

5 Wortverarbeitung

Katharina Spalek

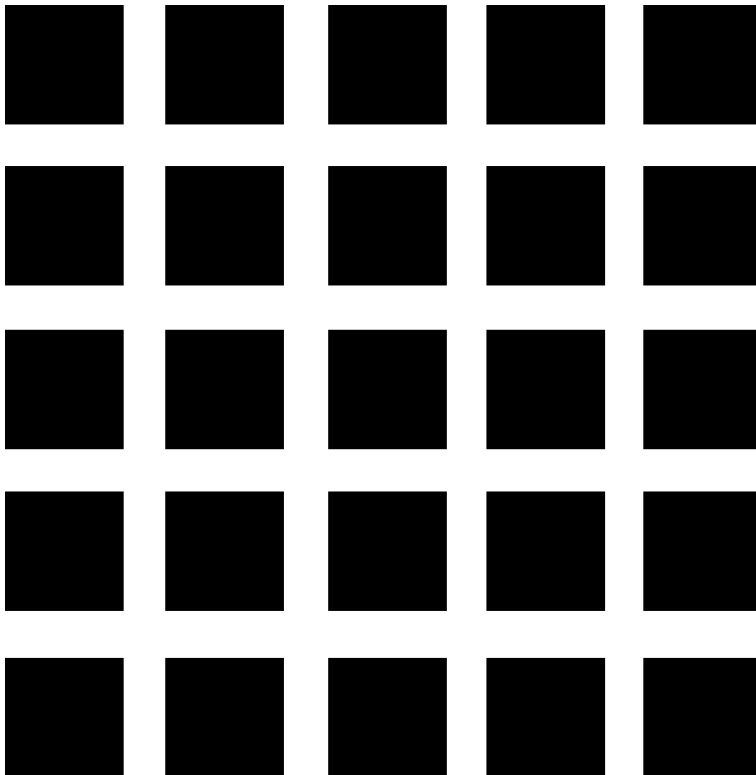


Abbildung 8: Optische Täuschung

Wieviele kleine dunkle Kreise sehen Sie auf den weißen Linien zwischen den schwarzen Rechtecken? Fangen Sie nicht an zu zählen – es gibt gar keine dunklen Kreise, sie sind nur das Produkt einer optischen Täuschung. Solche Gitter wie das hier gezeigte nennt man entweder Hermann-Gitter oder auch Hering-Gitter, da die beiden deutschen Physiologen und Hirnforscher Ludimar Hermann (1838–1914) und Ewald Hering (1834–1918) diese optische Täuschung bemerkt und beschrieben haben. Die Täuschung beruht auf dem folgenden Prinzip: Während der Verarbeitung des Nervenreizes werden vorhandene Kontraste je nach ihrer Position entweder verstärkt oder abgeschwächt. Dies geschieht automatisch und hängt mit der Verschaltung unserer Nervenzellen zusammen. Nervenzellen sind spezialisierte Rezeptoren, zum Beispiel für hell und dunkel. Jeder Rezeptor hemmt benachbarte Rezeptoren mit einem Bruchteil seiner eigenen Aktivierung. Dadurch wird die Aktivierungsenergie eines Rezeptors (und damit letztendlich die Intensität dessen, was wir sehen) nicht nur durch das Signal beeinflusst, sondern auch durch interne Verarbeitungsprozesse unseres Wahrnehmungsapparates.

Optische Täuschungen unterstreichen eine allgemeine Tatsache der Wahrnehmung, die auch für das Hören von Wörtern zutrifft. Was wir wahrnehmen, ist niemals ein exaktes Abbild des physikalischen Signals. Unser Wahrnehmungsapparat im Allgemeinen und das Sprachverarbeitungssystem im Besonderen filtern das Signal nach bestimmten Kriterien. Dabei entsteht aus dem Signal eine abstrakte Repräsentation, die die Grundlage des lexikalischen Zugriffs bildet. Darüber hinaus wird die Verarbeitung eines Sprachsignals direkt oder indirekt durch unser sprachliches Wissen beeinflusst. Zahlreiche empirische Befunde zeigen, wie unser Sprachverstehenssystem Informationen aus dem Signal herausfiltert, die für die Worterkennung irrelevant sind, und wie es Informationen kontextabhängig ergänzt, die im Signal nicht enthalten sind. Die Herausforderung für Modelle des Worterkennens ist es, den Prozess der Wortverarbeitung unter Berücksichtigung genau dieser Beobachtungen überzeugend zu erklären.

5.1 Lexikalischer Zugriff

5.2 Segmentierung von Wörtern

5.3 Worterkennen im Kontext

5.4 Modelle des Worterkennens

5.1 Lexikalischer Zugriff

Wenn wir ein Wort erkennen, so wissen wir, dass es sich um ein Wort unserer Muttersprache handelt, welcher Wortart es angehört, was es bedeutet und mit welchen anderen Wörtern es auftreten kann. Anders ausgedrückt, bedeutet das Erkennen eines Wortes, dass wir auf die Information zugreifen können, die zu diesem Wort in unserem mentalen Lexikon enthalten ist (→ KAPITEL 4.1). In vielen Beschreibungen des Wortverstehens werden drei Prozesse unterschieden – Zugriff, Auswahl und Integration. In der englischsprachigen Fachliteratur nennt man diese drei Prozesse meist *access* oder *initial contact*, *selection* und *lexical integration*. Beim Zugriff werden Einträge des Lexikons aktiviert, die mit dem wahrgenommenen Signal übereinstimmen. Im Normalfall bleibt nach einer anfänglichen Aktivierung mehrerer möglicher Kandidaten nur ein Kandidat übrig. Dieser Eintrag wird ausgewählt (Selektion). Nach erfolgreicher Selektion weiß man jedoch im Grunde nur, dass es ein Wort im Wortschatz gibt, das mit dem wahrgenommenen Signal ausreichend gut übereinstimmt. Die Aktivierung der Information, die mit dem selektierten Eintrag verbunden ist, geschieht erst während der lexikalischen Integration. Während diese Integration das Herzstück des eigentlichen Verstehens darstellt, sind auch lexikalischer Zugriff und Auswahl keine trivialen Prozesse. Diese beiden Vorgänge wurden in der psycholinguistischen Forschung sogar intensiver untersucht als die Integration.

Zugriff, Auswahl,
Integration

Psycholinguisten wollen zunächst verstehen, wie der initiale Kontakt mit dem Lexikon erfolgt. Wird das akustische Signal direkt auf Repräsentationen im Gedächtnis abgebildet? Die sogenannte *episodic lexicon theory* (Goldinger 1998) geht von solch einer direkten Abbildung des Signals auf einen Lexikoneintrag aus. Hierbei besteht der Lexikoneintrag eines Wortes aus Gedächtnisspuren jeder individuellen Realisation dieses Wortes. Das sprachliche Signal wird mit diesen Gedächtnisspuren verglichen (und kann seinerseits den lexikalischen Eintrag verändern). Die Mehrheit der Forscher nimmt dagegen an, dass eine zusätzliche Repräsentation zwischen dem Signal und dem Lexikon existiert, die sogenannte prälexikalische Repräsentation. Auf dieser Ebene werden abstrakte Einheiten aus dem Schallsignal extrahiert (→ KAPITEL 3.1), die dann den Input für die lexikalische Suche bilden. Der Vorteil einer solchen prälexikalischen Repräsentation besteht darin, dass die Variabilität des Sprachsignals verringert wird. Das Sprachsignal enthält viele Informationen, die für die Worterkennung nicht direkt wichtig sind. Eine akustische Realisierung eines Wortes ist niemals ab-

Episodic lexicon
theory

Prälexikalische
Repräsentation

solot identisch mit einer anderen akustischen Realisierung desselben Wortes. Eine Repräsentation, in der diese Vielzahl möglicher Schallsignale auf eine begrenzte Zahl von Phonemen oder Silben reduziert wird, kann daher die Effizienz der Wortverarbeitung erhöhen. Eine Alltagsbeobachtung stützt die Annahme einer solchen Zwischenrepräsentation: Wir erinnern uns zwar an gesprochene Worte, haben aber bereits nach kürzester Zeit keine Erinnerung mehr an Einzelheiten der akustischen Realisierung dieser Wörter (es sei denn, es ist dabei zu auffälligen Versprechern gekommen).

Phonem ...

Aus welchen Einheiten setzt sich eine prälexikalische Repräsentation zusammen? Hierzu gibt es verschiedene Vorschläge in der Fachliteratur. Ein häufig genannter Kandidat ist das Phonem. Für diese Annahme gibt es sowohl theoretische als auch empirische Gründe: Das Phonem ist die kleinste linguistische Einheit, die bedeutungsunterscheidend ist: sogenannte Minimalpaare wie *Latte* und *Ratte* haben eine andere Bedeutung und unterscheiden sich formal nur durch ein Phonem. Außerdem zeigt die kategoriale Wahrnehmung (→ KAPITEL 3.2), dass unser Wahrnehmungsapparat klangliche Unterschiede, die für die Phonemunterscheidung nicht relevant sind, nivelliert. Es gibt allerdings Beobachtungen, die sich weniger gut mit der Annahme vertragen, dass das Phonem die prälexikalische Einheit ist. Die Worterkennung leidet unter der akustischen Manipulation eines Lautes, auch wenn das manipulierte Merkmal linguistisch gesehen kein distinktives Merkmal ist, wenn sich also durch diese Manipulation die Zugehörigkeit des Lautes zu einer Phonemkategorie nicht verschiebt. Ein hiermit verwandtes Argument gegen das Phonem als prälexikalische Einheit ist die Beobachtung, dass akustische Unterschiede in der Realisierung desselben Phonems zur Worterkennung beitragen. Probanden in einem englischen Experiment konnten mit großer Sicherheit vorhersagen, ob sie das Wort *cap* oder das Wort *captain* hörten, noch bevor das /t/ im Signal erschienen war – die Länge des Vokals /a/ relativ zur Länge der Konsonanten /c/ und /p/ unterscheidet sich in *cap* und *captain* systematisch (Davis u. a. 2002). Dieser Längenunterschied trägt jedoch im Englischen nicht zur Phonemidentität bei. Wenn das sprachliche Signal auf ein Phonem abgebildet wird, geht diese Information also verloren. Da solche Unterschiede aber offensichtlich zum Erfolg der Worterkennung beitragen, kann es nicht sein, dass die prälexikalische Repräsentation sie vollständig herausfiltert. Einige Theorien nehmen daher an, dass der Kontakt zum Lexikon mithilfe von spektralen Repräsentationen gelegt wird. Hierbei werden Informationen über die Lage von Formanten, die Qualität des Burst (→ KAPITEL 3.1) und ähnliches aus dem Signal extrahiert. Aditi Lahiri und

... spektrale
Information ...

William Marslen-Wilson schlagen ebenfalls vor, dass die prälexikalische Einheit kleiner ist als ein Phonem (Lahiri/Marslen-Wilson 1991). In ihrem Modell dienen distinktive Merkmale wie Artikulationsmodus, Artikulationsstelle und Stimmbeteiligung als Input.

... oder distinktives
Merkmal als
prälexikalische
Einheit?

Während distinktive Merkmale und spektrale Eigenschaften kleiner sind als ein Phonem, gibt es auch Ansätze, nach denen die prälexikalische Einheit größer ist als ein Phonem. Jacques Mehler und seine Kollegen haben ein Experiment durchgeführt, in dem sie die Rolle der Silbe bei der Worterkennung untersucht haben (Mehler u. a. 1981). Probanden sollten eine sogenannte Fragmententdeckungsaufgabe durchführen. Dabei sollten sie so schnell wie möglich reagieren, wenn sie in einem Wort eine vorher spezifizierte Lautkombination wahrnahmen, z. B. *pa*. Die Probanden reagierten schneller, wenn die Lautkombination mit einer Silbe übereinstimmte, z. B. in *Pa-last*, als wenn sie nur Teil einer Silbe war, z. B. in *Pal-me*. Die Reaktion auf ein zu entdeckendes Fragment *pal* wurde ebenfalls getestet. Dieses Fragment wurde in *Pal-me* schneller entdeckt als in *Pa-last*. Dieser sogenannte Silbenpräferenzeffekt tritt nicht in jeder Sprache auf. Auch in den Sprachen, in denen er ursprünglich nachgewiesen wurde, lässt er sich nur bedingt replizieren und ist darum in der Literatur inzwischen umstritten. Dennoch gibt es andere Indizien dafür, dass die Silbe beim Worterkennen eine wichtige Rolle spielt. Die Psycholinguistin Anne Cutler und ihre Kollegen nehmen an, dass die metrische Struktur des akustischen Signals eine wichtige Rolle bei seiner Zerlegung in kleinere Einheiten spielt (Cutler u. a. 1986). Sie konnten zeigen, dass eine Sprache wie das Englische vor betonten Silben segmentiert wird, eine Sprache wie Französisch dagegen bei jeder Silbengrenze unabhängig von der Betonung. Diese metrischen Einheiten dienen dann zur Suche im mentalen Lexikon. Eine Segmentierung in Silben vor der lexikalischen Suche ist von Vorteil, weil die akustische Realisierung von Silben weniger variabel ist als diejenige von Phonemen.

Silbenpräferenzeffekt

Silben als
prälexikalische
Einheit

5.2 Segmentierung von Wörtern

Eine Herausforderung für unser Worterkennenssystem ist die Tatsache, dass Wörter meistens nicht in Isolation artikuliert werden, sondern in Kombination mit anderen Wörtern. Außerdem sind kürzere häufig in längere Wörter eingebettet, z. B. *Tee*, *Thema* und *Thematik* in *Mathematik*. Bevor es auf das Lexikon zugreifen kann, muss das System also erkennen, wo ein Wort beginnt. Dazu dienen dem erwachsenen Spre-

Metrik

cher ähnliche Hinweisreize wie dem Kind im Spracherwerb (→ KAPITEL 9.2). Jede Sprache hat typische metrische Muster. So werden im Deutschen Wörter oft auf der ersten Silbe betont. Obwohl es viele Ausnahmen gibt, kann diese Regel dem System einen ersten Anhaltspunkt hinsichtlich der Position der Wortgrenzen geben.

Phonotaktische Regeln

In jeder Sprache gibt es Kombinationen von Lauten, die innerhalb eines Wortes oder einer Silbe nicht vorkommen können. Im Deutschen kann z. B. /tʌ/ nicht in einer Silbe vorkommen – weder *tʌ noch *atʌ ist eine wohlgeformte Silbe des Deutschen. Über Silbengrenzen hinweg sind diese Kombinationen jedoch sehr wohl möglich, etwa in *At-las*. Andere Kombinationen sind auf bestimmte Silbentypen beschränkt. So kann eine deutsche Silbe zwar nicht mit /ʃ/ beginnen, wohl aber enden (z. B. *Wolf*). Umgekehrt ist die Kombination /ʃ/ nur am Wortanfang zulässig (z. B. *Flöte*), aber nicht am Wortende. Solche phonotaktischen Regeln können ebenfalls zur Segmentierung von Sprache verwendet werden.

Vokalharmonie

Wörter in Sprachen wie dem Türkischen und dem Finnischen unterliegen der sogenannten Vokalharmonie. Dabei dürfen innerhalb eines Wortes nur bestimmte Vokalklassen (z. B. hintere oder vordere Vokale) vorkommen. So heißt *Du bist Deutsche(r)* auf Türkisch *Alman-sın*, *Du bist Türke/in* dagegen *Türk-sün*. Häuser sind *ev-ler*, Jahre dagegen *yıl-lar*. Ein Wechsel der Vokalklasse kann daher eine Wortgrenze signalisieren, und Studien zeigen, dass Sprecher diese Information zur Segmentierung von Sprache nutzen (für das Finnische: Suomi u. a. 1997).

Lexikalisches Wissen

Im Gegensatz zu Kindern verfügen erwachsene Sprecher über ein Wortrepertoire ihrer Sprache. Viele Wörter, vor allem längere, müssen nicht zur Gänze gehört werden, bevor man sie erkennen kann. Es gibt einen Punkt, den sogenannten Diskriminationspunkt (*uniqueness point*), an dem sich ein Wort von allen anderen Wörtern einer Sprache unterscheidet. Alle Information, die diesem Punkt folgt, ist für den Prozess der Worterkennung redundant. Das lexikalische Wissen kann also dafür genutzt werden, eine Wortgrenze zu antizipieren.

Diskriminationspunkt

Eingebettete Wörter

Wie geht unser Sprachsystem mit eingebetteten Wörtern um? Warum wird nicht jedes Mal, wenn wir *Mathematik* hören, auch das Wort *Thema* aktiviert – oder wird es das vielleicht doch? Eine erste Hilfe leisten die Unterschiede in der phonetischen Realisierung (→ KAPITEL 5.1) von eingebetteten und freistehenden Wörtern. Außerdem spielt auch hier metrische Information eine Rolle. Anne Cutler und Dennis Norris (1988) haben untersucht, wie gut Probanden ein Wort erkennen, das in ein Nichtwort eingebettet ist (Cutler/Norris 1988). Die Erkennensleistung hängt davon ab, ob das eingebettete Wort in einer unbetonten Sil-

be auftritt oder in einer betonten. Unbetonte eingebettete Wörter werden häufig überhört. Der lexikalische Kontext hat ebenfalls einen Einfluss auf das Erkennen eingebetteter Wörter (McQueen u. a. 1994): Ein englisches Wort wie z. B. *mess* wurde häufiger in einem Nichtwort (→ KAPITEL 2.1) wie *nemes* wahrgenommen als in einem Nichtwort wie *domes*. Die Autoren argumentierten, dass *domes* ein möglicher Wortbeginn im Englischen ist (*domestic*) und daher das Erkennen von *mess* blockiert. Dennoch werden auch Wörter verarbeitet, die in längere Wörter eingebettet sind, wie verschiedene Autoren mit Priming-Experimenten (→ KAPITEL 2.1) gezeigt haben. *Captain* erleichtert über das eingebettete *cap* (Kappe, Mütze) das Erkennen des Wortes *hat* (Hut). Am Wortanfang eingebettete Wörter sind bessere Primes als am Wortende eingebettete Wörter. Es gibt aber auch Befunde, die zeigen, dass ein Wort wie *bone* (Knochen) in *trombone* (Posaune) durchaus ein Zielwort wie *ribs* (Rippen) primen kann.

5.3 Worterkennen im Kontext

Worterkennen ist ein kontextabhängiger Prozess. Man unterscheidet Effekte des lexikalischen Kontextes von Effekten des semantischen und syntaktischen Kontextes. Fragt man einen linguistischen Laien, wie man ein Wort erkennt, so wird er vermutlich antworten, dass man zunächst die Buchstaben identifiziert und dann aus ihnen das Wort zusammensetzt. Weist man ihn darauf hin, dass es neben der geschriebenen Sprache auch das gesprochene Wort gibt, so wird er vielleicht vermuten, dass man hier zunächst die einzelnen Laute (Phoneme) verarbeitet und daraus das Wort zusammensetzt.

Gegen diese intuitiv naheliegende Annahme sprechen jedoch einige empirische Befunde. Bereits James Cattell hat Ende des 19. Jahrhunderts beobachtet, dass kurze Wörter oft schneller benannt werden als einzelne Buchstaben (Cattell 1885). Dieser Effekt ist als Wortüberlegenheitseffekt in die Literatur eingegangen. Er widerlegt die Annahme, dass ein Wort Buchstabe für Buchstabe zusammengesetzt wird, denn dann müsste die Wortbenennungszeit mindestens der Summe der Benennungszeiten aller Buchstaben entsprechen. Gerald Reicher hat den Wortüberlegenheitseffekt systematisch untersucht (Reicher 1969). Er zeigte Probanden Buchstabenreihen, die entweder Wörter (z. B. *HAND*) oder keine Wörter ergaben (z. B. *ANHD*). Ein Buchstabe wurde nach sehr kurzer Zeit durch eine sogenannte Maske ersetzt, nämlich durch ein irrelevantes Zeichen wie das Prozentzeichen,

Kontextarten

Wortüberlegen-
heitseffekt

Buchstaben-
erkennung

sodass die Probanden nun *HAN%* oder *AHN%* sahen. Reicher fragte seine Probanden, welchen Buchstaben sie an der maskierten Stelle wahrgenommen hatten und ließ ihnen die Wahl aus zwei Alternativen. Diese Alternativen waren so gewählt, dass der Wortstatus gleich blieb, im Beispiel etwa D und G. Sowohl *Hand* und *Hang* sind existierende Wörter des Deutschen, *Ahnd* und *Abng* dagegen nicht. Reicher wollte vermeiden, dass die Probanden raten. Hätte er beispielsweise *HAND/HAN%* gezeigt und dann den Probanden die Wahl zwischen D und K gelassen, so wäre es durchaus möglich gewesen, dass ein Proband darum D wählt, weil nur dies ein existierendes Wort ergibt. Dank des klug gewählten experimentellen Aufbaus war diese Strategie nicht möglich. Die Probanden wählten deutlich häufiger den korrekten Buchstaben, wenn sie ihn in einem Wortkontext gesehen hatten als wenn sie ihn in einem Nichtwortkontext gesehen hatten. Buchstaben werden also innerhalb eines Wortes schneller und zuverlässiger erkannt als in einem Nichtwort.

Phonemrestauration

Ein ähnlicher Effekt tritt in der auditiven Modalität auf. Richard Warren spielte Probanden Tonaufnahmen von Wörtern vor (Warren 1970). Zuvor wurde ein Phonem aus dem Signal herausgeschnitten und durch ein Husten oder Rauschen ersetzt. Der lautliche Eindruck der Probanden war jedoch nicht der eines fehlenden Lautes. Stattdessen berichteten sie, dass während der Aufnahme gehustet wurde, dass das Wort jedoch intakt war. Dieser sogenannte Phonemrestaurationseffekt hängt sowohl von der Art des Signals als auch vom semantischen Kontext ab. Die akustische Illusion funktioniert besser für Konsonanten als für Vokale, und innerhalb der Konsonanten am besten für Frikative. Die Rolle des Kontextes hat Warren an einem Beispiel überzeugend demonstriert: Aus einem englischen Wort wurde der Beginn entfernt, sodass nur noch der Silbenreim (z. B. *eel* aus der Silbe *peel*) übrigblieb. Zu Beginn dieses Signales wurde ein Rauschen eingelegt. Der gleiche physikalische Reiz wurde nun in unterschiedlichen Kontexten dargeboten: *the %eel was on the orange*, *the %eel was on the shoe*, *the %eel was on the axle*, *the %eel was on the table*. Zuhörer behaupteten mit großer Sicherheit, im ersten Fall *peel* (Schale) gehört zu haben, im zweiten Fall *heel* (Absatz), im dritten Fall *wheel* (Rad) und im vierten Fall *meal* (Mahlzeit). Unser Sprachverstehenssystem ergänzt also Laute, die dem Kontext angemessen, aber nicht im Signal enthalten sind, und zwar so überzeugend, dass wir uns dieser ‚falschen‘ Wahrnehmung nicht verschließen können.

Ein weiterer Effekt, der zeigt, dass wir nicht erst Laute erkennen und daraus das Wort zusammenbauen, ist der sogenannte Ganong-Ef-

fekt (Ganong 1980). Es ist möglich, Laute zu erzeugen, die genau zwischen zwei Phonemen einer Sprache liegen. Bei zwei Phonemen, die sich nur hinsichtlich des Stimmeinsatzes (VOT) voneinander unterscheiden (z. B. /g/ und /k/) kann man diesen systematisch variieren, so dass man ein Kontinuum von Lauten erhält, deren Klangqualität zwischen /g/ und /k/ liegt. William Ganong führte die üblichen Aufgaben zum Testen kategorialer Wahrnehmung (→ KAPITEL 3.2) durch, ergänzte die Laute aber durch einen Kontext, der mit einem Endpunkt des Kontinuums ein Wort ergab (z. B. *Kissen*) und mit dem anderen Endpunkt ein Nichtwort (z. B. *Gissen*). Ganong beobachtete die charakteristische kategoriale Grenze in der Wahrnehmung, stellte aber fest, dass diese Grenze im Vergleich zur Wahrnehmung ohne Kontext in Richtung des Wortes verschoben war. Das heißt, Hörer nahmen länger ein /k/ wahr, bevor die Wahrnehmung zum /g/ hin umschlug. Man spricht hier von einem lexikalischen Bias – Hörer nehmen einen Laut, der vom Signal her mit zwei Phonemen gleich gut übereinstimmt, so wahr, dass sich im Kontext der anderen Laute ein Wort ergibt.

Ganong-Effekt

Die hier vorgestellten Effekte zeigen, dass Worterkennung nicht nur bottom-up, also signalgesteuert, geschieht, sondern dass unser lexikalisches Wissen auch top-down die Wahrnehmung beeinflusst. Modelle der Worterkennung müssen diese Beobachtungen erklären (→ KAPITEL 5.4). Generell gibt es zwei Möglichkeiten, wie mit dem Einfluss des lexikalischen Kontextes umgegangen werden kann: mit oder ohne Feedback. Fast alle Modelle des Worterkennens haben mehrere Repräsentationsebenen – mindestens eine Ebene für Phoneme und eine Ebene für Wörter. Phoneme haben bahnende Verbindungen zu den Wörtern, in denen sie vorkommen. Sobald ein Phonem erkannt ist, aktiviert es die mit ihm verbundenen Wörter. In Modellen mit Feedback aktivieren Wörter ihrerseits alle Phoneme, die in ihnen vorkommen. Daher kann ein Phonem über die Wortebene aktiviert werden, ohne dass es Aktivierung aus dem Signal erhält.

Feedback

Nicht alle Modelle nehmen Feedback an. Sie erklären die beschriebenen Effekte stattdessen, indem sie neben dem Lexikon ein Entscheidungsmodul ansetzen. Dieses Entscheidungsmodul integriert lexikalische und prälexikalische Information und kann so erklären, wieso die Wortebene Auswirkungen auf die Wahrnehmung von Phonemen haben kann.

Entscheidungsmodul

Bevor der Diskriminationspunkt eines Wortes erreicht ist, ist das sprachliche Signal mehrdeutig. Die niederländische Psycholinguistin Pienie Zwitserlood untersuchte mit dem sogenannten cross-modalen Primingparadigma (→ KAPITEL 2.1), ob der Kontext von vornherein un-

Mehrdeutigkeit

passende Wörter ausschließen kann, obwohl sie mit dem Signal übereinstimmen. Ihre niederländischen Probanden hörten Sätze der Form *Die Männer betrauten den Verlust ihres Kapi-*. An dieser Stelle kann das Wort noch *Kapitän* oder *Kapital* werden. Genau zu diesem Zeitpunkt wurde den Probanden visuell ein Wort dargeboten, für das sie eine lexikalische Entscheidung durchführen mussten. In kritischen Versuchsdurchgängen war das visuelle Wort entweder mit *Kapitän* verwandt (z. B. *Schiff*) oder mit *Kapital* (z. B. *Geld*). In einigen Versuchsdurchgängen wurden Kontexte verwendet, die die Bedeutung einschränkten, z. B. *Die Männer standen um das Grab herum. Sie betrauten den Verlust ihres Kapi-*. Dennoch beobachtete Zwitserlood, dass die lexikalische Entscheidung nicht nur für das visuell dargebotene Wort *Schiff*, sondern auch für das visuell dargebotene Wort *Geld* beschleunigt war, dass der lexikalische Zugriff also nicht durch die Kontextinformation eingeschränkt wurde (Zwitserlood 1989).

Homonymie

Polysemie

Gemeinhin versteht man unter Mehrdeutigkeit, dass ein Wort, auch wenn es vollständig gehört wurde, mehrere mögliche Bedeutungen hat. Man unterscheidet hier Homonyme und Polyseme. Die Bedeutungen eines Homonyms sind nicht miteinander verwandt. Häufig sind die Formen zweier Wörter aufgrund von Sprachwandelprozessen zu einer Form verschmolzen. Ein klassisches Beispiel von Homonymie ist *die Kiefer* vs. *der Kiefer*. Polyseme haben dagegen mehrere Bedeutungen, die sich von einer Grundbedeutung ableiten, wie *Blatt* für das Blatt einer Pflanze oder für ein Blatt Papier. In der Praxis sind Homonymie und Polysemie oft nicht ganz leicht zu trennen. Deshalb werden sie auch hier als eine Gruppe behandelt, obwohl es einige Experimente gibt, die andeuten, dass es durchaus Unterschiede in ihrer Verarbeitung gibt.

David Swinney hat sich in den 1970er- und 1980er-Jahren der Frage gewidmet, ob bei der Verarbeitung eines mehrdeutigen Wortes alle Bedeutungen aktiviert werden (z. B. Swinney 1979). Auch er verwendete dazu cross-modales Priming. Der gesprochene Kontextsatz disambiguierte ein mehrdeutiges Wort (d. h. er machte die Bedeutung eindeutig), z. B. *Er saß auf einer Bank*. Das visuell dargebotene Wort war in der kritischen Bedingung entweder mit der kontextadäquaten Bedeutung des Homonyms verwandt (*Tisch*) oder mit der kontextinadäquaten Bedeutung (*Geld*). Wurde das Wort direkt nach dem Prime dargeboten, so wurden beide Bedeutungen gleichermaßen geprimt. Wartete man jedoch ca. 250 ms mit der Darbietung des Wortes, so wurde nur noch die kontextadäquate Bedeutung geprimt. Swinney schloss daraus, dass beim lexikalischen Zugriff alle Bedeu-

tungen eines Wortes aktiv sind, unabhängig vom Kontext, dass aber der Kontext mit Verlauf von Zeit die unpassende Bedeutung hemmt. Patrizia Tabossi und ihre Kollegen haben in den 1980er- und 1990er-Jahren in mehreren Studien gezeigt, dass der Kontext auch einen direkten Einfluss auf die Aktivierung von Bedeutung haben kann (z. B. Tabossi 1988). Allerdings nur dann, wenn die beiden Bedeutungen des Homonyms nicht gleich häufig waren. In diesem Fall konnte ein Kontext, der sehr stark die dominante Bedeutung stützte, die Aktivierung der untergeordneten Bedeutung völlig unterdrücken.

5.4 Modelle des Worterkennens

Drei einflussreiche Modelle des Worterkennens bei gesprochener Sprache werden kurz vorgestellt (ein Modell zum Erkennen geschriebener Wörter wird in McClelland/Rumelhart 1981 beschrieben). Das Kohortenmodell stammt vom britischen Psycholinguisten William Marslen-Wilson und seinen Kollegen und hat drei große Entwicklungsphasen durchlaufen. In seiner ersten Form aus den 1970er-Jahren beginnt die Suche im Lexikon, sobald ein ausreichend großes Stück Signal (ca. 100–150 ms, d. h. 1–2 Phoneme) verarbeitet wurde. Mit diesem Input wird bottom-up eine sogenannte Wortanfangskohorte aktiviert: Alle Wörter des Lexikons, die mit diesen Phonemen beginnen, werden bereitgestellt. Je weiter sich das Signal in der Zeit entfaltet, desto kleiner wird die Kohorte, da Kandidaten, die mit dem Signal nicht mehr übereinstimmen, durch bottom-up-Inhibition (Hemmung) aus der Kohorte entfernt werden. Ein Wort wird erkannt, wenn sein Diskriminationspunkt erreicht ist. In der ersten Version des Kohortenmodells spielten der semantische und syntaktische Kontext eine wichtige Rolle, da unwahrscheinliche Kandidaten auch mittels top-down-Inhibition ausgeschlossen wurden. So wurde die Beobachtung erklärt, dass in einem günstigen Kontext ein Wort oft bereits vor seinem Diskriminationspunkt erkannt wird. Das Kohortenmodell misst dem Wortanfang ein sehr großes Gewicht bei, da ein lexikalischer Eintrag, der nicht mit dem ersten Phonem übereinstimmt, nicht in die Wortanfangskohorte aufgenommen wird. Dies hat den Nachteil, dass ein Wort, dessen initialer Laut inkorrekt artikuliert wurde, niemals erkannt werden dürfte – was ganz klar nicht der Fall ist. Ein weiterer Nachteil des frühen Kohortenmodells ist, dass der Kontext einen Wortkandidaten ganz ausschalten kann. Dies deckt sich nicht mit der Tatsache, dass wir auch kontextuell unpass-

Kohortenmodell

Wortanfangskohorte

sende Wörter wahrnehmen können. In einer Revision seines Modells in den 1980er-Jahren verzichtet Marslen-Wilson daher auf die top-down-Inhibition. Außerdem schloss er binäre Entscheidungen hinsichtlich der Kohortenzugehörigkeit aus. Stattdessen werden nun lexikalische Einträge mehr oder weniger stark aktiviert, je nachdem wie ähnlich sie dem Signal sind. Diese Ähnlichkeitsberechnung wird ständig aktualisiert und erlaubt auch Wörtern mit einer leichten Abweichung im Onset, zur Kandidatenmenge zu stoßen. Die neueste Version ist das distribuierte Kohortenmodell (z. B. Gaskell/Marslen-Wilson 1997), das sich radikal von seinen Vorgängern unterscheidet, da es keine Trennung zwischen semantischer und phonologischer Repräsentation annimmt, also keinen Unterschied zwischen Selektion und Integration macht. Der Input besteht aus phonetischen Merkmalen, die über eine Ebene versteckter Knoten (*hidden units*) direkt mit semantischen und phonologischen Repräsentationen verbunden sind.

TRACE

TRACE wurde 1986 von den amerikanischen Kognitionswissenschaftlern James McClelland und Jeffrey Elman vorgestellt (McClelland/Elman 1986). Der Name beruht darauf, dass die Modellstruktur gleichzeitig ein Verarbeitungssystem und eine Gedächtnisrepräsentation (*memory trace*) darstellt. TRACE steht in der Tradition des Konnektionismus und besteht aus einer Vielzahl untereinander verbundener Einheiten. Jede dieser Einheiten repräsentiert eine Hypothese über ein bestimmtes perzeptuelles Objekt zu einem bestimmten Zeitpunkt. Solch eine Hypothese kann beispielsweise lauten /k/, d. h., *dem Stück Schallsignal, das ich gerade wahrgenommen habe, entspricht ein /k/*. Die Einheiten sind in drei Ebenen organisiert, einer Merkmalebene, einer Phonemebene und einer Wortebene. Der Input in TRACE ist eine quasispektrale Repräsentation, die aus akustisch-phonetischen Merkmalen wie Burst oder Stimmeteiligung, besteht. Für jedes dieser Merkmale ist kodiert, wie prominent es im Input ist. Die Merkmalsknoten senden Aktivierung zu den Phonemen, mit denen sie über bahnende Verbindungen verknüpft sind. Teilweise ist die Phonemidentität erst festzustellen, wenn man beobachtet, wie sich die Merkmale über die Zeit hinweg entwickeln. TRACE berücksichtigt daher diese Entwicklung in seinen Berechnungen. Auch Phoneme und Wörter sind bahnend miteinander verbunden. Das Phonem /k/ hat beispielsweise Verbindungen zu allen Einheiten auf der Wortebene, in denen ein /k/ vorkommt, wie *Katze*, *Akte*, *Macke* und *Lack*. Wird /k/ aktiviert, so leitet es einen Bruchteil seiner Aktivierung über die bahnenden Verbindungen weiter zu den Wörtern, in denen es vorkommt (zu Aktivationsausbreitungsmodellen → KAPITEL 4.1). Innerhalb einer Ebene sind die Einheiten durch hemmen-

de Verbindungen miteinander verknüpft. Dies reflektiert die Tatsache, dass ein Teil des Signales, wenn er als /k/ identifiziert wurde, kein /t/ oder /f/ sein kann. Je stärker eine Repräsentation aktiviert ist, desto mehr schwächt sie alle Konkurrenten auf derselben Ebene ab. Aktivierung kann in beide Richtungen fließen. Haben zwei oder drei Phoneme mehrere Kandidaten auf der Wortebene aktiviert, so verstärken diese Wörter durch Feedback die Aktivierung aller Phoneme, die in ihnen vorkommen. Durch dieses top-down Feedback können sogar Phoneme aktiviert werden, die im Signal (noch) nicht vorgekommen sind.

SHORTLIST stammt aus der Feder von Dennis Norris (Norris 1994). Es ist in vielen Punkten TRACE sehr ähnlich, hat aber auch Aspekte des Kohortenmodells übernommen. Wortkandidaten werden durch bottom-up-Aktivierung ausgewählt. Diese Kandidaten werden in eine sogenannte SHORTLIST kopiert, innerhalb derer die Wörter über inhibierende Verbindungen verknüpft sind, sodass ein stark aktivierter Kandidat die anderen hemmen kann. Gleichzeitig wird immer wieder kontrolliert, wie gut die Wörter in der SHORTLIST mit dem sich stets weiter entfaltenden Signal zusammenpassen. Wörter, die nicht mehr mit dem Signal übereinstimmen, werden aus der Liste entfernt, es können aber im Gegensatz zu TRACE auch neue Kandidaten hinzugefügt werden, die z. B. aufgrund einer Abweichung in der Artikulation des Wortbeginns ursprünglich nicht berücksichtigt wurden, im weiteren Verlauf aber eine gute Übereinstimmung mit dem Signal zeigen. Ein weiterer Unterschied zu TRACE ist, dass SHORTLIST kein Feedback annimmt. Die neueste Version von SHORTLIST (Norris/McQueen 2008) könnte sich als wegweisend herausstellen. Sie bricht mit der Tradition von konnektionistischen Aktivationsausbreitungsmodellen und beruht stattdessen auf einem Theorem der Wahrscheinlichkeitstheorie, nämlich dem Bayes-Theorem. Dieses Theorem dient der Berechnung bedingter Wahrscheinlichkeiten: wie wahrscheinlich ist eine Hypothese, gegeben die Evidenz. In der Worterkennung wäre die Hypothese ein bestimmtes Wort, und die Evidenz das Sprachsignal. In der Formel zur Berechnung der bedingten Wahrscheinlichkeit ist auch die einfache Wahrscheinlichkeit der Hypothese enthalten, also die Wahrscheinlichkeit, dass die Hypothese wahr ist (bevor man die Evidenz kennt). Auf das Worterkennen übertragen kann man unter der einfachen Wahrscheinlichkeit, dass ein Wort auftritt, seine Frequenz und Kontextinformation subsumieren. Das heißt, dass hier auf rechnerische Weise top-down und bottom-up-Information elegant miteinander verbunden werden. Das Worterkennensmodell führt also ständig Wahrscheinlichkeitsberech-

SHORTLIST

Bayes-Theorem

nungen durch, an deren Ende das wahrscheinlichste Wort, gegeben Kontext und Signal, steht. In vielen Bereichen der Kognitionswissenschaft haben sich Bayes'sche Modelle in den letzten Jahren durchgesetzt, es bleibt abzuwarten, ob dies auch in der Wortverstehensforschung geschehen wird.

Neben allen Unterschieden haben die drei Modelle auch wichtige Gemeinsamkeiten: Worterkennen beginnt, sobald ein kleines Stück des Signals wahrgenommen wurde, und das Zielwort muss aus einer Menge möglicher Kandidaten ausgewählt werden.

Fragen und Anregungen

- Wieso nehmen viele Modelle des Worterkennens explizit oder implizit eine prälexikalische Repräsentationsebene an?
- Erörtern Sie die Vor- und Nachteile der Einheiten, die in der Literatur als prälexikalische Repräsentationen vorgeschlagen worden sind.
- Unser Sprachsystem verwendet sowohl sublexikalische Hinweise wie Metrik und phonotaktische Regelmäßigkeiten als auch lexikalische Hinweise zur Worterkennung. Stellen Sie sich ein Gespräch vor starkem Hintergrundgeräusch vor und überlegen Sie, ob hier die sublexikalische oder die lexikalische Information verlässlicher ist.
- Welche Unterschiede und Gemeinsamkeiten gibt es zwischen den Modellen des Worterkennens?

Lektüreempfehlungen

- **Angela Friederici (Hg.): Sprachrezeption**, Göttingen 1999. *Dieses Buch beschäftigt sich mit verschiedenen Fragestellungen der Sprachrezeption aus psycholinguistischer Sicht. Besonders relevant für die hier besprochenen Fragen sind die Kapitel 1–3.*
- **Delphine Dahan / James Magnuson: Spoken Word Recognition**, in: Gareth Gaskell (Hg.), *Oxford Handbook of Psycholinguistics*, Oxford 2007. *Übersichtsartikel, in dem lexikalischer Zugriff, Auswahl und Integration ausführlich behandelt werden. Außerdem wird ein Überblick über neu entstehende Forschungsfragen gegeben.*