

## **Zum Einfluß mnestischer, semantischer und konzeptueller Faktoren auf die Leistungen aphasischer Patienten im Token Test\***

Ewald Naumann, Stephanie Kelter und Rudolf Cohen

Sonderforschungsbereich 99 „Linguistik“ Universität Konstanz, Bundesrepublik Deutschland

### **Influence of Mnemonic, Semantic, and Conceptual Factors on the Performance of Aphasic Patients in the Token Test**

**Summary.** Three experimental modifications of the Token Test were presented to 40 aphasics with predominantly expressive disturbances and 40 control patients with diffuse brain damage. On each token a human figure was drawn (a man, woman, boy, or girl); the size, color, and form of the token were given in the instructions, or the depicted figure was given as the target either in addition to or instead of the size and form. The mere fact that a figure was depicted on the tokens did not affect the performance of the patients. However, when the figures were mentioned in the instructions, the aphasics' performance improved significantly, although it never reached the level of the controls. An analysis of the error types suggested that the poor performance of the aphasics in the Token Test is at least in part due to difficulties identifying complex stimuli when the features are separately named in the instructions.

**Key words:** Aphasia – Token test – Analytical information processing.

**Zusammenfassung.** Je 40 Aphasikern mit vorwiegend expressiven Störungen und Kontrollpatienten mit diffusen Hirnschädigungen wurden drei experimentelle Modifikationen des Token Tests vorgegeben, in denen die Plättchen des Tests mit verschiedenen Figuren beklebt waren und in den Anweisungen entweder wie üblich die Größe, Farbe und Form der Plättchen oder aber zusätzlich bzw. anstelle von Größe und Form die aufgeklebten Figuren ge-

---

\* Diese Untersuchung wurde von der Deutschen Forschungsgemeinschaft unterstützt. Unser Dank gilt den Ärzten, den Logopädinnen, dem Pflegepersonal und nicht zuletzt den Patienten der folgenden Kliniken: Abteilung Neurologie der Medizinischen Fakultät an der Rheinisch-Westfälischen Technischen Hochschule Aachen; Poliklinik für Stimm- und Sprachkranke der HNO-Klinik des Klinikums Steglitz der Freien Universität Berlin; Rheinische Landesklinik für Sprachgestörte Bonn; Neurologische Kliniken Dr. Schmieder, Allensbach und Gailingen

*Sonderdruckanforderungen an:* Ewald Naumann, Fachbereich I — Psychologie, Universität Trier, Schneiderhof, D-5500 Trier, Bundesrepublik Deutschland

nannt wurden. Die Beklebung der Plättchen hatte keinen bedeutsamen Einfluß auf die Leistungen. Demgegenüber führte die Nennung der Figuren in den Anweisungen bei den aphasischen Patienten zu signifikanten Leistungssteigerungen; allerdings wurde auch hier das Niveau der Kontrollpatienten nicht erreicht. Eine Fehleranalyse deutete darauf hin, daß das Versagen der Aphasiker weniger mit Schwierigkeiten bei der Speicherung der sprachlichen Information zusammenhängt, als vielmehr mit spezifischen Beeinträchtigungen bei der Identifikation von komplexen Reizen aufgrund sprachlich getrennt vorgegebener Einzelaspekte.

**Schlüsselwörter:** Aphasie – Token Test – Analytische Informationsverarbeitung.

## Einleitung

Der von De Renzi und Vignolo (1962) entwickelte Token Test (TT) hat sich in zahlreichen Untersuchungen als ein hervorragendes Verfahren zur Erfassung aphasischer Störungen erwiesen (z. B. Cohen et al. 1976; De Renzi u. Faglioni 1978; Orgass 1976a, b). So erfreulich die hohen Trefferquoten von über 80% unter diagnostischen Kriterien sein mögen, so irritierend sind sie doch unter psycho-pathologischen Gesichtspunkten: Es ist bis heute kaum gelungen herauszufinden, warum dieser simple Test so außerordentlich trennscharf ist.

Die naheliegende und zunächst von vielen Forschern vertretene Auffassung, der TT erfasse primär *Sprachverständnisstörungen*, erscheint — zumindest in dieser globalen Form — nicht mehr haltbar. Weder sind die Leistungsunterschiede zwischen überwiegend motorisch und überwiegend sensorische gestörten Aphasikern so bedeutsam, wie nach dieser These zu erwarten wäre (Cohen et al. 1976; Orgass 1976a, b; Poeck et al. 1972), noch ergaben sich in korrelationsstatistischen Analysen irgendwelche Hinweise darauf, daß der Test mit anderen Sprachverständnistests in einer engeren Beziehung stünde als mit Sprachproduktionsprüfungen (Cohen et al. 1976, 1980). Darüberhinaus erwies sich der Untertest V des TT, den man wegen seiner größeren linguistischen Komplexität oftmals als den eigentlich kritischen Teil angesehen hatte, keinesfalls als besonders trennscharf — im Gegenteil, er scheint der Gesamtvalidität sogar eher etwas abträglich zu sein (Woll et al. 1976). Schließlich differenziert der TT auch dann noch zwischen Aphasikern und nicht-sprachgestörten Hirngeschädigten, wenn er non-verbal vorgegeben wird, d. h., wenn dem Patienten anstelle der verbalen Anweisungen Prototypen von den zu identifizierenden Plättchen gezeigt werden (Brookshire 1978; Cohen et al. 1977).

In alternativen Erklärungsansätzen konzentriert man sich heute im Wesentlichen auf (a) die Anforderungen an das Kurzzeitgedächtnis, (b) den Einfluß semantisch-pragmatischer Faktoren, und (c) die Prozesse der konzeptuellen Verarbeitung.

Für die Hypothese, daß aphasische Patienten deshalb im TT versagen, weil sie besondere Schwierigkeiten haben, sich die zum Teil recht langen Anweisungen zu merken, läßt sich geltend machen, daß eine Reihe von Untersuchungen Hinweise auf eine reduzierte Gedächtnisspanne bei Aphasikern erbrachten (Albert 1976;

De Renzi u. Nichelli 1975; Goodglass et al. 1970; Saffran u. Marin 1975). Zum zweiten ist es plausibel anzunehmen, daß Aphasiker gerade über jene verbal-akustische Kodierungsform, die für das innere Memorieren längerer Sequenzen hilfreich ist, weniger gut verfügen als andere Hirngeschädigte. Die empirische Evidenz für den Zusammenhang zwischen den aphasischen Minderleistungen im TT und mnestischen Beeinträchtigungen ist jedoch gering (Brookshire 1978; Cohen et al. 1976; De Renzi et al. 1978; Lesser 1974).

Zahlreiche Autoren wiesen darauf hin, daß möglicherweise semantisch-pragmatische Faktoren wie z.B. die „mangelnde Redundanz“ und die „geringe Strukturiertheit des Testmaterials“ (Cohen et al. 1976; De Renzi u. Vignolo 1962), die „Künstlichkeit der Situation“ (Leischner 1974) oder die „Abstraktheit des Testmaterials“ (Kreindler et al. 1971; Lesser 1979) für die Minderleistungen der Aphasiker verantwortlich seien. Auch hier liegen bisher allerdings erst wenige einschlägige Untersuchungen vor. Daß die Redundanz der Mitteilung für die Verstehensleistungen von Aphasikern von Bedeutung ist, wurde zwar von Stachowiak et al. (1977) und Gardner et al. (1975) bei Texten bzw. Wörtern nachgewiesen; inwieweit diese Variable jedoch beim TT eine Rolle spielt, ist noch ungeklärt. Im Zusammenhang mit der „Abstraktheit“ und mangelnden Lebensnähe des Testmaterials konnten Kreindler et al. (1971) zeigen, daß Aphasiker ihre Leistungen im TT erheblich steigerten, wenn anstelle der „abstrakten“ Inhalte ‚Kreis‘ und ‚Viereck‘ konkrete Dinge wie ‚Haus‘ und ‚Blume‘ verwendet wurden. Demgegenüber ergaben sich aber in einer Untersuchung von Lesser (1979) keine Unterschiede zwischen den Leistungen in der Originalversion des TT und einer „multisensorischen“ Version, in der dem Patienten anstelle der bunten Plättchen runde und viereckige Stücke von aus dem Alltagsleben vertrauten Gegenständen wie z.B. einem Schwamm, Karton, Stoff usw. vorgelegt wurden.

In einer weiteren Hypothese wird davon ausgegangen, daß die Ursache der aphasischen Minderleistungen im TT im präverbalen, konzeptuellen Bereich zu suchen sei. Nur so sei zu verstehen, warum der TT offenbar genau jenen allgemeinen Faktor sprachlicher Fähigkeit repräsentiert, der sich in allen bisherigen faktorenanalytischen Untersuchungen als indifferent gegenüber den in der Aphasieforschung üblicherweise unterschiedenen Performanzbedingungen erwies (Cohen et al. 1976). Es wird angenommen, daß die Schwierigkeiten der Aphasiker wesentlich darin bestehen, die Einzelaspekte einer vorgegebenen Ganzheit systematisch abzufragen. Für die Annahme einer solchen kognitiven Störung spricht zum einen die bereits erwähnte Tatsache, daß der TT auch noch bei non-verbalen Vorgabe zwischen Aphasikern und anderen Hirngeschädigten diskriminiert, zum anderen die durchweg bedeutsamen Korrelationen zwischen dem TT und solchen non-verbalen Tests, in denen Aphasiker spezifische Ausfälle zeigen (vgl. Cohen et al. 1980). Inwieweit es sich dabei jedoch um die postulierte Beeinträchtigung bestimmter *analytischer* Fähigkeiten handelt, läßt sich auf der Basis der vorliegenden Untersuchungen allerdings noch nicht entscheiden.

Die drei Erklärungsansätze schließen einander nicht aus; es ist ja durchaus denkbar, daß der TT gerade deshalb so trennscharf ist, weil er verschiedene Aspekte der kognitiven Beeinträchtigungen aphasischer Patienten gleichzeitig erfaßt. Ziel der vorliegenden Arbeit war daher nicht so sehr, die einzelnen Hypothesen gegeneinander zu prüfen, als vielmehr, näheren Aufschluß über die

relative Bedeutsamkeit der verschiedenen Faktoren zu gewinnen. Zu diesem Zweck nahmen wir Abwandlungen am TT vor, die zwar nach allen drei Erklärungsansätzen zu Leistungssteigerungen führen sollten, jedoch in je unterschiedlichem Umfang. Die Abwandlungen bestanden im wesentlichen darin, daß die Plättchen mit Figuren — einem Mann, einer Frau, einem Jungen oder einem Mädchen — versehen waren und entweder anstelle der üblichen Anweisung oder zusätzlich zu dieser die entsprechende Figur auf dem Plättchen genannt wurde. Mit der letztgenannten Version, bei der in der Anweisung sowohl die Größe, Farbe und Form des Plättchens, als auch die Figur auf dem Plättchen genannt wurde (z. B. „Zeigen Sie mir bitte den kleinen blauen Kreis mit dem Mädchen!“), sollte geklärt werden, inwieweit eine Erhöhung der Redundanz der Anweisung die Lösung der Aufgabe erleichtert. Demgegenüber sollte die „konkrete“ Version, bei der *anstelle* der Größe und Form des Plättchens die Figur auf dem Plättchen genannt wurde (z. B. „Zeigen Sie mir bitte das blaue Mädchen!“), Aufschluß über die Bedeutsamkeit der Konkretheit des Materials, der Belastung des Kurzzeitgedächtnisses und der Beanspruchung konzeptuell-analytischer Fähigkeiten geben. Dabei gingen wir von den folgenden Überlegungen aus: Wenn — wie in der Hypothese der beeinträchtigten analytischen Fähigkeiten angenommen — die eigentliche Schwierigkeit der Aphasiker beim TT darin besteht, die in den Anweisungen einzeln genannten Merkmale (groß/klein; rot/grün/blau/gelb/weiß; Viereck/Kreis) mit den als Ganzheiten vorgegebenen Plättchen in Beziehung zu setzen, so sollten Aphasiker erheblich davon profitieren, wenn die Merkmale nicht einzeln, sondern bereits als komplexe Einheiten, etwa im Sinne eines „chunks“ (Miller 1956), vorgegeben werden. Eine derartige Chunkbildung ist bei den im Standard-TT verwendeten Merkmalen kaum möglich — es gibt kein Wort, das etwa ‚Kreis‘ und ‚groß‘ als komplexe Einheit repräsentieren würde. Wir mußten daher auf andere Inhalte übergehen. Uns erschienen die komplexen Einheiten ‚Mann‘, ‚Frau‘, ‚Bube‘ und ‚Mädchen‘ besonders günstig, da sie Kombinationen von je einem kontinuierlichen Merkmal (Alter) und einem dichotomen Merkmal (Geschlecht) entsprechend den ursprünglichen Merkmalen ‚Form‘ und ‚Größe‘ repräsentieren. Nun ist zu bedenken, daß die Vorgabe derartiger komplexer Einheiten (z. B. ‚Mann‘ anstelle von ‚(großes) Viereck‘; ‚Frau‘ anstelle von ‚(großer) Kreis‘) auch nach den anderen beiden Hypothesen zu Leistungssteigerungen führen müßten: Zum einen beinhalten die *chunks* eine Reduktion der Wörter in den Anweisungen, was für die Speicherung im Kurzzeitgedächtnis vermutlich von erheblicher Bedeutung ist. Zum zweiten sind Wörter wie „Mann“, „Frau“, „Bube“ und „Mädchen“ sicherlich weniger ‚abstrakt‘ und lebensfern als Wörter wie „Kreis“ und „Viereck“. Welcher der verschiedenen Faktoren für etwaige Leistungssteigerungen in welchem Umfang verantwortlich wäre, sollte deshalb anhand einer genaueren Fehleranalyse abgeschätzt werden. Leistungssteigerungen in den Untertests I und III, in denen nur die großen Plättchen vorgelegt werden, können allein auf semantisch-pragmatische Faktoren zurückzuführen sein, da hier die Verwendung der Wörter „Mann“ und „Frau“ anstelle von „Viereck“ und „Kreis“ keinerlei Vorteile im Sinne von chunks beinhaltet. Nur dann also, wenn die Leistungssteigerungen im Untertest II größer wären als die im Untertest I, und die im Untertest IV größer als die im Untertest III, könnte auf einen besonderen Beitrag mnestischer und/oder konzeptueller Faktoren ge-

geschlossen werden. Welcher dieser beiden Faktoren wiederum die größere Bedeutung hat, müßte sich an der Art der verbleibenden Fehler zeigen: Wenn die konzeptuelle Verarbeitung von Einzelaspekten eines globalen Ganzen die entscheidende Variable ist, so sollte allein die Fehlerzahl bezüglich der zu einer Einheit verschmolzenen Merkmale reduziert sein, nicht aber die Fehlerzahl bezüglich der weiterhin isoliert genannten Farbe. Ist hingegen die Gedächtnisbelastung, die Speicherkapazität, ausschlaggebend, so sollte die Chunkbildung im Sinne einer absoluten ‚Platzersparnis‘ zu einer allgemeinen, unspezifischen Verbesserung der Leistung führen. Das Ausmaß der Fehlerreduktion bezüglich des Merkmals ‚Farbe‘ wäre damit als Indikator für die Bedeutung mnestischer Faktoren anzusehen.

## Methodik

### *Probanden*

Die Untersuchung wurde an 40 Aphasikern und 40 nicht-aphasischen Kontroll-Patienten (je 27 Männer und 13 Frauen) durchgeführt. Es wurden nur solche Aphasiker in die Untersuchung aufgenommen, die nach dem übereinstimmenden Urteil des behandelnden Arztes und der Logopädin vorwiegend Störungen der Sprachproduktion bei gutem oder kaum gestörtem Sprachverständnis hatten. Bei den nicht-aphasischen Kontroll-Patienten handelte es sich um Patienten mit diffusen Hirnschädigungen; in ihren Krankengeschichten durften zu keinem Zeitpunkt Hinweise auf aphasische Störungen vermerkt worden sein. Es wurden ausschließlich Rechtshänder — beurteilt nach einer deutschen Version des Edinburgh-Inventory von Oldfield (1971) — mit Deutsch als Muttersprache im Alter zwischen 18 und 70 Jahren berücksichtigt.

Die beiden diagnostischen Gruppen wurden in je zwei gleich große Untergruppen geteilt, von denen jeweils eine die Abwandlungen des TT mit „konsistent“, die andere mit „inkonsistent“ beklebten Plättchen durchführte (vgl. ‚Testmaterial und Durchführung‘). Die insgesamt 4 Gruppen waren hinsichtlich ihres Alters ( $47,4 \leq \bar{x} \leq 49,9$  Jahre;  $F(3,76) = 0,17$ ) und hinsichtlich eines kombinierten Maßes für Schul- und Berufsausbildung ( $2,25 \leq \bar{x} \leq 2,45$ ;  $F(3,76) = 0,42$ ) gut vergleichbar. Sie entsprachen einander auch hinsichtlich ihrer Leistungen im Trail Making Test, Form A ( $1,84 \leq \bar{x}_{\log s} \leq 1,96$ ;  $F(3,76) = 0,66$ ), der zur Abschätzung des Ausmaßes der allgemeinen Leistungsbeeinträchtigung durch hirnanorganische Läsionen herangezogen wurde (Goldstein 1977; Hegenscheidt u. Cohen 1972; Reitan 1959).

Bei der Zusammenstellung der Gruppen wurde weder auf die Ätiologie des Hirnschadens, noch auf die Krankheitsdauer geachtet. Wie kaum anders zu erwarten, fanden sich erhebliche Unterschiede zwischen den Gruppen bezüglich der Ätiologie ( $\chi^2(6) = 31,6$ ;  $P < 0,01$ ); bei den beiden aphasischen Gruppen waren vasculäre Ursachen am häufigsten, bei den diffus hirngeschädigten Patienten ohne Aphasie hingegen traumatische Schädigungen. Die vier Gruppen unterschieden sich nicht wesentlich hinsichtlich der Krankheitsdauer ( $H(3) = 5,35$ ;  $P > 0,10$ ).

### *Testmaterial und Durchführung*

Jedem Probanden wurde — über drei Sitzungen verteilt — der TT in der Originalversion und in drei Experimentalversionen vorgegeben.

*Originalversion.* Zu Beginn der ersten Sitzung wurden jeweils die vier ersten Untertests des TT in der von Orgass (1976a, b) revidierten Fassung durchgeführt, wobei allerdings in Abweichung vom sonst Üblichen die Plättchen gleicher Farbe in der Vorlage jeweils untereinander lagen.

*Experimentalversionen.* In den beiden verbleibenden Sitzungen wurden die drei Experimentalversionen durchgeführt, wobei die Reihenfolge derselben für jeden Probanden zufällig bestimmt wurde.

In allen drei Experimentalversionen waren die Plättchen des Tests mit auf durchsichtigem Dia-Material fotografierten Zeichnungen von einem Mann, einer Frau, einem kleinen Jungen oder einem kleinen Mädchen beklebt. Die Plättchen waren gegenüber der Originalversion um den Faktor 1,5 vergrößert, so daß die aufgeklebten Figuren gut erkennbar waren. Bei der Zuordnung der Figuren zu den Plättchen wurden zwei verschiedene Versionen verwendet: eine „konsistente“ Beklebung und eine „inkonsistente“ Beklebung. Grund dafür war die Überlegung, daß die Merkmale der TT-Plättchen möglicherweise im Sinne konnotativer Bedeutungsaspekte mit den Figuren verknüpft sein könnten (z. B. ‚rund‘ als konnotativer Aspekt von ‚Frau‘ etc.) und somit die Identifikation der Plättchen in den Experimentalversionen leichter fallen könnte als in der Originalversion. Um diesen Effekt abschätzen zu können, wurde der einen Hälfte der Aphasiker bzw. der Kontroll-Probanden in den Experimentalversionen eine Plättchenvorlage geboten, in der alle großen Vierecke mit einem Mann, alle großen Kreise mit einer Frau, alle kleinen Vierecke mit einem Jungen und alle kleinen Kreise mit einem Mädchen beklebt waren (= „Konsistente Beklebung“), und der anderen Hälfte der Probanden jeweils eine Plättchenvorlage, bei der die Zuordnung der Figuren zur Größe bzw. Form der Plättchen unsystematisch war (= „Inkonsistente Beklebung“). Allerdings wurde auch bei der inkonsistenten Beklebung jede der 4 Figuren fünfmal verwendet — je einmal auf einem weißen, gelben, roten, grünen und blauen Plättchen.

Bei der konsistenten wie auch bei der inkonsistenten Beklebung lagen die Plättchen gleicher Farbe jeweils untereinander.

Abgesehen von den Unterschieden in der Art der Beklebung der Plättchen erhielten alle Probanden dieselben drei Experimentalversionen:

1. *Experimentalversion „Standard-Anweisung“*. Dem Probanden wurden die Anweisungen in der üblichen Art vorgegeben (vgl. Orgass 1976a,b).

2. *Experimentalversion „Redundante Anweisung“*. Dem Probanden wurden in den Anweisungen zusätzlich zu den Merkmalen Größe, Farbe und Form des Plättchen auch die Figur auf dem Plättchen genannt (z. B. „Zeigen Sie mir bitte den großen roten Kreis mit der Frau!“ (= Untertest II, konsistente Beklebung) oder „Zeigen Sie mir bitte den großen roten Kreis mit dem Buben!“ (= Untertest II, inkonsistente Beklebung)).

3. *Experimentalversion „Konkrete Anweisung“*. Dem Probanden wurden in den Anweisungen lediglich die Farbe des zu identifizierenden Plättchens und die aufgeklebte Figur genannt (z. B. „Zeigen Sie mir bitte die rote Frau!“ (= Untertest II, konsistente Beklebung) oder „Zeigen Sie mir bitte den roten Buben!“ (= Untertest II, inkonsistente Beklebung)).

In allen drei Experimentalversionen wurden nur die Untertests II, III und IV durchgeführt. Auf eine vollständige Durchführung des TT glaubten wir mit Rücksicht auf das Leistungsvermögen der Patienten verzichten zu müssen. Den aphasischen Patienten war es kaum zuzumuten, den für sie äußerst schwierigen TT insgesamt viermal innerhalb kürzester Zeit durchführen zu müssen. Nach den Ergebnissen der testanalytischen Arbeiten von Orgass (1976a,b) und Woll et al. (1976) erschien uns der Verzicht auf die Untertests I und V am ehesten möglich.

### Auswertung

Die Auswertung erfolgte nach der einfachen Richtig/Falsch-Methode, bei der dem Probanden für jede nicht vollständig richtig gelöste Aufgabe ein Fehlerpunkt angerechnet wird. Nur dort, wo es um den Vergleich bestimmter Fehlertypen ging, wurden falschen Teilreaktionen hinsichtlich Größe, Farbe und Form der Plättchen gesondert bestimmt.

### Ergebnisse

In guter Übereinstimmung mit allen bisherigen Befunden erweist sich auch in der vorliegenden Untersuchung der TT als außerordentlich sensibles Verfahren. Die Leistungen der beiden Aphasikergruppen in der Originalversion des TT (Plättchen unbeklebt; Standardanweisungen) sind mit Fehlermedianen von  $Mdn = 10,5$  (Streubereich: 0–38) bzw.  $Mdn = 11,0$  (Streubereich: 0–30) erheblich schlechter als

**Tabelle 1.** Mediane und Streubereiche (in Klammern) der Anzahl Fehler in den Experimentalversionen des TT (jeweils Untertests II-IV)

Art der Anweisung	Beklebung der Plättchen					
	konsistent			inkonsistent		
	Aphasiker	Hirngeschädigte ohne Aphasie	Mann-Whitney U-Test	Aphasiker	Hirngeschädigte ohne Aphasie	Mann-Whitney U-Test
Standard-Anweisung	9,50 (1-27)	0,33 (0-2)	$z = 5,13^*$	10,50 (0-26)	0,30 (0-2)	$z = 4,47^*$
Konkrete Anweisung	4,50 (0-21)	0,09 (0-5)	$z = 4,29^*$	4,50 (0-26)	0,13 (0-2)	$z = 4,20^*$
Redundante Anweisung	8,50 (0-21)	0,21 (0-3)	$z = 4,70^*$	4,20 (0-22)	0,24 (0-3)	$z = 4,84^*$

\*  $P < 0.001$ 

die der entsprechenden Kontrollgruppen von Hirngeschädigten ohne Aphasie, deren Fehlermediane bei  $Mdn = 0,60$  (Streubereich: 0-5) bzw.  $Mdn = 0,56$  (Streubereich: 0-3) liegen. Die Unterschiede zwischen den Aphasikern und den entsprechenden Kontrollpatienten sind nach dem Mann-Whitney *U*-Test hochsignifikant ( $z = 5,04$  bzw.  $z = 5,02$ ;  $P < 0.01$ ).

Wie aus Tabelle 1 zu ersehen ist, bleibt die hohe Diskriminationskraft auch in den drei Experimentalversionen erhalten: Mit  $z$ -Werten zwischen  $z = 4,20$  und  $z = 5,13$  sind die Unterschiede zwischen den aphasischen Probanden und den Kontroll-Probanden durchweg auf dem 0,1%-Niveau der Verlässlichkeit bedeutsam.

Der Einfluß der Beklebung der Plättchen wurde anhand intraindividuellder Vergleiche zwischen den Leistungen in den Untertests II bis IV der Originalversion und denen der Experimentalversion „Standard-Anweisung“ geprüft. Wie Wilcoxon-Tests zeigen, hat die Beklebung der Plättchen — sei sie nun „konsistent“ oder „inkonsistent“ — weder bei den aphasischen Patienten (konsistente Beklebung:  $T(17) = 76$ ; inkonsistente Beklebung:  $T(16) = 40$ ;  $P > 0.10$ ) noch bei den Kontroll-Patienten (konsistente Beklebung:  $T(6) = 9$ ; inkonsistente Beklebung:  $T(11) = 28$ ;  $P > 0.10$ ) einen signifikanten Einfluß auf die Leistungen.

Für den Vergleich der verschiedenen Experimentalversionen untereinander wurden die Fehlerwerte der Aphasiker und die der Kontroll-Patienten getrennt non-parametrischen zweifaktoriellen Quasi-Varianzanalysen mit der „Art der Beklebung“ als zweistufigem Faktor ohne Meßwiederholung und der „Art der Anweisung“ als dreistufigem Faktor mit Meßwiederholung unterzogen (vgl. Bradley 1968).

Die Ergebnisse der Analyse für die Aphasiker zeigen, daß die Art der Beklebung — wie nach den zuvor referierten Ergebnissen kaum anders zu erwarten — keinen bedeutsamen Einfluß auf die Leistungen hatte ( $z = 0,42$ ;  $P > 0.10$ ). Demgegenüber erweist sich die Art der Anweisung als signifikant ( $\chi^2(2) = 19,6$ ;  $P < 0.001$ ). Wie nachfolgende Wilcoxon-Tests zeigen, machten die Aphasiker

**Tabelle 2.** Mediane und Streubereiche (in Klammern) der Fehlerwerte der Aphasiker ( $n=40$ ) bezüglich der Größe, der Farbe und der Form der Plättchen bei den Experimentalversionen „Standard-Anweisung“ und „Konkrete Anweisung“ (jeweils Untertests II–IV)

Art der Anweisung	Fehlerart		
	Größe (Maximum = 30)	Farbe (Maximum = 50)	Form (Maximum = 50)
Standard-Anweisung	1,17 (0–10)	4,50 (0–24)	5,50 (0–22)
Konkrete Anweisung	0,33 (0–9)	3,79 (0–29)	0,45 (0–10)

ker bei der „redundanten Anweisung“ wie auch bei der „konkreten Anweisung“ erheblich weniger Fehler als bei der Standard-Anweisung ( $T(19)=26$  bzw.  $T(19)=1$ ;  $P<0.01$ ). Die „redundante Anweisung“ ist jedoch *nicht*, wie man vermuten könnte, die leichteste: Bei der „konkreten Anweisung“ wurden signifikant weniger Fehler gemacht als bei der „redundanten Anweisung“ ( $T(19)=39$ ;  $P<0.05$ ). Die Interaktion der Faktoren „Art der Beklebung“ und „Art der Anweisung“ ist nicht bedeutsam ( $H(2)=2,02$ ;  $P>0.10$ ).

Bei den diffus Hirngeschädigten, bei denen sich durchweg nur sehr geringe Fehlerzahlen finden, läßt sich weder für die „Art der Beklebung“ noch für die „Art der Anweisung“ ein signifikanter Einfluß nachweisen ( $z=0,12$  bzw.  $\chi^2(2)=1,50$ ;  $P>0.10$ ). Auch die Interaktion ist nicht signifikant ( $H(2)=1,49$ ;  $P>0.10$ ).

Um weitere Information darüber zu erhalten, warum die Nennung der Figuren in den Anweisungen den Aphasikern die Identifikation der Plättchen so erleichterte, wurden die Fehler dieser Patienten bei der Standard-Anweisung und bei der „konkreten Anweisung“ genauer analysiert. Dabei wurden die Daten aller 40 Aphasiker zusammengefaßt, da sich die Art der Beklebung nicht als bedeutsam erwiesen hatte.

Zunächst wurden die Leistungen bei den beiden Anweisungsarten getrennt für Untertest III, bei dem die „konkrete Anweisung“ lediglich die Ersetzung der Wörter „Viereck“ und „Kreis“ durch die ‚konkreteren‘ Wörter „Mann“ und „Frau“ implizierte, und für den Untertest IV, bei dem die „konkrete Anweisung“ darüberhinaus eine Chunkbildung beinhaltete (z. B. „Mann“ anstelle von „großes Viereck“), verglichen. Untertest II wurde nicht berücksichtigt, da hier die Fehlerzahlen insgesamt zu gering waren (Mediane  $\leq 0,82$ ).

Es zeigte sich, daß die Aphasiker sowohl im Untertest III als auch im Untertest IV bei der „konkreten Anweisung“ weniger Fehler machten als bei der Standard-Anweisung (Untertest III:  $Mdn=1,50$  gegenüber 2,79; Untertest IV:  $Mdn=2,50$  gegenüber 5,00). Die Unterschiede sind in beiden Fällen nach dem Wilcoxon-Test signifikant (Untertest III:  $z=2,67$ ; Untertest IV:  $z=4,60$ ;  $P<0.01$ ). Ein Vergleich des Ausmaßes der Fehlerreduktion in den beiden Untertests ergab jedoch, daß die Leistungssteigerungen bei der „konkreten Anweisung“ gegenüber der Standard-Anweisung im Untertest IV erheblich größer sind als im Untertest III ( $z=3,22$ ;  $P<0.01$ ).

In einer zweiten Analyse wurde geprüft, ob die Nennung der Figuren in der „konkreten Anweisung“ allein zu einer Reduktion der Größen- und Formfehler



geführt hatte, oder ob bei dieser Art der Anweisung auch die Zahl der Farbfehler vermindert werden konnte. Tabelle 2 zeigt die Mediane und Streubereiche für die Größen-, Farb- und Formfehler der aphasischen Patienten bei der Standard-Anweisung und bei der „konkreten Anweisung“. Die Werte basieren jeweils auf den Summen der Fehler in den Untertests II bis IV. Bei der „konkreten Anweisung“ wurden Verwechslungen von männlich/weiblich als „Formfehler“ und Verwechslungen von kindlich/erwachsen als „Größenfehler“ definiert.

Intraindividuelle Vergleiche mit dem Wilcoxon-Test zeigen, daß den Aphasikern bei der „konkreten Anweisung“ sowohl signifikant weniger Größenfehler ( $T(17)=29,5$ ;  $P<0.05$ ), als auch signifikant weniger Formfehler ( $T(19)=0$ ;  $P<0.01$ ) unterlaufen als bei der Standard-Anweisung. Demgegenüber unterscheiden sich die beiden Anweisungsarten nicht bedeutsam hinsichtlich der Zahl der Farbfehler ( $T(17)=41,5$ ;  $P>0.10$ ).

## Diskussion

Wie in allen bisherigen Untersuchungen erwies sich auch in der vorliegenden Arbeit der TT als ein äußerst trennscharfes Verfahren zur Erfassung spezifisch aphasischer Störungen. Dieses Ergebnis ist insofern bemerkenswert, als nur Aphasiker mit geringen oder nicht erkundbaren Sprachverständnisstörungen in die Untersuchung aufgenommen worden waren. Daß selbst diese Patienten so erheblich schlechter abschnitten als die Kontrollgruppe diffus Hirngeschädigter, verstärkt die Zweifel an der noch immer weit verbreiteten Auffassung, der Test erfasse nur rezeptive Störungen.

Betrachtet man die Ergebnisse für die verschiedenen Experimentalversionen, in denen die Plättchen des Tests mit Zeichnungen von menschlichen Figuren beklebt waren, so muß man zunächst feststellen, daß die Modifikation des *Test-materials*, d. h. die Beklebung der Plättchen an sich, offenbar ohne Bedeutung ist: In der Experimentalversion, in der die Plättchen zwar beklebt, die Anweisungen jedoch dieselben wie die der Originalversion waren, waren die Leistungen der Patienten nicht anders als in der Originalversion selbst. Dies ist für die Bedingung der inkonsistenten Beklebung nicht weiter verwunderlich. Nicht so trivial ist dieses Ergebnis jedoch für die Konsistenz-Bedingung, in der eine perfekte Korrelation zwischen den Merkmalen der Plättchen selbst (groß/klein; rund/eckig) und den aufgeklebten Figuren (erwachsen/kindlich; weiblich/männlich) bestand. Es wäre durchaus denkbar gewesen, daß die damit verbundene klarere Struktur der Vorlage und die bessere Unterscheidbarkeit der Plättchen den Patienten die Orientierung und damit das Auffinden der gesuchten Plättchen erleichtern würde. Offenbar konnten auch die konnotativ-assoziativen Beziehungen (z. B. „rund“ — „Frau“; „klein“ — „Bube/Mädchen“), die sich aus der systematischen Zuordnung ergaben, nicht genutzt werden. Dieser Befund ist möglicherweise im Zusammenhang mit jenen zu sehen, die darauf hindeuten, daß Aphasiker nicht nur Schwierigkeiten beim Verständnis der denotativen, sondern auch der konnotativen Bedeutung von Wörtern haben (Ammon et al. 1977; Doehring u. Swisher 1972; Gardner u. Denes 1973).

Im Unterschied zu den Modifikationen der Plättchenvorlage erwiesen sich die Abwandlungen der Testanweisungen als bedeutsam für die Leistungen der Aphasiker. Sowohl in der Experimentalversion mit „redundanten Anweisungen“ als auch in der Experimentalversion mit „konkreten Anweisungen“ waren die Leistungen der Aphasiker signifikant besser als in der Experimentalversion mit Standard-Anweisungen.

Daß die „redundante Anweisung“, in der zusätzlich zu den üblichen Angaben über Größe, Farbe und Form des Plättchens auch die aufgeklebte Figur genannt wurde, den Aphasikern die Lösung der Aufgabe erleichterte, könnte zunächst als Beleg dafür angesehen werden, daß tatsächlich die mangelnde Redundanz der Standard-Anweisungen die wesentliche Ursache für das Versagen der Aphasiker im TT ist. Gardner et al. (1975) und Stachowiak et al. (1977) konnten zeigen, daß die Redundanz für das Verstehen von Wörtern und zusammenhängenden Texten bei Aphasikern von erheblicher Bedeutung ist und es wäre durchaus einleuchtend, wenn dieser Faktor auch beim TT eine ausschlaggebende Rolle spielte. Unsere Ergebnisse sprechen insgesamt jedoch nicht dafür: Zum einen machen Aphasiker auch bei der „redundanten Anweisung“ noch deutlich mehr Fehler als diffus Hirngeschädigte, zum anderen sind die Leistungen der Aphasiker bei der „redundanten Anweisung“ *schlechter* als bei der „konkreten Anweisung“, in der nur die Farbe der Plättchen und die aufgeklebten Figuren genannt wurden. Bezogen auf die Leistungen bei der „konkreten Anweisung“ scheint die Erhöhung der Redundanz also eher nachteilig für die Patienten! Allerdings ist dabei zu berücksichtigen, daß in unserer Experimentalversion die Erhöhung der Redundanz gleichzeitig eine Verlängerung der Testanweisung beinhaltete.

Wie ist nun zu erklären, daß die Nennung der aufgeklebten Figuren in den Anweisungen den Aphasikern die Aufgaben so sehr erleichterte? Zum einen scheinen hier semantisch-pragmatische Faktoren eine Rolle zu spielen. Die Ergebnisse des Untertests III zeigen, daß die Patienten schon dort von den „konkreten Anweisungen“ profitierten, wo diese lediglich die Ersetzung der inhaltsarmen Wörter „Viereck“ und „Kreis“ durch die gehaltvolleren Wörter „Mann“ und „Frau“ implizierten. Dieses Ergebnis stimmt recht gut mit dem Befund von Kreindler et al. (1971) überein, die ebenfalls erhebliche Leistungssteigerungen von aphasischen Patienten in einer konkreten Version des TT feststellten, in der „Kreis“ und „Viereck“ durch „Blume“ und „Haus“ ersetzt waren. Es scheint also, daß für Aphasiker Faktoren wie die „Konkretheit“ oder „Lebensnähe“ tatsächlich von Bedeutung sind. Offensichtlich gilt dies aber nicht uneingeschränkt: Lesser (1979) konnte bei einer Modifikation des TT, bei der die Farben der Plättchen durch verschiedene Materialien wie Stoff, Gummi, Schokolade ersetzt waren, keinerlei Fehlerreduktion bei den aphasischen Patienten finden. Dies könnte darauf hindeuten, daß größere Konkretheit und Lebensnähe keinen Nutzen haben, wenn lediglich einzelne Aspekte ersetzt werden, ohne daß dadurch gleichzeitig die Reize als Ganzes „bedeutungsträchtiger“ und damit semantisch besser diskriminierbar werden. Daß die semantische Diskriminierbarkeit für Aphasiker von wesentlicher Bedeutung ist, vermuteten schon Goodglass et al. (1966), die zur Erklärung der Tatsache, daß Aphasikern verschiedene semantische Kategorien so unterschiedlich schwer fallen, darauf verwiesen, daß „... these (colours, letters, geometric forms, numbers) are well-demarcated

semantic groups, whose members are not very distinct from each other with respect to experiential connotations. In contrast, object names and actions each represent a potentially limitless variety of connotations." (S. 86).

Die vorliegende Untersuchung zeigt jedoch, daß solche semantisch-pragmatische Faktoren nicht die einzig relevanten sind. Die Tatsache, daß die Leistungssteigerungen bei der „konkreten Anweisung“ im Untertest IV erheblich größer waren als im Untertest III, belegt, daß die Chunkbildung für die Aphasiker eine ausschlaggebende Rolle spielte. Da diese Chunkbildung eine Reduktion der Wörter in den Anweisungen beinhaltetete („Mann“ anstelle von „großes Viereck“ etc.) könnte man zunächst geneigt sein, die Leistungssteigerungen auf die geringeren Anforderungen an das Kurzzeitgedächtnis zurückzuführen. Die Ergebnisse der detaillierten Fehleranalyse sprechen jedoch dagegen: Die Verringerung der Fehler bei der „konkreten Anweisung“ (gegenüber der Standard-Anweisung) betraf allein die Fehler bezüglich der zu einer Einheit verschmolzenen Merkmale ‚Größe‘ und ‚Form‘; die Zahl der Farbfehler blieb hingegen unverändert. Wäre die Belastung des Kurzzeitgedächtnisses der ausschlaggebende Faktor, so hätte man im Sinne einer allgemeinen „Platzersparnis“ eine unspezifische Fehlerreduktion erwarten müssen. Daß Aphasiker nur dort Fehler vermeiden konnten, wo zusammenfassende Codes für die verschiedenen Einzelmerkmale geboten wurden, deutet auf eine tieferliegende Beeinträchtigung im präverbalen, konzeptuellen Bereich hin. Es scheint, daß die Patienten Schwierigkeiten dabei haben, einzeln genannte Merkmale mit den als globale Ganzheiten vor ihnen liegenden Plättchen in Beziehung zu setzen. In diesem Sinne wäre es nicht unplausibel, die Ursache für die hohe Diskriminationskraft des TT darin zu suchen, daß dieser Test die primär von der Funktionstüchtigkeit der dominanten Hemisphäre abhängigen analytischen Fähigkeiten zur gedanklichen Isolierung und klassifikatorischen Erfassung einzelner Merkmale prüft.

## Literatur

- Albert ML (1976) Short-term memory and aphasia. *Brain Lang* 3:28–30
- Ammon KH, Moerman C, Guleac JD (1977) Aphasic defective perception of connotative meaning of verbal items which have no denotative meaning. *Cortex* 13:453–457
- Bradley JV (1968) Distribution-free statistical testing. Prentice-Hall, New York
- Brookshire RH (1978) A Token test battery for testing auditory comprehension in brain-injured adults. *Brain Lang* 6:149–157
- Cohen R, Kelter S, Engel D, List G, Strohn H (1976) Zur Validität des Token-Tests. *Nervenarzt* 47:357–361
- Cohen R, Kelter S, Schäfer B (1977) Zum Einfluß des Sprachverständnisses auf die Leistungen im Token-Test. *Z Klin Psychol* 6:1–14
- Cohen R, Kelter S, Woll G (1980) Analytical competence and language impairment in aphasia. *Brain Lang* (in press)
- De Renzi E, Vignolo LA (1962) The Token test: A sensitive test to detect receptive disturbances in aphasia. *Brain* 85:665–678
- De Renzi E, Nichelli P (1975) Verbal and non-verbal short-term memory impairment following hemispheric damage. *Cortex* 11:341–354
- De Renzi E, Faglioni P (1978) Normative data and screening power of a shortened version of the Token test. *Cortex* 14:41–49
- De Renzi E, Faglioni P, Previdi H (1978) Increased susceptibility of aphasics to a disorder task in the recall of verbal commands. *Brain Lang* 6:14–21

- Doehring BG, Swisher LP (1972) Disturbances of connotative meaning in aphasia. *J Comm Dis* 5:251–258
- Gardner H, Albert ML, Weintraub S (1975) Comprehending a word: The influence of speed and redundancy on auditory comprehension in aphasia. *Cortex* 11:155–162
- Gardner H, Denes G (1973) Connotative judgements by aphasic patients on a pictorial adaptation of the semantic differential. *Cortex* 9:183–196
- Golden CJ (1977) Validity of the Halstead-Reitan neuropsychological battery in a mixed psychiatric and braininjured population. *J Consult Clin Psychol* 45:1043–1051
- Goodglass H, Gleason JB, Hyde MA (1970) Some dimensions of auditory language comprehension in aphasia. *J Speech Hear Res* 13:595–606
- Goodglass H, Klein B, Carey P, Jones K (1966) Specific semantic word categories in aphasia. *Cortex* 2:74–89
- Hegenscheidt M, Cohen R (1972) Zur Erfassung der Umstellungsfähigkeit bei hirnanorganisch geschädigten Personen. *Z Klin Psychol* 1:1–20
- Kreindler A, Gheorghita N, Voinescu I (1971) Analysis of verbal reception of a complex order with three elements in aphasics. *Brain* 94:375–386
- Leischner A (1974) Die neuropsychologisch-hirnpathologische Untersuchung. *Arch Psychiat Nervenkr* 219:53–77
- Lesser R (1974) Verbal comprehension in aphasics: An English version of three Italian tests. *Cortex* 10:247–263
- Lesser R (1979) Turning tokens into things: linguistic and mnestic aspects of the initial sections of the Token test. In: Boller F, Dennis M (eds) *Auditory comprehension: clinical and experimental studies with the Token test*. Academic Press, New York
- Miller GA (1956) The magical number seven, plus or minus two: some limits of our capacity for processing information. *Psychol Rev* 63:81–87
- Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia* 9:97–113
- Orgass B (1976a) Eine Revision des Token Tests. I. Vereinfachung der Auswertung, Itemanalyse und Einführung einer Alterskorrektur. *Diagnostica* 22:70–87
- Orgass B (1976b) Eine Revision des Token Tests. II. Validitätsnachweis, Normierung und Standardisierung. *Diagnostica* 22:141–156
- Poeck K, Kerschensteiner M, Hartje WA (1972) A quantitative study on language understanding in fluent and non-fluent aphasia. *Cortex* 8:299–304
- Reitan RM (1959) *A manual for the administration and scoring of the trail making test*. Indiana University Press, Indiana
- Saffran EM, Marin OS (1975) Immediate memory for word lists and sentences in a patient with deficient auditory short-term memory. *Brain Lang* 2:420–433
- Stachowiak FS, Huber W, Poeck K, Kerschensteiner M (1977) Text comprehension in aphasia. *Brain Lang* 4:177–195
- Woll G, Naumann E, Cohen R, Kelter S (1976) Kreuzvalidierung der Revision des Token Tests durch Orgass. *Diagnostica* 22:157–162

Eingegangen am 13. Februar 1980