

## Sprachliche Musterbildung und Mustererkennung

Von Willem J. M. LEVELT (Nijmegen)

Mit 8 Abbildungen



*W. J. M. Levelt*

Psycholinguistische Forschung beschäftigt sich zu einem bedeutenden Teil mit der Untersuchung sprachlicher Mustererkennung und Musterbildung. Sprachliche Kommunikation ist eine artspezifische Fähigkeit des *Homo sapiens*. Psycholinguisten untersuchen, wie wir Sprache produzieren und verstehen und auf welche Weise Kinder diese Fähigkeiten erwerben. Gesprochene Sprache kann zunächst einmal als akustisches Ereignis aufgefaßt werden, das sich durch eine hochkomplexe Dynamik sowohl im Frequenz- als auch im Intensitätsbereich auszeichnet. Die Zahl wohlgeformter sprachlicher Muster ist unbeschränkt. Als wohlgeformt bezeichnen wir ein sprachliches Muster, das Sprachbenutzer als mögliches Muster in ihrer jeweiligen Sprache ansehen.

Ein erster – wenn auch nicht der interessanteste – Grund für die Unbegrenztheit sprachlicher Muster ist der, daß das sprachliche Signal auf einer Anzahl von Dimensionen kontinuierlich variieren kann. Wenn ein und dieselbe Äußerung in zwei verschiedenen akustischen Kontexten gesprochen wird, werden zwei unterschiedliche Sprachsignale erzeugt, und die Menge möglicher Kontexte ist unbegrenzt. Ein schon interessanterer Aspekt ist der, daß der Vokaltrakt verschiedener Menschen nicht völlig identisch ist. Ein und dieselbe Äußerung, von zwei verschiedenen Personen gesprochen, beispielsweise einem Mann und einer Frau, wird damit auch zu unterschiedlichen Sprachsignalen führen. Im allgemeinen bereitet es Menschen keine Schwierigkeiten, diese unterschiedlichen Sprachsignale als Token derselben Äußerung zu verstehen. Dies ist ein wichtiger Aspekt, wenn wir bedenken, daß automatische Sprachverarbeitungssysteme noch immer größte Schwierigkeiten haben, eine solche Leistung zu vollbringen.

Der tiefere Grund für die Unbegrenztheit sprachlicher Muster ist allerdings im generativen Charakter der Sprache zu finden. Für jede wohlgeformte sprachliche Sequenz können wir eine längere Sequenz finden, die ebenfalls wohlgeformt ist. Dieser generative Charakter ist morphologischer und syntaktischer Natur. Morphologische Generativität wird besonders deutlich in Sprachen wie dem Deutschen. *Bahn* ist ein Wort, *Bahnübergang* ist ein Wort, *Bahnübergangssicherung* ist ein Wort, *Bahnübergangssicherungsanlage* ist ein Wort, und es existiert keine prinzipielle Beschränkung für die Anzahl der Konstituenten solcher Nominalkomposita. Syntaktische Generativität ermöglicht es, für jeden wohlgeformten Satz einen längeren, ebenfalls wohlgeformten Satz zu bilden. Nehmen Sie beispielsweise den wohlgeformten Satz: *Wenn Hans sagt, daß es regnet, lügt er*. Daraus läßt sich der längere und ebenfalls wohlgeformte Satz *Wenn Hans sagt, daß Peter sagt, daß es regnet, lügt er* bilden. Dies können wir unendlich weiterführen: *Wenn Hans sagt, daß Peter sagt, daß Gretchen sagt, daß es regnet, lügt er* und so weiter. Grammatiken sind unter dieser Perspektive genaunommen rekursive Systeme, die solche infiniten Mengen von sprachlichen Mustern erzeugen.

Vor dem Hintergrund der Unbeschränktheit sprachlicher Muster können damit Spracherkennungs- und Sprachproduktionstheorien, die eine einfache Abbildung auf bzw. von gespeicherten sprachlichen Mustern (*template matching*) beinhalten, verworfen werden. Denn es ist nicht der Fall, daß wir eine endliche Menge möglicher Sätze erworben und im Gedächtnis gespeichert haben. Wir besitzen weder eine endliche Menge artikulatorischer Satzmuster, die wir bei passender Gelegenheit produzieren, noch verfügen wir über eine endliche Menge akustischer

Satzmuster, die wir angemessen zu interpretieren gelernt haben, wenn wir sie in unserer Umwelt wahrnehmen.

Die Basis dieses generativen Systems ist jedoch finit. Die Elemente, die die Bausteine komplexer Satzmuster darstellen, bilden ein endliches Vokabular. Zunächst gibt es das Vokabular der Phoneme. Phoneme sind die Einheiten im Lautsystem einer Sprache. Den meisten von uns sind sie als Vokale und Konsonanten bekannt. In jeder Sprache existiert nur eine beschränkte Menge dieser Phoneme, ihre Zahl schwankt etwa zwischen 25 und 40. Neben dem Vokabular der Phoneme ist das der lexikalischen Einheiten zu nennen. Dieses bezeichnen wir als das *mentale Lexikon*. Das mentale Lexikon ist der Speicher derjenigen Wörter, die häufig von Sprachbenutzern gebraucht werden. Sein Umfang ist nicht genau zu bestimmen. Wie wir soeben gesehen haben, ist das Lexikon einer Sprache selbst ein generatives System. Zum Beispiel sind alle Zahlwörter des Deutschen Bestandteil des deutschen Lexikons, und allein von diesen Wörtern gibt es schon eine unendliche Menge. Dies bedeutet, daß unser finites Gehirn sie nicht alle speichern kann. Nur Wörter, die häufiger gebraucht werden, sind gespeichert, zusammen mit ihrer Bedeutung, ihren syntaktischen Eigenschaften und ihrer Lautform. Schätzungen über die Größe des mentalen Lexikons eines Erwachsenen gehen auseinander. Aber wir können mit Sicherheit von einigen tausend oder sogar einigen zehntausend Wörtern ausgehen.

Ein erster Schritt bei jeder sprachlichen Mustererkennung oder Mustererzeugung besteht aus dem Zugriff auf diese Wörter im mentalen Lexikon. Beim Spracherkennen muß der Hörer das Sprachsignal in Segmente zerlegen, die möglichen Wörtern entsprechen. Kann eine Übereinstimmung gefunden werden zwischen einem Teil des sprachlichen Signals und der Lautform eines der Wörter im mentalen Lexikon? Diesen Prozeß nennen wir *lexikale Aktivierung*. Durch das Sprachsignal werden ein oder mehrere Wörter im mentalen Lexikon aktiviert. Sobald genau ein lexikaler Eintrag gefunden ist, dessen Muster dem wahrgenommenen Lautmuster entspricht, bezeichnen wir das betreffende Wort als *ausgewählt* bzw. *erkannt*; seine Bedeutung (die im mentalen Lexikon spezifiziert ist) kann nun integriert werden in den Kontext der übrigen bereits erkannten Wörter der Äußerung.

Bei der Sprachproduktion handelt es sich um einen vergleichbaren Prozeß. Wenn wir etwas bestimmtes ausdrücken wollen, dann werden wir solche Wörter wählen, deren Kombination eine angemessene Äußerung ergibt. Zum Beispiel, wenn ich ausdrücken möchte, daß ich ein Schaf sehe, muß ich das Wort *Schaf* in meinem mentalen Lexikon auswählen. Zunächst können allerdings einige weitere Wörter aktiviert werden, wie beispielsweise *Ziege* oder *Hammel*. Aber zu einem bestimmten Zeitpunkt wird das Wort *Schaf* ausgewählt. Sobald dies geschehen ist, wird die phonologische Form dieses Wortes aktiviert und das entsprechende artikulatorische Muster erzeugt. Im folgenden wird deutlich werden, wie wichtig die Unterscheidung zwischen Aktivierung und Auswahl ist. Sowohl bei der Wahrnehmung als auch der Produktion von Sprache können Wörter aktiviert werden, ohne auch ausgewählt zu werden. Wie wir sehen werden, verhalten sich aktivierte Wörter anders als ausgewählte.

Zunächst möchte ich einige Aspekte der Worterkennung skizzieren. Danach werde ich auf ähnliche Aspekte der Wortproduktion eingehen. Und schließlich werde ich argumentieren, daß die Übereinstimmungen in beiden Bereichen Hinweise darauf geben, daß ein und dasselbe lexikale System sowohl bei der Erzeugung als auch bei der Erkennung lexikaler Muster beteiligt ist.

### *Aktivierung und Auswahl bei der Worterkennung*

Das Erkennen gesprochener oder geschriebener Wörter ist zum größten Teil ein automatischer, reflexartiger Prozeß. Der Erkennensprozeß wird gesteuert durch den sprachlichen Reiz und kann nicht bewußt verhindert werden. Dieses einfache Phänomen ist der Grund dafür, daß wir der folgenden Aufforderung so schwer folgen können:

Denken Sie jetzt bitte nicht an einen Würfel.

Die Worterkennung ist jedoch *nicht allein* von den Charakteristika des (visuellen oder akustischen) Reizes abhängig. In Abbildung 1 finden Sie zwei Sätze, die durch Tintenflecke beschmutzt sind. Beide Sätze sind jedoch noch gut lesbar: *Eine Abteilung zu leiten macht viel Spaß* und *Kriege bringen entsetzlich viel Leiden in die Welt*.



Abb. 1 Effekt des Bedeutungskontextes bei der Worterkennung. Beim Lesen dieser Sätze werden die beschmutzten Wörter „leiten“ und „leiden“ trotz physischer Identität korrekt unterschieden.

Das Wort *leiten* im ersten Satz ist wegen der Tintenflecke visuell völlig identisch mit dem Wort *Leiden* im zweiten Satz. Trotzdem werden sie als unterschiedliche Wörter wahrgenommen, und dies kann nur durch den Satzkontext bedingt sein.

Psycholinguisten sind mit großem Aufwand der Frage nachgegangen, in welcher Weise reizabhängige *Bottom-up*- und wissensabhängige *Top-down*-Prozesse bei der Worterkennung ineinandergreifen. Beispielsweise wurde an unserem Institut nachgewiesen, daß der semantische und syntaktische Satzkontext das Erkennen eines Wortes erleichtern oder auch erschweren kann, abhängig davon, wie gut das Wort in den betreffenden Kontext paßt. Dieser erleichternde oder erschwerende Effekt entsteht jedoch nicht unmittelbar, sondern erst, nachdem die reizgesteuerte Verarbeitung schon erheblich fortgeschritten ist (ZWITSERLOOD 1989).

Ich möchte solche Einflüsse von Wissensstrukturen auf die Worterkennung hier nicht weiter vertiefen, sondern meine Ausführungen auf das Erkennen isoliert dargebotener Wörter beschränken. Insbesondere werde ich auf die Frage eingehen, in welcher Weise das Erkennen einzelner Sprachlaute mit dem Erkennen des gesamten Wortes zusammenwirkt. Unmittelbar einleuchtend ist, daß bei der Worterkennung Information von der Ebene der Sprachlaute auf die Ebene der Wörter weitergeleitet wird. Aber gibt es auch einen gegenläufigen Prozeß, in dem Information von der Ebene der Wörter auf die Ebene der Sprachlaute zurückfließt (vgl.: Abb. 2)? Konnektionisten nehmen jedenfalls an, daß das der Fall ist (ELMAN and MCCLELLAND 1988).

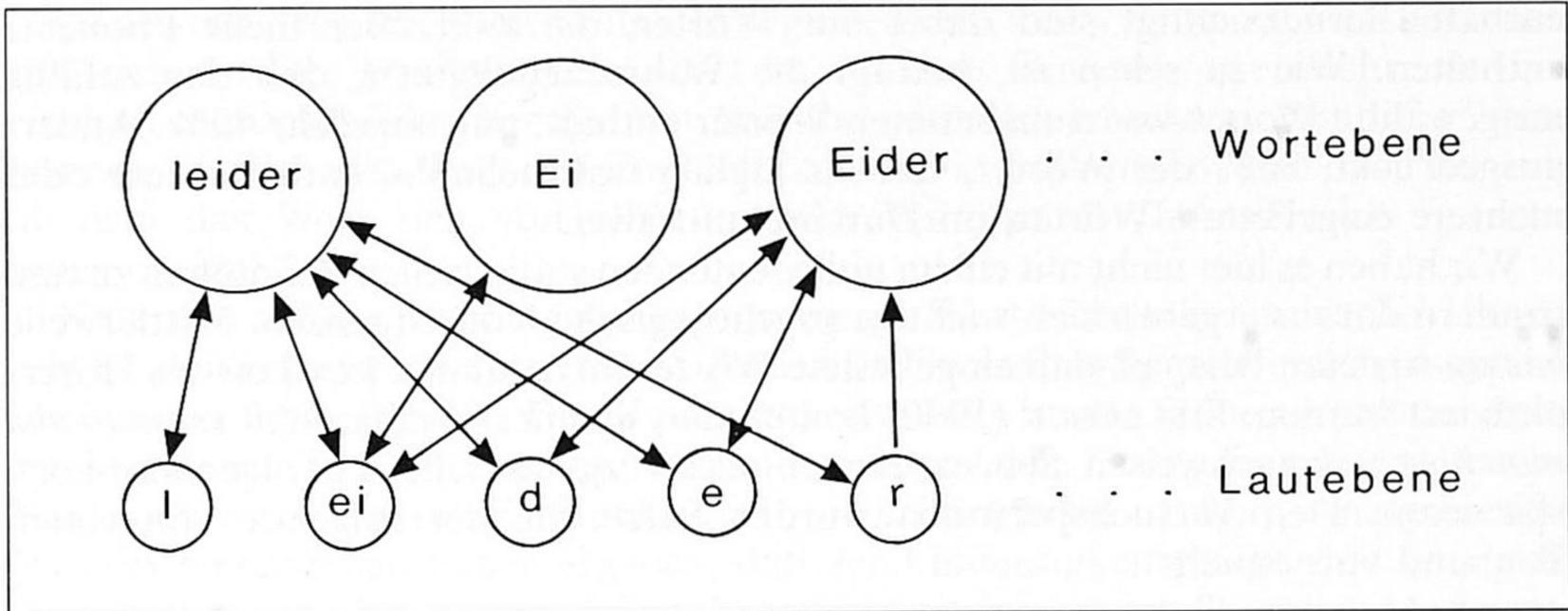


Abb. 2 Möglicher Informationsfluß zwischen Lautebene und Wortebene bei der Worterkennung

Um diesen Aspekt deutlicher zu machen, sollte ich zunächst auf die Begriffe Aktivierung und Auswahl eingehen.

Viele Wörter beinhalten andere Wörter. Das gesprochene Wort *leider* beispielsweise enthält die Wörter *Lei*, *Leid*, *Ei*, *Eid* und *Eider*. Eine solche Einbettung von Wörtern in anderen Wörtern ist eher die Regel als die Ausnahme. Wenn Sie zufällig ein Wort aus einem beliebigen gesprochenen niederländischen Text wählen, dann beträgt die erwartete Anzahl eingebetteter Wörter 1,847, also ungefähr zwei. Abbildung 3 stellt dies anschaulich dar. Dargestellt ist die Wahrscheinlichkeit  $Pr(W)$ , daß ein zufällig aus einem Text ausgewähltes Wort  $W$  eingebettete Wörter

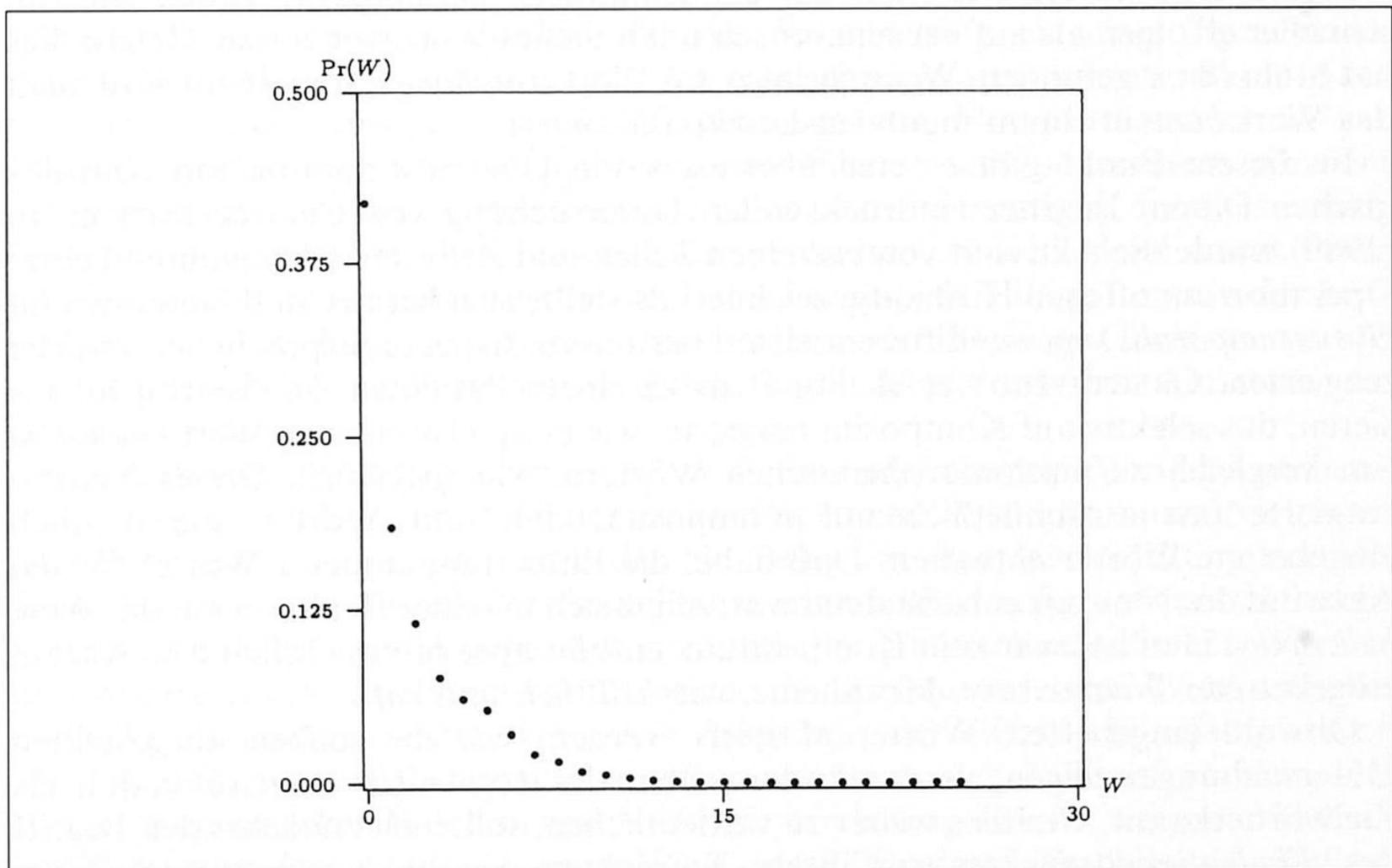


Abb. 3 Auftretenswahrscheinlichkeit  $Pr(W)$  eines Wortes mit  $W$  eingebetteten Wörtern (im Niederländischen).  $E(W) = 1,847$

enthält. Berücksichtigt sind dabei nur Wörter, die zwei oder mehr Phoneme enthalten. Wie zu sehen ist, beträgt die Wahrscheinlichkeit, daß das zufällig ausgewählte Wort *keine* eingebetteten Wörter enthält, nur ungefähr 40%. Anders ausgedrückt, 60% der Wörter, die wir täglich wahrnehmen, enthalten ein oder mehrere eingebettete Wörter, im Durchschnitt zwei.

Wir haben es hier nicht mit einem unbedeutenden statistischen Phänomen zu tun, sondern daraus ergeben sich wichtige psychologische Konsequenzen. Mittlerweile wissen wir zum Beispiel, daß eingebettete Wörter im mentalen Lexikon des Hörers aktiviert werden. SHILLCOCK (1990) konnte dies in einer sogenannten *cross-modal matching task* nachweisen. Sein experimentelles Vorgehen läßt sich folgendermaßen skizzieren. Den Versuchspersonen wurden Sätze wie der folgende von einem Tonband vorgespielt:

He carefully placed the trombone on the table.

^  
RIB

Während die Versuchspersonen den Satz hören, wird zu einem bestimmten Zeitpunkt ein visueller Stimulus dargeboten. Dieser visuelle Reiz kann entweder ein Wort (wie *tip*) oder ein Nichtwort (wie *sef*) sein. Aufgabe der Versuchsperson ist, durch Drücken einer Taste anzugeben, ob der visuelle Stimulus ein Wort oder ein Nichtwort ist. Diese Aufgabe nennen wir lexikale Entscheidungsaufgabe. Dabei wird die Zeit zwischen dem Beginn der Darbietung des visuellen Reizes und dem Tastendruck gemessen. Der Trick in SHILLCOCKS Experiment besteht nun darin, auch Wörter als visuellen Stimulus darzubieten, die einem eingebetteten Wort semantisch ähnlich sind. Zum Beispiel ist das Wort *bone* eingebettet in das Wort *trombone*. Wenn das eingebettete Wort im mentalen Lexikon des Hörers aktiviert wird, dann sollte die Reaktion auf ein semantisch ähnliches Testwort, wie *rib*, schneller erfolgen als auf ein semantisch unähnliches Wort, wie *house*. Genau dies hat SHILLCOCK gefunden. Wenn Sie also das Wort *trombone* hören, dann wird auch das Wort *bone* in Ihrem mentalen Lexikon aktiviert.

In diesem Punkt gibt es eine überraschende Übereinstimmung mit neurologischen Daten. In einer eindrucksvollen Untersuchung von CREUTZFELDT et al. (1989) wurde die Aktivität von einzelnen Zellen und Zellverbänden während einer Operation am offenen Hirn aufgezeichnet. Es stellte sich heraus, daß Neuronen im *Gyrus temporalis superior* differentiell auf bestimmte Aspekte gesprochener Sprache reagierten. CREUTZFELDT et al. konnten bei einem Patienten ein Neuron lokalisieren, das selektiv auf Komposita reagierte, wie beispielsweise das Wort *horseshoe* (im Vergleich zu monomorphemischen Wörtern, wie *spaghetti*). Dieses Neuron reagierte fast ausschließlich auf Komposita, d. h. auf Wörter, die deutlich eingebettete Wörter enthalten. Daß dabei die Einbettung anderer Wörter für die Aktivität des Neurons entscheidend war, zeigte sich in seiner Reaktion auf das Wort *helicopter*. Dies ist zwar kein Kompositum, enthält aber offensichtlich eine Anzahl eingebetteter Wörter bzw. Morpheme, wie *hell*, *heli* und *cop*.

Obwohl eingebettete Wörter aktiviert werden, wie die soeben dargestellten Untersuchungen zeigen, werden sie dennoch in der Regel nicht *ausgewählt*, d. h. als Zielwort erkannt. Um dies weiter zu verdeutlichen, sollte ich zunächst den Begriff des *Eindeutigkeitpunktes* einführen. Betrachten Sie noch einmal das Wort *trombone*. Im Englischen gibt es Tausende von Wörtern, die mit dem Phonem /t/ beginnen. Immerhin noch einige Hundert beginnen mit /tr/. Ungefähr zehn

beginnen mit /tro/. Nicht mehr als vier beginnen mit /trom/. Und nur zwei beginnen mit /tromb/. Der Eindeutigkeitspunkt ist mit dem Phonem /o/ erreicht; nur ein einziges englisches Wort beginnt mit /trombo/, nämlich *trombone*. Auf diese Weise können wir für jedes Wort den Eindeutigkeitspunkt bestimmen, also das Phonem, ab dem das Wort sich von allen anderen Wörtern der betreffenden Sprache unterscheidet.

Natürlich sollte ein isoliert wahrgenommenes Wort nicht erkannt werden, bevor sein Eindeutigkeitspunkt erreicht ist. Andernfalls wäre das Sprachverstehenssystem hochgradig fehleranfällig. Der Eindeutigkeitspunkt ist der frühestmögliche Zeitpunkt, zu dem ein Wort überhaupt erkannt werden *kann*. Es ist aber eine empirische Frage, ob es zu diesem Zeitpunkt auch *tatsächlich* erkannt wird. Analysen von Worterkennenszeiten haben ergeben, daß der Eindeutigkeitspunkt eine sehr gute Annäherung an den tatsächlichen Erkennungspunkt darstellt (vgl.: MARSLEN-WILSON und TYLER 1981).

Kehren wir zu unserem Beispielwort *trombone* zurück. Dieses Wort kann erkannt werden, d. h. *als Zielwort ausgewählt* werden, sobald das Segment /trombo/ verarbeitet worden ist. Unabhängig davon wird das eingebettete Wort *bone* aktiviert. Wir müssen aus diesem Grunde Aktivierung und Auswahl sorgfältig unterscheiden. Eingebettete Wörter können aktiviert werden, bevor und nachdem das Zielwort ausgewählt worden ist (d. h. als Zielwort erkannt worden ist).

Lassen Sie mich zu meiner Ausgangsfrage zurückkehren: Gibt es einen Informationsfluß von der Ebene der Wörter auf die Ebene der einzelnen Sprachlaute? Wenn ja, schließt sich folgende Frage an: Wird die Verarbeitung der Sprachlaute eines Wortes erleichtert, wenn das Wort aktiviert ist, oder muß es zuvor auch ausgewählt worden sein?

Eine kürzlich an unserem Institut durchgeführte Untersuchung von FRAUENFELDER et al. (1990) gibt eindeutige Antworten auf diese Fragen. Zunächst läßt sich tatsächlich ein lexikaler Einfluß auf die Verarbeitung einzelner Sprachlaute nachweisen. Allerdings setzt dies die Auswahl eines Wortes voraus. Bloße Aktivierung ist nicht ausreichend. Das Experiment von FRAUENFELDER et al. sah folgendermaßen aus. Die Versuchspersonen führten eine sogenannte *phoneme monitoring task* aus. Sie wurden angeleitet, eine Taste zu betätigen, sobald sie einen bestimmten Sprachlaut, beispielsweise ein /p/, hörten. Anschließend wurde ihnen ein Tonband mit einer langen Liste von Items vorgespielt. Einige Items in der Liste enthielten den Ziellaut /p/, beispielsweise *Pagina* oder *Bioskop*. Gemessen wurde, wie schnell die Versuchspersonen durch Tastendruck auf das Auftreten des Ziellautes reagierten.

Die Liste enthielt auch Nichtwörter, wie *Pafime* oder *Deoftop*. Wird das *phoneme monitoring* davon beeinflusst, ob der Ziellaut in einem Wort oder einem Nichtwort auftritt? Wenn dies zuträfe, wäre das ein Beleg für eine Rückwirkung von der Wortebene auf die Ebene der Sprachlaute innerhalb des Sprachverarbeitungssystems. Zunächst einmal zeigte sich ein deutlicher Unterschied zwischen Wörtern und Nichtwörtern. Im Durchschnitt wurde der Ziellaut in einem Wort schneller entdeckt als in einem Nichtwort. Allerdings war der Effekt vollständig davon abhängig, in welcher Position innerhalb eines Items der Ziellaut erschien. Der Unterschied zwischen Wörtern und Nichtwörtern betrug 164 ms, wenn der Ziellaut am Ende eines Items auftrat, wie in *Bioskop* (vs. *Deoftop*). Erschien der Ziellaut allerdings am Anfang eines Items, wie in *Pagina* (vs. *Pafime*), war kein Unterschied zu beobachten.

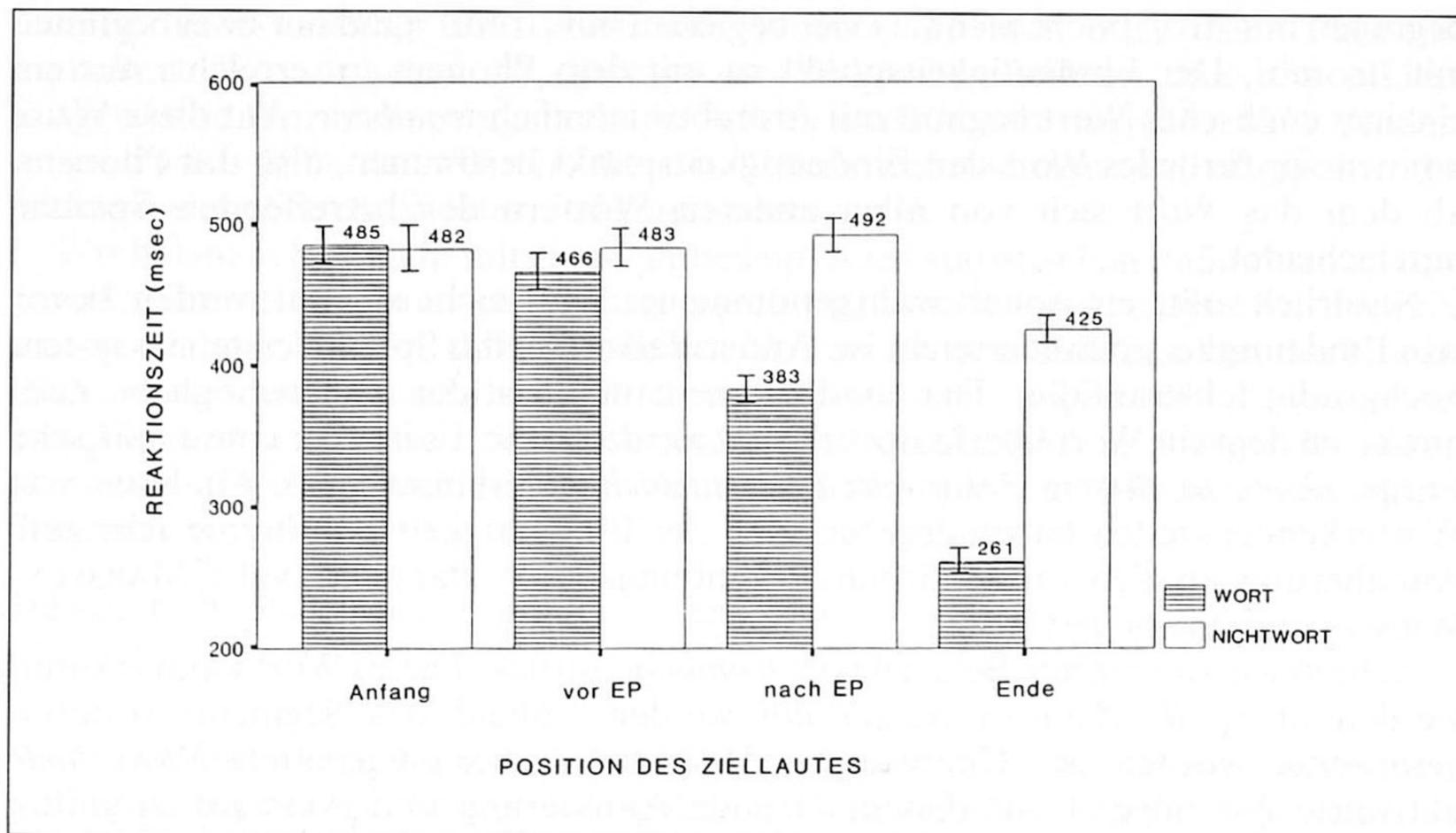


Abb. 4 Reaktionszeit für Ziellaute in Wörtern und Nichtwörtern in Abhängigkeit von ihrer Position: Itemanfang, vor Eindeutigkeitspunkt (EP), nach EP, und Itemende (nach: FRAUENFELDER et. al. 1990)

Wichtig in diesem Experiment waren auch die Reaktionszeiten für Ziellaute, die in der Mitte eines Items erschienen, wie in *Olympiade*. Ob es einen Unterschied zwischen Wörtern und Nichtwörtern gab, war in solchen Fällen abhängig davon, ob der Ziellaut vor oder nach dem Eindeutigkeitspunkt auftrat. Und dies ist für unsere Fragestellung unmittelbar relevant. Wir wissen aus einer Reihe von Untersuchungen, daß das Zielwort aktiviert ist, *bevor* der Eindeutigkeitspunkt erreicht ist. Doch offensichtlich erleichtert dies nicht das Erkennen des Zielphonems. Nur diejenigen Ziellaute, die *nach* dem Eindeutigkeitspunkt auftreten, werden leichter erkannt. Anders ausgedrückt, der Erleichterungseffekt ist abhängig davon, ob ein Wort ausgewählt worden ist, nicht aber davon, ob es aktiviert worden ist.

Wie verhält es sich mit den eingebetteten Wörtern? Sie sollten keinen Erleichterungseffekt bewirken, da sie zwar aktiviert, nicht aber ausgewählt werden. Judith HENSTRA hat kürzlich an unserem Institut genau dies auch gefunden (unveröffentlichte Diplomarbeit). In einem ihrer Experimente manipulierte sie die Aktivationsstärke eines eingebetteten Wortes. Ich möchte ihr Vorgehen an einem Beispiel erläutern. Nehmen Sie als Zielwort *Barock*. Es beinhaltet das Wort *Rock*. Wir wissen, daß ein Wort zusätzlich aktiviert wird, wenn ihm ein semantisch ähnliches Wort vorangeht. Wenn wir also vor dem Zielwort *Barock* das Wort *Kleid* darbieten, erhält das eingebettete Wort *Rock* zusätzliche Aktivierung. Wenn wir aber vor dem Wort *Barock* das Wort *Harke* präsentieren, wird die Aktivationsstärke des eingebetteten Wortes *Rock* nicht erhöht. In HENSTRAS Experiment hörten die Versuchspersonen eine Liste gesprochener Wortpaare, wie *Kleid – Barock* oder *Harke – Barock*. Aufgabe der Versuchspersonen war, eine Taste zu drücken, wenn sie im zweiten Wort den Ziellaut /k/ hörten. Wenn, so die Hypothese, die Aktivierung eines eingebetteten Wortes das Erkennen seiner konstituierenden Sprachlaute erleichtert, dann sollte das /k/ in *Kleid – Barock* schneller als in *Harke – Barock* erkannt werden. Die Ergebnisse des Experiments zeigen jedoch,

daß dies nicht der Fall ist<sup>1</sup>. Das wiederum, so vermute ich, hängt damit zusammen, daß das eingebettete Wort zwar aktiviert, aber nicht ausgewählt worden ist.

Lassen Sie mich diesen ersten Teil über Aspekte der Worterkennung zusammenfassen. Das Erkennen eines Wortes erleichtert offensichtlich das Erkennen seiner konstituierenden Sprachlaute. Aber dafür reicht die bloße Aktivierung des Wortes nicht aus. Nur wenn das Wort auch *ausgewählt* worden ist, erleichtert es die nachfolgende Verarbeitung der Sprachlaute. Aus konnektionistischer Perspektive ist dieser Befund nur schwer zu erklären. Entgegen den Behauptungen von ELMAN und MCCLELLAND (1988) leiten aktivierte Wortknoten keine Aktivierung zu Phonemknoten, d. h. Knoten auf der Ebene der Sprachlaute, zurück. Die dargestellten Befunde ergeben ein deutlich anderes Bild. Sobald der Hörer ein Wort ausgewählt, d. h. erkannt hat, wird die Gedächtnisspur des zugehörigen Lautmusters verfügbar. Dieses Lautmuster unterstützt anschließend das *phoneme monitoring*. Wir wissen einfach, daß *Barock* den Laut /k/ enthält. CUTLER und NORRIS (1979) vermuten, daß es bei der *phoneme monitoring task* einen Wettlauf zwischen zwei Prozessen gibt. Der erste ist gesteuert durch die akustischen Charakteristika des Reizes. Sie hören den Reiz /k/ im akustischen Muster /barok/. Der zweite Prozeß ist wissensgesteuert. Sobald Sie erkannt haben, daß das betreffende Wort *Barock* ist, können Sie ableiten, daß es den Laut /k/ enthält. Derjenige Prozeß, der den Wettlauf gewinnt, bestimmt die Schnelligkeit der Reaktion. Der wissensgesteuerte Prozeß kann allerdings erst nach dem Eindeutigkeitspunkt einsetzen. Bloß aktivierte Wörter haben keinen Zugriff auf ihre jeweilige Lautform.

Im folgenden möchte ich mich dem Aspekt der Mustererzeugung, also der Produktion gesprochener Sprache, zuwenden. Wie wir sehen werden, zeichnet sich in diesem Bereich ein ganz ähnliches Bild ab.

### *Aktivierung und Auswahl bei der Wortproduktion*

In der Einleitung habe ich zwei Stufen bei der Produktion von Wörtern unterschieden. Ich möchte das am Beispiel einer Bildbenennungsaufgabe verdeutlichen. Stellen Sie sich einmal kurz vor, was geschieht, wenn jemand eine einfache Zeichnung, in der beispielsweise ein Schaf dargestellt ist, benennt. Die Zeichnung wird zunächst eine Anzahl von möglichen Zielwörtern aktivieren, neben dem Zielwort *Schaf* vielleicht auch bedeutungsähnliche Wörter wie *Ziege*. Zu einem bestimmten Zeitpunkt wählt der Sprecher das Zielwort *Schaf* aus. Damit ist der erste Schritt vollzogen, der Prozeß der *lexikalen Auswahl*. Während eines zweiten Schritts wird auf die Lautform des Zielwortes zugegriffen und ein artikulatorisches Muster erstellt (/schaf/). Diesen Schritt bezeichnen wir als *phonologische Enkodierung*.

---

<sup>1</sup> HENSTRA fand einen Effekt eingebetteter Wörter am Wortanfang, wie in *Raster* (eingebettet: *Rast*). Meine Interpretation dieses Befundes ist, daß – möglicherweise durch Priming induziert – das eingebettete Wort nicht nur aktiviert, sondern auch ausgewählt worden ist. Diese Auswahl ist fehlerhaft und muß natürlich korrigiert werden, sobald weitere Sprachlaute verarbeitet werden. Der ebenfalls beobachtete Effekt für einzelne Wörter, die in Nichtwörtern eingebettet sind, könnte in gleicher Weise durch deren Auswahl bedingt worden sein; in diesen Fällen gibt es kein (anderes) Wort, das hätte ausgewählt werden können. Wenn diese Vermutung zutrifft, sollte kein Effekt zu finden sein für ein zweites in ein Nichtwort eingebettetes Wort (wie in *kiloop*, das zwei niederländische Wörter enthält, nämlich *kilo* und *loop*). In einem solchen Falle sollte nur das erste, nicht aber das zweite Wort ausgewählt werden. Genau dies hat HENSTRA in einem weiteren Experiment auch gefunden.

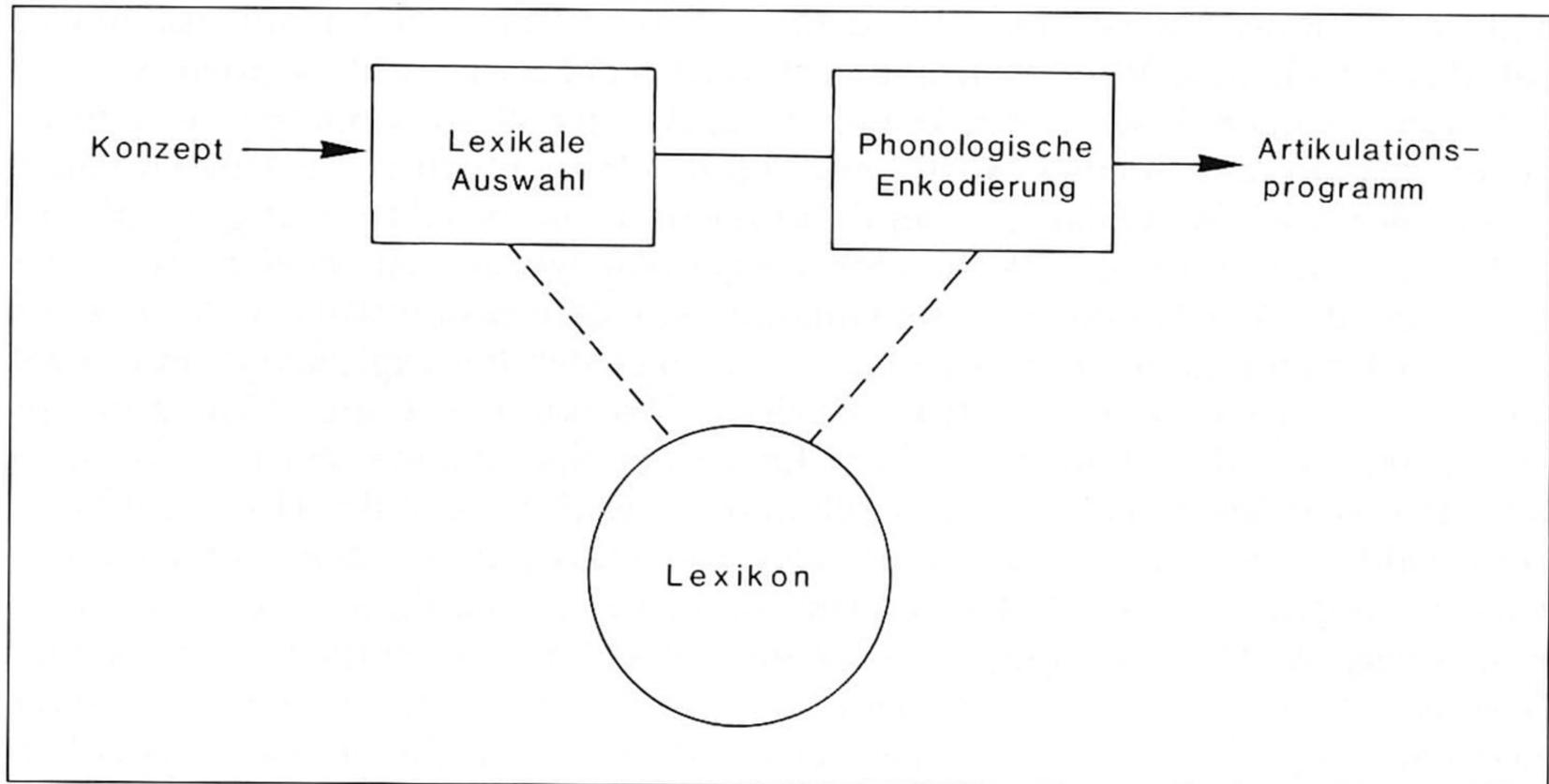


Abb. 5 Zwei Stufen beim lexikalen Zugriff in der Sprachproduktion: lexikale Auswahl und phonologische Enkodierung

zung (vgl.: LEVELT 1989). Sobald dieses Muster verfügbar ist, kann das Wort ausgesprochen werden (vgl.: Abb. 5).

Wenden wir uns nun erneut den Aspekten Aktivierung und Auswahl zu. Wie wir zuvor gesehen haben, können beim Worterkennen nur ausgewählte Wörter auf ihre gespeicherte Lautform zugreifen. Gilt dies auch für den Bereich der Wortproduktion? Wenn ja, dann sollten nur ausgewählte Wörter, in unserem Beispiel das Wort *Schaf*, phonologisch spezifiziert werden, nicht aber Wörter, die bloß aktiviert worden sind, wie *Ziege*. Insbesondere sollte der Sprecher keine artikulatorischen Muster für solche Wörter erstellen.

Dieser Punkt ist stark umstritten; konnektionistische Theorien der Sprachproduktion sagen genau das Gegenteil vorher. Solchen Theorien zufolge leitet jeder Wortknoten, der aktiviert wird, seine Aktivierung auf die Ebene der Sprachlaute weiter, unabhängig davon, ob das betreffende Wort auch ausgewählt worden ist oder nicht (DELL 1986). Vor kurzem haben wir dazu eine umfangreiche Studie veröffentlicht (LEVELT et al. 1991), deren Ergebnisse, so meinen wir, völlig eindeutig sind. Lassen Sie mich zunächst einmal unser experimentelles Vorgehen darstellen.

Aufgabe unserer Versuchspersonen war die Benennung von einfachen Bildern. Sie sahen nacheinander eine umfangreiche Serie von Bildern, und sollten jedes so schnell wie möglich benennen. In einigen Durchgängen mußten sie eine zweite Aufgabe ausführen. Diese zweite Aufgabe war eine lexikale Entscheidungsaufgabe. Kurz nach dem Erscheinen des Bildes hörten die Versuchspersonen einen akustischen Stimulus, der entweder ein Wort, wie *tip*, oder ein Nichtwort, wie *sef*, war. Immer wenn die Versuchspersonen einen akustischen Reiz hörten, sollten sie so schnell wie möglich durch Drücken einer von zwei Tasten angeben, ob es ein Wort ist oder nicht. Gemessen haben wir die Zeitspanne zwischen dem Beginn des akustischen Stimulus und dem Tastendruck.

Was wir zunächst mit dieser Aufgabe erfassen können, ist die phonologische Aktivierung des Zielwortes. Nehmen wir noch einmal an, daß die Versuchsperson das Bild eines Schafes sieht. Sie bereitet unmittelbar die Benennung des Bildes vor.

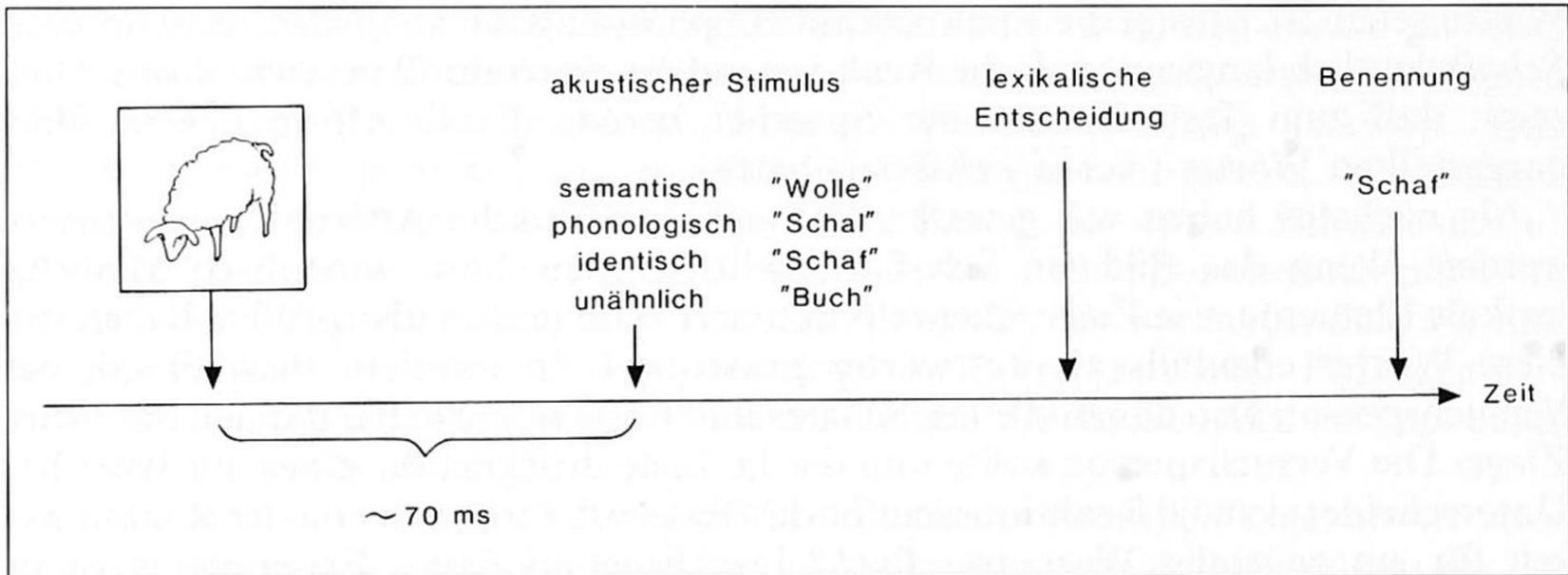


Abb. 6 Schematische Darstellung der kombinierten Benennungs-/lexikalen Entscheidungsaufgabe

Kurz nach dem Beginn der Bilddarbietung (ungefähr 70 ms danach) hört die Versuchsperson das Wort *Schal*, das dem Zielwort *Schaf* phonologisch ähnlich ist. Die Versuchsperson drückt nun die Ja-Taste, da *Schal* ein Wort ist, und wir messen, wieviel Zeit zwischen dem Beginn des Testworts *Schal* und dem Tastendruck vergangen ist. Wenn die Versuchsperson zu dem Zeitpunkt, zu dem sie das Testwort hört, bereits auf die Lautform des Zielwortes zugegriffen hat, dann sollte dies die Geschwindigkeit der lexikalen Entscheidung für das Wort *Schal* beeinflussen. Um dies zu testen, haben wir auch ein neutrales Testwort als Kontrollbedingung verwendet. Dieses neutrale Testwort kann beispielsweise *Buch* sein, denn es ist dem Wort *Schaf* phonologisch nicht ähnlich. In Abbildung 7 sind die Ergebnisse eines solchen Experiments dargestellt.

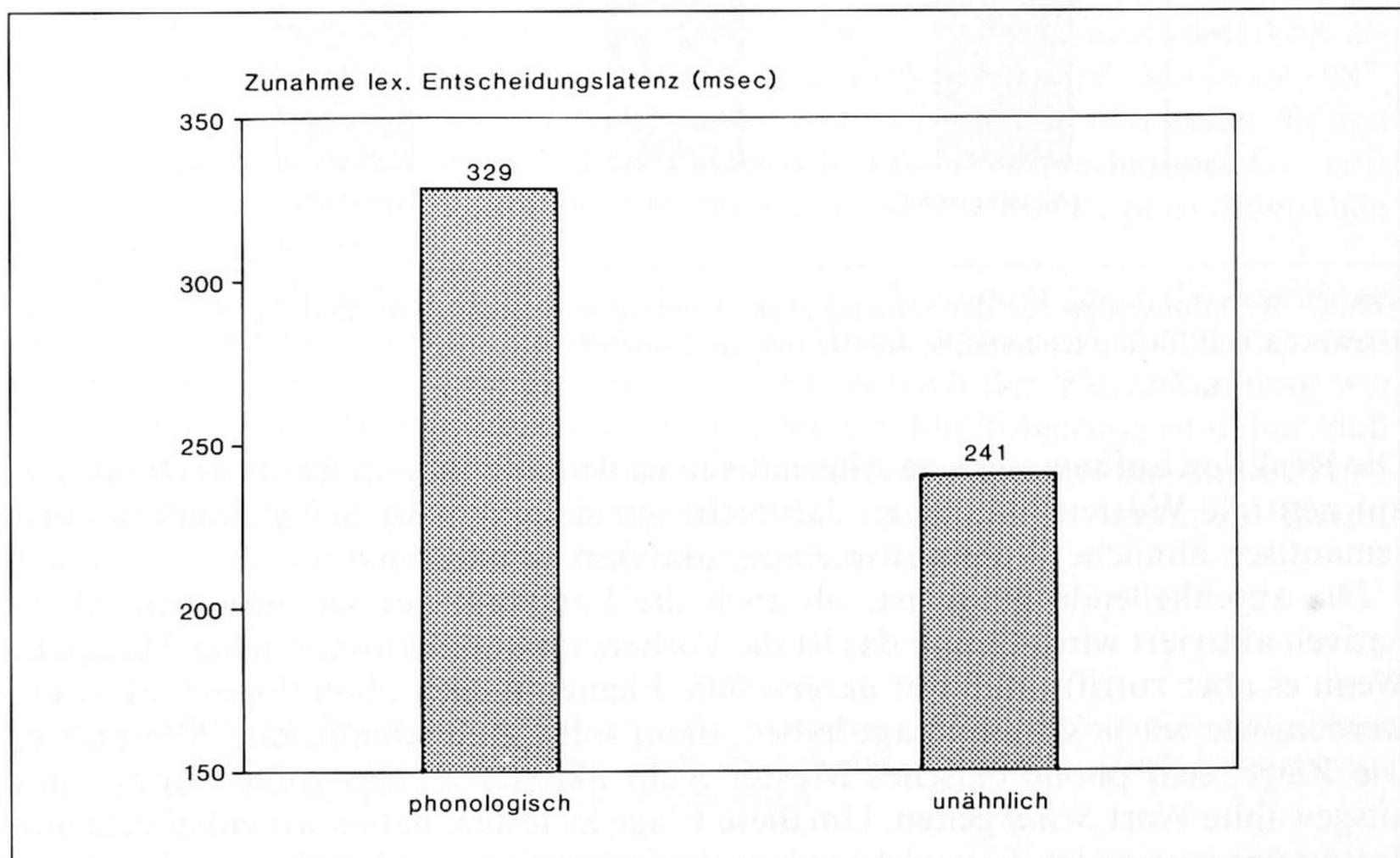


Abb. 7 Reaktionszeiten für dem Zielwort phonologisch ähnliche und unähnliche Testwörter

Wie zu sehen ist, erfolgt die Reaktion auf ein phonologisch ähnliches Testwort (wie *Schal*) deutlich langsamer als die Reaktion auf ein neutrales Wort (wie *Buch*). Dies zeigt, daß zum Testzeitpunkt der Sprecher bereits die Lautform des im Bild dargestellten Wortes (/schaf/) aktiviert hatte.

Als nächstes haben wir getestet, ob auch semantische Alternativen aktiviert werden. Wenn das Bild ein Schaf darstellt, werden dann semantisch ähnliche lexikale Elemente, wie *Ziege*, ebenfalls aktiviert? Um dies zu überprüfen, haben wir diese Wörter ebenfalls als Testwörter präsentiert. In unserem Beispiel sah die Versuchsperson also das Bild eines Schafes und hörte unmittelbar danach das Wort *Ziege*. Die Versuchsperson sollte nun die Ja-Taste drücken, da *Ziege* ein Wort ist. Unterscheidet sich die Reaktionszeit für das Testwort *Ziege* nun von der Reaktionszeit für ein neutrales Wort, wie *Buch*? Das Ergebnis dieses Experiments ist in Abbildung 8 dargestellt.

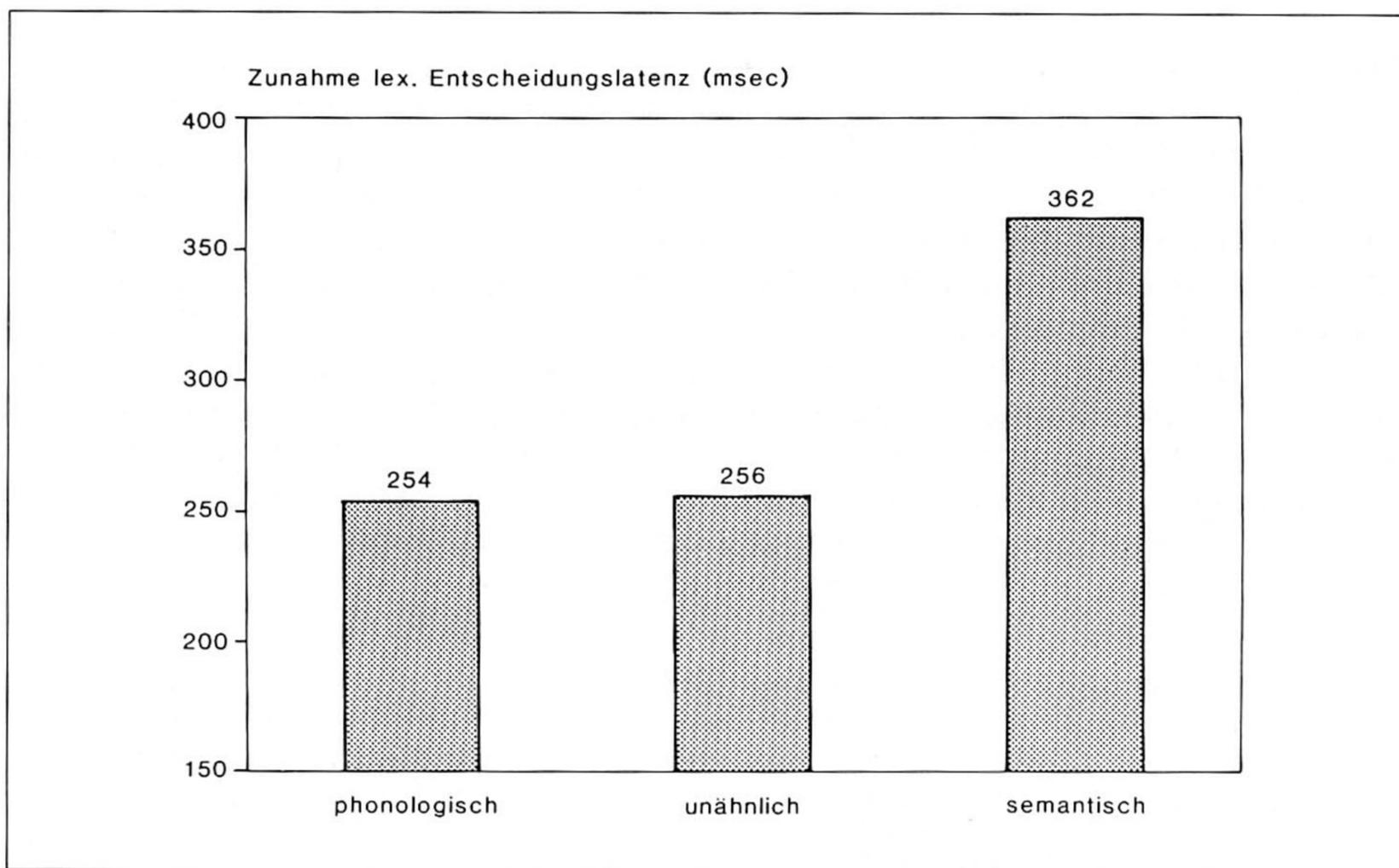


Abb. 8 Reaktionszeiten für der semantischen Alternative phonologisch ähnliche und unähnliche Testwörter und für die semantische Alternative als Testwort

Die Reaktion auf semantische Alternativen ist deutlich langsamer als die Reaktion auf neutrale Wörter. Dies zeigt, daß nicht nur das Zielwort *Schaf*, sondern auch semantisch ähnliche Wörter, wie *Ziege*, aktiviert worden sind.

Die abschließende Frage ist, ob auch die Lautform der semantischen Alternativen aktiviert wird. Genau das ist die Vorhersage konnektionistischer Theorien. Wenn es aber zutrifft, daß nur *ausgewählte* Elemente auch phonologisch aktiviert werden, wie wir ja vorausgesagt haben, dann sollte eine semantische Alternative, wie *Ziege*, sein phonologisches Muster nicht aktivieren. Das sollte nur für das ausgewählte Wort *Schaf* gelten. Um diese Frage zu testen, haben wir entsprechende Testwörter verwendet. Ein solches phonologisches Testwort für *Ziege* ist beispielsweise *Zierde*. Die Versuchsperson sah also das Bild eines Schafes. Ungefähr 70 ms

später hörte sie das Testwort *Zierde* und drückte die Ja-Taste, weil *Zierde* ein Wort ist. Wenn nun *Ziege* phonologisch aktiviert worden ist, dann erwarten wir längere Reaktionszeiten für das Testwort *Zierde* als für ein neutrales Testwort. Das Ergebnis dieses Experimentes ist ebenfalls in Abbildung 8 dargestellt. Der Befund ist eindeutig: Es gibt keinerlei Unterschied. Wir haben also keinen Hinweis darauf gefunden, daß aktivierte semantische Alternativen auch phonologisch aktiviert werden. Die Schlußfolgerung daraus ist, daß auch in der Sprachproduktion nur *ausgewählte* lexikale Elemente ihre phonologische Form aktivieren. Entgegen den Vorhersagen konnektionistischer Theorien reicht die bloße Aktivierung eines lexikalischen Elementes dafür nicht aus.

Lassen Sie mich schließen mit einigen Spekulationen über die Frage, ob wir über zwei mentale Lexika verfügen, eines für das Worterkennen und ein anderes für die Wortproduktion, oder ein einziges, das beiden Funktionen dient.

### *Ein einheitliches System für Worterkennen und Wortproduktion?*

In psycholinguistischen Kreisen gibt es deutlich unterschiedliche Ansichten bezüglich der Frage, ob wir über ein oder zwei mentale Lexika verfügen. ALLPORT (1984) hat argumentiert, daß, wenn wir über zwei mentale Lexika verfügten, es möglich sein sollte, aphasische Patienten zu finden, die eine selektive Schädigung eines der beiden Lexika aufweisen. Gibt es, so seine Frage, Patienten, deren Fähigkeit zur Produktion von Wörtern eingeschränkt ist, nicht aber deren Fähigkeit zum Verstehen von Wörtern? Umgekehrt, lassen sich Patienten mit intakter Wortproduktion aber geschädigtem Wortverstehen finden? ALLPORTS Sichtung der relevanten Literatur ergab keinerlei Hinweise auf solche Patienten.

Unabhängig von derartigen Überlegungen haben Konnektionisten argumentiert, daß es ein einheitliches Lexikon gäbe, auf das sowohl beim Sprachverstehen als auch bei der Sprachproduktion zugegriffen werde. Sie fassen dieses Lexikon als Netzwerk auf, in dem bidirektional Aktivierung weitergeleitet wird (MACKAY 1987, DELL 1988). Aber gerade diese letzte Annahme ist falsch; sie läßt sich weder für den Verstehensprozeß noch für den Produktionsprozeß aufrechterhalten. Es trifft einfach nicht zu, daß aktivierte lexikale Einheiten Aktivierung zu phonologischen Einheiten weiterleiten.

Diese falsche Modellvorhersage ist allerdings kein Argument gegen die Annahme eines einheitlichen lexikalischen Systems. Das eindrucksvollste Merkmal der dargestellten Befunde ist, daß sie gleichermaßen im Bereich der Worterkennung wie auch im Bereich der Wortproduktion gelten. Meine Schlußfolgerung ist daher, daß diese neuen Befunde aus unserem Institut die Annahme eines einheitlichen Lexikons unterstützen. Daß es nicht konnektionistischer Art ist, sollte uns nicht davon abschrecken, diesen wichtigen Aspekt weiter zu verfolgen.

### *Danksagung*

Ich danke Herrn Jörg JESCHENIAK herzlich für die Übersetzung aus dem englischen Original, Herrn Dr. Harald BAAYEN und Frau Frauke HELWIG für die der Abbildung 3 zugrunde liegende Berechnungen und Frau Inge TARIM für die Anfertigung der Zeichnungen.

## Literatur

- ALLPORT, D. A.: Speech production and comprehension: One lexicon or two? In: PRINZ, W., and SANDERS, A. F. (Eds.): *Cognition and motor processes*, pp. 209–228. Berlin, Heidelberg: Springer 1984
- CREUTZFELDT, O., OJEMANN, G., and LETTICH, E.: Neuronal activity in the human temporal lobe I. Responses to speech. *Exp. Brain Res.* 77, 451–475 (1989)
- CUTLER, A., and NORRIS, D.: Monitoring sentence comprehension. In: COOPER, W. E., and WALKER, E. C. T. (Eds.): *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, pp. 113–134. Hillsdale, NJ: Erlbaum 1979
- DELL, G.: A spreading activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321 (1986)
- DELL, G.: The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *J. Mem. Lang.* 27, 124–142 (1988)
- ELMAN, J. L., and MCCLELLAND, J. L.: Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *J. Mem. Lang.* 27, 143–165 (1988)
- FRAUENFELDER, U., SEGUI, J., and DIJKSTRA, T.: Lexical effects in phonemic processing: Facilitatory or inhibitory? *J. Exp. Psychol.: Hum. Perc. Perform.* 16, 77–91 (1990)
- LEVELT, W. J. M.: *Speaking: From intention to articulation*. Cambridge, MA.: MIT Press 1989
- LEVELT, W. J. M., SCHRIEFERS, H., VORBERG, D., MEYER, A. S., PECHMANN, T., and HAVINGA, J.: The time course of lexical access in speech production: A study of picture naming. *Psych. Rev.* 98, 122–142 (1991)
- MARSLEN-WILSON, W., and TYLER, L.: Central processes in speech understanding. *Phil. Trans. R. Soc. Lond. Series B* 295, 317–322 (1981)
- MACKAY, D.: *The organization of perception and action*. New York: Springer 1987
- SHILLCOCK, R.: Lexical hypotheses in continuous speech. In: ALTMANN, G. (Ed.): *Cognitive models of speech processing*, pp. 24–59. Cambridge, MA.: MIT Press 1990
- ZWITSERLOOD, P.: The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32, 25–64 (1989)

Prof. Dr. Willem J. M. LEVELT  
MPI für Psycholinguistik  
Wundtlaan 1  
NL 6525 XD, Nijmegen