

Bedienungsanleitung

SPRACHEINGABE -MODUL 690 032.4

zum KLEINCOMPUTER robotron Z 9001
robotron KC 8511
robotron KC 87

Inhalt

1.	Verwendungszweck	3
2.	Lieferumfang und Ergänzungsteile	3
3.	Technische Daten	3
4.	Wirkungsweise	5
4.1.	Hardware	5
4.2.	Software	5
5.	Hinweise für die Inbetriebnahme	6
5.1.	Einbauanleitung	6
5.2.	Inbetriebnahme	6
5.3.	Arbeit mit dem Sprechzeug SP 75	6
6.	Arbeit mit dem SPRACHEINGABE-MODUL	6
6.1.	Grundlagen	6
6.2.	Bedienung	7
6.3.	Anlegen eines neuen Wortschatzes	8
6.4.	Arbeiten mit vorhandenem Wortschatz	9
6.5.	Wiedererkennung	10

1/87 a

Jt 1130/87

VEB ROBOTRON-MESSELEKTRONIK > OTTO SCHÖN < DRESDEN
Lingnerallee 3, Postschließfach 211, Dresden, DDR-8012

- 3 -

1. Verwendungszweck

Der SPRACHEINGABE-MODUL 690 032.4 (im folgenden Text SEM genannt) wandelt analoge Signale (z.B. Sprache über Mikrofon) in digitale Signale um, die vom Kleincomputer weiterverarbeitet werden können. Im Falle der Spracheingabe dient der Modul der Wiedererkennung von einzeln gesprochenen Worten von einem systembekannten Sprecher. Er kann als sogenannte "phonetische Tastatur" überall dort eingesetzt werden, wo herkömmliche Tastaturen auf Grund der Einsatzbedingungen nicht geeignet oder wenig effektiv sind. Solche Einsatzfälle sind beispielsweise Arbeitsplätze, an denen die Hände und der Gesichtssinn des Bedieners durch andere Aufgaben gebunden sind (z. B. grafische und Mikroskopiearbeitsplätze) bzw. Beleuchtung, Witterungsverhältnisse oder Verschmutzung den Betrieb einer Tastatur erschweren.

2. Lieferumfang und Ergänzungsteile

2.1. Lieferumfang

1 SPRACHEINGABE-MODUL 690 032.4 ZAK-Nr. 138 21 78 005 000419

mit:

1 SPRECHZEUG SP 75 Sachnummer: 536 536.8
(vorn VEB Funkwerk Kölleda - modifiziert)

1 Magnetbandkassette R 0113 mit Software zum Betreiben des
SPRACHEINGABE-MODULs Sachnummer: 536 537.6

1 Bedienungsanleitung

1 Garantieurkunde

2.2. Ergänzungsteile (nicht im Lieferumfang enthalten)

- RAM-ERWEITERUNGS-MODUL 690 003.5 ZAK-Nr. 138 21 78 005 000136 (ist für das Betreiben des SPRACHEINGABE-MODULS notwendig)

- Dytimische Hör-Sprach-Garnitur DHS 66,
ZAK-Nr. 137 34 75 008 002027, vom VEB Funktechnik Leipzig bzw. deren einzelnes Mikrofon DM 1911 können als Ersatz für das zum Lieferumfang gehörende Sprechzeug SP 75 eingesetzt werden (siehe Abschnitt 3, Schalterstellung auf Modul).

3. Technische Daten

Wortschatzgröße 50 Wortklassen, sind maximal möglich
(abhängig vom Sprechverhalten, da die Anzahl der Referenzmuster im Speicher begrenzt ist) Eine geringere Wortschatzgröße erhöht die Erkennungswahrscheinlichkeit (optimale Wortschatzgröße bis 25 Wortklassen)

Anzahl der Referenzmuster	max. 200 (im Arbeitsspeicher des Kleincomputers abgelegt; begrenzt durch Rechenzeit und Programmaufbau)
Wortlänge	0,2 bis 1,8 s (Wortlängen unter 0,2 s werden als Störgeräusch gewertet und unterdrückt ; längere Worte werden nach 1,8 s abgebrochen)
Reaktionszeit nach Wortende	0,2 s
Pause zwischen 2 Worten	minimal 0,2 s

Bei der Arbeit mit 20 Wortklassen ergibt sich für den geübten Sprecher folgende Erkennungswahrscheinlichkeit:

- Erkennungsquote bei kooperativem Sprechverhalten : > 90 %
- Rückweisung und Fehlererkennung verteilen sich auf ca. gleiche Teile der verbleibenden ca. 10 %

Herausgeführte Leitungen zum Steckverbinder X2 des Anschlußkabels :

X2:1	Eingang für Mikrofon
X2:2	Masse
X2:3	Ausgang für Hörer
X2:4	12 V Betriebsspannung für Mikrofonvorverstärker des SP 75

Verwendete Portadressen

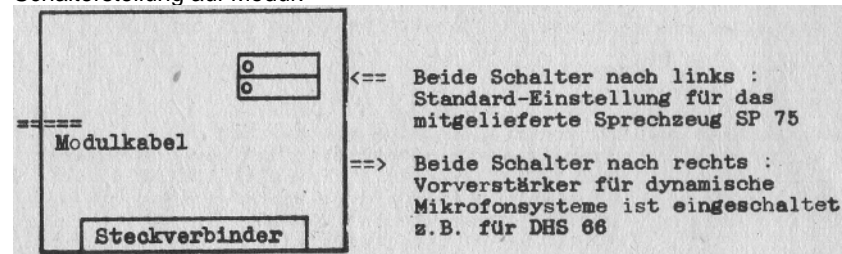
CTC :	E0H bis E3H
LED ein :	E4H (bis E7H)
LED aus :	E8H (bis EBH)

Belegter Speicherbereich

Programm SPRACHE1:
6000 H bis 6E8FH (bzw. bis 7CFFH einschließlich max. Wortschatz)

Programm SPRACHE2:
9800H bis A68FH (bzw. bis B50CH einschließlich max. Wortschatz)

Schalterstellung auf Modul:



4. Wirkungsweise

4.1. Hardware

Das analoge Signal des Mikrofons der Hör-Sprech-Garnitur gelangt über das Modulkabel zur Eingangsschaltung des SPRACHEINGABE-MODULS 690 032.4 . Hier ist das Zuschalten eines Vorverstärkers für dynamische Mikrofonsysteme möglich. Das zum Modul mitgelieferte Sprechzeug SP 75 besitzt ein piezokeramisches Mikrofon mit eigenem Vorverstärker. Dieser Vorverstärker wird über die im SEM bereitgestellten 12 V mit Betriebsspannung versorgt. Beim Betreiben des SEM mit dem mitgelieferten Sprechzeug SP 75 ist das Abschalten des modulinternen Vorverstärkers (entspricht dem Auslieferungszustand, siehe Abschnitt 3) notwendig. Das Eingangssignal wird nach einem einstellbaren Spannungsteiler einer Verstärkerstufe zugeführt. Nach der Verstärkung wird das Signal durch einen Hoch- und Tiefpaß aufgespalten. Die Schnittfrequenz der Frequenzweiche liegt bei 1. kHz. In anschließenden Begrenzerstufen werden, für Hoch- und Tiefpaß getrennt, TTL-gerechte Signale erzeugt und einem CTC-Baustein U 857 übergeben. über diesen CTC-Baustein ist vom Rechner des Kleincomputers die Abfrage des gewandelten Analogsignals möglich.

4.2. Software

Zum SPRACHEINGABE-MODUL 690 032.4 wird die Magnetbandkassette R 0113 mitgeliefert. Auf dieser Kassette sind 2-Programme zum Betreiben des Moduls auf unterschiedlichen Speicheradressen vorhanden. Das Programm SPRACHE1 wird in den Speicherbereich der 1. RAM-Erweiterung ab 4000H geschrieben (Programm SPRACHE2 in den Bereich der 2.RAM-Erweiterung). Für den Interruptbetrieb des CTC-Bausteines im SEM ist dieser Speicherbereich notwendig (Interruptbehandlungsprogramm muß außerhalb des Grund-RAMs des Kleincomputers geladen sein).

Die beiden CTC-Kanäle für Hoch- und Tiefpaß arbeiten als Zähler mit der Zeitkonstante 1, d.h., bei jedem Nulldurchgang des Analogwertes wird durch Hoch- oder Tiefpaß ein Interrupt ausgelöst. Der Interruptabstand ist ein Maß für die momentane Frequenz des Analogsignals. Die zeitliche Folge dieser Frequenzen stellt nach interner Umrechnung und Normierung ein REFERENZMUSTER dar, welches im Lernvorgang nach Zuweisung einer KLASSENNUMMER als MUSTER im WORTSCHATZ (im Arbeitsspeicher des Kleincomputers) abgelegt werden kann. Beim Wiedererkennen wird dieses Referenzmuster zur Wortsuche nach bestimmten Ähnlichkeitsmerkmalen herangezogen. Dazu muß der entsprechende Wortschatz im Arbeitsspeicher des Kleincomputers vorhanden sein.

Das Programm SPRACHE1 (oder SPRACHE2) ermöglicht:

- Laden und Abspeichern eines Wortschatzes
- Lernen (ohne oder mit vorhandenem Wortschatz)
- Korrekturen im Wortschatz (Umbenennen oder Löschen von Wortklassen)
- Wiedererkennen von Worten

5. Hinweise für die Inbetriebnahme

5.1. Einbauanleitung

Vom Modulschacht des Kleincomputers ist die Abdeckung zu entfernen. Die Zugenlastung 690 018.0 (gehört nicht zum Lieferumfang) wird in die seitlichen Halterungsschlitze eingerastet. Bei ausgeschaltetem Kleincomputer wird der SPRACHEINGABE-MODUL 690 032.4 auf den 1. Steckplatz (unmittelbar hinter der Tastatur; ist für die Interruptbehandlung notwendig) gesteckt. Auf einen beliebigen anderen Steckplatz ist ein RAM-MODUL 690 003.5 mit Moduladresse 4000H zu stecken (bei Programm SPRACHE2 ist zusätzlich RAM-Speicher ab der Adresse 8000H notwendig). Der verbleibende Speicherbereich bzw. Platz im Modulschacht des Kleincomputers kann nach Belieben genutzt werden.

5.2. Inbetriebnahme

Nach dem Anschließen des Sprechzeuges SP 75 kann das Steuerprogramm aus dem Betriebssystem-Modus mit dem Kommando:

CLOAD SPRACHE1 (nicht selbst startend) oder
SPRACHE1 (automatischer Programmstart)

in den RAM des Kleincomputers eingeladen werden.

5.3. Arbeiten mit dem Sprechzeug SP 75

Die Erkennungswahrscheinlichkeit des Spracheingabe-Systems hängt ganz wesentlich von einer reproduzierbaren Schallaufnahme ab. Im Falle der Spracheingabe werden an deren Reproduzierbarkeit hohe Anforderungen gestellt. Der Mikrofonabstand ist demzufolge bei der LERN- und ERKENNER-Phase wie folgt einzustellen:

Abstand Mikrofon -- Mund: 1 bis 2 cm
Mikrofon ca. 30 Grad aus der direkten Schallrichtung weggedreht

6. Arbeit mit dem SPRACHEINGABE-MODUL

6.1. Grundlagen

Bevor mit dem SEM gearbeitet werden kann, muß ein sogenanntes LERNEN absolviert werden. Dabei müssen alle Worte mehrmals vorgesprochen werden. Das SPRACHE-Programm legt diese Worte als MUSTER im WORTSCHATZ an. Infolge von Ausspracheunterschieden werden je Wort mehrere Muster angelegt. Sie bilden eine KLASSE und haben eine gemeinsame KLASSENNUMMER.

Als Klassennummern sind die Zahlen 1 bis 50 zugelassen.
Der Wortschatz kann insgesamt maximal 200 Muster aufnehmen.

Der Wortschatz ist in 5 voneinander unabhängige TEILWORTSCHATZE mit den Nummern 1 bis 5 eingeteilt.

Diese Teilwortschätze können verwendet werden, um:

- unterschiedliche Aussprachevarianten eines Sprechers (Lernen an verschiedenen Tagen) aufzunehmen oder
- voneinander unabhängige Teilwortschätze für verschiedene Sprecher anzulegen (maximale Wortschatzgröße beachten !)

Beim Lernen wird in dem betreffenden Teilwortschatz ermittelt, wie oft ein Muster verwendet worden ist. Wenig benutzte Muster (Versprecher, Fremdgeräusche) werden nach einiger Zeit durch das SPRACHE-Programm selbständig gelöscht (für den Bediener nicht direkt bemerkbar).

Bei der Wiedererkennung ist anzugeben, ob im Gesamtwortschatz bzw. in welchem Teilwortschatz gesucht werden soll (siehe folgende Abschnitte).

6.2. Bedienung

Nach dem Start des Programms SPRACHE1 (oder SPRACHE2) erscheint auf dem Bildschirm folgendes Menü:

Musteranzahl im Wortschatz	0
Eingestellter Teilwortschatz	1

- A Wortschatz laden
- B Lernen
- C Wortschatz abspeichern
- D Teilwortschatz ändern
- E Klassen umbenennen
- F Klassen löschen
- G Wortschatz löschen
- STOP Programmende

Kennbuchstabe:

Nach Drücken des entsprechenden Kennbuchstabens und Quittieren mit **ENTER** (außer bei **STOP**-Tastendruck) wird das gewünschte Teilprogramm aufgerufen.

Alle Eingaben erfolgen im Bildschirmdialog. Es kann jederzeit, mit **☐**, **DEL** und **CLLN** korrigiert werden, bis **ENTER** als Abschluß gedrückt wird.

Durch **STOP**-Tastendruck kann bei jeder Eingabe der momentane Programmabschnitt abgebrochen und in die jeweils übergeordnete Programmebene zurückgekehrt werden.

6.3. Anlegen eines neuen Wortschatzes

Nach dem Start des SPRACHE-Programms (1 oder 2) ist standardmäßig der Teilwortschatz 1 eingestellt. Soll der Lernvorgang in einem anderen Teilwortschatz erfolgen (z.B. Teilwortschatz 3), so kann dieser wie folgt umgestellt werden:

Bildschirmausschrift	Eingaben
Kennbuchstabe:	D ENTER
Teilwortschatz 1	3 ENTER
Im Menü erscheint:	Teilwortschatz 3

Nach dem Aufsetzen des Sprechzeugs SP 75 (siehe Abschnitt 5.3) kann der eigentliche LERN-Vorgang beginnen. Es ergibt sich folgender Bildschirmdialog:

Bildschirmausschrift	Eingaben	Bemerkungen
Kennbuchstabe:	B ENTER	Teilprogramm LERNEN starten
Teilwortschatz 1		Anzeige zur Kontrolle
Sprechen:	1. Wort über Mikrofon eingeben	
Erkannte Klasse: 0		d.h., keine gültige Wortklasse erkannt
	(z.B.) 1 ENTER	Zuweisung der Klassennummer 1
Sprechen:	nächstes Wort über Mikrofon eingeben	
Erkannte Klasse: 0	(z.B.) 2 ENTER	Zuweisung der Klassennummer 2

Nach dem Sprechen eines Wortes wird die erkannte Klassennummer angezeigt. Sie kann korrigiert (durch Eingabe von 'Zahl' ENTER) oder bestätigt werden (nur ENTER). Die Eingabe von 0 ENTER bewirkt das Ignorieren des zuletzt gesprochenen Wortes (z.B. Störgeräusche, Versprecher usw).

Die Zuweisungen der Klassennummern müssen sehr gewissenhaft erfolgen. Jede falsch zugewiesene bzw. falsch bestätigte Klassennummer führt zu einem Fehler im Wortschatz. Das betreffende Wort kann nicht mehr eindeutig einer bestimmten Klasse zugeordnet werden. Die Folge sind erhöhte Fehlerkennungsraten.

Im Fehlerfall kann eine Wortschatzkorrektur durch Löschen einzelner Klassen (Kennbuchstabe F aus dem Menü aufrufen) erfolgen.

Die vorgesehenen Worte werden, wie im Beispiel gezeigt, mehrmals nacheinander über das Mikrofon des Sprechzeugs in den SEM eingegeben. Worte, die schon in der Lernphase ständig zu Verwechslungen führen, sind ungeeignet (sprecherabhängig) und sind durch geeignete andere Worte zu ersetzen.

Durch STOP-Tastendruck läßt sich der Lernvorgang an beliebiger Stelle unterbrechen. Bei der Rückkehr in das Menü wird die aktuelle Musteranzahl des Wortschatzes (Summe über alle Teilwortschatze) angezeigt.

Der Lernvorgang kann beendet werden, wenn eine zufrieden stellende Erkennbarkeit erreicht ist und die Musteranzahl im Wortschatz einen stabilen Wert angenommen hat bzw. leicht rückläufig ist.

Der angelernte Wortschatz kann auf Tonbandkassette wie folgt ausgeladen werden:

Bildschirmausschrift	Eingaben	Bemerkungen
Kennbuchstabe:	C ENTER	
Filename:	(z.B.) W O R T E 1 ENTER	
start tape	ENTER	Wortschatz wird abgespeichert

Eine Kontrolle des abgespeicherten Wortschatzes ist über Verify wie bei anderen Programmen möglich.

6.4. Arbeiten mit vorhandenem Wortschatz

Nach dem Start des Programms SPRACHE1 (oder SPRACHE2) kann ein vorhandener Wortschatz wie folgt eingeladen werden:

Bildschirmausschrift	Eingaben	Bemerkungen
Kennbuchstabe:	0 ENTER	
Filename:	(z.B.) W O R T E 1 ENTER	
start tape	ENTER (bei Vorton auf Tonbandkassette) drücken;	nach Programmende erfolgt ein Rücksprung in das Menü

Nach dieser Handlung steht der Wortschatz zur Nutzung bereit. Ein Nachlernen (Anpassen des Wortschatzes an die momentanen Umgebungsbedingungen und die Tagesform des Sprechers) ist jedoch unbedingt zu empfehlen. Dieses Nachlernen erfolgt über Kennbuchstaben B, wie in Abschnitt 6.3 beschrieben (Achtung: zuvor Teilwortschatz wie beim Lernen einstellen!). Der aktualisierte Wortschatz kann erneut abgespeichert werden (siehe Abschnitt 6.3).

6.5. Wiedererkennung

Das Programm SPRACHE1 (bzw. SPRACHE2) enthält die Routine zum Suchen des Erkennungsergebnisses. Das entsprechende Programm 1 oder 2 ist in den Arbeitsspeicher zu laden. Nachdem über daß Menü des SPRACHE-Programms der gewünschte Wortschatz geladen wurde und ein Nachlernvorgang (siehe Abschnitt 6.4) erfolgte, muß das SPRACHE-Programm durch erneuten STOP-Tastendruck verlassen und in das Betriebssystem zurückgekehrt werden. Von hier aus kann das Anwenderprogramm gestartet werden.

Ein Aufruf des Erkennerprogramms setzt sich aus folgenden Teilen zusammen:

- Einstellen des Teilwortschatzes, in dem gesucht werden soll

TW=255 (bzw. FFH) = Standardwert und bedeutet Suchen im gesamten Wortschatz

TW=1 (bis TW=5) bedeutet, daß in Teilwortschatz 1 (bzw. 2 oder 3 ... 5) nach dem Erkennungsergebnis gesucht wird.

- Aufrufen des Erkennerprogramms A oder B

Erk.-Progr. A : Startadresse in Zelle 76 (bzw. 4CH)
 Nach Aufruf dieses Erkennerprogramms wird auf ein gültiges Wort gewartet, d.h., bei Rückweisung (Klassennummer 0) wird ein BEEP-Signal ausgegeben und weiter auf ein gültiges Wort gewartet.

Erk.-Progr. B : Startadresse in Zelle 78 (bzw. 4EH)
 Kommt nach jedem erkannten Wort mit der Klassennummer (auch mit Klassennummer 0) zurück.

- Abholen des Erkennungsergebnisses von Adresse 81 (bzw. 51H)
- Fehlerauswertung durch Testen von Bit 7 des Erkennungsergebnisses (Bit 7 = 1 -> Abbruch der Erkennung durch STOP)

Anwenderbeispiele für einen Erkenneraufruf:

a) in BASIC:

```

10 TW=1: POKE 80, TW      ;! Teilwortschatz einstellen
20 !                      ;( nur bei Bedarf)
30 CALL DEEK(76)         ;! Aufruf des Erkennerprogramms A
40 ER=PEEK(81)           ;! Erkennungsergebnis
50 IF ER>127 THEN GOTO xxx: ! Fehlerbehandlung

```

b) in Assemblersprache:

```

LD      A, 1
LD      (50), A           ; Teilwortschatz 1 einstellen
                          ; ( nur bei Bedarf)

LD      HL, MARK
PUSH    HL               ; RETURN- Adresse
LD      HL, (4CH)
JMP     (HL)             ; Aufruf des Erkennerprogramms A
MARK:   LD      A, (51H)  ; Erkennungsergebnis
        BIT     7, A
        JPNZ   FEHLER   ; Fehlerbehandlung

```