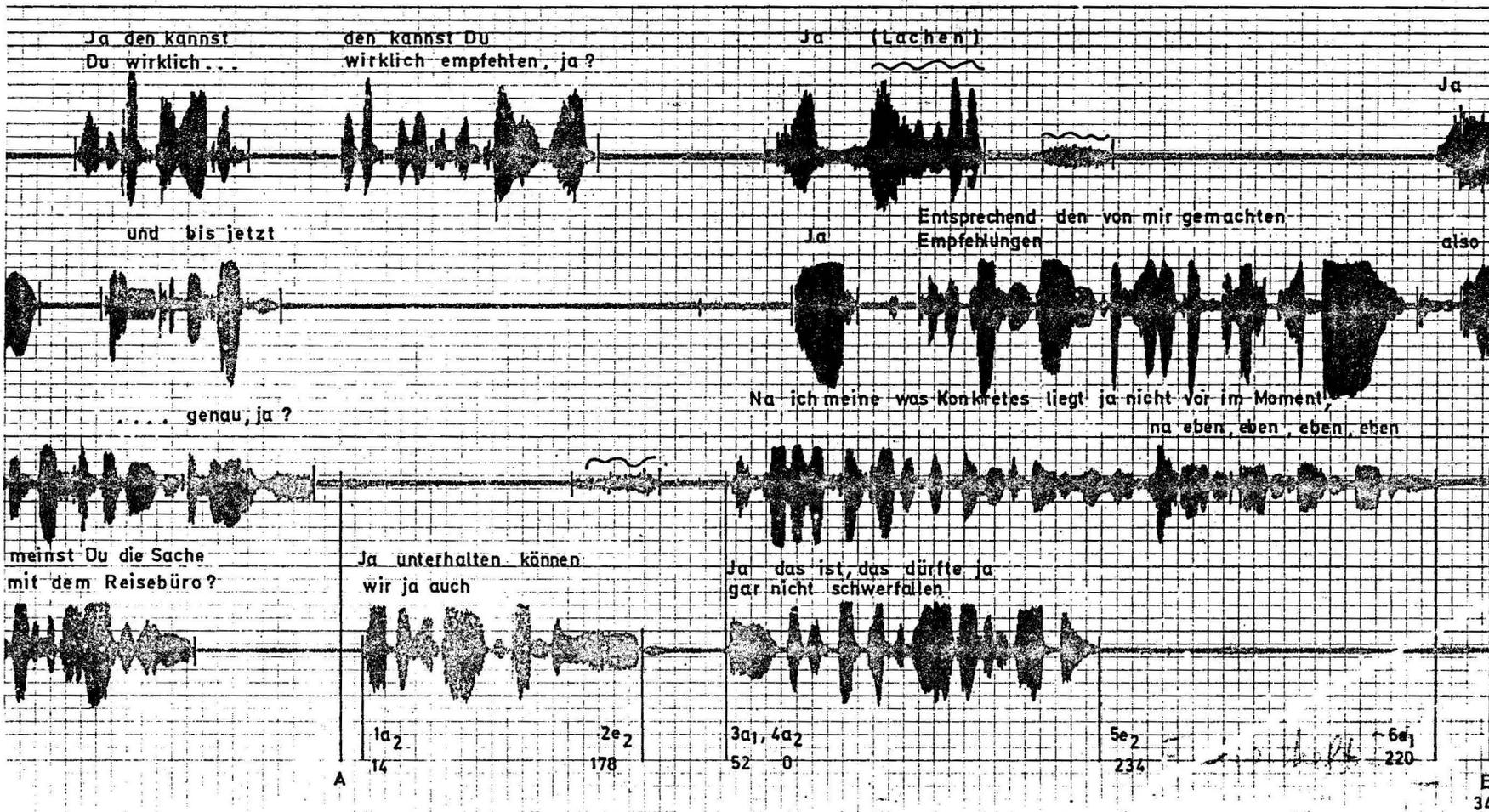
Tab. I

Vergleich von Oszillogramm und Ausgangsspannung
des Sprachdetektors

(Gespräch 3/49 Sprecher 2, Zeiteinheit 10 ms)

Man sieht bereits an dieser kurzen Tabelle, daß der Detektor zwar die gleichen Sprechabschnitte und Pausen "erkennt" hat wie sie die Handauswertung ergab, daß aber die Sprechabschnitte kürzer und die Pausen länger ausfallen. Dies ist z.T. durch das Prinzip bedingt. So wird beim Detektor wegen der Silbepausenunterdrückung keine Pause unter 200 ms = 20 Einheiten ausfallen, während man bei der Handauswertung durchaus solche Pausen findet. (In Tab. I z.B. die Pausen 80 ms und 40 ms) Außerdem kann der menschliche Auswerter auch dort noch Sprache im Oszillogramm erkennen, wo sie ins Rauschen "eintaucht", während sich der Detektor aufgrund der Voreinstellung der Schwelle dann für eine Pause entscheidet.

Es wird sich jedoch durch Einbau der erwähnten weiteren Baugruppe zur Überbrückung der Pausen infolge welliger Hüllkurven und durch noch sorgfältigere Einstellung aller Schwellen und Zeiten eine noch weitergehende Übereinstimmung der automatisch und manuell ermittelten Sprachabschnitte und -pausen erreichen lassen. Dazu wären gegebenenfalls weitere Untersuchungen erforderlich.



Rastereinheit 100 ms. A, E Anfang und Ende des codierten Abschnittes.
a, e Anfang und Ende eines Sprechabschnittes.

Indices 1, 2 für Sprecher 1, 2.

Vorzahl 1-6 fortlaufende Nummer des Ereignisses

Zahl unter jedem Ereignis bezeichnet Abstand vom Vorhergehenden in 10 ms Einheiten

Bild 8 Schleifenzillogramm von zwei Zwiesgesprächen

3. Verschlüsselung der Sprechereignisse

3.1 Kennzahlen für Beginn, Ende und Abstand

Eine zweckmäßige Verschlüsselung des Zeitverhaltens der Sprachenergie zweier Sprecher besteht in der Einführung des Ereignisbegriffes für jede Änderung des Zustandes auf einer der Leitungen. Dabei kann man auch aus dem Kontext erkenntliche Ursachen der Zustandsänderungen in die Verschlüsselung mit einbeziehen und damit den Übergang zur Untersuchung semantischer Information vollziehen.

So wurde bei der manuellen Verschlüsselung mitberücksichtigt, ob der Sprecher mit einer Frage abschließt oder mit einer Antwort beginnt, ob er unartikulierte Laute von sich gibt usw. Da die Ergebnisse mit der automatischen Auswertung verglichen werden sollen, wurde diese feinere Unterscheidung jedoch nicht ausgenutzt. Sie soll auch nicht beschrieben werden.

Als Ereignisse gelten also der Sprechbeginn und Sprechende der Sprecher 1 und 2:

- a₁ Sprecher 1 beginnt zu sprechen
- e₁ Sprecher 1 beendet sein Sprechen
- a₂ Sprecher 2 beginnt zu sprechen
- e₂ Sprecher 2 beendet sein Sprechen

Anfang A und Ende des verschlüsselten Abschnittes sind durch Sonderzeichen gekennzeichnet. (Anstelle der mnemotechnisch günstigen Buchstaben werden im Rechner Integerzahlen verwendet). Der Abstand vom vorhergehenden Ereignis in 10 ms-Einheiten wird dem Ereignis nachgestellt. Die Verschlüsselung des in Bild 8 vorliegenden Abschnittes lautet also A 1a₂ 14, 2e₂ 178, 3a₁ 52, 4a₂ 0, 5e₂ 234, 6e₁ 220, E 34.

Die hier vorgetzten laufenden Nummern dienen der späteren Erläuterung eines Sortierprogrammes.

Die beschriebene Verschlüsselung ist zweckmäßig:

1. Sie läßt sich einfach ausführen; insbesondere ist es wesentlich einfacher, kleine Abstände zwischen zwei benachbarten Ereignissen auszumessen, als den Abstand eines jeden Ereignisses vom Gesprächsanfang zu ermitteln.
2. Die Abstandsangaben wachsen nicht mit wachsender Gesprächslänge, so daß sich keine Probleme beim Speichern im Rechner ergeben. Bei der Weiterverarbeitung der Abstandsangaben anstelle der fortlaufenden Zeit ergeben sich keine Nachteile.

3.2 Ermittlung der Zeitabschnitte

Hat man für jedes Ereignis zwei Integerzahlen gespeichert, nämlich Ereignisart und Abstand vom vorhergehenden Ereignis, dann lassen sich durch logische Verknüpfungen die interessierenden Zeitabschnitte ermitteln. Werden z.B. die Zeiten des Sprechens von Teilnehmer 1 gewünscht, so stellt man im Programm die Frage nach dem Kennzeichen a_1 und summiert von seinem Auftreten an alle Abstandsangaben bis einschließlich der zum nächsten Zeichen e_1 gehörigen. Die Zeiten des Gleichzeitigsprechens werden ermittelt, indem man nach zwei aufeinanderfolgenden Kennzeichen $a_1 a_2$ oder $a_2 a_1$ sucht und die hinter dem nun folgenden Ereignis - notwendigerweise ein e_1 oder e_2 - stehende Abstandsangabe als Dauer eines Abschnittes von Gleichzeitigsprechen entnimmt. Die Anzahl und die Verteilungsfunktion der so gewonnenen Abschnitte werden berechnet.

3.3 Manuelle Verschlüsselung

Die manuelle Verschlüsselung des Zwiegespräches geht aus von einem Schleifenoszillogramm der Sprechwechselspannungen. In Bild 8 ist ein Ausschnitt aus zwei laufenden Zwiegesprächen zu sehen. Durch Vergleich des Oszillogramms mit Bandaufnahmen der Gespräche gelingt es, die Bedeutung der oszillographierten Sprechabschnitte zu ermitteln. Die ge-

naue Zuordnung einzelner Silben zu Abschnitten des Oszillogramms kann mit Hilfe eines Speicheroszillographen vorgenommen werden. Bei der vorliegenden Auswertung wird sie nicht benötigt.

3.4 Automatische Verschlüsselung

Die Einrichtung zur Sprachverschlüsselung sollte die binären Signale am Ausgang des jedem Sprecher zugeordneten Sprachdetektors so umformen und speichern, daß nachträglich alle gewünschten Sprechabschnitte wie Pausen, Gleichzeitigsprechen usw. nach Anzahl und Verteilungsfunktion daraus zu gewinnen sind. Die für die manuelle Verschlüsselung verwendete Bezeichnung der "Ereignisart" - wie Beginnen und Enden der Sprechaktivität von Teilnehmer 1 oder 2 - und des "Abstandes" eines Ereignisses vom vorhergehenden hatte sich als zweckmäßig erwiesen. Es waren außerdem alle die weitere statistische Verarbeitung betreffenden Programme für diese Art der Verschlüsselung bereits erstellt. Deshalb sollte als Endergebnis der automatischen Verschlüsselung des Vorhandenseins oder Fehlens der Sprachenergie im System beider Sprecher ebenfalls die Darstellung durch "Ereignisart" und "Abstand" gewählt werden.

Zur Lösung dieser Aufgabe bieten sich zwei verschiedene Verfahren an:

1. Die direkte Verschlüsselung des "Ereignisses" - gewonnen durch eine Logikschaltung aus den Potentialen der Sprachdetektoren - und des "Abstandes" vom vorhergehenden Ereignis, der mit einem rasch rückstellbaren Zähler gemessen werden kann. Beide Schlüsselzahlen werden von einem schnellen Locher übernommen.
2. Laufende Abtastung der beiden Sprachdetektorausgänge im 2,5; 5 oder 10 ms-Takt und fortlaufende Ablochung der sich ergebenden binären Information auf je vier Spuren eines 8-Kanallochstreifens, so daß die Abschnitte des Spre-

chens oder Nichtsprechens als gelochte oder nicht gelochte Strecken des Streifens erscheinen; daran anschließend mit einem Digitalrechner Umschlüsselung in die gewünschte Darstellung.

3.4.1 Direkte Verschlüsselung

Die "Ereignisse" werden durch die Flanken des Ausgangspotentials an den Sprachdetektoren bestimmt. Ist dieses während des Sprechens positiv, so bedeutet z.B. eine positive Flanke am Ausgang von Detektor 1 "Teilnehmer 1 beginnt zu sprechen" = a_1 , eine negative am Ausgang von Detektor 2 "Teilnehmer 2 hört mit Sprechen auf" = e_2 usw.

Bei jedem Ereignis wird der Zählerinhalt in einen Pufferspeicher gegeben, der Zähler selbst wird rückgestellt und neu gestartet. Während der neuen Zählperiode wird die Schlüsselzahl des soeben eingetretenen "Ereignisses" und der Inhalt des Pufferspeichers, nämlich der "Abstand" vom vorhergehenden auf einen Lochstreifen übertragen. Der Vorteil dieses Verfahrens liegt in seiner Einfachheit, der Nachteil darin, daß es vorkommen kann, daß der zeitliche Abstand zwischen zwei Ereignissen kürzer ist als die Zeit, die zum Ablochen der Kennziffer und der u.U. mehrstelligen Abstandzahl des vorhergehenden Ereignisses benötigt wird.

Wird z.B. zur Abstandsmessung ein Zähltakt von 10 ms verwendet und folgt nun nach einem Abstand t_1 zwischen zwei Ereignissen von $t_1 \geq 100$ ms ein solcher von $10 \leq t_2 < 20$ ms Dauer, wie es z.B. beim fast gleichzeitigen Einsetzen beider Sprecher nach einer längeren Pause leicht eintreten kann, so benötigt ein Locher zum Abarbeiten der mindestens zweistelligen Zahl aus dem Pufferspeicher mit 100 Z/s mindestens 20 ms, nach $t_2 < 20$ ms wird jedoch der neue Abstand - eine 10 ms-Einheit - in den Pufferspeicher geschoben und die dort vorhandene Information während des Auslesens zerstört. Man kann allerdings die Wahrscheinlichkeit, dieser Art Fehler beliebig klein zu machen, wenn man mehrstufige Pufferspeicher mit einer genügenden Stufenanzahl vorsieht, so-

lange die "Ereignisrate" kleiner ist als die Lochergeschwindigkeit Z/sec dividiert durch die mittlere Zahl der pro Ereignis abzulochenden Ziffern. Nach dem heutigen Stand der Technik wäre dieses Verfahren vorzuziehen.

3.4.2 Abtastung der Trigger, direkte Ablochung und anschließende Umschlüsselung

Zur praktischen Verwirklichung wurde dieser Weg als Diplomaufgabe gestellt. Die Aufgabe lautete, die an den Ausgängen der Sprachdetektoren anliegenden Potentiale mit einem 5- oder 10ms-Takt abzutasten und die sich ergebende binäre Information in einen 8-Kanal-Lochstreifen zu überführen.

Jedem Sprachkanal waren vier Lochreihen zuzuordnen, auf denen bei angesprochenem Detektor eine fortlaufende Lochung erscheinen sollte. Die Aufgabe wurde von BÖTTCHER 1968 gelöst /1/.

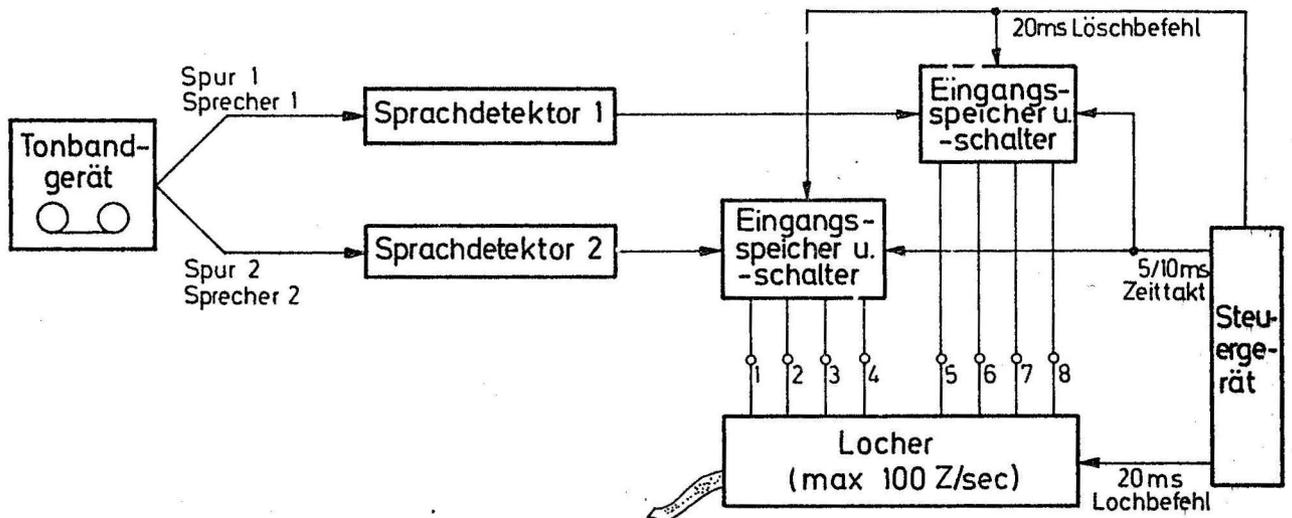


Bild 9
Sprachcodiereinrichtung Blockschaltbild

Das Blockschaltbild der Sprachverschlüsselungseinrichtung zeigt Bild 9. Die Lochergrundfrequenz wurde auf 50 Z/sec

festgesetzt. Nach jedem Lochertakt werden die Eingangsspeicher zurückgesetzt. In den folgenden 20 ms bis zum nächsten Lochertakt werden diejenigen der 4 Bit der Eingangsspeicher nacheinander gesetzt, für die im Augenblick eines Zeittaktes ein Signal vom Sprachdetektor vorlag. Bei 5 ms-Zeittakt und angesprochenem Detektor 1 werden z.B. nacheinander die Speicher 5, 6, 7 und 8 gesetzt, beim 10 ms-Zeittakt und angesprochenem Detektor 2 zunächst Speicher 1 und 2 und dann gleichzeitig Speicher 3 und 4. Beim Ablochen der ganzen Reihe, ausgelöst durch den Lochbefehl, erscheint dort ein Loch, wo vorher ein Speicher gesetzt worden war.

Das paarweise Auftreten von Löchern bei 10 ms-Abtastung wird zur Sicherung gegen Fehler verwendet.

Bild 10 zeigt einen Lochstreifen und das Ergebnis der Umschlüsselung. Man überprüft leicht ihre Richtigkeit: Anfang der Zeitzählung ist die Zeile, in der die Anfangsmarke des Gesamtabschnittes (einzelnes Loch) steht. Vier Zeilen und ein Lochpaar später erfolgt "Beginn des Sprechens von Teilnehmer 1". Es erscheint also als Ereignis die Schlüsselzahl $1 \hat{=} a_1 \hat{=} \text{"Sprecher 1 beginnt zu sprechen"}$ und als Abstand die $9 \hat{=} 90 \text{ ms} \hat{=} 4.20 \text{ ms}$ für vier Zeilen und 10 ms für ein Lochpaar usw.

Das Umschlüsselungsprogramm verwertet die als Dualzahlen gelesenen Lochkombinationen als Zustandskennzahlen. Von den $2^8 = 256$ möglichen Lochkombinationen können im Rahmen der Abtastung einige nicht auftreten, u.a. weil durch Stör- und Silbenpausenunterdrückung bestimmte Mindestabstände zwischen einigen Ereignissen eingehalten werden. Jede Lochkombination wird gedeutet. - Ist sie zugelassen, so beschreibt sie einen Zustand, der mit dem Vorzustand verglichen wird. Hat sich der Zustand geändert, so liegt ein Ereignis vor. Aus der jeweiligen Zustandskennzahl und dem Vorzustand ergibt sich nicht nur die Art des Ereignisses, sondern auch

ein Zeitwert gemäß der Stellung des Ereignisses innerhalb der Zeile. Im Rechnerprogramm wird die dem Zeitwert entsprechende Zahl und die Ereigniskennziffer ermittelt und ausgegeben. Bei nicht zugelassenen Lochkombinationen wird lediglich der Zeitzähler um $20 \text{ ms} \hat{=} 2$ Einheiten weitergesetzt. Gleichzeitig wird das Programm angehalten und eine Fehlermeldung ausgedruckt, nach der der Operateur entscheiden kann, ob weiterzulesen oder abzurechnen ist.

Die Abfrage der gelesenen Zustandskennzahlen auf Zugehörigkeit zur Liste der zugelassenen erfolgt in der Reihenfolge der Wahrscheinlichkeit ihres Auftretens. Damit werden die kleinen Abfrageschleifen am häufigsten durchlaufen.

Eine weitere in /1/ nicht verwirklichte Vereinfachung des Programms hätte darin bestanden, eine Abfrage: "Ist die soeben gelesene Zahl gleich der vorhergehenden?" vorzusehen. Da gleiche Lochkombinationen im Lochstreifen sehr häufig aufeinanderfolgen, hätte dies zu einer erheblichen Zeitersparnis geführt. Es ist dann nämlich pro gelesener Zeile des Lochstreifens nur der Zeitzähler um 2 Einheiten entsprechend 20 ms weiterzuschalten; auf eine Erkennung der Zustandskennzahl und alle weiteren Entscheidungen kann verzichtet werden, da ein "Ereignis" nicht vorliegen kann.

Das Umschlüsselprogramm für eine Abtastung im 5 ms -Zeitakt erwies sich wegen der hohen Zahl von Abfragen für die zur Verfügung stehende Rechenanlage Z 25 als zu umfangreich.

Bei 10-ms -Abtastung sind nur $2^4 = 16$ Kombinationen von Lochpaaren innerhalb der Verschlüsselung der Triggerzustände möglich. Mit den Sprachtriggern der Sprachcodiereinrichtung und dem für 10 ms -Takt geschriebenen Rechnerprogramm wurde die Auswertung der vorhandenen Gespräche einwandfrei durchgeführt.

(Die mit automatischer Auswertung erhaltenen Verteilungsfunktionen sind in Bild 29 bis Bild 35 dargestellt und können mit denjenigen aus manueller Auswertung Bild 22 bis Bild 28 verglichen werden).

4. Algorithmus zur Bestimmung der Ereignisse am "fernen Leitungsende aus am "nahen" gemessenen

Bisher war die Rede von zwei Sprechern, die man sich, da nichts Näheres darüber ausgesagt worden war, so nahe beieinander vorstellt, daß der Zeitbegriff keiner weiteren Erläuterung bedurfte. Diese Verhältnisse ändern sich bei Verbindungen über eine große Entfernung.

4.1 Zustandekommen und Wirkung der Laufzeit

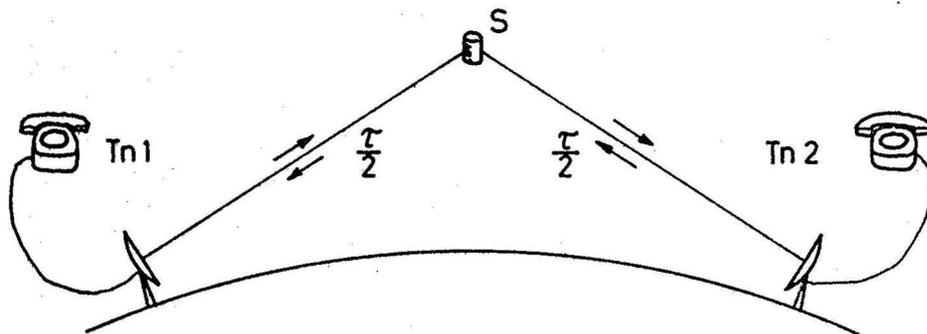


Bild 11: Entstehung der Laufzeit τ bei Übertragung eines Ferngesprächs über einen Satelliten

Bild 11 zeigt das Zustandekommen von Übertragungslaufzeit bei einer Fernsprechverbindung, die über einen Satelliten geführt wird. Die Funkwelle breitet sich mit Lichtgeschwindigkeit aus. Bei einem Synchronsatelliten ergibt sich eine einfache Laufzeit von $\tau = \text{ca } 250 \text{ ms}$; der genaue Wert hängt von der Lage der Bodenstationen und der Länge der Kabelverbindungen zu diesen ab.

Durch die Übertragungslaufzeit kann es nun geschehen, daß beide Fernsprechteilnehmer gleichzeitig zu sprechen beginnen, ohne es zu wissen. Sie sprechen dann zwar gleichzeitig, bemerken dies jedoch zunächst nicht. Es ist daher zweckmäßig, den Zeitbegriff auf einen Ort zu beziehen, etwa den Standort eines Teilnehmers. Die Zustände des

Vorhandenseins oder Fehlens von Sprachenergie und der Zeitpunkt ihres Eintretens werden ebenfalls an einem konkreten Ort gemessen. Die Zeitangabe für das Eintreten der Ereignisse sind nur im Zusammenhang mit der dazugehörigen Ortsangabe sinnvoll.

4.2 Definition der Zeitangaben

Wir definieren unter der Annahme, daß nur die Teilnehmer Tn 1 und Tn 2 anwesend sind und kein Rückhören möglich ist, die Zustände beim Sprechen wie folgt:

Sprechen von Teilnehmer 1

ist das Vorhandensein von Sprechwechselspannung hinter dem Mikrofon von Tn 1;

als (verzögerte) Wirkung davon:

Hören von Teilnehmer 2

ist das Vorhandensein von Sprechwechselspannung am Telefon von Tn 2 .

Sprechen von Teilnehmer 2

ist das Vorhandensein von Sprechwechselspannung am Mikrofon von Tn 2;

als (verzögerte) Wirkung davon:

Hören von Teilnehmer 1

ist das Vorhandensein von Sprechwechselspannung am Telefon von Tn 1 .

Die Begriffe Gleichzeitigsprechen und Gleichzeitigschweigen werden wie folgt festgelegt:

Gleichzeitigsprechen

ist Sprechen eines Teilnehmers, während er gleichzeitig etwas hört.

Gleichzeitigschweigen

ist Schweigen eines Teilnehmers, während er gleichzeitig nichts hört.

Die Ereignisse a_1, a_2, e_1, e_2 bedeuten jetzt:

Am nahen Leitungsende:

Teilnehmer 1 spricht: a_1 beginnt e_1 endet

Teilnehmer 1 hört : a_2 beginnt e_2 endet

am fernen Leitungsende.

Teilnehmer 2 hört : a_1 beginnt e_1 endet

Teilnehmer 2 spricht: a_2 beginnt e_2 endet

Wo keine Verwechslung möglich ist, möge es gestattet sein, die alten Bezeichnungen aus 3.1 S. 20 beizubehalten.

4.3 Unterschied zwischen Verbindung über Satellitenstrecke und über Versuchsanordnung

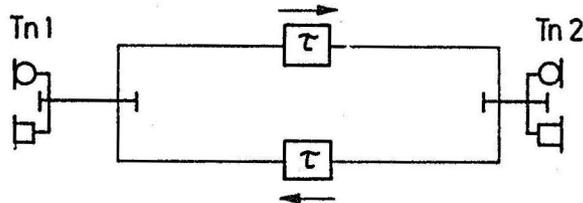


Bild 12: Ersatzschaltbild der Verbindung über Satelliten

Bild 12 zeigt das Ersatzschaltbild einer Verbindung, bei der Hin- und Rückwege über einen Fernmeldesatelliten geführt werden. In der Versuchsanordnung wurde die doppelte Laufzeit in einen Zweig geschaltet. Das ist billiger als zwei Laufzeit erzeugende Einrichtungen in jeden Weg zu schalten und bietet gleichzeitig den Vorteil, in der Verbindung von Tn 2 zu Tn 1 kein Gerät zu haben, das die Qualität der Übertragung durch Dämpfung, Klirren, Brummen usw. abhängig von der eingeschalteten Laufzeitstufe beeinträchtigen könnte, so daß etwaige Unterschiede in der Beurteilung dann möglicherweise gar nicht aufgrund von Laufzeit zustande kämen.

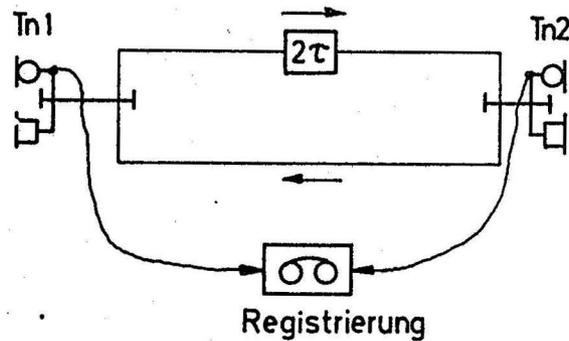
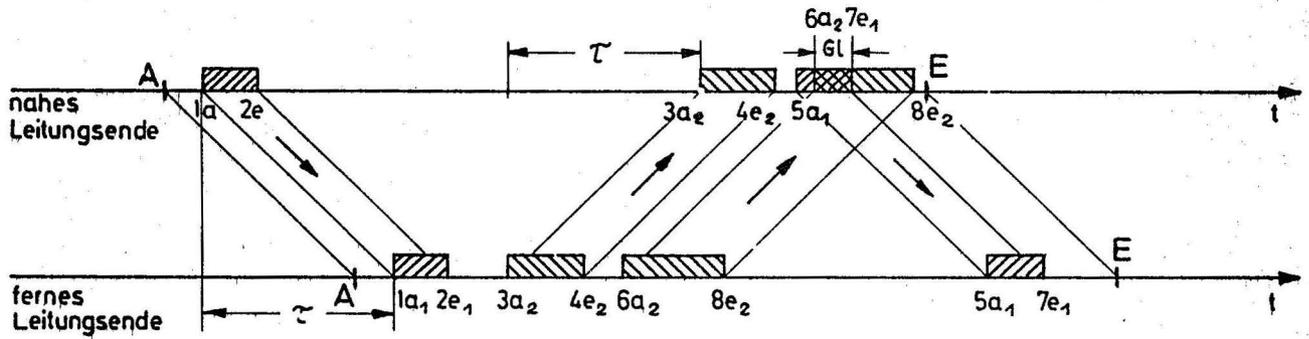
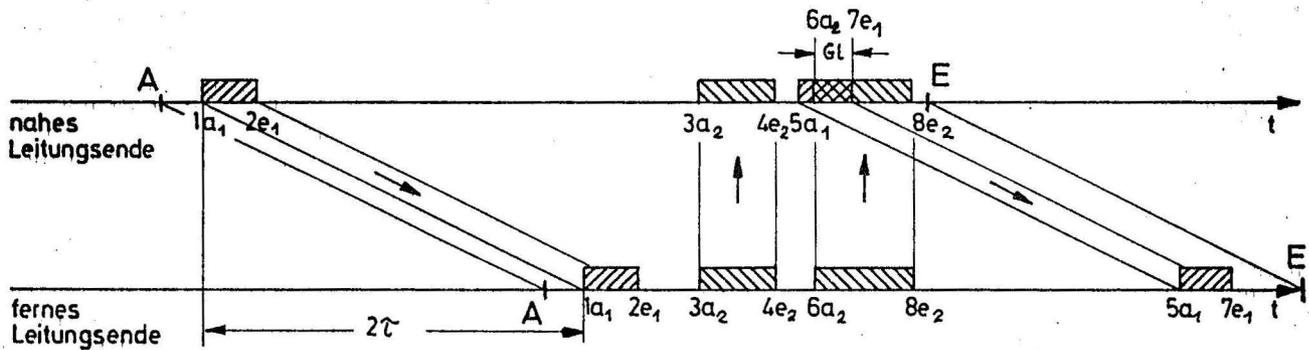


Bild 13: Ersatzschaltbild der Versuchsanordnung

Der Vergleich der am Platz von Tn 1 und Tn 2 abgegebenen Fragebogen läßt einen Rückschluß auf nicht laufzeitbedingte Einflußfaktoren auf das Urteil zu. Die Verbindung Tn 2 zu Tn 1 dient dabei als Referenzverbindung gegenüber derjenigen von Tn 1 zu Tn 2. Diese Untersuchung wurde vom Verfasser vorgenommen. Das Ergebnis ist in /6/ S. 14 zu finden. Im wesentlichen ergab sich: 11,6 % der Versuchsteilnehmer bemerkten Klirren in der Verbindung von Tn 1 zu Tn 2. Der Unterschied zu 4,1 % der Teilnehmer der Gegenrichtung, die Klirren bemerkten, ist signifikant mit $\alpha < 0,01$. Da jedoch kein signifikanter Unterschied zwischen den bei verschiedenen Laufzeitstufen abgegebenen Urteilen gefunden wurde (/7/ S. 114), ist es nicht notwendig, diesem Störeffekt weiter nachzugehen.



a) Verhältnisse auf Satellitenverbindung (Bild 11 u. 12)



b) Verhältnisse auf simulierter Verbindung (Bild 13)

- | | | | |
|--------|-----------------------|----------------------|----------------------------|
| | nahes Leitungsende | fernes Leitungsende | |
| //// | Teilnehmer 1 spricht, | Teilnehmer 2 hört | a_1 beginnt, e_1 endet |
| \\\\\\ | Teilnehmer 1 hört | Teilnehmer 2 spricht | a_2 " , e_2 " |
| Gl | Gleichzeitigsprechen | | |

Bild 14: Sprechabschnitte und -pausen auf Leitungen mit Laufzeitverzögerung.

Bild 14 zeigt Sprechpausen und -abschnitte bei den Teilnehmern, wie sie sich in einer Fernverbindung über einen Satelliten (Bild 11, 12) oder in der Versuchseinrichtung (Bild 13) ergeben könnten.

Es ergeben sich für den stationären Fall am nahen Leitungsende nach beiden Schaltungen die gleichen Abschnitts-Pausenverteilungen (vergl. Bild 14 a mit Bild 14 b für nahes und fernes Leitungsende getrennt).

Für das ferne Leitungsende ist die Verteilung der Abschnitte und Pausen um den Wert τ in der Versuchsverbindung gegenüber der Satellitenverbindung verschoben, was aber ohne Belang bleibt, da den Teilnehmern die absolute Zeit nicht zur Verfügung steht.

Die Abschnitt-Pausenverteilungen am nahen Leitungsende ist jedoch verschieden von derjenigen am fernen. So findet man ein Gleichzeitigsprechen im Beispiel Bild 14 nur am nahen Leitungsende.

Nach Bild 13 wurden die Sprechwechselspannungen an beiden Mikrofonen aufgenommen, das entspricht einer Registrierung am nahen Leitungsende. Es ist nun von Interesse, die Verhältnisse am fernen Leitungsende kennen zu lernen. Sie können rechnerisch aus den gemessenen Werten bestimmt werden.

4.4 Rechnerische Bestimmung der Ereignisse am fernen Leitungsende

Es mögen die Ereignisse von Bild 14 in der auf S. 20 angegebenen Weise verschlüsselt werden. Man nehme an, daß 1 mm einer Zeiteinheit entspreche. Dann ergibt sich für das nahe Leitungsende:

A	1a ₁	2e ₁	3a ₂	4e ₂	5a ₁	6a ₂	7e ₁	8e ₂	E Ereignis
	5	7	58	10	3	2	5	8	2 Abstand

Tab. II a) Ereignisse am nahen Leitungsende

Behält man die alte Reihenfolge der Ereignisse auch am fernen Leitungsende bei, dann ergeben sich aus Bild 14 folgende Abstände:

A	1a ₁	2e ₁	3a ₂	4e ₂	5a ₁	6a ₂	7e ₁	8e ₂	E Ereignis
	5	7	8	10	53	-48	55	-42	52 Abstand

Tab. II b) Ereignisse am fernen Leitungsende

Die Vertauschung der zeitlichen Reihenfolge verschiedener Ereignisse äußert sich hier in den negativen Abständen.

Schreibt man die Ereignisse am fernen Leitungsende in der Reihenfolge ihres Auftretens an, so erhält man aus Bild 14:

A	1a ₁	2e ₁	3a ₂	4e ₂	6a ₂	8e ₂	5a ₁	7e ₁	E Ereignis
	5	7	8	10	5	13	35	7	10 Abstand

Tab. II c) Ereignisse am fernen Leitungsende

Die Ereignisse sind für das nahe Leitungsende fortlaufend nummeriert worden, und diese Nummern sind für das ferne Leitungsende beibehalten worden. Sie treten dort in vertauschter Reihenfolge auf. Das zeigt noch einmal an, daß es durch die Laufzeiteinwirkung zu lokalen Vertauschungen in der Reihenfolge der Ereignisse gekommen ist.

Wenn man die Ereignisse am fernen Leitungsende rechnerisch bestimmen will, so hat man zunächst die durch Laufzeit bewirkte Verschiebung eines Teils der Ereignisse zu betrachten und hinterher die richtige zeitliche Reihenfolge aller Ereignisse herzustellen. Es ist also eine Verschiebe- und ein Sortieralgorithmus zu programmieren.

4.4.1 Verschiebealgorithmus

Ogleich die Verschiebung nach Bild 12 und die nach Bild 13 gemäß Bild 14 a und b zum gleichen Ergebnis führt, unterscheiden sich doch die Rechenschritte in beiden Fällen etwas voneinander, und es werde daher nur noch auf Bild 13 bzw. Bild 14 b bezug genommen.

Die Verschiebung betrifft alle Ereignisse, die dem Teilnehmer 1 zugeordnet sind, die also den Index 1 tragen. Sie sind um $2\tau = 50$ Einheiten zu verschieben. Man kann die Anfangs- (A) und Endmarke (E) mitverschieben, wie es hier getan wurde, und betrachtet sie dann als zu Teilnehmer 1 gehörig. Man könnte sie auch als zu Teilnehmer 2 gehörend betrachten und müßte sie dann festhalten.

Um die Ereignisse mit dem Index 1 um $2\tau = 50$ Einheiten zu verschieben, könnte man den zu jedem Ereignis gehörenden Zeitwert auf der t-Achse ermitteln und zu allen Zeitwerten mit dem Index 1 2τ addieren. Das ist unnötig aufwendig. Bei der hier gewählten Verschlüsselung ist nicht der zu jedem Ereignis gehörende Zeitwert, sondern der zeitliche Abstand dieses Ereignisses vom vorhergehenden festgehalten, und man wird daraus nicht die absoluten Zeitwerte berechnen, um die Verschiebung durchzuführen, sondern überlegen, welche Wirkung die Laufzeit auf Abstände hat.

Man erkennt in Bild 14 b), daß sich

1. der Abstand um $2\tau = 50$ Einheiten verkürzt, wenn ein Ereignis den Index 2 trägt und das vorhergehende den Index 1, also zu Teilnehmer 1 gehört und eben erst nach

- der Zeit 2τ bei Teilnehmer 2 eintritt.
2. der Abstand um 2τ verlängert, wenn das Ereignis den Index 1 trägt und das vorhergehende den Index 2.
 3. der Abstand unverändert bleibt, wenn ein Ereignis den gleichen Index trägt wie das vorhergehende.

Mit diesen drei Regeln kann man die Tabelle II b direkt aus II a gewinnen: Die Abstände des $1a_1$ von A und des $2e_1$ von $1a_1$ ändern sich nicht - nach 3. (man beachte, daß A wie A_1 und E wie E_1 zu lesen sind). Der Abstand des $3a_2$ von $2e_1$ verkürzt sich von 58 auf 8 um $2\tau = 50$ nach 1. Der Abstand des $4e_2$ von $3a_2$ bleibt wiederum nach 3. unverändert, während sich der Abstand des $5a_1$ von $4e_2$ um 2τ von 3 auf 53 verlängert nach 2. usw.

4.4.2 Sortieralgorithmus

Verwendetes Verfahren - Beschreibung und Einordnung

Die aus dem Verschiebealgorithmus hervorgegangenen Daten z.B. Tab. II b enthalten die Ereignisse nicht in ihrer zeitlichen Reihenfolge. Zur Weiterverarbeitung sind jedoch zeitlich geordnete Listen z.B. II c erforderlich. Sie zu erstellen, ist die Aufgabe des Sortieralgorithmus.

Das im folgenden zu beschreibende Verfahren wurde ohne Kenntnis der im Anhang beschriebenen Tauschverfahren gefunden.

Gleichwohl gehört es zu ihnen. Es enthält gegenüber dem in /13/ beschriebenen zwei Abwandlungen. Die erste betrifft die Reihenfolge des Durchlaufs der Liste, die zweite das Sortierkriterium.

Einen Hinweis darauf, daß die abgewandelte Reihenfolge des Durchlaufs bekannt ist, fand der Verfasser in RECHENBERG Prog. II. 12-25 Vorlesungsscript. Dort wird von der Notwendigkeit, eine soeben getauschte Zahl unter die bereits vorhandenen einzuordnen, gesprochen, so daß ein vollständig geordneter Listenteil entsteht - erst dadurch würde das Verfahren effektiv. Es folgt dann die Beschreibung des Algorithmus.

In diesem Bericht wird an zwei Beispielen gezeigt, daß das Sortieren in einem Durchgang bei maximaler Unordnung ebensoviel Vergleiche benötigt wie das bei WALNUT beschriebene Verfahren des Sortierens in mehreren Durchgängen. während es bei weitgehend vorgeordneten Listen in der Tat wesentlich effektiver ist als das Verfahren nach /13/.

Die Abwandlung des Sortierkriteriums ergab sich aus der speziellen Problemstellung der Aufgabe. Sie kann auch in anderen Fällen vorteilhaft sein. Es ist nicht sehr wahrscheinlich, daß sie bekannt ist. Wegen des Umfangs der durchzusehenden Literatur wurde darauf verzichtet, dies genau zu prüfen.

4.4.2.1 Auswechseln in einem Durchgang

Das Verfahren nach /13/ wird zunächst dahingehend abgewandelt, daß an die Stelle mehrfacher Durchgänge ein einziger Durchgang tritt. Es sei an Bild 16 erläutert:

10	10	10	10	10	10	10	13	
5	5	<u>5</u>	13	13	13	13	10) vergleichen
13	13	13	5	<u>5</u>	7	7	7	= auswechseln
7	7	7	7	7	5	5	5	

Bild 16

Sortieren durch Auswechseln in einem Durchgang.

Von unten beginnend werden die Zahlen miteinander verglichen. Jede beim Fortschreiten nach oben neu hinzukommende Zahl, die größer ist als die unter ihr stehende, fügt sich der unter ihr stehenden geordneten Teilliste an. Ist die neu hinzukommende Zahl jedoch kleiner als die unter ihr stehende z.B. $5 < 13$ in Spalte 2, wird getauscht. Nach der Vertauschung wird nun die neu hinzugekommene Zahl (5) mit der jetzt unter ihnen stehenden (7) verglichen (Spalte 4) und danach gegebenenfalls erneut getauscht und erneut mit einer darunterstehenden verglichen, bis die neu hinzugekommene Zahl richtig eingeordnet ist - eventuell auf dem letzten Listenplatz. Die untere Teilliste ist dann vollständig geordnet. Erst jetzt wird die nächste Zahl einbezogen (10 in Spalte 6). Hat man schließlich die an der Spitze der Liste

stehende Zahl eingeordnet, so sind alle Zahlen geordnet. Der Algorithmus bietet beim Programmieren keine Schwierigkeiten.

Im vorliegenden Beispiel ergeben sich 5 Vergleiche; das ist ein Vergleich weniger als bei dem Verfahren mit mehreren Durchgängen. (Anhang) Bei maximaler Unordnung in diesem Beispiel (6 Inversionen) benötigt das Verfahren des Ordnen in einem Durchgang ebenfalls 6 Vergleiche.

Für maximal ungeordnete Daten hat das Sortieren in einen Durchgang keinen Vorteil gegenüber dem Sortieren in mehreren Durchgängen.

Das Sortieren in einem Durchgang ist jedoch dem mehrfachen Durchgehen überlegen, wenn die Daten weitgehend geordnet sind und nur "lokale" Vertauschungen enthalten:

10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
7	7	7	7	7	7	7	7	7	7	7	7	<u>7</u>	9	9	9
9	9	9	9	9	9	9	9	9	9	9	9	9	7	<u>7</u>	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	7
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
1	1	<u>1</u>	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	1	<u>1</u>	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	<u>2</u>	1	1	1	1	1	1	1	1	1	1	1

) vergleichen = austauschen

Bild 17

Ordnen weitgehend geordneter Daten in einem Durchgang

Das Ordnen der vorgelegten 10 Zahlen, die zwei lokale Vertauschungen enthalten, benötigt in einem Durchgang 12 Vergleiche. Der Verlauf ist Bild 17 zu entnehmen. -

Beim Ordnen in mehreren Durchgängen ergeben sich im ersten Durchgang 9 Vergleiche. Das Ergebnis des 1. Durchganges zeigt Bild 18.

Der zweite Durchgang betrifft nur neun Zahlen, enthält also 8 Vergleiche, das Ergebnis ist die richtig geordnete Folge der Zahlen 1-9. Der dritte Durchgang mit 7 Vergleichen wird noch ausgeführt, es ergeben sich jedoch keine Vertauschungen mehr, so daß das Verfahren damit am Ende angekommen ist.

Insgesamt waren also $(9+8+7)=24$ Vergleiche auszuführen. Das sind doppelt soviel, wie beim Verfahren in einem Durchgang. Hier zeigt sich deutlich die Überlegenheit des Sortierens in einem Durchgang.

10	10	10
7	9	9
9	7	8
8	8	7
6	6	6
5	5	5
4	4	4
1	3	3
3	1	2
2	2	1

Bild 18

Ordnen der Daten nach Bild 17 in mehreren Durchgängen

Die aus der Sprachstatistik stammenden Daten bestehen aus längeren Abschnitten, in denen Sprache und Sprechpausen eines Sprechers aufeinanderfolgen. Diese Abschnitte werden durch den Verschiebealgorithmus nicht umgeordnet. Nur an den Stellen, an denen es zum Wechselgespräch kommt, kann durch Laufzeiteinfluß die zeitliche Reihenfolge der Ereignisse vertauscht werden. Das sind lokale, u.U. über zwei oder drei Ereignisse hinübergreifende Umordnungen in einer Folge weit überwiegend geordneter Ereignisse. Die tatsächlichen Verhältnisse sind also noch günstiger als im Beispiel an Bild 17 gezeigt, so daß der Vorteil bei Verwendung des mitgeteilten Sortierverfahrens anstelle des in /13/ angegebenen noch größer ist.

4.4.2.2 Differenzen als Sortierkriterium

Die zweite Abwandlung des bekannten Sortierens durch Vertauschen liegt in der Anwendung der Differenzen als Sortierkriterium. Dies ist hier zweckmäßig, da die Zeitwerte nicht selbst, sondern nur ihre Differenzen als Abstände vorliegen.

Das Verfahren werde an folgendem Beispiel eingeführt.

Gegeben sind die Zeitangaben 0, 10, 5, 13, 7, 15. Sie mögen von oben nach unten in steigender Reihenfolge angeordnet werden, und zwar zunächst unter Verwendung des bereits erläuterten Tauschverfahrens in einem Durchgang mit der Maßgabe, jetzt zu tauschen, wenn nach dem Vergleich eine größere Zahl oberhalb einer kleineren steht. Den Vorgang zeigt Bild 19.

0	0	0	0	0	0	0	0	0	0	0)
10	10	10	10	10	10	<u>10</u>	5	5	5	5)
5	5	5	5	5	5	5	10	<u>10</u>	7	7)
13	13	<u>13</u>	7	7	7	7	7	7	10	10)
7	7	7	13	13	13	13	13	13	13	13) vergleichen
15	15	15	15	15	15	15	15	15	15	15	= vertauschen

Bild 19

Herstellung einer aufsteigend geordneten Folge durch Vertauschen benachbarter Zahlen in einem Durchgang

Tatsächlich liegen nicht, wie in den bisherigen Beispielen, einzelne Zahlen vor, die zu sortieren sind, sondern Zahlenpaare aus der Schlüsselzahl der "Ereignisart" und der des Abstandes" dieses Ereignisses vom vorhergehenden. Wir werden diese Zahlenpaare in Zukunft als Informationseinheiten bezeichnen.

Die Folge der Ereignisse beginnt mit dem Anfang A des verschlüsselten Abschnittes, dem man zweckmäßigerweise den Zeitwert 0 zuordnet, und endet mit dem Ende des verschlüsselten Abschnittes E, zu dem der größte Zeitwert der Folge gehört.

In Bild 20 sind die in Bild 19 gezeigten Zeitwerte mit der Schlüsselzahl für die Ereignisart und laufender Nummer und mit der Abstandsangabe vom vorhergehenden Ereignis ergänzt worden.

(Laufende Nr.)	Ereignisart	(Zeitangabe)	Zeitl. Abstand vom vorhergehenden Ereignis
./.	A	0	./.
1	a_2	10	10
2	a_1	5	- 5
3	e_2	13	8
4	e_1	7	- 6
./.	E	15	8

Bild 20

Ungeordnete Folge von Informationseinheiten.

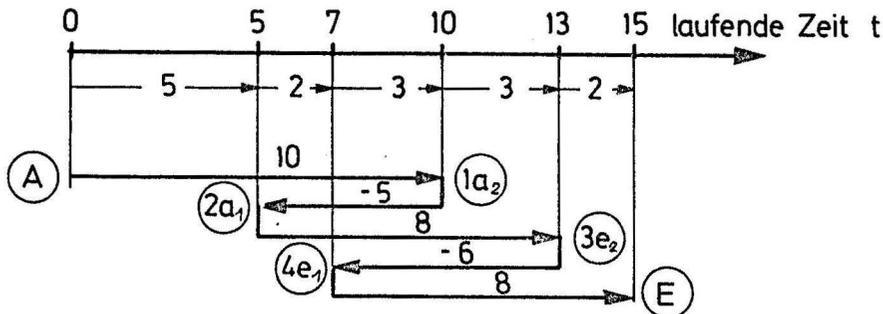


Bild 21: Diagramm zur Folge der Informationseinheiten A, $1a_2$, 10, $2a_1$, -5, $3e_2$, 8, $4e_1$, -6, E, 8

Die Ereignisse sind ihren Zeitwerten zugeordnet.

Die Pfeilkette gibt die Reihenfolge in der Ausgangsliste (Bild 20) an. Die Länge eines Pfeils stellt den Abstand des an der Pfeilspitze bezeichneten Ereignisses vom vorhergehenden dar. Das Vorzeichen der Abstandsangabe ist positiv, wenn der zugehörige Pfeil in die positive Richtung der t-Achse (nach rechts) zeigt, und negativ, wenn er entgegengesetzt gerichtet ist.

Die praktisch verwendeten Informationseinheiten besitzen die eingeklammerten Spalten: "laufende Nummer" und "Zeitangabe" nicht.

Daß die Informationseinheiten nicht in ihrer zeitlichen Reihenfolge geordnet sind, erkennt man, wie schon auf S. 34 gesagt, in der Zeile "zeitlicher Abstand" an den negativen Vorzeichen. Daraus folgt, daß man die Abstände als Sortierkriterium verwenden kann. Wie das Sortieren anhand der Abstände geschieht, werde an Bild 21 erläutert.

Der Sortiervorgang in der Liste Bild 20 beginnt wie derjenige in Bild 19 am Schluß der Liste. Die Abstandsangaben werden von hinten beginnend bis zu einem negativen Vorzeichen durchgegangen. Als ersten Abstand mit negativem Vorzeichen findet man im Beispiel (Bild 20, 21) den Abstand -6 , der zu $4e_1$ gehört. In Bild 21 erkennt man, daß der an der Spitze mit $4e_1$ gekennzeichnete Pfeil 6 Einheiten lang ist und nach links zeigt. Er zeigt in die negative Zeitrichtung, weil die Ereignisse an seinem Ende $3e_2$ und an seiner Spitze $4e_1$ in der Liste Bild 20 in falscher zeitlicher Reihenfolge stehen. Sie müssen ausgetauscht werden. Diese Vertauschung entspricht genau derjenigen der Zeitwerte 7 und 13 in der Liste 19. Allgemeiner gesprochen sind die zu einem negativen Abstand gehörenden Ereignisse miteinander zu vertauschen. Der Abstand selbst wird dadurch positiv. Praktisch ist in der Informationseinheit der Abstand einem Ereignis zugeordnet. Die zu einem negativen Abstand gehörende Schlüsselzahl muß mit derjenigen der davorstehenden Informationseinheit getauscht werden. Das negative Vorzeichen der Abstandsangabe verschwindet.

Aus Bild 21 erkennt man, welche Veränderungen der Abstände sich durch die Vertauschung ergeben. Wenn der Abstandspfeil, der von $3e_2$ nach $4e_1$ zeigte, umgedreht wird, bleibt die Länge des umgedrehten Pfeils erhalten. Der geschlossene Linienzug der Pfeile, der die Reihenfolge in der nun entstandenen Liste angeben soll, führt jedoch von $2a_1$ nunmehr nach $4e_1$ und nicht mehr nach $3e_2$ und der Pfeil, der bei E endet, geht nun von $3e_2$ aus. Diese Folgerung wäre auch direkt aus der Liste Bild 20 nach dem Vertauschen von $3e_2$ mit $4e_1$ zu ziehen gewesen, da auf $2a_1$ nunmehr $4e_1$ und auf $3e_2$ E folgt. Für die Länge der an den umgedrehten anschließenden Abstandspfeile bedeute dies, daß beide um die Länge des ursprünglich umzudrehenden Pfeiles verkürzt werden. Sie können dabei prinzipiell ihrerseits die Richtung wechseln. Zur Liste Bild 20 zurückgekehrt heißt das, daß die beiden dem negativen Abstand benachbarten Distanzen um den Betrag des negativen Abstandes zu vermindern sind. Sie können dabei ihrerseits negativ werden. Damit sind die zu einer Vertauschung gehörenden Schritte beschrieben. Sie lassen sich praktisch in einem Unterprogramm mit wenigen Befehlen erfassen:

1. Aufsuchen eines negativen Abstandes
2. Tauschen der Schlüsselzahl in der zugehörigen Informationseinheit mit der davorstehenden
3. Vorzeichenumkehr beim aufgesuchten Abstand selbst, Addition des negativen Abstandes zu beiden Nachbarabständen.

Programmiert man das Austausch-Sortierverfahren mit einem Durchgang vom Ende der Liste beginnend, dann wäre-nach dem Auffinden und Beseitigen eines negativen Abstandes-zur Abfrage nach der nächsten negativen Differenz zunächst ein Schritt auf das Ende der Liste hin zu tun, ehe man, falls kein negativer Abstand gefunden wird, wieder auf den Anfang der Liste zuläuft.

Als abschließendes Beispiel werde die Liste aus Bild 20 nach diesem Verfahren geordnet:

Ereign.	Abst.	Ereign.	Abst.	Ereign.	Abst.	Ereign.	Abst.
A	./.	A	./.	A	./.	A	./.
1 a ₂	10	<u>1 a₂</u>	10	2 a ₁	5	2 a ₁	5 ← 8.
2 a ₁	- 5	2 a ₁	- 5 ← 5.	<u>1 a₂</u>	5	4 e ₁	2
<u>3 e₂</u>	8	4 e ₁	2 ← 4.	4 e ₁	- 3 ← 6.	1 a ₂	3
4 e ₁	- 6 ← 2.	3 e ₂	6	3 e ₂	6	3 e ₂	3 ← 7.
E	8 ← 1.	E	2 ← 3.	E	2	E	2

i. ← i. Abfrage nach negativem Vorzeichen = Austausch

Bild 22 Ordnen einer Liste in einem Durchgang anhand der Differenzen als Sortierkriterium

Die Zahl der Vorzeichenabfragen in Bild 22 ist gleich derjenigen der Vergleiche benachbarter Zahlen in Bild 19, nämlich 8. Die Richtigkeit des Ergebnisses überprüft man durch Vergleich der geordneten Liste mit Bild 21, dem die geordnete Ereignisfolge und die zugehörigen Abstände ohne weiteres zu entnehmen sind.

Das verwendete Sortierverfahren gehört zu Klasse der Austauschsortierverfahren. Für das vorliegende Problem weitgehend geordneter Daten, die nur lokale Vertauschungen aufweisen und bei denen als Sortierkriterium die Abstände gegeben sind, bietet es folgende Vorteile:

- 1) Es benötigt sehr wenig Speicherplatz und für weitgehend geordnete Daten, wie sie hier vorliegen, besonders wenig Rechenzeit.
- 2) Das Abarbeiten der Liste in einem Durchgang ergibt eine erheblich kleinere Anzahl von Vergleichen, als man bei mehrfachem Durchgang der Liste erhalten würde
- 3) Im vorliegenden Material stehen nur Differenzen als Sortierkriterium zur Verfügung. Dadurch tritt an die Stelle des Vergleichs von zwei Zahlen die Kontrolle des Vorzeichens einer Zahl. Dies kann zur Beschleunigung des Verfahrens ausgenutzt werden und fällt ins Gewicht, wenn die Zahl der Vergleiche wesentlich größer als die der Vertauschungen ist.

Der gefundene Sortieralgorithmus ist für das vorliegende Problem äußerst zweckmäßig.

4.5 Ergebnisse der manuellen und der automatischen Auswertung:

Als Ergebnisse werden hier die Verteilungsfunktionen der verschiedenen Zeitabschnitte Bild 22 - 35 vorgelegt. Einige der Darstellungen finden sich in /7/ und sind dort diskutiert. Hier sei vor allem auf den Unterschied zwischen manueller (Bild 22 - 28) und automatischer (Bild 29 - 35) Auswertung hingewiesen. Die Verteilungsfunktionen aus letzterer verlaufen bei etwas kürzeren Zeitwerten als diejenigen aus manueller Auswertung. So liegen die Mittelwerte der Sprechabschnitte nach Bild 23 bei etwa 2 Sekunden gegenüber 1,1 Sekunden in Bild 30 und die Mittelwerte der Sprechpausen nach Bild 24 bei 2,5 sec gegenüber 1,9 sec nach Bild 31. Pausen unter 200 ms kommen bei automatischer Auswertung nicht vor (Bild 31). Schließlich findet man die Mittelwerte der Zeiten des Gleichzeitigsprechens im Bild 25 u. 27 bei ca 0,6 sec. gegenüber ca 0,3 sec in Bild 32 und 34. Die Zeiten des Gleichzeitigschweigens sind in allen Fällen ca 0,6 sec. im Mittel (Bild 26, 28, 33, 35).

Im Prinzip zeigen die manuell und automatisch ausgewerteten Kurven die gleiche Tendenz, so sind die Sprechabschnitte kürzer als die Sprechpausen. Die Zeiten des Gleichzeitigsprechens- und schweigens sind kürzer als beide.

Durch die in 2.4.3 S. 10 angegebene Verbesserung und genauere Dimensionierung des Sprachdetektors ließe sich sicherlich eine noch weitergehende Übereinstimmung der manuell und automatisch erstellten Verteilungsfunktionen erreichen.

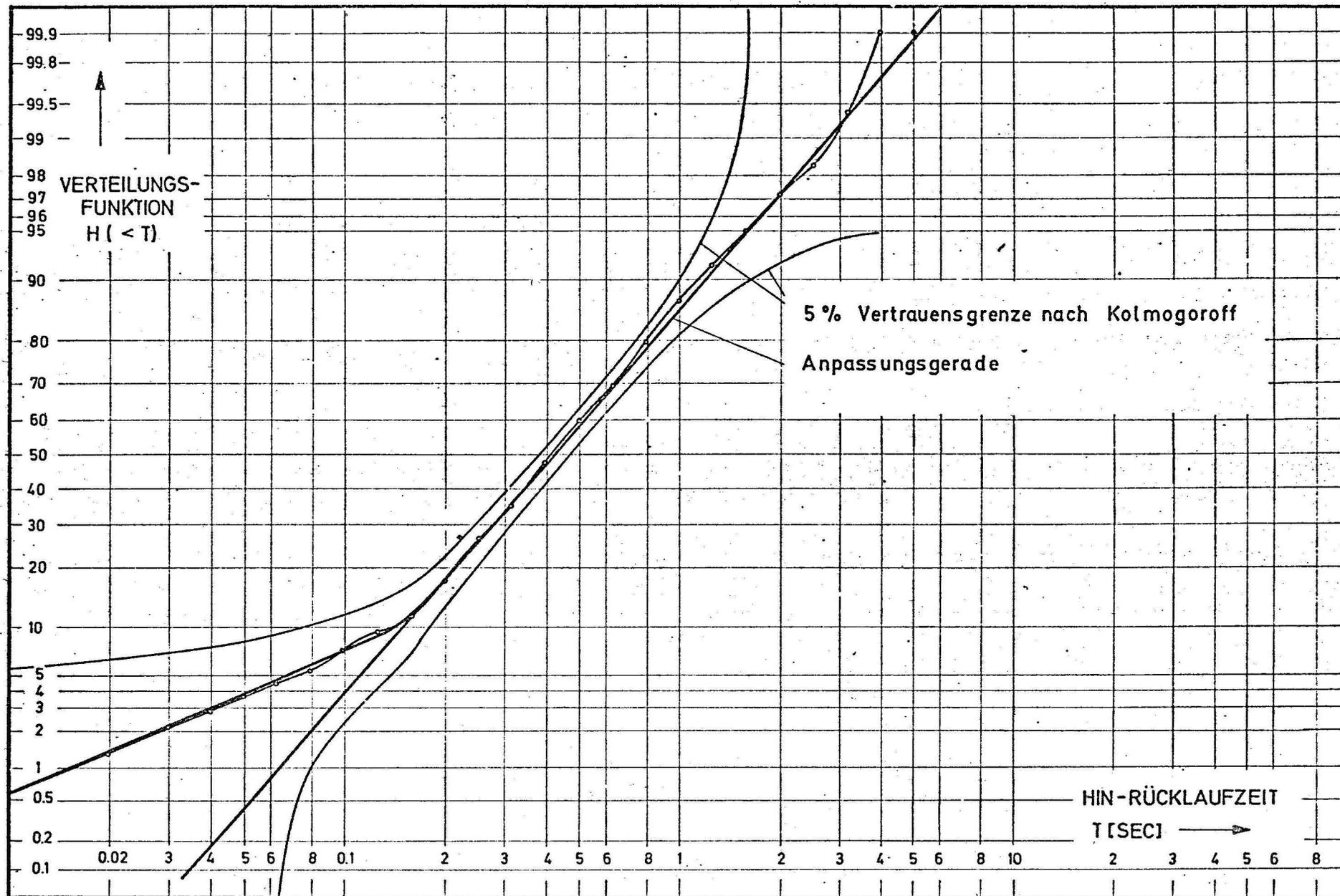


Bild 22: VERTEILUNGSFUNKTION DER ZEITEN DES GLEICHZEITIGSPRECHENS BEI $37 \times 2\frac{1}{4}$ MIN DES GESPRÄCHS ZWEIER PERSONEN, GEMESSEN VON HAND IM OSZILLOGRAMM VOM „NAHEN“ LEITUNGSSENDE. HIN-RÜCKLAUFZEIT 620 MSEC.

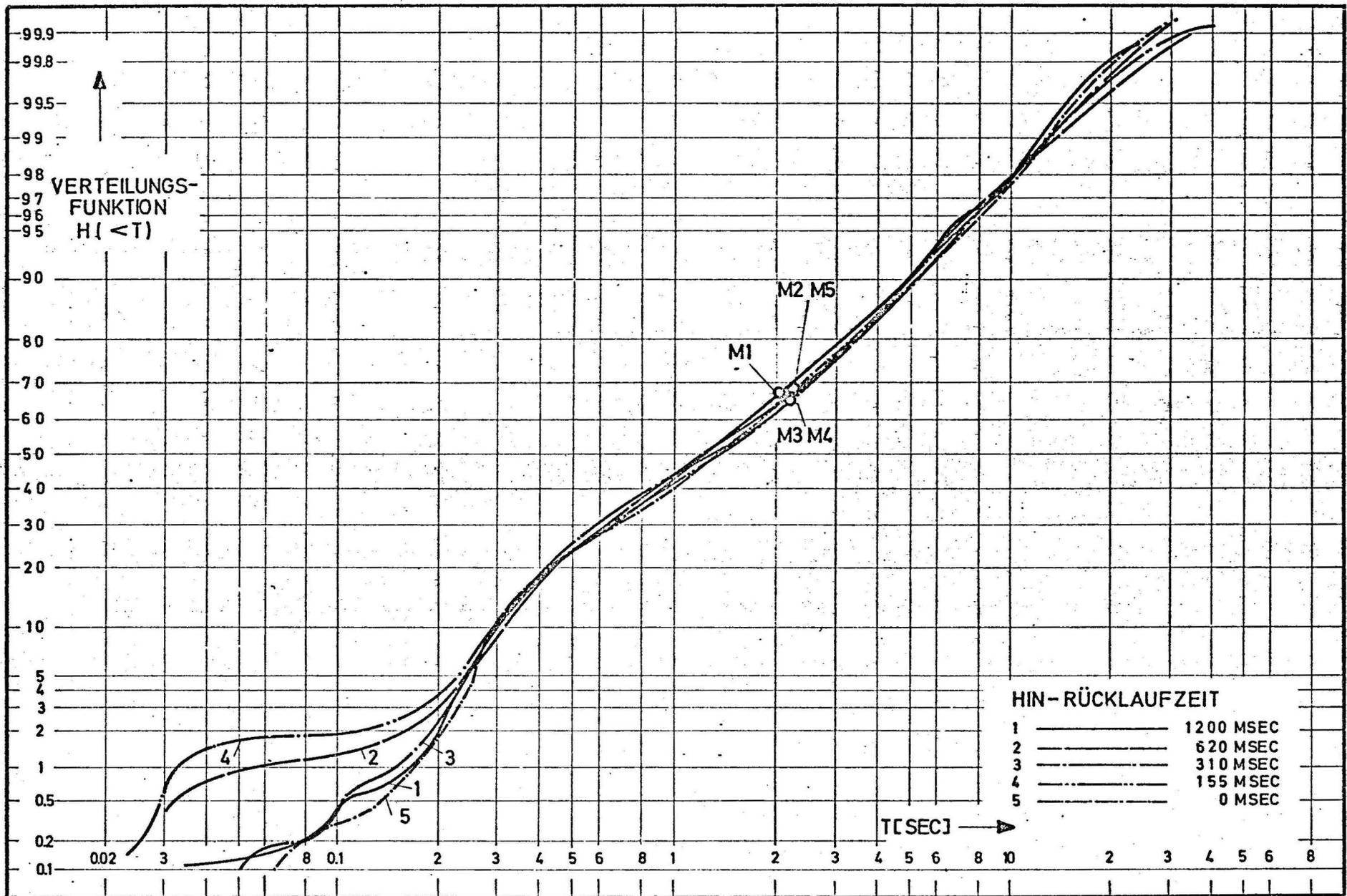


Bild 23: VERTEILUNGSFUNKTIONEN DER SPRECHABSCHNITTE BEI $74 \times 2 \frac{1}{4}$ MIN SPRACHE JE LAUFZEITSTUFE. ZEITWERTE VON HAND IM OSZILLOGRAMM GEMESSEN. HIN-RÜCKLAUFZEIT 0.....1200 MSEC.

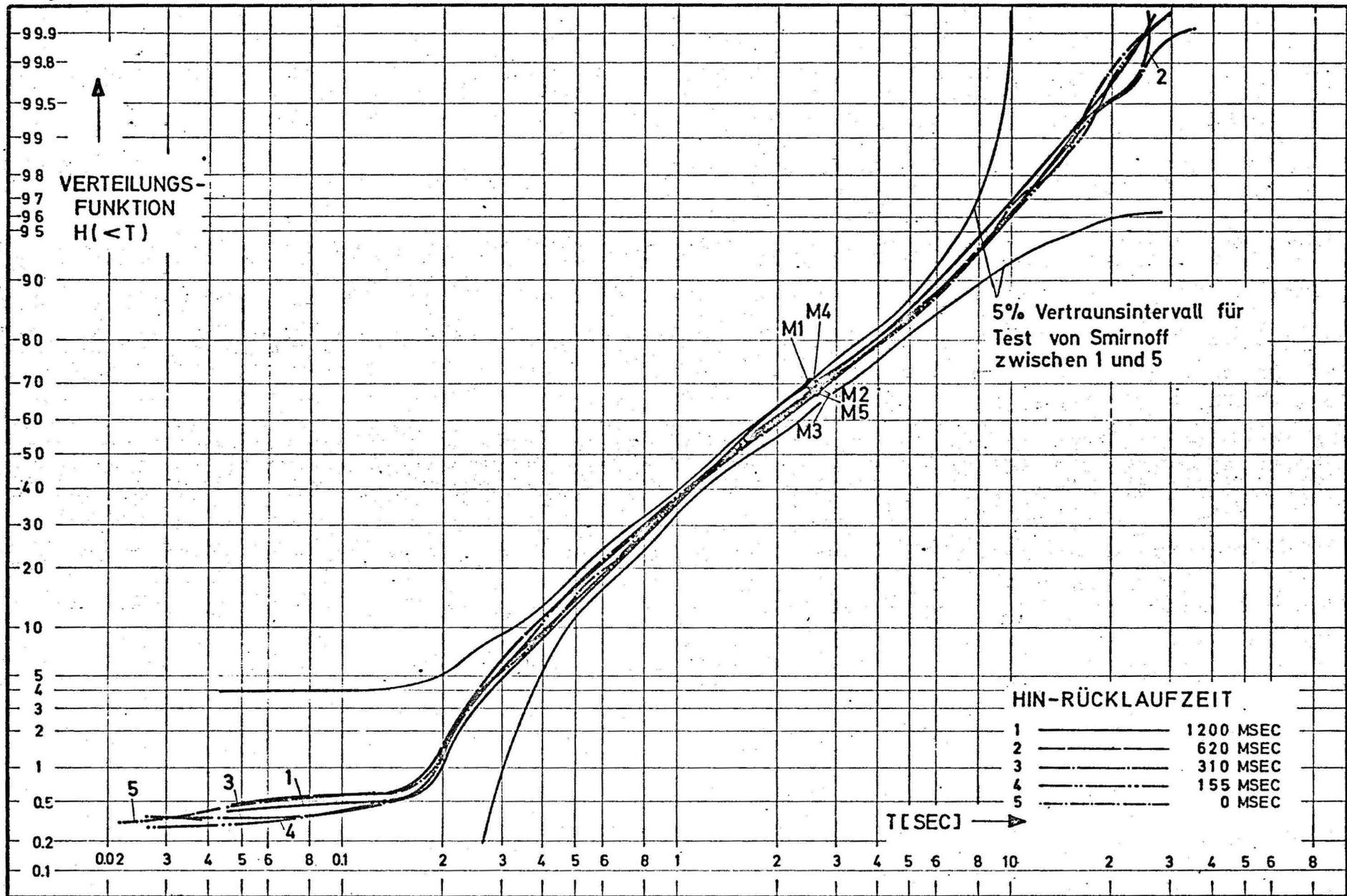


Bild 24: VERTEILUNGSFUNKTIONEN DER SPRECHPAUSEN BEI $74 \times 2 \frac{1}{4}$ MIN SPRACHE JE LAUFZEITSTUFE. ZEITWERTE VON HAND IM OSZILLOGRAMM GEMESSEN. HIN-RÜCKLAUFZEIT 0.....1200 MSEC.

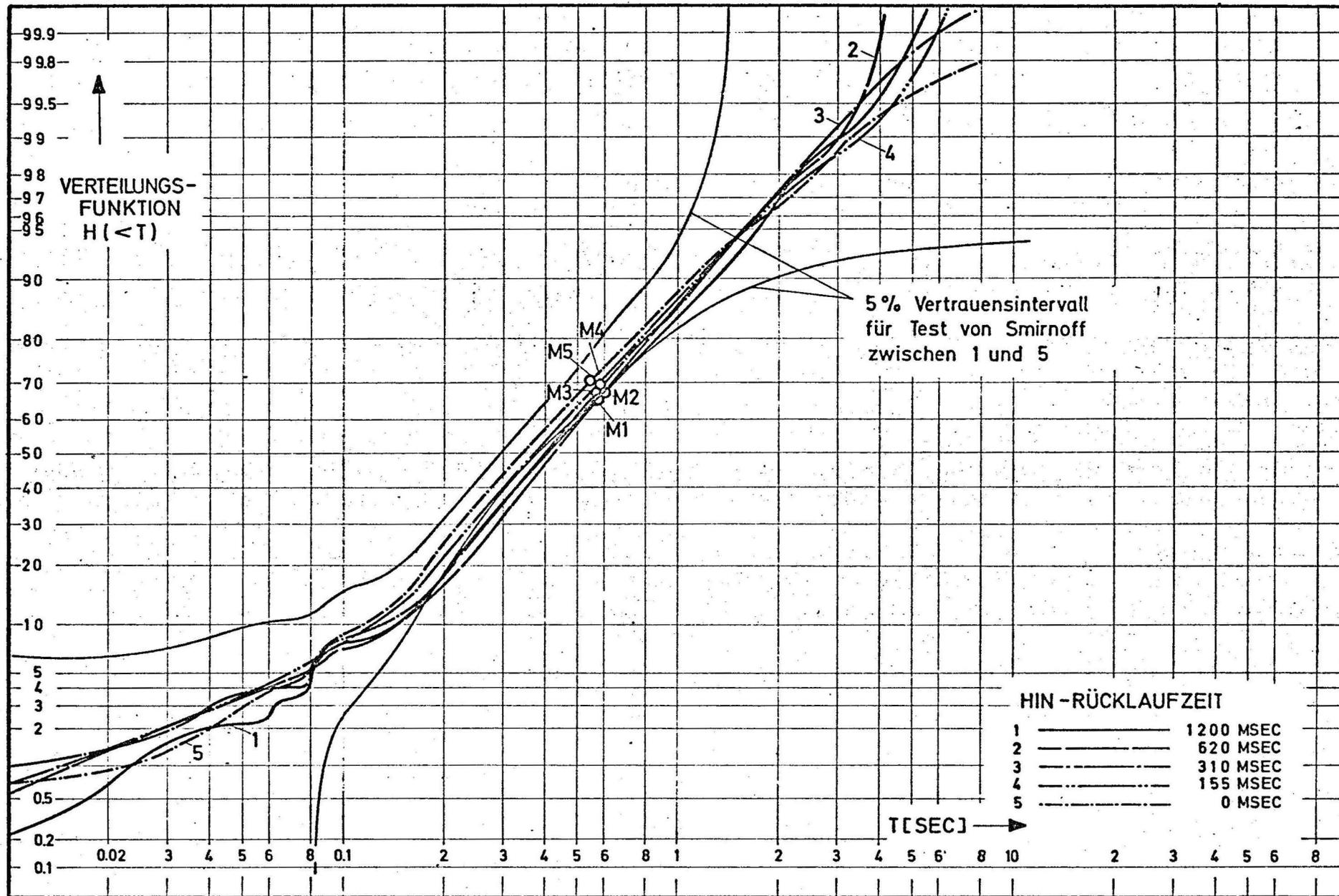


Bild 25: VERTEILUNGSFUNKTIONEN DER ZEITEN DES GLEICHZEITIGSPRECHENS BEI $37 \times 2 \frac{1}{4}$ MIN DES GESPRÄCHS ZWEIER PERSONEN JE LAUFZEITSTUFE, GEMESSEN VON HAND IM OSZILLOGRAMM VOM „NAHEN“ LEITUNGSENDE. HIN-RÜCKLAUFZEIT 0.....1200 MSEC.

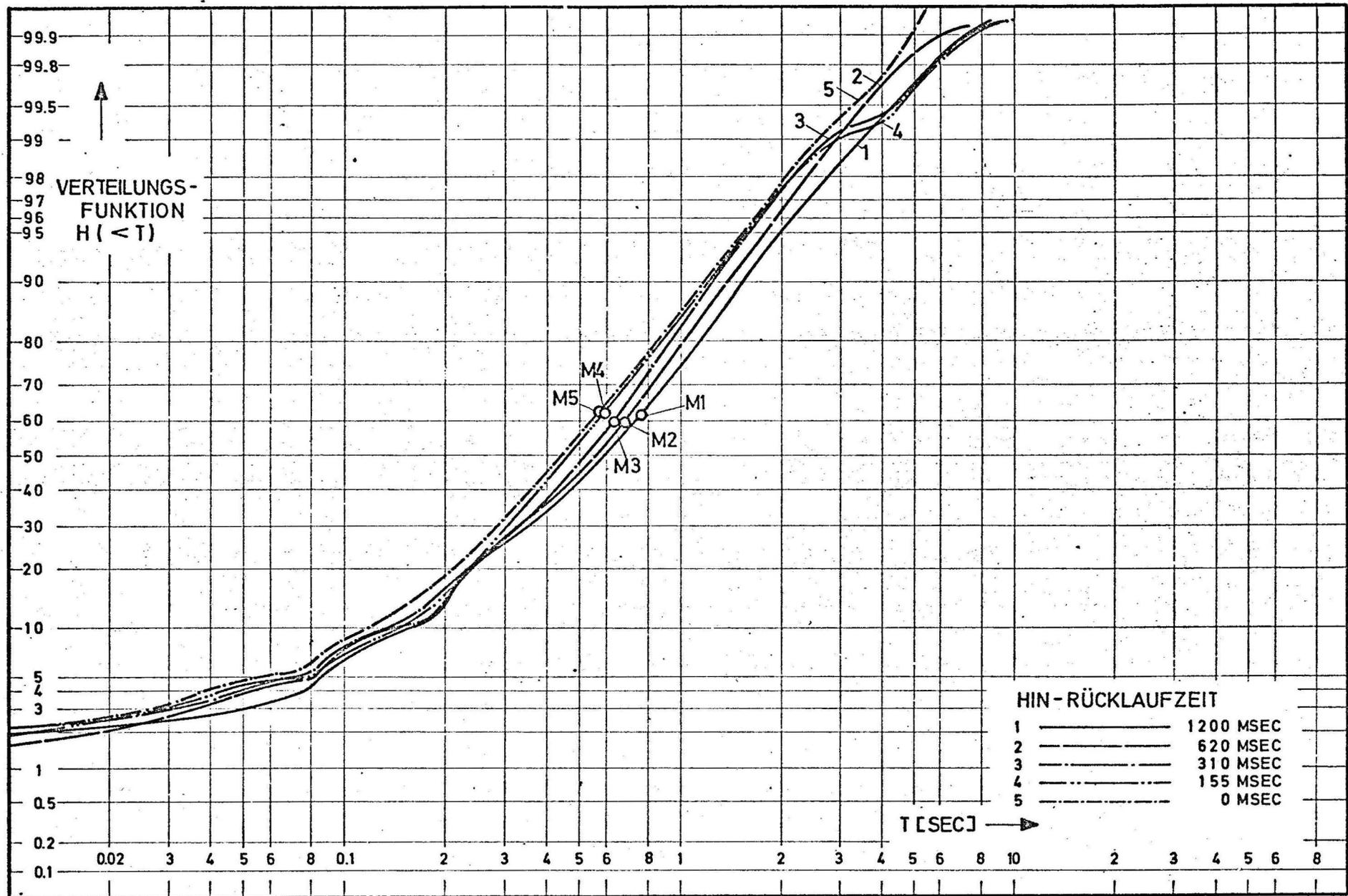


Bild 26: VERTEILUNGSFUNKTIONEN DER ZEITEN DES GLEICHZEITIGSCHWEIGENS BEI $37 \times 2 \frac{1}{4}$ MIN DES GESPRÄCHS ZWEIER PERSONEN JE LAUFZEITSTUFE, GEMESSEN VON HAND IM OSZILLOGRAMM VOM „NÄHEN“ LEITUNGSENDE.
HIN - RÜCKLAUFZEIT 0.....1200MSEC.

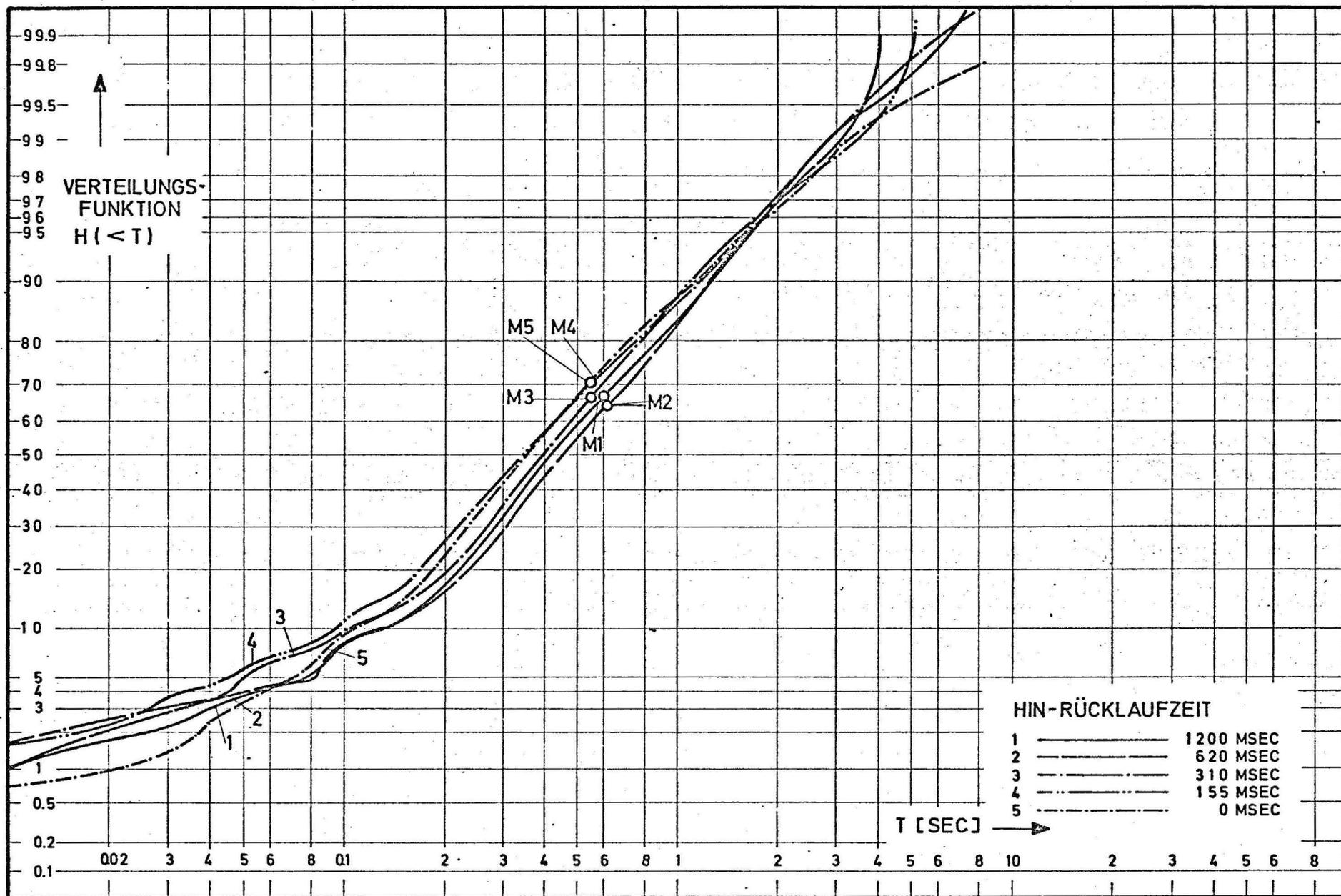


Bild 27: VERTEILUNGSFUNKTIONEN DER ZEITEN DES GLEICHZEITIGSPRECHENS BEI $37 \times 2 \frac{1}{4}$ MIN DES GESPRÄCHS ZWEIER PERSONEN JE LAUFZEITSTUFE. ERRECHNET FÜR DAS „FERNE“ LEITUNGSSENDE AUS DEN ERGEBNISSEN FÜR DAS „NAHE“ LEITUNGSSENDE BEI HANDAUSWERTUNG VOM OSZILLOGRAMM. HIN-RÜCKLAUFZEIT 0.....1200 MSEC.

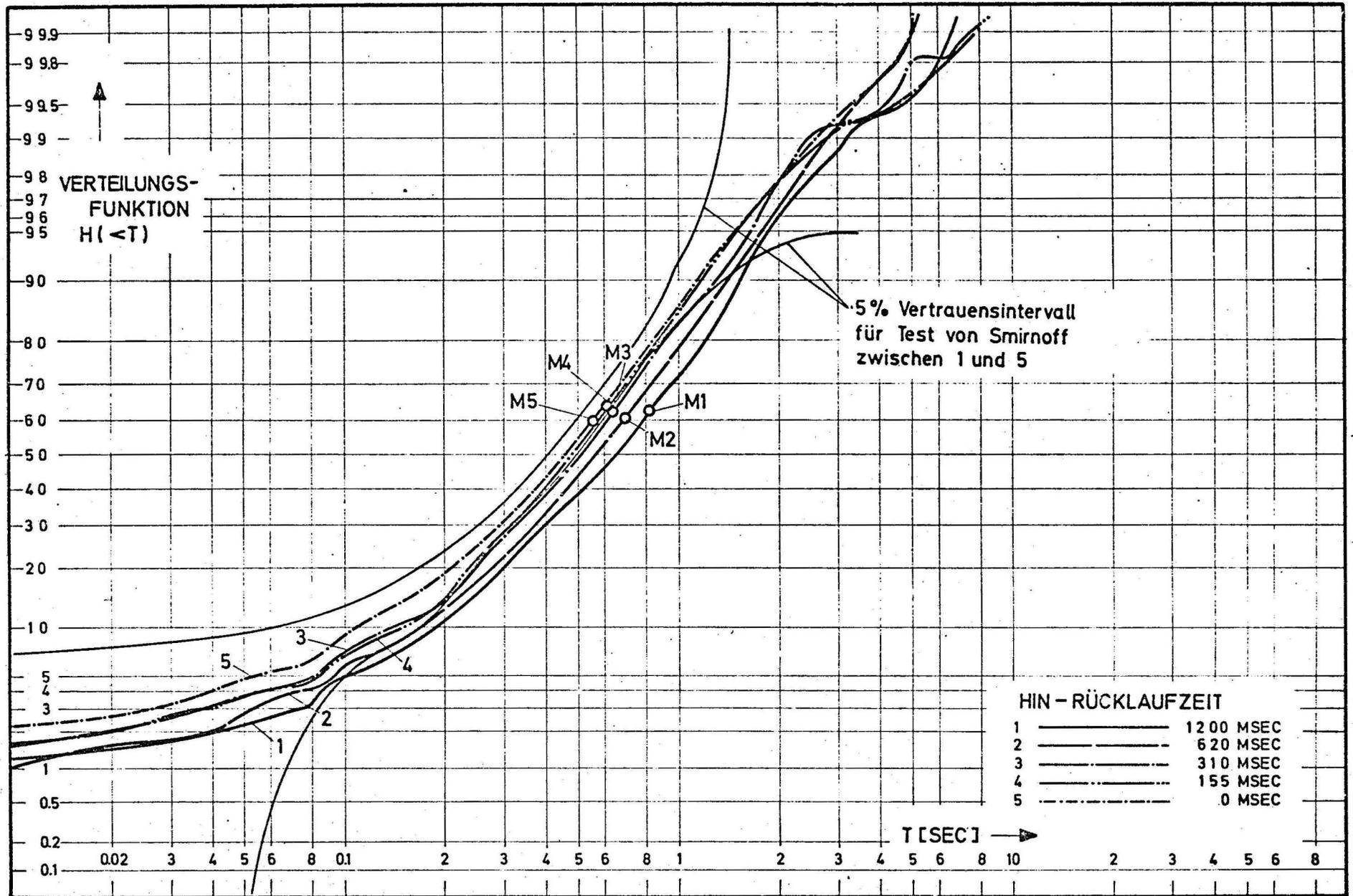


Bild 28: VERTEILUNGSFUNKTIONEN DER ZEITEN DES GLEICHZEITIGSCHWEIGENS BEI $37 \times 2 \frac{1}{4}$ MIN DES GESPRÄCHS ZWEIER PERSONEN JE LAUFZEITSTUFE, ERRECHNET FÜR DAS „FERNE“ LEITUNGSENDE AUS DEN ERGEBNISSEN FÜR DAS „NAHE“ LEITUNGSENDE BEI HANDAUSWERTUNG VOM OSZILLOGRAMM. HIN-RÜCKLAUFZEIT 0....1200 MSEC.

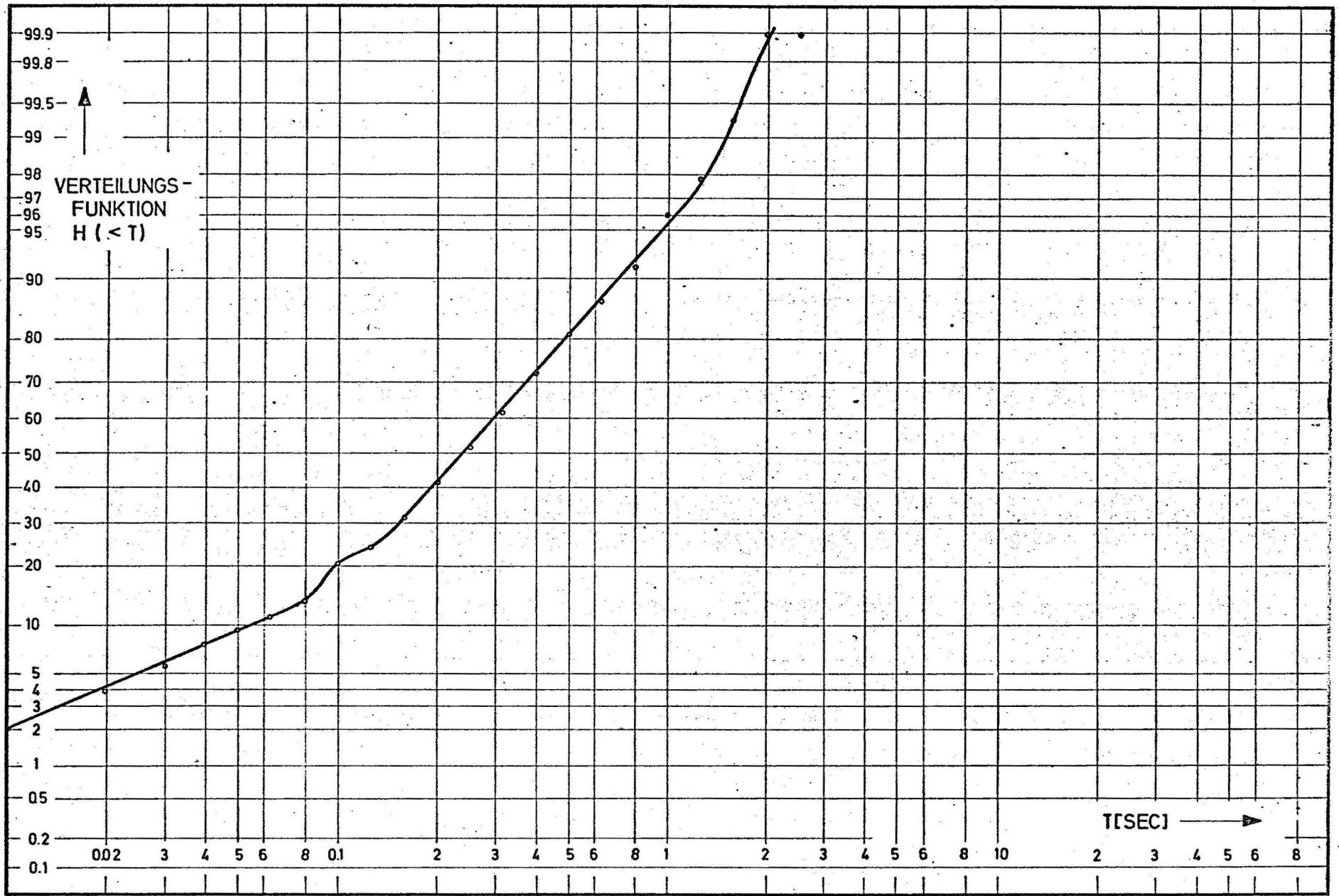


Bild 29: VERTEILUNGSFUNKTION DER ZEITEN DES GLEICHZEITIGSPRECHENS BEI $37 \times 21/4$ MIN DES GESPRÄCHS ZWEIER PERSONEN, ERRECHNET FÜR DAS „FERNE“ LEITUNGSENDE AUS DEN ERGEBNISSEN FÜR DAS „NAHE“ LEITUNGSENDE BEI MASCHINELLER ABSCHNITTSERKENNUNG. HIN-RÜCKLAUFZEIT 1200 MSEC.

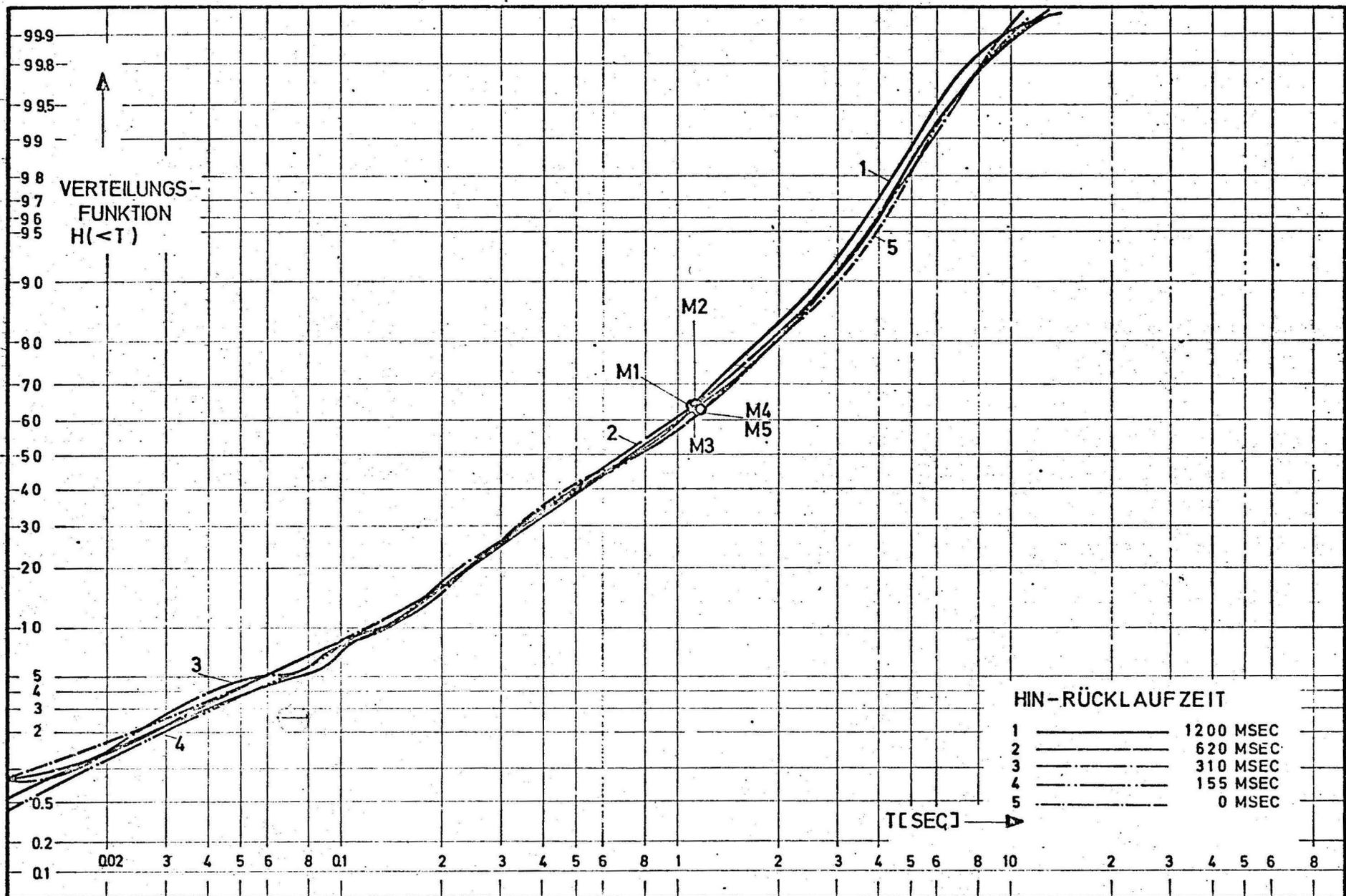


Bild 30: VERTEILUNGSFUNKTIONEN DER SPRECHABSCHNITTE BEI $74 \times 2 \frac{1}{4}$ MIN SPRACHE JE LAUFZEITSTUFE. ZEITWERTE SELBSTTÄTIG GEMESSEN BEI MASCHINELLER ABSCHNITTSERKENNUNG. HIN-RÜCKLAUFZEIT 0.....1200MSEC.

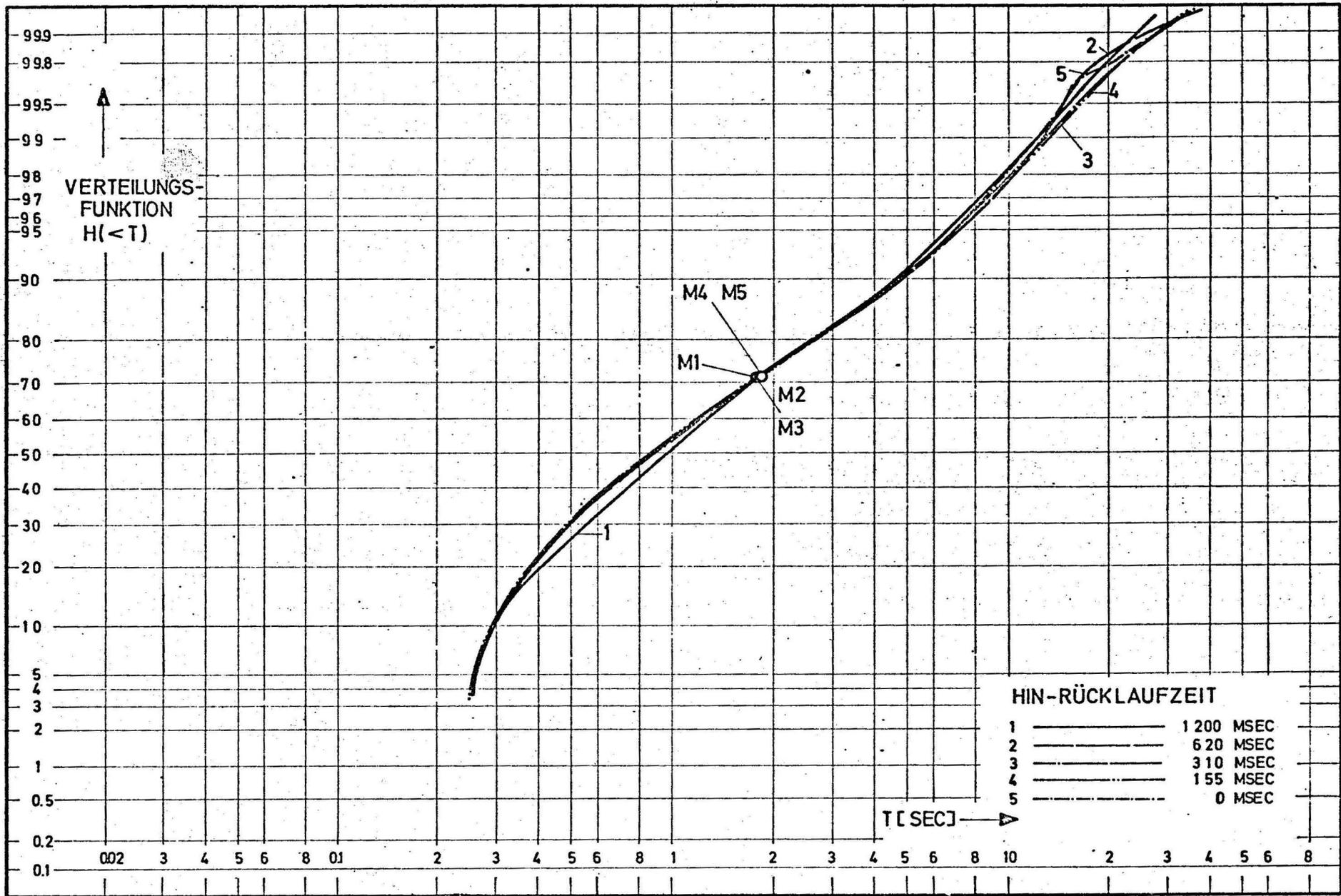


Bild 31: VERTEILUNGSFUNKTIONEN DER SPRECHPAUSEN BEI $74 \times 2 \frac{1}{4}$ MIN SPRACHE JE LAUFZEITSTUFE. ZEITWERTE SELBSTTÄTIG GEMESSEN BEI MASCHINELLER ABSCHNITTSERKENNUNG. HIN-RÜCKLAUFZEIT 0.....1200MSEC.

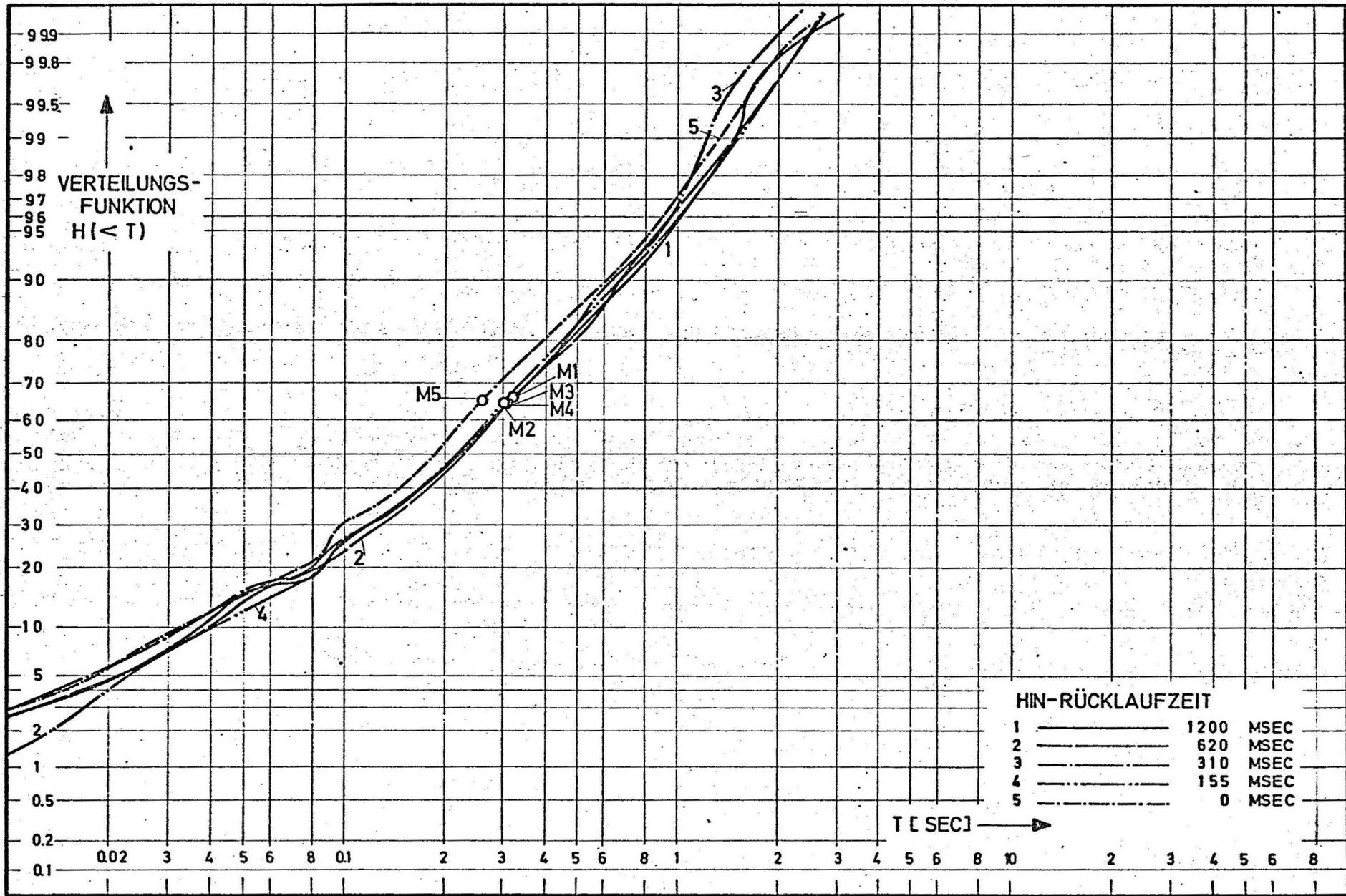


Bild 32: VERTEILUNGSFUNKTIONEN DER ZEITEN DES GLEICHZEITIGSPRECHENS BEI $37 \times 2 \frac{1}{4}$ MIN DES GESPRÄCHS ZWEIER PERSONEN JE LAUFZEITSTUFE, SELBSTTÄTIG GEMESSEN AM „NAHEN“ LEITUNGSSENDE BEI MASCHINELLER ABSCHNITTSERKENNUNG. HIN-RÜCKLAUFZEIT 0.....1200MSEC.

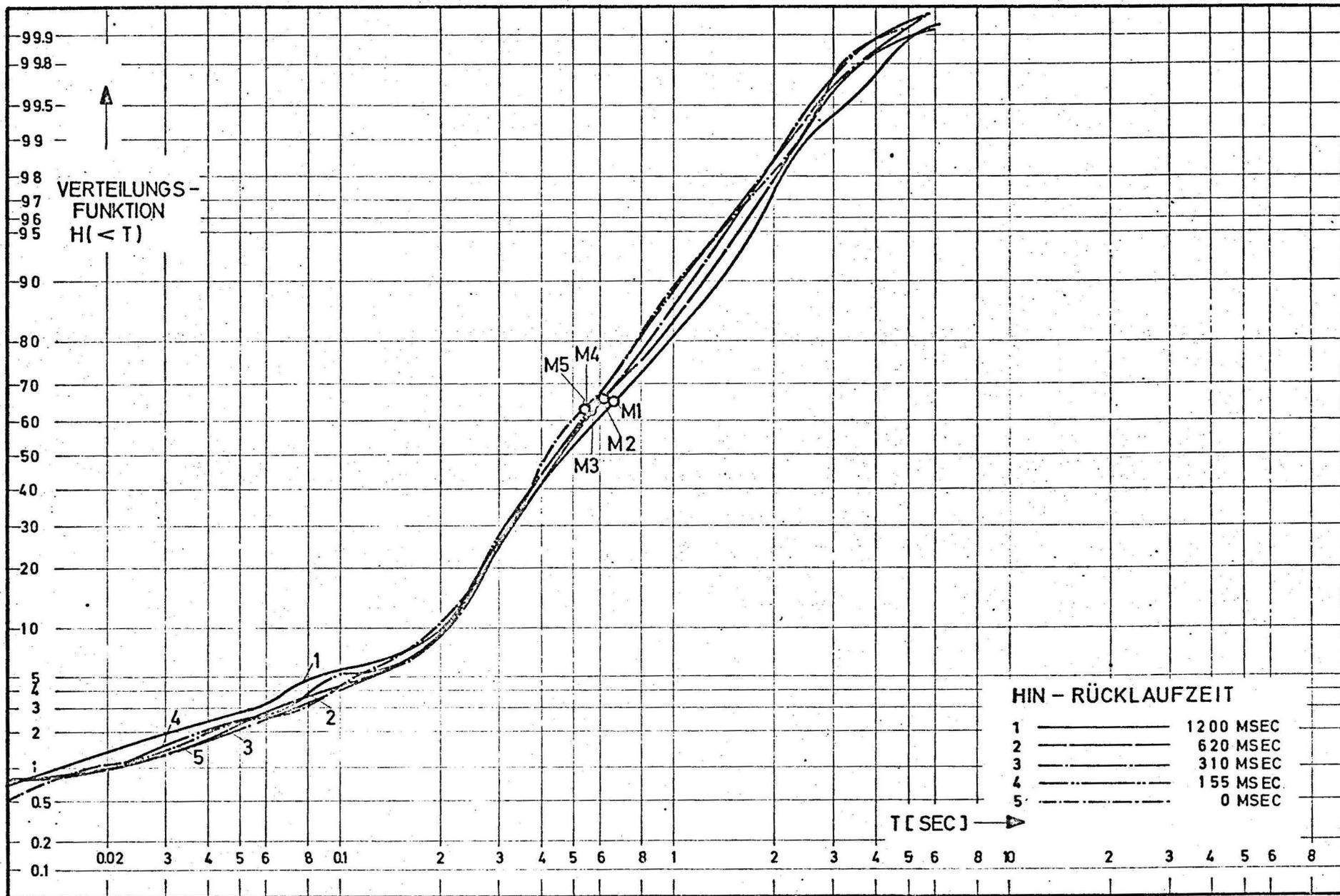


Bild 33: VERTEILUNGSFUNKTIONEN DER ZEITEN DES GLEICHZEITIGSCHWEIGENS BEI $37 \times 2 \frac{1}{4}$ MIN DES GESPRÄCHS ZWEIER PERSONEN JE LAUFZEITSTUFE, SELBSTTÄTIG GEMESSEN AM „NAHEN“ LEITUNGSSENDE BEI MASCHINELLER ABSCHNITTS-ERKENNUNG. HIN-RÜCKLAUFZEIT 0...1200MSEC.

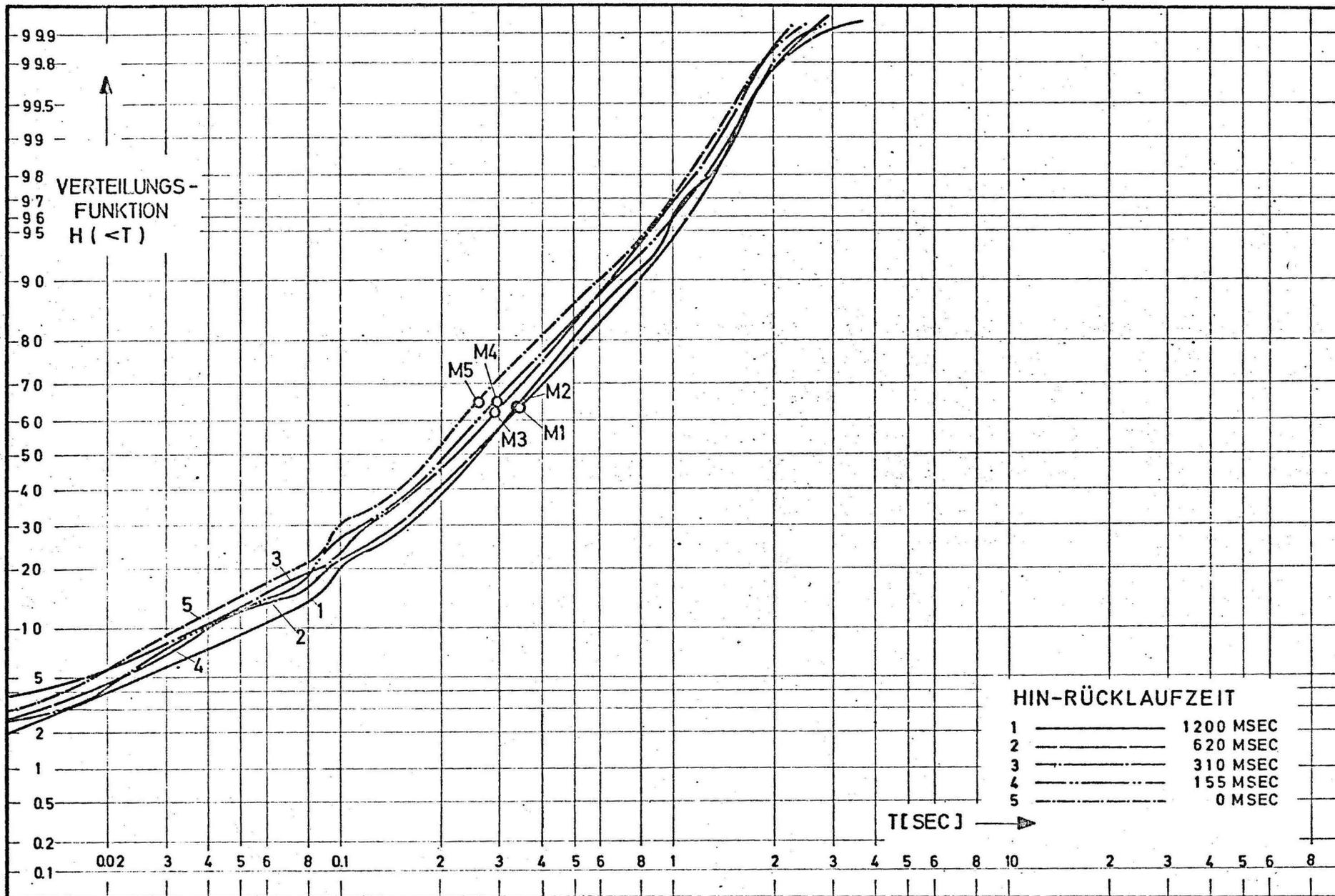


Bild 34: VERTEILUNGSFUNKTIONEN DER ZEITEN DES GLEICHZEITIGSPRECHENS BEI $37 \times 2\frac{1}{4}$ MIN DES GESPRÄCHS ZWEIER PERSONEN JE LAUFZEITSTUFE, ERRECHNET FÜR DAS „FERNE“ LEITUNGSSENDE AUS DEN ERGEBNISSEN FÜR DAS „NAHE“ LEITUNGSSENDE BEI MASCHINELLER ABSCHNITTSERKENNUNG. HIN-RÜCKLAUFZEIT 0.....1200 MSEC.

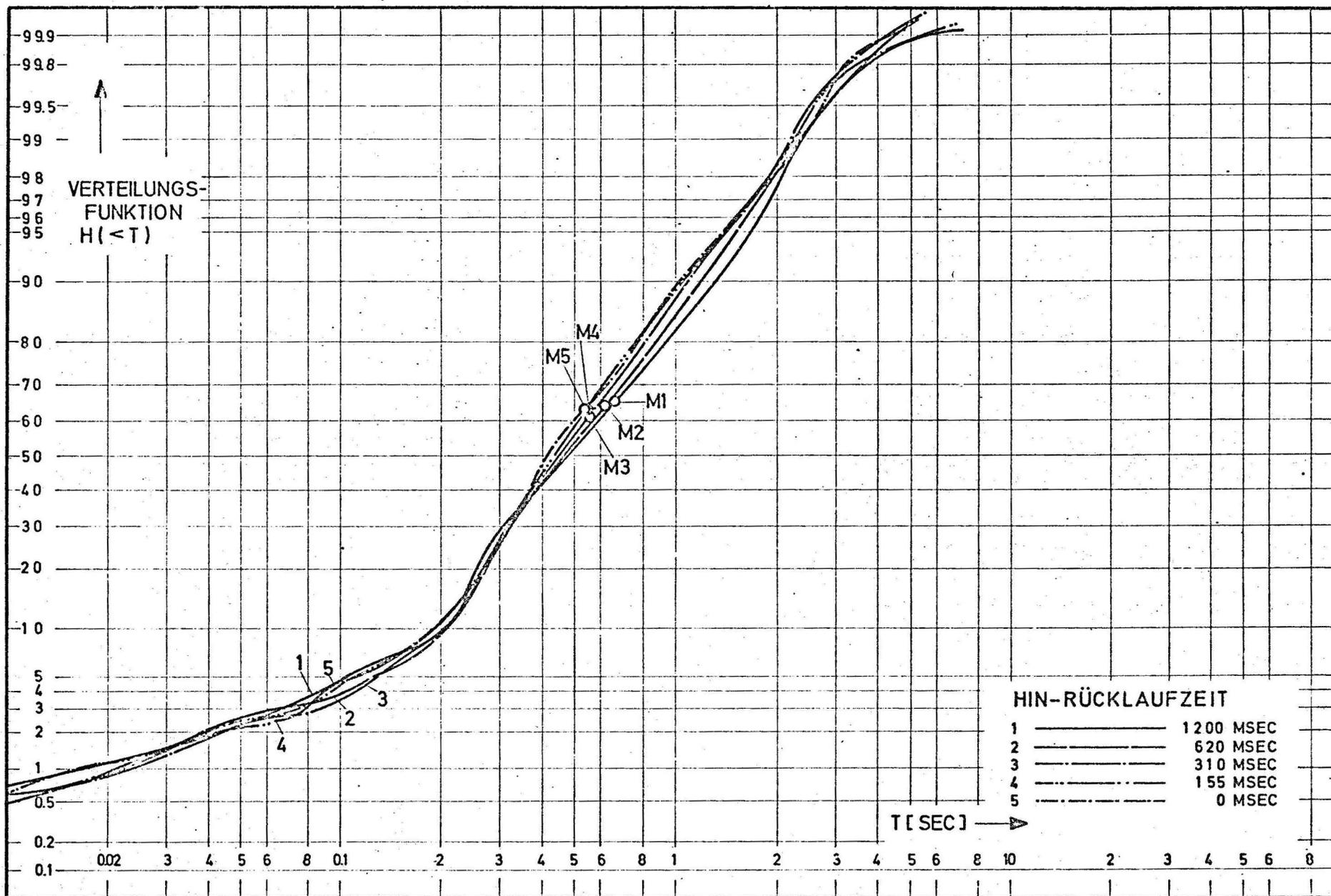


Bild 35: VERTEILUNGSFUNKTIONEN DER ZEITEN DES GLEICHZEITIGSCHWEIGENS BEI $37 \times 2 \frac{1}{4}$ MIN DES GESPRÄCHS ZWEIER PERSONEN JE LAUFZEITSTUFE, ERRECHNET FÜR DAS „FERNE“ LEITUNGSSENDE AUS DEN ERGEBNISSEN FÜR DAS „NAHE“ LEITUNGSSENDE BEI MASCHINELLER ABSCHNITTSERKENNUNG. HIN-RÜCKLAUFZEIT 0.....1200 MSEC.

Anhang

Sortierverfahren

Zweck dieser kurzen Darstellung ist es, die Einordnung des im Bericht mitgeteilten Verfahrens zu erleichtern.

Man unterscheidet externes und internes Sortieren. Zu den externen Sortierverfahren gehört z.B. das für große Datenmengen sehr effektive Mischen von Listen, auch Trennen und Mischen /9/ oder Verfahren der "Collation" /12/ genannt, das nach /9/ von v. NEUMANN entwickelt und begründet wurde.

Fast alle internen Sortierverfahren meint z.B. WALNUT /13/ auf drei Verfahren zurückführen zu können: auf Auswahl, Auswechseln und Suchen.

Auswahl:

Das "replace sorting" /13/ S. 258, das Auswahl-sortieren ist am bekanntesten. Die Liste der Informationseinheiten wird durchgemustert, und die Einheit mit dem kleinsten oder größten Sortierkriterium wird bei n Zahlen in $(n-1)$ Vergleichen ermittelt. Sie wird an den Schluß oder Anfang der Liste gesetzt. Dann wird das Verfahren auf die verbleibenden $n-1$ Informationseinheiten angewendet usw. Man benötigt hierbei insgesamt $n \cdot (n-1) / 2$ Vergleiche. Der Speicherplatzbedarf ist gering. Sortiert man nur die Adressen der Informationseinheiten, so benötigt man nur zwei Zusatzzellen, in die das zu vergleichende Sortierkriterium und seine Adresse übernommen werden. Das Verfahren ist bei kurzen Listen bis ca 100 Einheiten zweckmäßig /10/ S. 12 - 25, bei längeren Listen ist es jedoch zu langsam.

Auswechseln

Das "exchange sorting", Sortieren durch Auswechseln sei etwas ausführlicher besprochen, da das schließlich verwendete Verfahren prinzipiell dazugehört. Man vergleicht bei

diesem Verfahren jeweils zwei benachbarte Sortierkriterien und vertauscht sie gegebenenfalls, um sie in die "richtige" Reihenfolge zu bringen.

Mit Hilfe der Algebra läßt sich das wie folgt beschreiben: HORNFECK /8/ S. 31: "Intuitiv glaubt man zu wissen, daß man jede Anordnung der Zahlen $1, 2, \dots, n$, durch sukzessives Vertauschen von je zweien herstellen kann. Dies bestätigt der Satz: Jede Permutation $f \in S_n, n \geq 2$, läßt sich als Produkt $t_k t_{k-1} \dots t_1$ von Transpositionen schreiben.

Sind $f = t_k t_{k-1} \dots t_1$ und $f = t'_1 t'_{1-1} \dots t'_1$ zwei derartige Darstellungen, so sind k und 1 entweder beide gerade oder beide ungerade".

Eine solche Darstellung im von uns gewünschten Sinne liegt vor, wenn jedes t_k die Vertauschung zweier vorher benachbarten Zahlen bewirkt.

An einem Beispiel aus /13/ S. 263 sei dies erläutert. Die Aufgabe lautet: Die Zahlen $10, 5, 13, 7$ sind in absteigender Reihenfolge zu ordnen.

Es sei $f = \begin{pmatrix} 5 & 7 & 10 & 13 \\ 10 & 5 & 13 & 7 \end{pmatrix}$ die Permutation, die die Elemente in die Ausgangsreihenfolge $10, 5, 13, 7$ überführt, und $h = \begin{pmatrix} 5 & 7 & 10 & 13 \\ 13 & 10 & 7 & 5 \end{pmatrix}$ diejenige, die sie in die fallend geordnete Endreihenfolge $13, 10, 7, 5$ überführt. Gesucht ist diejenige Permutation $t = \dots t_2 t_1$, die die Gleichung $h = t \cdot f$ erfüllt (die Multiplikation der Permutationen werde ihrem Charakter als Abbildung gemäß von recht nach links ausgeführt) und die derart in die Transpositionen $t'_i t'_{i-1} \dots$ zerlegbar ist, daß $g_1 = t_1 f$, $g_2 = t_2 g_1$ usw. .. ist, wobei t_1 zwei in f benachbarte Elemente vertauscht, t_2 zwei in g_1 benachbarte usw.

Die Lösung lautet im vorgelegten Beispiel:

$$t_1 = (5 \ 13) \quad \text{denn} \quad g_1 = t_1 \cdot f$$
$$\begin{pmatrix} 5 & 7 & 10 & 13 \\ 10 & 13 & 5 & 7 \end{pmatrix} = (5 \ 13) \begin{pmatrix} 5 & 7 & 10 & 13 \\ 10 & 5 & 13 & 7 \end{pmatrix}$$

entsteht aus f durch Vertauschen der benachbarten Ziffern 5 und 13, und $t_2 = (10\ 13)$ vertauscht die nun in g_1 und $g_2 = t_2 g_1$ benachbarten Ziffern 10 und 13. Schließlich überführt $t_3 = (5\ 7)$ die Permutation g_2 in die gewünschte $h = t_3 g_2$, indem die in diesen benachbarten Ziffern 5 und 7 vertauscht werden.

$h = t_3 \cdot t_2 \cdot t_1 \cdot f = t \cdot f$, worin $t_3 t_2 t_1 = (5\ 7)(10\ 13)(5\ 13) = t$ auch geschrieben werden kann als $t = \begin{pmatrix} 5 & 7 & 10 & 13 \\ 10 & 5 & 13 & 7 \end{pmatrix}$, denn 5 wird durch t_1 in 13 überführt und 13 durch t_2 in 10, also wird 5 in 10 überführt usw. In diesem Beispiel ist die gesuchte Permutation t gleich der Ausgangspermutation $t = f$, und sie läßt sich durch drei Transpositionen jeweils benachbarter Ziffern darstellen.

Bezeichnet man bei allen möglichen Vergleichen von n Zahlen - das sind $n(n-1)/2$ Vergleiche, d.h. bei $n = 4$, $4 \cdot 3/2 = 6$ Vergleiche - jeden Fall, in dem eine größere Zahl vor einer kleineren steht, als Inversion, dann enthält die Ausgangsreihenfolge 10 5 13 7 drei Inversionen, nämlich $10 > 5$, $10 > 7$, $13 > 7$, die Endreihenfolge jedoch alle 6 möglichen. Die Zahl der fehlenden drei Inversionen ist gleich der Zahl der hier verwendeten Transpositionen t_3, t_2, t_1 .

Wird als Resultat eine fallende Reihenfolge gewünscht, so ist die Zahl der Transpositionen mit der Eigenschaft, nur jeweils zwei benachbarte Zahlen zu vertauschen, gleich der Zahl der Inversionen.

Als Beispiel sei die Zahlenreihe 5 7 3 4 aufsteigend zu ordnen. Sie enthält vier Inversionen $5 > 3$, $5 > 4$, $7 > 3$, $7 > 4$. Mit den gleichen Bezeichnungen wie oben findet man:

$$f = \begin{pmatrix} 3 & 4 & 5 & 7 \\ 5 & 7 & 3 & 4 \end{pmatrix}$$

$$h = \begin{pmatrix} 3 & 4 & 5 & 7 \\ 3 & 4 & 5 & 7 \end{pmatrix} = e \text{ (Einheitselement der Permutationsgruppe)}$$

$$h = e = t f = f f \quad \text{und} \quad t = f = (5\ 4)(4\ 7)(3\ 5)(3\ 7) =$$

$$t_4 t_3 t_2 t_1 .$$

Eine andere Zerlegung in Transpositionen ist z.B.

$t = (4\ 7)(3\ 5) = t'_2\ t'_1$. Diese Transpositionen vertauschen jedoch nicht benachbarte Elemente, ihre Anzahl 2 ist nicht gleich derjenigen der Inversionen, jedoch ist sie gerade und gehorcht damit dem zweiten Teil des mitgeteilten Satzes aus /8/.

Beschreibung des Sortierens durch Auswechselln nach /13/

Erster Durchgang

US	10	10	10	10	13
)	13
	5	5	13	13	10
)	13		
	13	13	5	5	5
)				
	7	7	7	7	7

Zweiter Durchgang

S	13	13	13	[13]
)
US	10	10	10	10
)
	5	7	7	7
)	7		
	7	5	5	[5]

Dritter Durchgang

S	13	[13]	[13]
	10	10	10
))
US	7	7	7
)		
	5	[5]	[5]

-) Vergleich
- = Auswechselln
- US Unsortiert
- S Sortiert

Bild 15

Sortieren durch Auswechselln in mehreren Durchgängen

Die Vorgehensweise erklärt sich aus Bild 15.

Im ersten Durchgang wird die Zahl auf dem letzten Platz (7) mit dem auf dem vorhergehenden (13) verglichen. $13 > 7$, daher findet kein Austausch statt. Wenn der tiefer stehende Zahlenwert größer als der davorstehende ist, wird ausgetauscht. In dieser Weise werden alle benachbarten Sortierkriterien miteinander verglichen, bis die Spitze der Liste erreicht ist. War bei dem letzten Durchgang mindestens ein Austausch vorgenommen worden, so folgt ein weiterer Durchgang.

Man erkennt, daß die Zahl der Durchgänge und damit der Aufwand des Verfahrens davon abhängt, wieweit die Daten bereits geordnet sind, daß es also bei weitgehend geordneten Daten besonders vorteilhaft ist. Da am Ende eines Durchganges die größte Zahl stets an der Spitze der noch nicht geordneten Liste steht, kann man deren Länge nach jedem Durchgang um einen Platz verkürzen. Dadurch entfallen in Bild 15 die drei eingeklammerten Vergleiche.

Suchen:

Beim dritten Verfahren, dem Suchsortieren, bedient man sich zweier Listen. Jede Zahl der unsortierten Liste wird in die zweite Liste gebracht und mit Hilfe binärer Suchalgorithmen zwischen den dort schon vorhandenen Zahlen eingeordnet. Alle Zahlen, die kleiner als die neue Zahl sind, werden dann einen Platz hinuntergerückt. Das Verfahren benötigt sehr wenig Vergleiche, jedoch Platz für eine neue Liste und viel Verschiebungen zwischen den Speicherplätzen.

Vergleich der Verfahren

Aus der hier gegebenen Übersicht läßt sich in ganz überschlägiger Abschätzung sagen, daß Suchsortierverfahren beim Vorhandensein binärer Unterprogramme nützlich sind, bei großen Datenmengen jedoch wegen der zahlreichen Verschiebungen ausscheiden. Das Sortieren durch Auswahl eignet sich ebenfalls nur für kurze Listen.

Die Zahl der Vertauschungen bei Auswahl und Tauschverfahren ist gleich der Zahl der Inversionen oder deren Differenz zu $n(n-1)/2$. Obwohl Vertauschungen wesentlich mehr Zeit erfordern als Vergleiche, ist ihre Anzahl meist gering gegen die der Vergleiche. Man beschränkt sich bei Aufwandsabschätzungen meist auf die Vergleiche.

Die Zahl der Vergleiche beim Auswahlverfahren ist stets $n \cdot (n-1)/2$. Beim Tauschverfahren liegt sie zwischen $n(n-1)/2$ und $n-1$ je nach vorher vorhandener Ordnung, bei weitgehender Ordnung also erheblich unter derjenigen des Auswahlverfahrens.

Literaturverzeichnis

- /1/ Böttcher, K. Erfassung und Auswertung von Meßdaten.
Diplomarb. am Inst. f. Fernmeldetechnik
der TUB, ausgeg. 12.12. 67
- /2/ Brady, P.T. A Technique for Investigating ON-OFF
Patterns of Speech
BSTJ Bd. 44 , 1965, S. 2 ff
- /3/ Brady, P.T. A Statistical Basis for Objective
Measurement of Speech-Levels
BSTJ Bd. 44, 1965, S. 1453 ff
- /4/ Chapin, Ned. Einführung in die elektronische Daten-
verarbeitung
Oldenbourg, Wien 1967
- /5/ Giloi-Lauber Analogrechnen
Springer Berlin 1963, S. 59, 80
- /6/ Höhne, H.D. Influence of Long Transmission Delay
and Reverberation on the Telephone
Conversation of Testpersons. In Reports
on the 4th Int. Symp. on Hum. Fact
in Telephony
VDE Verlag Berlin, 1970, S. 7 - 28
- /7/ Höhne, H.D. Einfluß der Übertragungslaufzeit auf
die Kommunikation von Versuchspersonen.
Dissertation TUB eingereicht 1970.
- /8/ Hornfeck Algebra
1969, de Gruyter, S. 31
- /9/ Lutz, T.u. Hauf, V. Programmierfiabel
Franck'sche Vol., Stuttg. 1969, S. 236
- /10/ Rechenberg, P. Scriptum der Vorlesung Prog. II WS 70/
S. 12 - 22 bis 12 - 33
- /11/ Stevens, S. Handbook of Experimental Psychologie.
1967, Wiley Inc. New York, S. 1059

/12/ Thüring, B.

Einführung in die Methoden der Programmierung kaufmännischer u. wissenschaftlicher Probleme für elektronische Rechenanlagen.

Göller, Baden-Baden, 1957

/13/ Walnut, F.K.

Introduction to Computer Programming and Coding.

Prentice-Hall Inc., S. 258 - 274, 1968

/14/ Zerboni, H,v.

Entwurf und Aufbau einer Einrichtung zur Erkennung des Vorhandenseins von Sprache.

Studienarbeit am Inst. f. Fernmelde-technik der TUB, ausgeg. 1.9. 67

