

Diplomarbeit :

MEG-Korrelate der Wortverarbeitung und
ihre Automatische Erkennung

Erstbetreuer: Prof. Dr. G. Dogil,
IMS, Universität Stuttgart.

Zweitbetreuer: PD Dr. Dr. F. Pulvermüller,
Soz. wiss. Fakultät, Universität Konstanz.

Diplomand: Ramin Assadollahi,
IMS, Universität Stuttgart.

Abgabe: 10. Juni 1999

1. Die Einleitung	5
2. Die Theorien.	7
2.1. Modultheorie	7
2.2. Kortikal motivierte Netzwerkmodelle	13
3. Die Hypothesen.....	17
3.1. Modultheorie	17
3.2. Kortikal motivierte Netzwerkmodelle	18
4. Die Stimuli.....	21
4.1. Stimulus-Wörter.....	22
4.2. Fragebogen.....	25
4.2.1. Gestellte Fragen.....	26
4.2.2. Bewertungen.....	27
4.2.3. Statistik.....	28
5. Die Aufnahme.....	30
5.1. Versuchsperson.....	30
5.2. Stimuli.....	30
5.3. Aufgabe.....	30
5.4. Versuchsablauf.....	31
5.5. Daten Akquisition.....	32
6. Das Screening	33
7. Die Datenverarbeitung.	36
7.1. Offline-Reduktion des Rauschens.....	36
7.2. Epochengenerierung.....	36
7.3. Mittelungen.....	37

7.4. Baseline-Korrektur.....	38
7.5. Weitere Artefakt-Behandlung.....	38
7.6. Festlegung der Zeitscheiben.....	40
7.7. Suche nach den Foci.....	42
7.8. Berechnung der Foci-Daten.....	43
8. Die Statistik.....	43
8.1. Statistik für Silben-/Frequenzpaare.....	45
8.1.1. Global Field Power	45
8.1.2. Foci > 100fT	46
8.2. Statistik für Wortklassenvergleiche	48
8.2.1. Global Field Power	48
8.2.2. Foci > 100fT	49
8.3. Weitere Ergebnisse	52
9. Die Diskussion.	54
10. Das Netz.	63
10.1. Einleitung.....	63
10.2. Daten.	67
10.3. Training.....	68
10.4. Generierung der Karte.....	69
10.5. Der Erkennungsvorgang.	70
10.6. Erstes Ergebnis.....	71
10.7. Diskussion.	72
10.8. Weitere Architekturen.....	72
10.9. Endergebnisse.	74
10.10. Diskussion.	75

11. Die Literatur.....	77
12. Das Abbildungsverzeichnis.	79

1. Die Einleitung.

"Fast jeden Satz, den wir sagen, sagen wir zum ersten Mal.", schrieb einmal Hans Kamp, Professor für formale Logik und Sprachphilosophie an der Universität Stuttgart.

Die Frage nach der Entstehung einer solchen, unglaublichen Vielfalt unseres Geistes hat schon immer die Menschheit beschäftigt. Viele Theorien und Mythen über den menschlichen Geist haben wir kommen und gehen sehen und immer wieder andere unternehmen den Anlauf, eines der letzten großen Rätsel der Menschheit zu lösen.

Doch erst in den letzten 30 Jahren haben Forscher einige Instrumente in die Hand bekommen, um auf die Aufgabe unter physiologisch meßbaren Gesichtspunkten anzugehen. Damit konnte man erstmals den Geist über seine Sinne mit verschiedenen Signalen konfrontieren und messen, welche physiologischen Antworten sein Gehirn über diese Meßgeräte gibt. Man stellte dabei fest, daß sich die Antworten abhängig von der Stimulationsart unterschieden und daß viele Hirnantworten sich über mehrere Stimulationen hinweg gleichartig wiederholen. Von der großen Struktur dieser Antworten kann man inzwischen sagen, daß sie sich sogar interindividuell sehr ähnlich sind.

Verschiedene Theorien werden heute in der Forschung diskutiert, die diese Hirnantworten erklären sollen. Dabei machen einige Theorien Voraussagen über bestimmte neuronale Architekturen und andere wollen aus der kognitiven Architektur heraus argumentieren.

Fakt ist dennoch, daß bestimmte Stimulationen bestimmte Hirnantworten verursachen: Die Forschung zur Wortverarbeitung im menschlichen Gehirn, versucht Faktoren zu bestimmen, welche linguistische Eigenschaften der Wörter sind und die Hirnantworten auf diese Wörter modulieren. Beispielsweise stehen hier Silbenanzahl oder auch die Häufigkeit des Wortes im Alltag zur Diskussion.

Da Wörter aufgrund ihrer physikalischen Eigenschaften (Buchstabenanzahl, Silbenlänge) sehr kurz sind, ist es von Nöten, Meßgeräte zur Untersuchung des Verarbeitungsprozesses heranzuziehen, die eine hohe zeitliche Auflösung haben. Die größte Rolle spielen dabei das EEG (Elektroenzephalogram) und das MEG (Magnetoenzephalogram).

Wenn sich also Hirnantworten auf verschiedene Wörter unterscheiden, ähnlich wie sie sich akustisch unterscheiden, dann sollten sie auch verschiedene Signale in diesen Meßgeräten erzeugen.

Mit den Kenntnissen aus der automatischen Spracherkennung, der algorithmisierten Klassifikation von Signalstücken also, könnte sich folglich ein Erkenner für MEG-Signale konstruieren lassen.

Dazu wurde in der vorliegenden Arbeit eine Versuchsperson einem sogenannten memory-task unterzogen: Sie sollte sich eine Liste von Wörtern merken, welche dann im Experiment visuell dargeboten wurden. Bei bestimmten Wörtern, die seltener erschienen als andere sollte sie eine Taste drücken. Die Hirnantworten auf die häufig präsentierten Wörter wurden dann genauer analysiert. Da die Wörter bestimmte, in einer Voruntersuchung bestätigte Eigenschaften hatten (Silbenzahl, Frequenz, Wortklasse, Modalität, etc.) konnten die Ableitungen auf Unterschiede in Abhängigkeit dieser Eigenschaften untersucht werden.

Da diese Unterschiede tatsächlich vorhanden waren und statistisch untermauert werden konnten, bestand Anlaß zur Hoffnung, ein künstliches Neuronales Netz könne die einzelnen Ableitungen klassifizieren. Diese Hoffnung bestätigte sich, denn das Netz konnte überzufällig Kategorisierungen in Wortklasse und sogar in einzelne Wörter vornehmen.

2. Die Theorien.

Was ist die Architektur der Verarbeitung von menschlicher Sprache und wie sieht ihr neurobiologisches Substrat aus? Zwei Architekturen stellen die beiden Extreme dar: Der Modularismus und der Konnektionismus.

Der Modularismus geht das Problem der geistigen Verarbeitungsmechanismen top-down an und versucht aus den Erfordernissen, die Probleme und Lösungen stellen, den Aufbau zu beschreiben. Dabei werden immer speziellere Bearbeitungs-Module formuliert, die immer speziellere Aufgaben lösen können.

Der Konnektionismus geht dagegen bottom-up vor und beschreibt aufgrund neurobiologischer Beobachtungen wie Synapsenaufbau, Zellverschaltung, Reaktionszeiten etc., Verarbeitungsmechanismen, die zunächst sehr grundlegend sind, aber durch größere Architekturen auch größere Phänomene modellieren. Nimmt man noch den morphologischen Aufbau des Gehirns zu den Motivationen der Modellierung hinzu, so kommt man zu kortikal motivierten Netzwerkmodellen.

2.1. Modultheorie

In "The Modularity Of Mind" stellt Fodor 1983 ein Konzept vor, welches die Verarbeitungsstufen bei geistigen Aufgaben modellieren könnte. Wenn man eine Aufgabe in der Tradition der klassischen künstlichen Intelligenz in Probleme und Lösungen, deren Teilprobleme und Teillösungen aufspaltet, kann man einen seriellen Prozeß formulieren, der aus dem Lösen kleiner, spezieller Aufgaben besteht. Zur Lösung dieser Aufgaben, die vielleicht sogar in manchen Fällen aus dem einfachen Übersetzen eines Codes in einen anderen bestehen können, schlägt Fodor sogenannte Module vor. Diese Spezialisten für bestimmte Aufgaben können durch ihre serielle Anordnung eine größere Aufgabe lösen.

Die Anordnung geschieht in Kanälen, die Daten zwischen den Modulen hin- und herschicken. Ein Modul kann mit mehreren anderen Modulen verbunden sein, aber dennoch ist zu einem Zeitpunkt der Verarbeitung eines bestimmten Datums nur ein Modul aktiv.

Caplan beschreibt 1992 Eigenschaften eines Sprachsystems, das auf diesen Ideen beruht.

Im Bezug auf Sprache stellt er dabei vier Punkte als besonders bedeutend heraus:

1. Das Sprachsystem ist keine monolithische Struktur, sondern besteht aus einzelnen, spezialisierten Modulen; jedes Modul benötigt eine bestimmte Art von Eingangsdaten und gibt eine bestimmte Art von Daten aus. Der Syntaxprozessor würde daher ausschließlich Wortrepräsentationen als Input nehmen, aber keine akustischen. (Fodor 1983: informationally encapsulated)
2. Sprachprozesse laufen unweigerlich und automatisch ab: Das Lesen eines Wortes und der Zugriff auf seine Bedeutung gehen automatisch von statten, selbst wenn es der Aufgabe nicht zuträglich ist. (z.B. Stroop-Task: lesen des Worts "rot", das in blauer Farbe geschrieben ist.) Fodor 1983: Parsing ist ein Reflex.
3. Sprachprozesse laufen unbewußt ab, die Repräsentationen auf den unteren Ebenen sind dem Bewußtsein nicht zugänglich.
4. Sprachprozesse laufen schnell und mit hoher Genauigkeit ab.

Zu einem Minimalsystem gehört ein semantisch-konzeptuelles System, welches mit dem perzeptuellen System interagiert und den Zugriff wie auch die Organisation von Weltwissen ermöglicht.

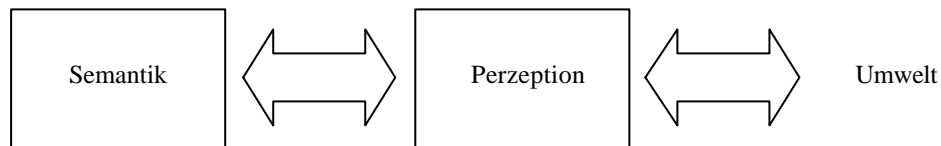


ABBILDUNG. EIN KLEINES, MODULARES SPRACHSYSTEM.

Wortbedeutungen im Semantikmodul könnten durch Features realisiert sein, eine Klassenzugehörigkeit wäre durch den Vergleich mit dem Klassenprototypen zu ermitteln. Es gibt verschiedene Theorien darüber, ob es einen einzigen, amodalen Speicher für die semantischen Daten gibt, oder ob mehrere modalitätsspezifische Speicher existieren.

Modalitäten spielen auch auf der Ebene der Wortverarbeitung eine Rolle. Um die perzeptuellen Daten in Semantik umzusetzen, wird ein Lexikon eingesetzt. Ein solches Lexikon könnte eine Art Übersetzer sein, oder auch eine Tabelle, in der nachgesehen wird, zu welchen z.B. phonetischen Codes ein semantischer Code gehört. Monsell (1987) betrachtet das Lexikon als universellen Codeübersetzer, der zwischen Lese-, Schreib-, Hör-, Sprech- und Konzeptcodes hin- und herübersetzt.

Einen Vorschlag, wie die Übersetzung bzw. Erkennung eines Wortes aus den geometrischen Bestandteilen oder auch Bestandteilen aus anderen Modalitäten geschehen könnte, lieferte Morton im Jahre 1969.

Nach Morton (1969, 1970) sammelt ein sogenanntes Logogen Evidenzen (Features) für die Erkennung eines bestimmten Wortes. Sind genügend Evidenzen vorhanden, dann wird das Wort als erkannt ausgegeben. Dies geschieht über einen Schwellwertmechanismus - liegt die Anzahl aller beobachteten, für dieses Wort charakteristischen Evidenzen über einem gewissen Wert, der Schwelle, so gilt das Wort als erkannt. Um Frequenzeffekten Rechnung zu tragen, wird diese Schwelle bei häufiger Benutzung des Logogens gesenkt. Ein lexikalischer Zugriff erfolgt direkt und parallel über alle Wörter, denn alle Logogene

erhalten gleichzeitig alle Eingangsdaten, doch nur dasjenige, welches diese Daten als wichtig für die Erkennung des Wortes ansieht, wird eine Ausgabe tätigen. In der ursprünglichen Version ist das Logogen unabhängig von der Modalität, es können also visuelle und auditorische Daten zusammen für die Erkennung eines Wortes sorgen. Das Logogen ist jedoch nicht verantwortlich für die Verarbeitung semantischer Features, so daß die Aufteilung des obigen Systems gewahrt bleibt. Dennoch sind die Verbindungen zwischen einem Logogen und der Semantik bidirektional, um die Information aus dem semantischen Kontext für eine eventuelle Disambiguierung mit in Betracht ziehen zu können.

Doch wie ist das Lexikon aufgebaut? Es stellt sich die Frage, ob das Hören eines Wortes zu einem Zugriff auf das auditive Lexikon führt, oder ob es lediglich ein einziges modalitäts-unspezifisches Lexikon gibt, das das Gehörte bearbeitet. Wenn man Modalitätsabhängigkeit postuliert, kann man die konstituierenden Module verschieden aufteilen. Für die vier Verarbeitungsaspekte Lesen, Hören, Schreiben, Sprechen könnte es somit vier spezielle Lexika geben, die jeweils eine dieser Modalitäten ausschließlich bedienen. Man könnte jedoch auch nur zwei Lexika fordern, es würde sich dann die Frage nach der Aufteilung ergeben. Eine Möglichkeit wäre, die Lexika nach gesprochen und schriftlich aufzuteilen (gesprochen: Hören, Sprechen; geschrieben: Lesen, Schreiben), die andere wäre, sie nach Ein- und Ausgabe zu kategorisieren (Eingabe: Lesen, Hören; Ausgabe: Schreiben, Sprechen).

Viele Arbeiten argumentieren für und gegen die eine oder andere Sicht. Zum Beispiel gibt es Evidenzen, daß Spracheingabe- und Sprachausgabepfade verschieden sind. Dies kann man schließen aufgrund von PET-Untersuchungen (Petersen, Fox, Posner, Mintun & Raichle 1989), die ergaben, daß das Hören von Wörtern andere Kortextareale aktiviert als das Vorlesen oder das Wiederholen von Wörtern. Wenn Worterkennung und

Wortproduktion dennoch interferieren, dann geschieht dies auf sub-lexikalischer Verarbeitungsebene, weil beispielsweise der gleiche phonetische Kurzzeitspeicher (Puffer) benutzt wird.

Die Definition des Logogen-Modells von Morton macht die Vorhersage, daß zusätzliche Information aus einer anderen Modalität die Erkennung eines Wortes erleichtert, da mehr Evidenzen für das Erreichen des Schwellwerts bewertet werden können. Dieses Modell würde für ein einziges Lexikon sprechen. Da man aber in Priming- und Bildbenennungsaufgaben (Winnik & Daniel, 1970) gezeigt hat, daß der modalitätsübergreifende Primingeffekt ist kleiner als der Primingeffekt innerhalb einer Modalität, sah sich Morton 1979 gezwungen, zwei getrennte Logogene (ein visuelles und ein auditives) zu fordern. Weitere Forschung führte zu einer weiteren Unterteilung in Ein- und Ausgabe-Logogene, womit die anfangs vorgeschlagenen vier Lexika formuliert wären. Diese Theorie wird durch neuropsychologische Daten von hirngeschädigten Patienten gestützt. Da es z.B. regionale Hirnschädigungen gibt, die den Ausfall des Lesens nicht aber der anderen drei Modalitäten verursachen, wird angenommen, daß sich dort das visuelle Inputlexikon befindet. Als generelles Vorgehen in der Modultheorie kann bezeichnet werden, daß wenn eine Läsion einer bestimmten Hirnregion einen funktionellen Ausfall nach sich zieht, dieser Hirnregion der Sitz des Moduls für die Funktion zugeschrieben wird.

Im Falle von vier speziellen Lexikonmodulen muß auch die Perzeption in ihre Teile aufgespaltet werden, die dann die spezifischen Eingabedaten für die modalitätsspezifischen Lexika liefern. Eine Verfeinerung dieser Architektur schlugen Ellis und Young 1988 vor:

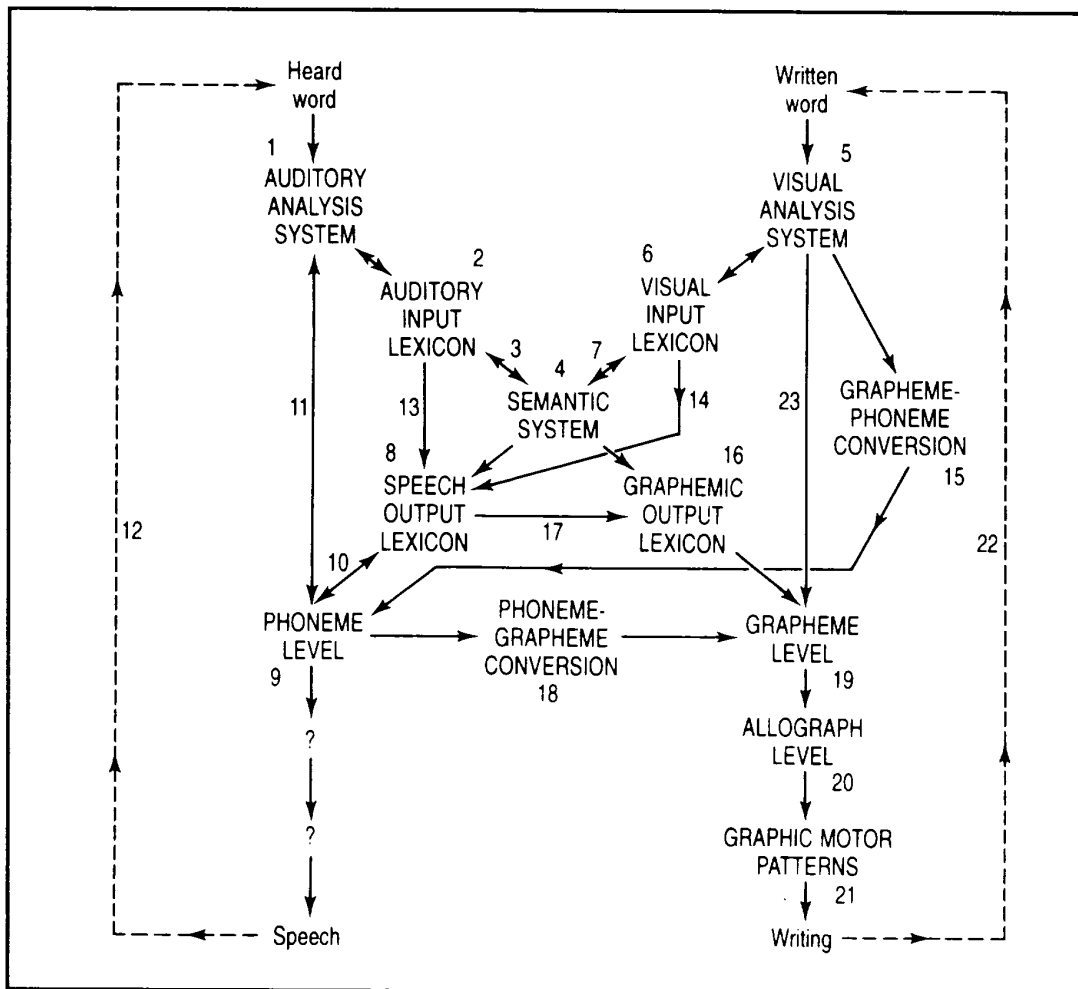


ABBILDUNG: DAS ELLIS & YOUNG MODELL VON 1988.

Bei Verarbeitung von geschriebenen Wortmaterial wird im visuellen Analysesystem zunächst die Zeichenkette in ihre geometrischen Bestandteile zerlegt, es folgt ein Zugriff auf das visuelle Lexikon und dann kann die semantische Verarbeitung beginnen. Dies erfordert also die serielle Aktivierung von drei verschiedenen Modulen, die stufenweise Daten für das jeweils folgende Modul liefern.

2.2. Kortikal motivierte Netzwerkmodelle

Der Ansatz der Konnektionisten sieht nicht das Modul als kleinste verarbeitende Einheit vor, sondern das Neuron. Ein Neuron empfängt Daten durch seine Dendriten, verarbeitet diese eingehende Information und gibt gegebenenfalls über das Axon Daten aus. Formaler beschrieben, werden Eingangsdaten in Zahlen gewichtet (mit einer Zahl, dem Gewicht, multipliziert), aufsummiert und mit einem Schwellenwert verglichen. Liegt die Summe über dem Schwellenwert, so wird eine Zahl ausgegeben:

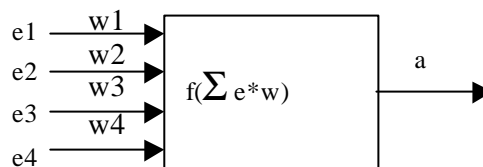


ABBILDUNG. EIN FORMALES NEURON (E EINGANGSWERT, W GEWICHT, A AUSGANGSWERT, F AUSGABEFUNKTION).

Dies erinnert stark an das Logogen-Modell von Morton. Man könnte Evidenzen als Zahlen codieren: Evidenz vorhanden bedeutet 1, Evidenz nicht vorhanden bedeutet 0. Setzt man diejenigen Gewichte, die wichtig für die Erkennung des Wortes sind auf den Wert 1 und alle anderen auf den Wert 0, so kann man eine Ausgabefunktion f definieren:

Falls, die Summe der Eingangswerte $e >$ Schwellwert s , dann ist $a = 1$ sonst 0.

Somit wäre ein formales Logogen beschrieben, wenn man für $a=1$ definiert, daß das Wort als erkannt ausgegeben wird und sonst nicht. Die Funktionsweise neuronaler Netze ist also auf dieser Ebene sehr ähnlich.

Um solche Neuronen lernen zu lassen, muß noch eine Lernregel definiert werden, die dem Neuron sagt, was es im Falle einer Missklassifikation tun soll, um bei erneuter Präsentation die richtige Antwort zu liefern. Ein solche Lernregel hat Donald Hebb 1949 definiert:

**Die synaptische Eigenschaft (verstärken oder hemmen) steigt
bei gleichzeitiger prä- und postsynaptischer Aktivität.**

(einfachste Form: $\Delta w = k \cdot e \cdot a$, k Konstante)

Diese Lernregel kommt heute noch in abgewandelter Form (Widrow&Hoff oder Delta-Regel) in modernen künstlichen neuronalen Netzen vor. Die Abwandlung besteht darin, daß die Veränderung der synaptischen Eigenschaft (i.e. des Gewichts) ein Ende hat, wenn die berechnete Ausgabe der vorgelegten entspricht.

Die Hebb'sche Regel läßt sich paraphrasieren: die Verbindung zweier Neuronen verstärkt sich, wenn sie gleichzeitig aktiv sind und damit steigt die Wahrscheinlichkeit, daß wenn eines der beiden aktiviert ist, auch das andere ko-aktiviert wird. Dies gilt nicht nur für zwei sondern auch für mehr Neuronen, so daß diese Eigenschaft zu sogenannten Cell Assemblies führt. Cell Assemblies sind Verbände von Neuronen, die untereinander stark verschaltet sind. Wenn genügend Neuronen einer Cell Assembly aktiviert sind, dann "zündet" sie und alle anderen werden auch aktiv. Damit kann fehlende Information ergänzt werden und eine Erkennung stattfinden. Cell Assemblies sind in ihrer ursprünglichen Definition auf einen Ort beschränkt, heute gibt es jedoch Anlaß, über sogenannte transkortikale Cell Assemblies nachzudenken (Pulvermüller&Mohr 1996), die sich über den ganzen Kortex erstrecken und so verschiedene Modalitäten umfassen können.

Nach der Hebb'schen Regel sollten also diejenigen Neuronen des Kortex mit einander stark verschaltet sein, die gleichzeitig aktiv sind. Wird beispielsweise ein Wort gehört, während gleichzeitig ein bestimmtes Objekt gesehen wird und dies häufig und regelmäßig, so werden Neuronen der Kortexareale, die bei der phonetischen Repräsentation involviert sind und solche der Areale, die für die visuelle Repräsentation des Objekts verantwortlich

sind, stärker mit einander verschaltet. Dies hat zur Folge, daß beim Hören des Wortes in vielen Fällen auch der parieto-okzipitale Kortex ko-aktiviert wird und umgekehrt. Für den umgekehrten Weg genügt es dabei, wenn das Objekt verdeckt ist oder in leicht abgewandelter Form präsentiert wird - durch das Zünden der jeweiligen Cell Assembly wird die Information automatisch ergänzt.

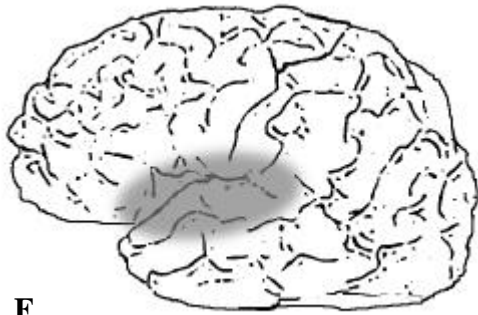
Selbst für die phonetische Repräsentation ist dieses Konzept anzuwenden. Das Aufbauen einer internen Repräsentation von Lauten und Silben geschieht hierbei durch die Koaktivierung von motorischen Kortexteilen (Zunge, Kehlkopf etc.) mit den auditorischen Kortizes, die aktiviert werden, wenn der selbst produzierte Laut gehört wird. Eine solche phonetische Repräsentation umfaßt also Sensorik wie auch Motorik im perisylvischen Bereich, die Produktion und die Perzeption sind in einer Repräsentation verschmolzen. Mit diesem Ansatz können Effekte von Aphasien erklärt werden, die für andere Theorien schwer erklärbar sind. Eine Broca-Aphasie, die die Produktion von Sprache beeinträchtigt, hat in den meisten Fällen auch bestimmte Symptome einer Wernicke-Aphasie, die die Perzeption von Sprache verschlechtert. Mit einer starken Verschaltung der beiden Areale durch Cell Assemblies, die sich über beide Areale erstrecken, kann dieser Effekt erklärt werden. Denn wird die Broca-Region verletzt, so wird ein Teil einer solchen Assembly zerstört und damit ist es für sie in beiden Modalitäten schwerer, den Schwellenwert für eine Zündung zu erreichen (Pulvermüller & Preissl, 1994).

Auf höherer Ebene macht die Theorie der Cell Assemblies eine fundamental andere Definition des Begriffs Semantik möglich. Die Semantik eines Wortes liegt in der Koaktivierung von Arealen, die mit der physiologischen Umsetzung oder Funktion des Wortes zu tun haben. Sie ist also für jedes Wort direkt und individuell lokalisierbar kodiert. Wörter einer Wortgruppe aktivieren dann ähnliche Areale, wohingegen Wörter anderer

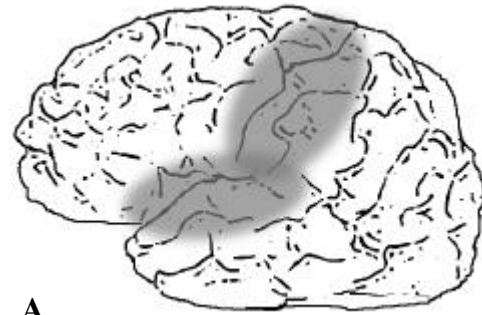
Gruppen andere Areale aktivieren. Evidenzen dafür wurden mehrfach erbracht (z.B. Pulvermüller, Lutzenberger & Birbaumer 1995).

Funktionswörter, die ausschließlich syntaktische Funktion innehaben, die ja nicht auf den Körper bezogen kann, werden im perisylvischen Bereich repräsentiert. Dieser Bereich ist, wie angedeutet, für die phonetische Repräsentation Wörtern verantwortlich.

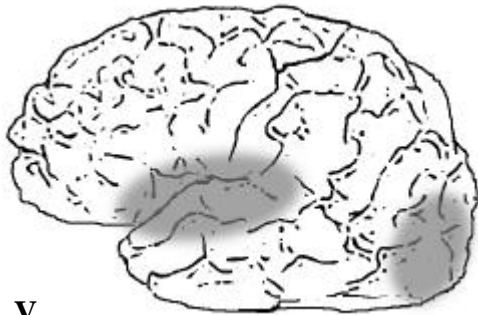
Die Bedeutung eines Wortes, das mit einer Körperbewegung verbunden wird, läßt sich beschreiben als die gemeinsame Aktivierung vom Motorkortex und der perisylvischen Region, da das Wort beim Erlernen häufig mit einer Bewegung gehört oder gesprochen wurde. Entsprechend hat ein Nomen, das mit visuellen Eigenschaften versehen ist eine Koaktivierung des perisylvischen Bereich mit Teilen des Okzipitallappens. Da solche Koaktivierungen auch über mehrere Modalitäten hinweg geschehen können, ruft die Aktivierung eines Wortes mit sowohl visueller als auch motorischer Assoziation sowohl Aktivierung in Teilen des visuellen Systems wie auch in Teilen des motorischen Systems hervor:



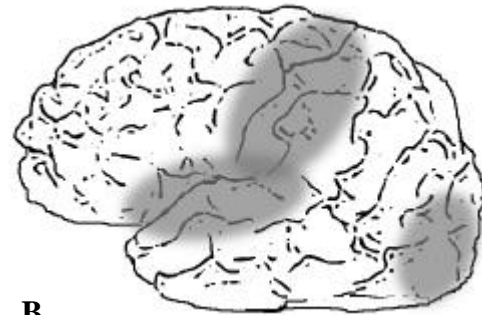
F



A



V



B

ABBILDUNG. AKTIVATIONEN FÜR VERSCHIEDENE WORTKLASSEN (F FUNKTIONSWÖRTER, A VERBEN MIT MOTORISCHER ASSOZIATION, V NOMEN MIT VISUELLER ASSOZIATION, B NOMEN MIT VISUELLER UND MOTORISCHER ASSOZIATION)

Im Gegensatz dazu besagt die modularistische Definition von Semantik, daß ein Code, der zunächst aus physiologischen Daten gewonnen wird, übersetzt wird in einen Code, der die Bedeutung des Wortes in sich trägt. Die Semantik eines Wortes ist dabei nicht lokalisierbar, wohl aber das Modul, welches für die Prozessierung von Semantik verantwortlich ist.

3. Die Hypothesen.

3.1. Modultheorie

Hypothesen, die im Rahmen einer modularistischen Theorie formuliert werden können:

Da die Semantik eines Wortes in dieser Sicht unabhängig ist vom Ort der Repräsentation, sollten sich Wortklassen- und Wortunterschiede nicht in der kortikalen Distribution unterscheiden. Bei allen Wörtern wird das Semantikmodul (Ellis&Young 88) in Gang gesetzt.

Durch das serielle Prozessieren der Information von physiologischen Daten zu Semantik sollten physiologische Entsprechungen von Prozessen im Lexikon geschehen bevor physiologische Entsprechungen von Prozessen in der Wort-Semantik zu beobachten sind.

Wenn die Hirnantworten zwischen den Wörtern unterscheiden, dann sollten die Unterschiede zwischen Wörtern verschiedener lexikalischer Kategorien früher erscheinen als bei die Unterschiede auf Wörter, die der selben lexikalischen Kategorie angehören, sich aber stark in der Bedeutung unterscheiden.

Vermutlich könnte das visuelle Analysesystem bei zweisilbigen Wörtern länger aktiv sein als bei einsilbigen.

Wenn man einen speziellen Suchalgorithmus postuliert, der das Lexikon nach den Häufigkeiten der Wortbedeutungen ordnet (Hogaboam & Perfetti 1975), so könnten Wörter, die häufig benutzt werden, schneller gefunden werden als selten benutzte. Dies hätte zur Folge, daß dieses Modul bei hochfrequenten Wörtern kürzer aktiv ist als bei niederfrequenten.

3.2. Kortikal motivierte Netzwerkmodelle

Wie im vorigen Kapitel erläutert, müßte ein Verb mit motorischen Assoziationen aufgrund der Hebb'schen Lernregel einige Zeit nach der Präsentation Aktivationen im linken, temporalen Kortex (Wernicke, Broca, phonetische Repräsentation) aber auch links und rechts Zentral (zentrale Furche, motorischer Kortex) erzeugen. Ein Nomen mit visuellen

Assoziationen sollte hingegen keine Aktivierung im motorischen Kortex hervorrufen, wohl aber im visuellen. Es wäre daher zusätzlich zur links perisylvischen Aktivierung eine zentro-okzipitale Aktivierung zu erwarten. Ein Nomen, dessen Bedeutung beide Modalitäten umfaßt, sollte folgerichtig alle drei Areale aktivieren, wohingegen ein Funktionswort, das weder motorische noch visuelle Assoziationen hervorruft, ausschließlich links perisylvisch aktivieren sollte.

Mit diesen Hypothesen lassen sich also Vorhersagen über den Ort oder die Orte der Repräsentation eines Wortes treffen. Wie steht es aber mit der Stärke der Aktivität?

Wenn ein starkes externes Signal zu einer starken Aktivierung führt, so fließen viele Ionen, die dann wiederum eine höhere, meßbare Feldstärke erzeugen - soweit ein biokybernetischer Ansatz. Neurobiologisch betrachtet, sollte die Aktivierung durch zweisilbige Wörter wegen der höheren Anzahl der Kontrastpunkte, die die Ränder der Buchstaben ausmachen, höher sein als die Aktivierung, die durch ein einsilbiges Wort generiert wird. Dieser physikalische Unterschied sollte also zumindest bei niederen Arealen, die zuerst vom Stimulus aktiviert werden, sehr frühzeitig nach Stimulus-onset meßbar sein.

Auch eine Vorhersage über den Einfluß der Worthäufigkeiten läßt sich treffen. Dazu sind zwei Ansätze möglich: Ein informationstheoretischer und ein neurodynamischer. Die Informationstheorie (Shannon 1948) besagt, daß ein Reiz, der selten auftritt mehr Unsicherheit beseitigt als ein häufiger Reiz. Je mehr Unsicherheit beseitigt wird desto höher ist der Informationsgehalt einer Nachricht (des Reizes).

Das Konzept des Informationsgehalts läßt sich für die Unterscheidung von durch Wörter evozierten Potentialen ausnutzen. Wird nämlich ein Wort sehr häufig präsentiert, so wird seine Repräsentation weiter ausgebaut. Dies führt bei weiteren Präsentationen zu einer Verringerung der Beseitigung von Unsicherheit und damit zu einer Verringerung des

Informationsgehalts. Angenommen, die Aktivierung im Kortex ist proportional zum Informationsgehalt, so würde dies bedeuten, daß ein Wort mit hoher Frequenz eine geringere Aktivierung auslöst und umgekehrt.

Der neurodynamische Ansatz, der zur Erklärung wiederum Hebb'sche Cell Assemblies heranzieht, geht davon aus, daß die Repräsentation eines hochfrequenten Wortes eine stärkere verschaltete Cell Assembly bedeutet als die eines niederfrequenten. Die Hebb'sche Regel zeigt, daß die häufige Präsentation eines Wortes zu häufigen Ko-aktivierungen der konstituierenden Neuronen führt und damit zu stärkeren Verbindungen.

Die Bedeutung der starken inneren Verschaltung einer Cell Assembly liegt in der erleichterten Reizweiterleitung. Wenn man davon ausgeht, daß ein bestimmtes Maß an Aktivität immer vorhanden ist (Spontanaktivierung), dann hat eine stark verschaltete Cell Assembly eine höhere (aber dennoch unterschwellige) Grundaktivierung als eine leicht verschaltete. Bei einer Zündung durch externe Stimulation wird daher die stark verschaltete Assembly weniger zusätzliche Aktivierung verursachen, da die maximale Aktivierung der Cell Assembly schneller erreicht wird. Doch erst diese zusätzliche Aktivierung wird durch die Meßgeräte erfaßt und daher zeigt eine Messung eine schwächere Aktivierung des Kortex bei hochfrequenten Wörtern als bei niederfrequenten Wörtern:

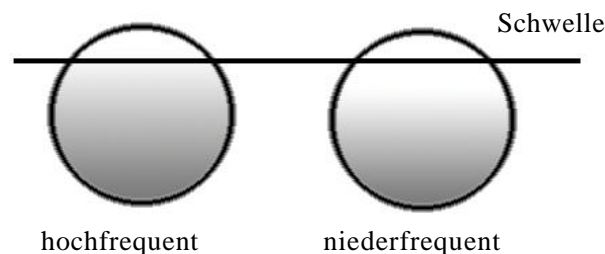


ABBILDUNG. CELL ASSEMBLIES ALS REPRÄSENTATIONEN FÜR HOCH- UND NIEDERFREQUENTE WÖRTER. (DUNKEL: GRUNDAKTIVATION, HELL: ZUSÄTZLICHE AKTIVATION BEI ZÜNDUNG)

Beide Ansätze, der informationstheoretische wie auch der neurodynamische sagen also eine geringere Aktivierung bei hochfrequenten Wörtern als bei niederfrequenten Wörtern voraus.

Vermutlich würde eine solche Unterscheidung später geschehen als die Unterscheidung zwischen ein- und zweisilbigen Wörtern, da für den ersten Effekt die Repräsentation des Wortes zunächst aktiviert werden muß, wohingegen für den zweiten die vorhandene physikalische Begebenheit genügt.

4. Die Stimuli.

Um die obengenannten Hypothesen bezüglich der Repräsentation (Lokalität) und der Eigenschaften (Silben- und Frequenzeffekt) von Wörtern zu untersuchen, wurde zunächst eine Matrix formuliert, die die in diesem Rahmen gesteckten hypothetischen Eigenschaften vollständig nachbildet. Eine solche Tabelle hat einen dreidimensionalen Aufbau (Wortklasse x Frequenz x Silbenzahl). Als Wortklassen wurden die vier modalitätsabhängigen Stufen:

- Funktionswort (function - amodal)
- motorisches Verb (action - unimodal)
- visuelles Nomen (visual - unimodal)
- visuelles und motorisches Nomen (both - bimodal)

vorgesehen. Da nur die binären Unterscheidungen zwischen ein- und zweisilbigen und hoch- und niederfrequenten Wörtern benutzt werden sollten, kann man die Kombinationen auch optisch zweidimensional formulieren:

	function	action	visual	both
1 high				
1 low				
2 high				
2 low				

Somit konnten zusätzlich wortklassenspezifische Frequenz- und Silbenunterschiede untersucht werden oder auch der globale Unterschied zwischen ein- und zweisilbigen Wörtern.

4.1. Stimulus-Wörter

Es wurde zunächst eine Liste mit möglichen Wörtern aus dem Celex-Korpus gemacht. Dabei wurde bei den Frequenzen darauf geachtet, daß die Wörter mit niedriger Frequenz eine Mannheim Frequenz zwischen 50 - 300 haben und die hochfrequenten zunächst einmal bei 500 anfangen.

Diese Tabelle ist basiert nun auf Frequenzen von 200 bzw. 250 vs. 1400 bzw. 1800, um den Abstand zwischen hoch- und niederfrequent möglichst groß zu machen.

	Funkt. (function)	mot. Verben (action)	vis. Nomen (visual)	vis.+mot. Nomen (both)	mittlere Mannheim-Frequenz
1/low	dank	sprang	Hund	Lauf	~200
1/high	wo	liegt	Tag	Hand	~1800
2/low	obgleich	schlagen	Bühne	Finger	~250
2/high	jedoch	gehen	Augen	Gewalt	~1400

Im Gleichklang mit dem allgemeinen Vorgehen, Frequenzen in Vorkommen in einer Million Wörter Text anzugeben, werden nachfolgende Tabellen die Frequenzen nicht mehr in Mannheimfrequenzen (Vorkommen in sechs Millionen Wörtern) angeben.

Es waren jedoch auch weitere Änderungen an den Wörtern selbst vorzunehmen. Gewalt hat jedoch nur die Frequenz 731 mußte daher aus der Tabelle genommen werden. Das Wort wo hat eine zu geringe Buchstabenanzahl und mußte daher durch ein anderes ersetzt werden.

Informelle Umfragen bei Studienkollegen zeigten, daß Hand und Finger oft auf den eigenen Körper (zentrale Furche) bezogen werden und Augen eher auf fremde Augen (visueller Kortex), allerdings nicht bei Kontaktlinsen- und Brillenträgern, weil diese anscheinend sehr oft über ihre eigenen Augen nachdenken und das Wort folglich mit den eigenen Augen assoziieren.

Im Verlauf der Auswahl der Stimuli wurde weiterhin die Forderung gestellt, daß die Morphemzahl der Wörter einer Reihe gleich sein solle. Auch sollte die Anzahl der Buchstaben noch genauer abgestimmt werden. Es wurde daher eine weitere Version erstellt. Zumeist zog das Auswechseln eines Wortes auch das Wechseln anderer Wörter mit sich, so daß die Tabelle nun völlig anders aussah:

	function	action	visual	both
1/low	plus	gehn	Boot	Kinn
To.fr.	87	79	86	66
Le.fr.	87	7302	155	67
1/high	seit	---	Tag	Hand
To.fr.	1748		1737	1879
Le.fr.	1748		5174	2830
2/low	südlich	betrat	Schafe	Zähnen
To.fr.	52	88	54	61
Le.fr.	52	217	83	199
2/high	bevor	setzte	Abend	Haltung
To.fr.	607	658	670	624
Le.fr.	607	2208	790	630

Hier mußte nun Abend entfallen, da es nur ein Morphem enthält. Zur Illustration sind hier auch die Lemma-Frequenzen (Le.fr.) dargestellt, es wurde jedoch weiterhin ausschließlich nach Token-Frequenz (To.fr.) gematcht.

In einer dritten Version wurde die Morphemzahl besser abgestimmt und es wurde darauf geachtet, daß kein Wort die Betonung auf der zweiten Silbe hat.

	function	action	visual	both
1/low	plus	holt	Boot	Kinn
To.fr.	15	15	15	12
1/high	seit	tun	Tag	Hand
To.fr.	308	236	301	329
2/low	deinen	tanzen	Schafe	Glieder
To.fr.	11	14	10	11
2/high	jener	zeigen	Zimmer	Zeitung
To.fr.	123	133	130	123

Die nachfolgend beschriebene Fragebogen-Untersuchung hat ergeben, daß Boot und Kinn vertauscht werden müssen, da sie von den Befragten umgekehrt eingeschätzt wurden.

Damit konnte nun die finale Version der Stimuli beschrieben werden:

	function	action	visual	both
1/low	plus	holt	Kinn	Boot
To.fr.	15	15	12	15
1/high	seit	tun	Tag	Hand
To.fr.	308	236	301	329
2/low	deinen	tanzen	Schafe	Glieder
To.fr.	11	14	10	11
2/high	jener	zeigen	Zimmer	Zeitung
To.fr.	123	133	130	123

Damit hatte die Tabelle alle Voraussetzungen:

- Ausgegliche Frequenzen
- Ausgegliche Silbenzahl
- Alle Betonungen auf der ersten Silbe
- Möglichst gut abgestimmte Buchstabenzahl
- Ausgegliche Morphemzahl

Im Anschluß wurde die Liste der Dummies gemacht, die auch Wörter anderer Sprachen enthalten konnte, denn die Dummies sollten ja nicht ausgewertet werden:

Plus - Plum, -, Holt - Holz, Holm, Boot - Brot, Boom, Kinn - Zinn, Kind, Seit - Zeit, Seil, Tun - Fun, nun, Tag - sag, mag, Hand - Hanf, Hund, Deinen - keinen, seinen, Tanzen - Ranzen, Wanzen, Schafe - Schale, Schabe, Glieder - Flieder, -, Jener - jeder, jenes, Zeigen - Reigen, Feigen, Zimmer - Dimmer, nimmer, Zeitung - Leitung,-

Für einige Wörter konnte kein zweiter Dummie gefunden werden. Es wurde jedoch befunden, daß die Dummie-Menge groß genug sei.

4.2. Fragebogen

Die Stimulus-Generierung war somit vorerst abgeschlossen und es konnte nun ein Fragebogen entwickelt werden, der diese Kategorisierung statistisch und empirisch rechtfertigt.

Es sollte überprüft werden, ob sich die vorgeschlagenen Wortkategorien zumindest kognitiv und subjektiv von einander unterscheiden. Geht man davon aus, daß diese äußerliche, mentale Unterscheidung ihre neurobiologische Entsprechung hat, dann sollte sie auch für oder gegen die aufgestellten Hypothesen sprechen. Für die Wortklassen wurden drei Hauptfaktoren als kritisch angesehen:

1. Abstraktheit
2. Anteil an Visuellem
3. Anteil an Motorischem

Weiterhin war auch von Interesse, ob die CELEX-Frequenzen den Einschätzungen der Personen entsprechen, oder nicht. Die für das Matching benutzten CELEX-Frequenzen sind bezogen auf schriftliche Kommunikation wie Zeitungen, Bücher etc.. Es erscheint sinnvoll, die Frequenzen für das Schriftliche heranzuziehen, da die Wörter visuell präsentiert und damit gelesen werden. Es wurde eine Differenzierung der erfragten Frequenzen in "persönlich" und "allgemein" gemacht. Als zusätzliche Bewertung wurden die mit dem jeweiligen Wort assoziierte Emotion und ihre Ausprägung abgefragt.

4.2.1. Gestellte Fragen

Im Folgenden die einzelnen Fragen, die Erklärungen in Klammern wurden selbstverständlich nicht auf dem Fragebogen abgedruckt und sind hier zum besseren Verständnis angegeben:

- "Wie abstrakt ist das Wort für Sie?" (Abstraktheitsgrad)
- "Wie stark erinnert Sie das Wort an eine Bewegung, die Sie selbst durchführen können?" (motorische Assoziation)
- "Wie stark erinnert Sie das Wort an einen Gegenstand, den Sie visuell wahrnehmen können?" (visuelle Assoziation)
- "Wie häufig verwenden Sie das Wort?" (Frequenz, persönlich)
- "Wie häufig, glauben Sie, kommt das Wort im Alltag (andere Personen, Zeitung, TV etc.) vor?" (Frequenz, allgemein)
- "Spricht Sie das Wort eher an oder eher nicht?" (emotionale Intensität)

- "Verbinden Sie mit dem Wort eher eine angenehme oder eine unangenehme Empfindung?" (emotionale Valenz)

4.2.2. Bewertungen

Die Bewertungen geschahen für jedes Wort einzeln, wobei der Befragte Werte von 1 (nicht oder wenig) bis 10 (sehr) angeben konnte. Die Durchschnitte der 10 Befragten für die drei Hauptinteressen motorisch, visuell und abstrakt ergaben folgende Tabellen:

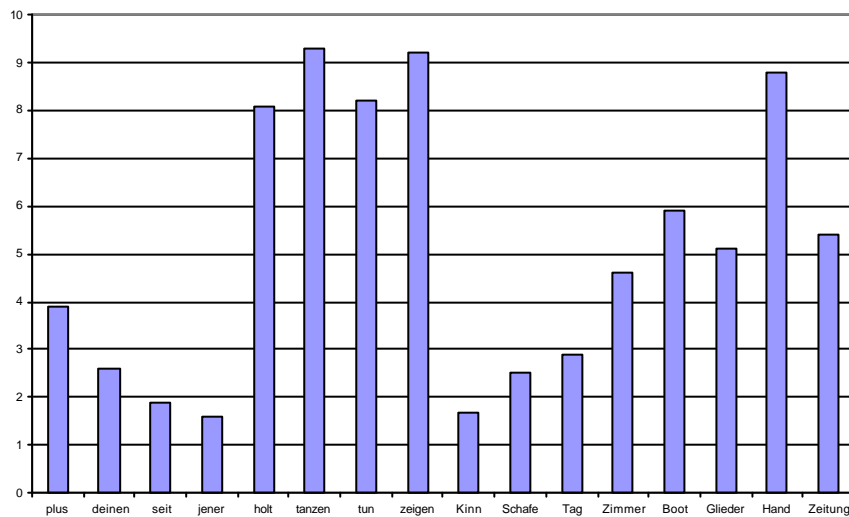


ABBILDUNG. EINSCHÄTZUNGEN, MOTORISCH.

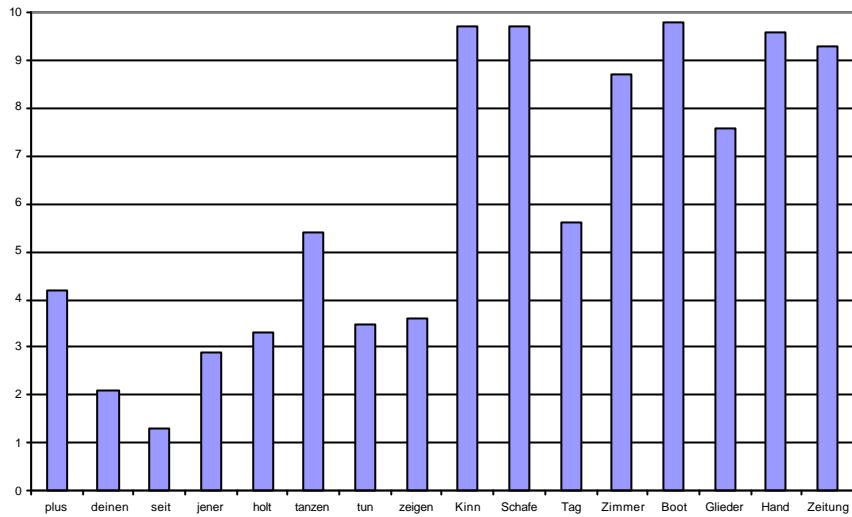


ABBILDUNG. EINSCHÄTZUNGEN, VISUELL.

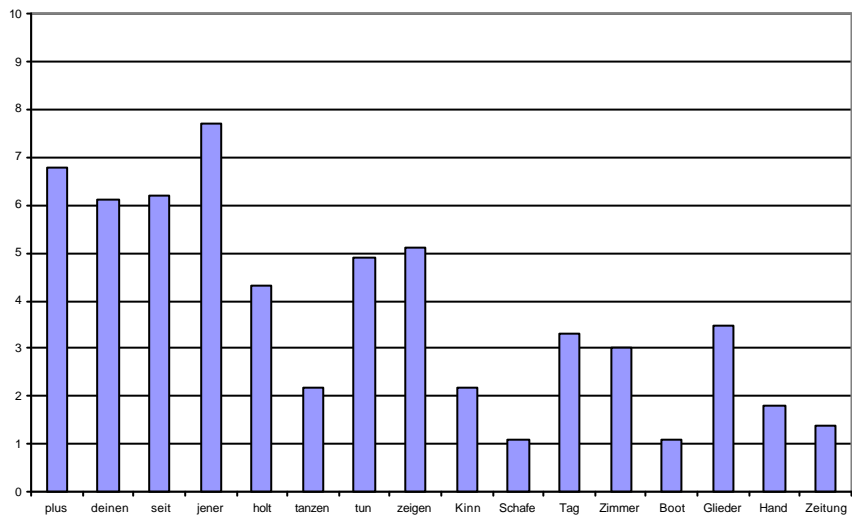


ABBILDUNG. EINSCHÄTZUNGEN, ABSTRAKT

4.2.3. Statistik.

Diese Bewertungen könnten also ein Hinweis auf Richtigkeit der von uns hypostasierten Worteigenschaften sein. Eine Statistik sollte den Eindruck noch weiter untermauern. Die folgenden Ergebnisse zeigen unter anderem eine signifikante Interaktion ($F(3,27) = 45.5$, $p < 0.00001$) zwischen der Assoziation (motorische und visuelle Bewertungen) und den Wortklassen:

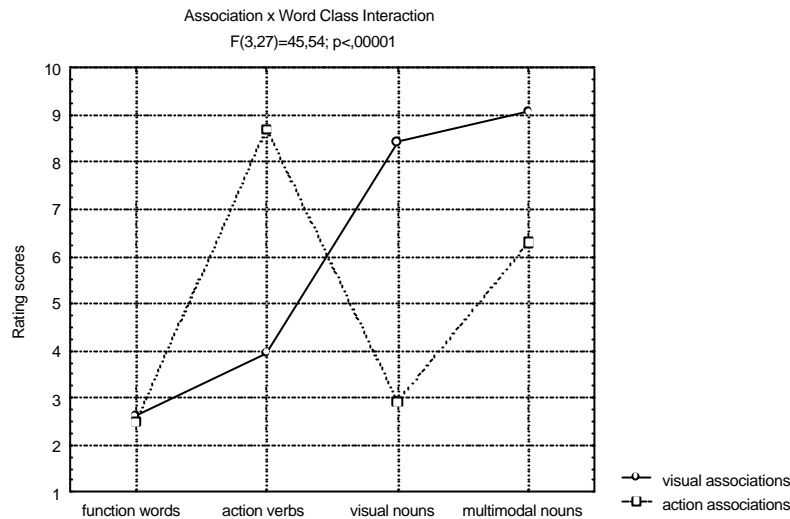


ABBILDUNG. DIE BEWERTUNGEN VON VISUELL UND MOTORISCH FÜR DIE WORTKLASSEN.

Diese Graphik zeigt, wie deutlich sich die Funktionswörter von den anderen Wörtern abheben, indem ihre Bewertungen sowohl für visuelle als auch für motorische Assoziationen sehr niedrig sind. Es ist auch erkennbar, daß sich die motorischen Verben (action) und die multimodalen (motorisch und visuell, both) Nomen in der Bewertung der motorischen Assoziation signifikant von den anderen unterscheiden und umgekehrt, die Klasse both und die Klasse visual starke visuelle Assoziationen hervorrufen. Post-hoc Tests dazu ergaben :

	visuell		motorisch	
	F(1,9)	p <	F(1,9)	p <
FA		n.s.	87.9	0.000006
FV	78.2	0.00001		n.s.
FB	103.0	0.000003	22.7	0.001
AV	47.7	0.00007	83.5	0.000008
AB	51.7	0.00005	13.9	0.005
VB	5.5	0.47	32.3	0.0003

Auch die einzelnen Interaktionen der Wortklassen mit den Faktoren Silbenzahl ($F(3,27) = 10.9, p < 0.00007$) und Frequenz ($F(3,27) = 3.4, p < 0.03$) respektive ihrer Dreifachinteraktion ($F(3,27) = 5.7, p < 0.003$) waren signifikant. Da hier nur die Bestätigung der Wortklassenunterschiede und ihrer semantischen Eigenschaften von

Interesse war, wird an dieser Stelle von näheren Erläuterungen zu den Silben- und Frequenzeffekten Abstand genommen.

5. Die Aufnahme.

5.1. Versuchsperson

Es wurde für den Hauptversuch eine bezahlte rechtshändige (ermittelt nach der Oldfield-Methode) weibliche Testperson mit 13 Jahren Schulausbildung (Abitur) herangezogen, die Deutsch als Muttersprache hat und keine linkshändigen Personen in ihrer Familie. Ihr Alter betrug zur Zeit der Ableitung 21 Jahre. Sie war keine Brillenträgerin und wies keine neurologischen Erkrankungen oder Anzeichen von Drogenkonsum auf.

5.2. Stimuli

Die obengenannten Stimuli wurden für den Versuch eingesetzt.

5.3. Aufgabe

Der Versuchsperson wurden zunächst alle Wörter wie auch der Versuchsablauf vor der ersten Ableitung gezeigt und sie wurde angewiesen sich die Wörter zu merken.

Während den Ableitungen war die Aufgabe der Versuchsperson, das präsentierte Wort zu lesen, ohne es akustisch oder auch nur motorisch zu realisieren. Falls dieses Wort zu den normalen Stimuli gehörte, war nichts weiter zu tun. Gehörte das Wort jedoch zu der Klasse der seltenen Wörter (Dummies), war eine Taste mit dem linken Zeigefinger (rechte, nicht sprachdominante Hemisphäre) zu drücken. In den Pausen war das Fixationskreuz in der Mitte des Bildschirms zu fixieren.

Die Versuchsperson wurde instruiert, nicht zu blinzeln und keine unnötigen Augen- oder Kopfbewegungen zu machen.

5.4. *Versuchsablauf*

Um eine genügend große Anzahl an Daten für die Statistiken und das neuronale Netz zu gewinnen, wurden innerhalb von 4 Wochen 12 Sitzungen. Jede Sitzung bestand aus 3 Teilen (sog. Runs). Ein Run bestand aus 20 Blöcken, jeder dieser Blöcke umfaßte alle 16 Stimuli und 2 der seltenen Wörter.

Zwischen den Sitzungen lagen zwei bis fünf Tage Pause. Bei den Ableitungen konnte die Versuchsperson die Länge der Pause zwischen den Runs und den einzelnen Blöcken selbst bestimmen.

Innerhalb der Blöcke erfolgte die Präsentation in zufälliger Reihenfolge, um Gewöhnungs- und Lerneffekte zu verhindern. Die Wörter wurden für 100ms tachistoskopisch, wobei das Interstimulusintervall zwischen 1.5-2.1s variierte. Die tachistoskopische Präsentation macht Augenbewegungen, die durch Ionenbewegung in den Augenmuskeln große Störsignale erzeugen würden, weniger wahrscheinlich. Die Pause zwischen den Stimulus-Präsentationen soll eine Überlappung der verschiedenen Hirnaktivitäten so weit wie möglich ausschließen. Sie variiert, um die CNV (Contingent Negative Variation, Erwartungspotential) auszuschließen.

Die visuellen Daten wurden auf einem LCD-Projektor außerhalb der MEG-Ableitungskammer erzeugt und über Umlenkspiegel in die Kammer auf eine Mattscheibe projiziert, um elektromagnetische Störungen verursacht durch die Bildröhre eines Monitors auszuschließen.

Die Größe der in weißen Großbuchstaben in der Mitte präsentierten Wörter überschritt 7x3cm nicht. Zeitlich zwischen den Wörtern erschien auf der 1.4m entfernten Mattscheibe in der Mitte von den Augen der VP ein weißes Kreuz, welches die Versuchsperson fixieren sollte.

5.5. Daten Akquisition

Die Ableitung geschah sitzend in dem Ganzkopfmagnetometer MAGNES 2500 WH der Firma Biomagnetic Technologies Inc., BTI.

Aufgezeichnet wurden die Daten kontinuierlich, das heißt, über die Stimulus-Grenzen hinweg, um in der anschließenden Vorverarbeitung in einzelne Epochen zerlegt zu werden. Die Abtast-Rate betrug 508,61 Hz mit einer Bandbreite von 100 Hz, wobei während der Ableitung ein 0.1Hz Hochpass-Filter angewandt wurde. Die Dauer eines kontinuierlich aufgezeichneten Runs betrug 1000s bzw. 17min. Neben den MEG Signalen wurden auch EKG (Elektrokardiogramm) und EOG (Elektrookulogramm, horizontal und vertikal) aufgezeichnet. Die EOG-Aufzeichnungen dienten zur Zurückweisung von Artefakten, zum Beispiel der Aussonderung von Epochen, während derer die VP mit den Augen geblinzelt hat. Das EKG wurde aufgezeichnet, um den Einfluß des Herzmuskels zu beurteilen, der obwohl weit von den SQUIDS entfernt, aufgrund seiner Größe ein potentiell starkes Störsignal darstellen könnte.

Ein weiterer Kanal zeichnete die sogenannten Trigger auf, die durch den Stimulationscomputer mit der Stimulationssoftware ERTS V3.25 generiert wurden. Die Trigger sind ein elektrisches Signal, welches kurzzeitig angelegt wird, um die Präsentation des Stimulus in den Daten zu markieren. Da man einzelnen Wörtern verschiedene Trigger zuordnen kann, fällt es leicht, bei der Vorverarbeitung die Daten einzelner Wörter zusammenzufassen. Die ERTS-Software ist auch dafür verantwortlich, die Reaktion der

Versuchsperson aufzuzeichnen. Dabei wird ein weiterer Trigger auf dem Triggerkanal gesetzt, der die Reaktion der Versuchsperson eindeutig kennzeichnet.

Anzumerken ist noch, daß das MEG weitere Referenzkanäle besitzt und aufzeichnet, die nicht das biomagnetische Signal der Versuchsperson aufzeichnen, sondern das Hintergrundrauschen. Diese Referenzkanäle werden in zwei Stufen zur Erhöhung des Signal-Rausch-Verhältnisses eingesetzt. Die erste Stufe geschieht online, d. h. während der Ableitung und die zweite Stufe wird in der Datenverarbeitung durch Software vollzogen.

6. Das Screening.

Für das Screening wurde die BTI-Software Dataeditor benutzt, welche die verschiedenen Kanäle der gemittelten Epochen anzeigen kann. Darüber hinaus ist es möglich, die Kanäle mehrerer Ableitungen wie Folien übereinander zu legen, um Abweichungen mit dem bloßen Auge zu sondieren. Dazu wurden die Mittelungen über alle Ableitungen herangezogen, denn wie weiter unten noch ausgeführt, verbessert jede weitere Mittelung das Signal-Rausch-Verhältnis und läßt somit Unterschiede deutlicher hervortreten.

Einige Beispiele sollen nun folgen:

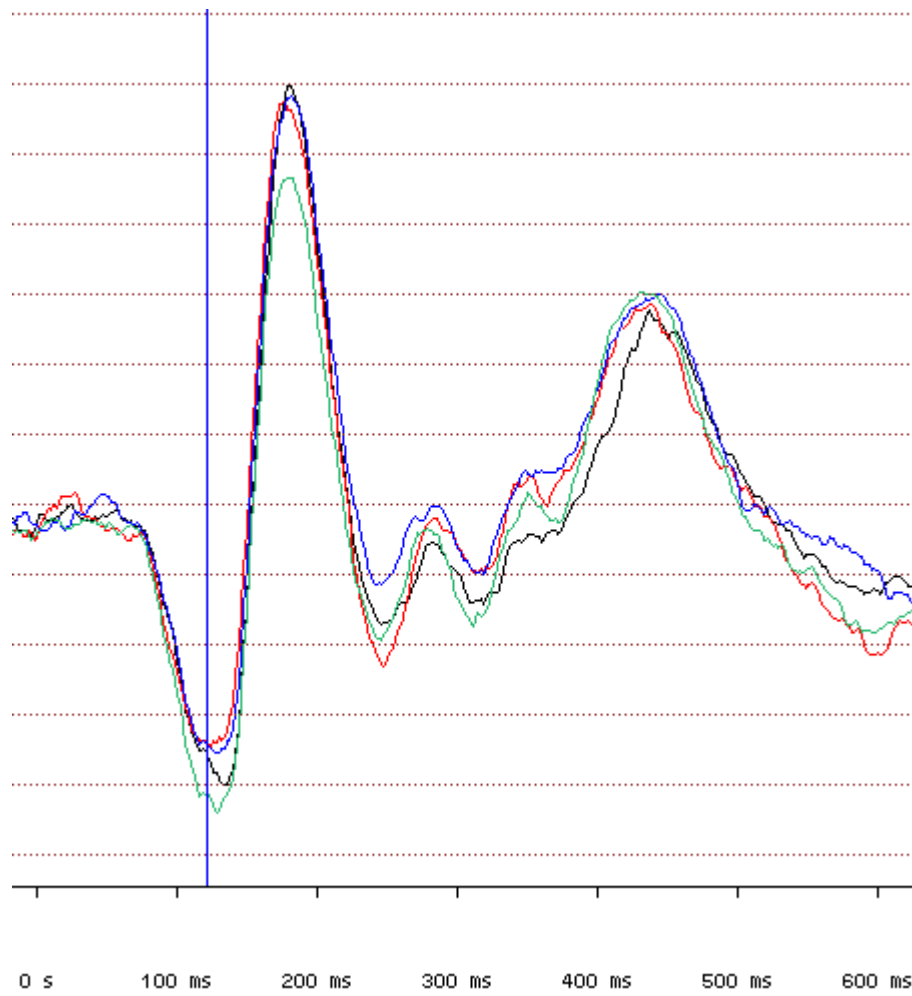


ABBILDUNG. WORTKLASSENUNTERSCHIEDE: BOTH (1. V. U.) UND FUNCTION (2. V. U.)
 HEBEN SICH FRÜZEITIG VON VISUAL (3. V. U.) UND ACTION (4. V. U.) AB.

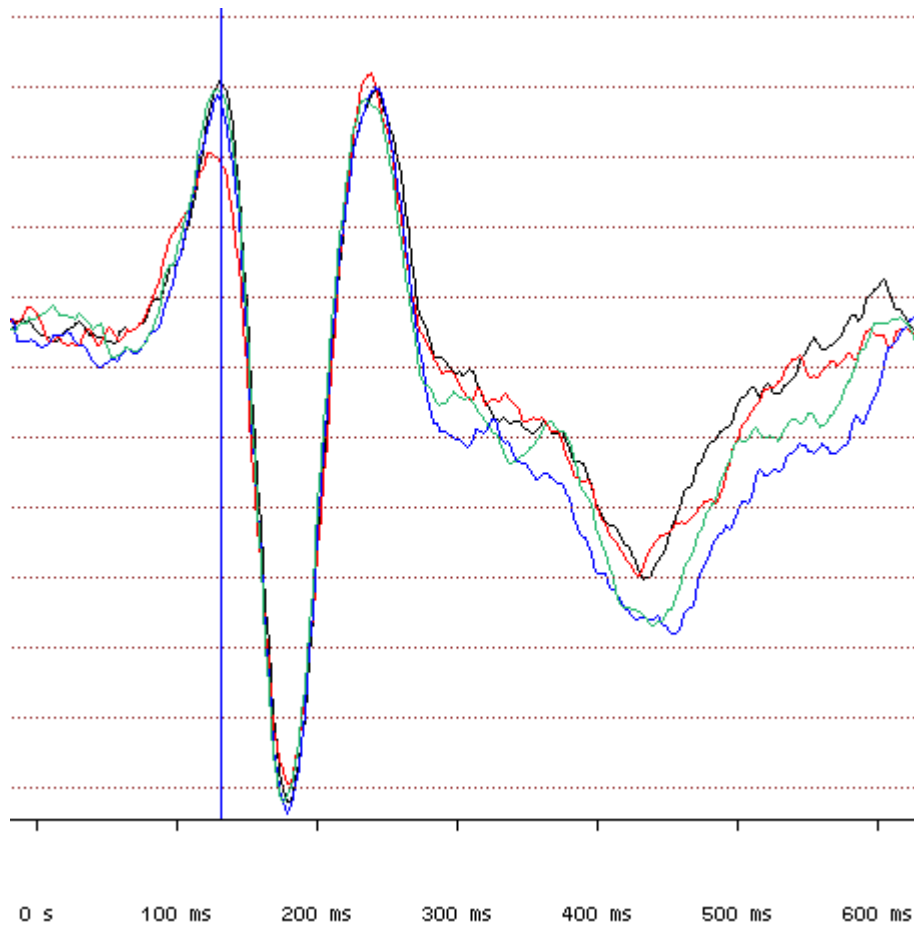


ABBILDUNG. WORTKLASSENUNTERSCHIEDE: ACTION (1. V. U.) HEBT SICH VON DEN DREI ANDEREN KATEGORIEN AB.

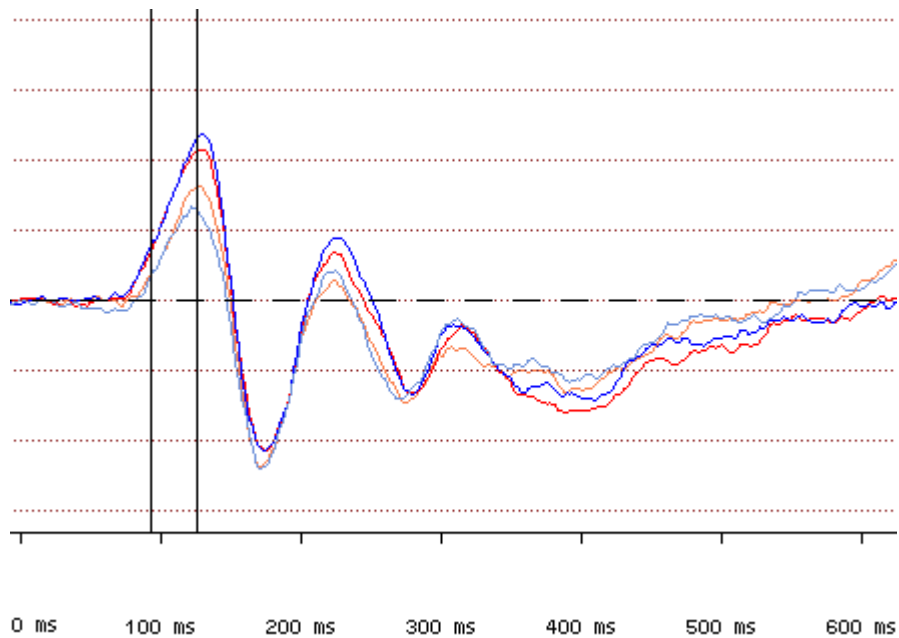


ABBILDUNG. SILBEN-/FREQUENZUNTERSCHIEDE: ZUNÄCHST SPALTEN SICH DIE SILBEN AUF, DANN DIE FREQUENZEN. (V. U. N. O.: H1, L1, L2, H2)

Es scheint, als gäbe es bestimmte Zeitfenster, in welchen spezifische Eigenschaften besser hervortreten als in anderen. Auch scheinen bestimmte Wortklassen sich auf verschiedenen Kanälen stärker abzuheben als andere. Die würde für eine verschieden verteilte Lokalisation der Wörter sprechen.

Auch hat es den Anschein, als würde sich der Einfluß der Silbenzahl zeitlich vor dem Einfluß der Worthäufigkeit im Signal widerspiegeln.

Um diese ersten Eindrücke genauer zu beleuchten, mußten also genauere Analysen gemacht werden. Dazu wurden die Daten in eine Form gebracht, die die statistische und weitere Verarbeitung möglich macht.

7. Die Datenverarbeitung.

7.1. Offline-Reduktion des Rauschens

Die kontinuierlich aufgezeichneten Rohdaten wurden einer Artefaktreduktion unterzogen, um das Hintergrundrauschen, welches auf den Referenzkanälen des MEG aufgezeichnet wird vom Nutzsignal zu eliminieren. Dabei werden globale Gewichte für jeden Run errechnet. Aufgrund der Korrelation von Daten und Referenzkanälen wird anschließend das gewichtete Referenzsignal von den Nutz-Kanälen des MEG subtrahiert. Diese Berechnung wurde mit der BTI Standardsoftware Noisereducer durchgeführt.

7.2. Epochengenerierung

Nun konnten die kontinuierlichen Daten mittels der BTI-Software Boolean Averager in Epochendaten geschnitten werden, was das Datenvolumen reduzierte, da nun alle Epochen von der Stimulusasynchronie befreit in gleich lange Stücke gebracht wurden. Jede Epoche bestand aus 1500ms MEG-Signal, 300ms Baseline vor dem Trigger und 1200ms danach.

Dabei wurden zusätzlich all jene Epochen ausgesondert, bei welchen auf den Stimulustrigger ein Tastendruck folgte. Damit wurden alle Epochen ausgeschlossen, die einen Dummie (seltenes Wort) enthielten und alle, bei denen fälschlich die Taste gedrückt wurde.

7.3. Mittelungen

Die bewährte Methode, Mittelungen über viele Einzelableitungen eines Stimulus (Ereigniskorrelierte Potentiale) zu betrachten, sollte auch hier angewendet werden. Mittelungen werden dazu benutzt, die Hintergrundaktivität aus dem Signal zu entfernen und sich auf die relevanten, stimulus-getriggerten Aktivitätszeichen zu beschränken. Je mehr Mittelungen über ein EKP gemacht werden desto deutlicher treten die Charakteristika hervor:

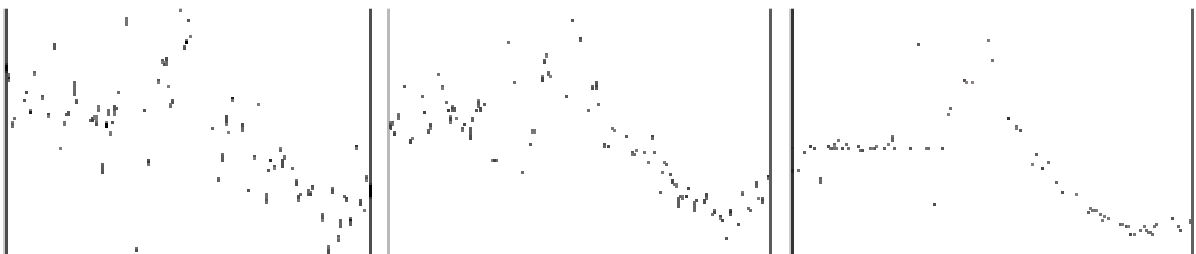


ABBILDUNG. VERSCHIEDENE MITTELUNGEN V.L.N.R.: EIN RUN, EINE SESSION, ALLE
Das entsprechende BTI-Programm Averager errechnet Mittelungen für Datensätze, die aus Epochen mit dem gleichen Stimulus bestehen. Darüber hinaus kann auch hier ein Kriterium für die Auslassung von Artefakten definiert werden. Es lassen sich für die MEG-Kanäle Werte definieren, die bestimmen, wie weit das Minimum und das Maximum in einer Epoche auseinander liegen dürfen, um noch als nicht-kontaminierte Epoche gelten zu dürfen. Der Standardwert ist hier $3pT$, welcher zunächst auch herangezogen wurde. Analog gibt es dieses Kriterium auch für die EOG-Kanäle mit dem Standardwert $120mV$.

Nun konnten Mittelungen für einzelne Runs, einzelne Tage (3 Runs) und globale Mittelungen (alle Runs) erzeugt werden und dies geschah für die einzelnen Wörter ($16 = A/B/F/V \times H/L \times 1/2$), die Wortklassen ($4 = A/B/F/V$), die Silben/Frequenz-Paare ($4 = H/L \times 1/2$) und für alle Wörter ($1 = \text{Wort, nicht-Dummie}$) zusammen.

7.4. Baseline-Korrektur

Alle Mittelungen wurden im Anschluß einer sogenannten Baseline-Korrektur unterzogen. Alle Werte eines Kanals werden im Zeitraum von 100 Millisekunden vor dem Stimulus-Beginn aufaddiert und durch ihre Anzahl geteilt. Diese Mittelung über einen Zeitraum ergibt dann eine Zahl, um die die gesamte Wellenform verschoben wird, so daß der Durchschnitt aller Punkte im Zeitraum von -100 bis 0ms den Wert Null ergibt. Damit sind dann die Mittelungen für verschiedene Wörter, Wortklassen etc. alle mit einander vergleichbar.

7.5. Weitere Artefakt-Behandlung

Die oben erwähnte Software zur Mittelung von EKP gibt nach jedem Aufruf unter anderem den Artefakt-Anteil aus. Für alle gemachten Mittelungen wurde das jeweilige Artefakt-Anteil ermittelt und für alle Runs ausgegeben. Es ergab sich, daß sehr viele Epochen als Artefakt gewertet wurden, die Ausbeute an artefaktfreien Epochen also sehr gering war (oft kleiner als 30%). Eine genauere Untersuchung der artefaktbehafteten Epochen ergab, daß zumeist die beiden äußeren Sensoren-Ringe (siehe Graphik) des Magnetoenzephalographen die Artefakt-Grenze überschritten. Diese sind besonders sensibel für die darunterliegenden Nacken- und Gesichtsmuskeln und haben deshalb hohe Amplituden. Daher wurde das Artefakt-Kriterium von $3pT$ zwischen Minimum und

Maximum in einer Epoche auf 5pT erhöht und gleichzeitig beschlossen, die beiden äußeren Ringe bei weiteren Auswertungen nicht mehr in Betracht zu ziehen:

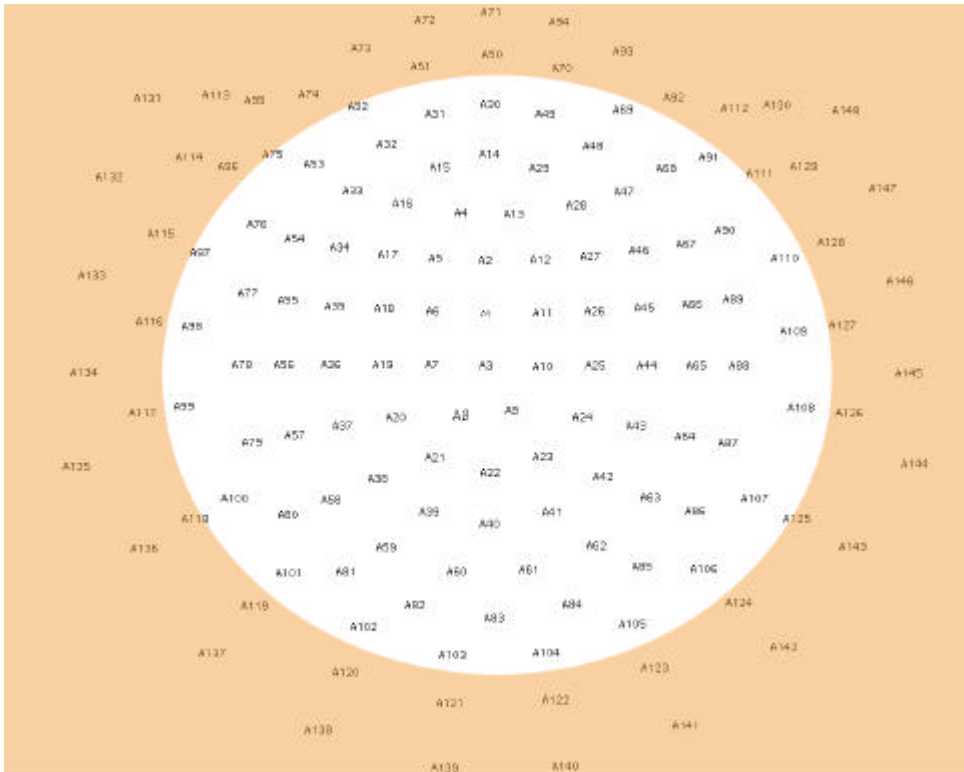


ABBILDUNG. DIE BEIDEN ÄUSSEREN RINGE WURDE WEGEN ZU HOHER ARTEFAKTRATE VON WEITEREN AUSWERTUNGEN AUSGESCHLOSSEN.

Das Kriterium kann bei Versuchspersonen individuell festgelegt werden, da die Signalgröße - und damit auch die Grösse der Artefakte - von Faktoren wie Kopfgrösse und -position im Verhältnis zu den SQUIDS abhängt.

Zusätzlich wurde das Kriterium für die EOG-Känale von 120mV auf 100mV verschärft, um Augenartefakte sicher ausschließen zu können.

Nach dieser grundlegenden Entscheidung mußten die Daten vollständig neu gerechnet werden.

Die resultierenden Daten hatten akzeptable Werte und es wurde beschlossen, diejenigen Runs bei weiteren Auswertungen außer Acht zu lassen, die ein Artefakt-Anteil höher als

60% aufwiesen. Aufgrund dieser Entscheidung fielen 7 Runs raus und es blieben 28 Runs bzw. 11 Versuchstage übrig.

7.6. Festlegung der Zeitscheiben

Zwei Betrachtungen der Worteigenschaften sollten statistisch näher untersucht werden. Zum einen wäre es interessant, den Einfluß der Wortkategorie auf den kortikalen Ort und die Zeit der Unterscheidung zwischen den Klassen function, action, both und visual genauer zu kennen. Zum anderen sind die Einflüsse Silbenanzahl und Frequenz von großem Interesse. Bei der Festlegung der Zeitbereiche wurden die einzelnen Daten über alle Runs hinweg gemittelt.

Nach erstem Screening mittels Betrachten der Daten in einem Darstellungsprogramm (Dataeditor) waren bereits einige Zeitbereiche aufgefallen, die Abweichungen in den Signalen zeigten.

Es wurde ein globales Maß der Differenz zwischen Paaren von Bedingungen verwendet. Um damit die ersten, vorläufigen Bestimmungen der Zeitfenster zu untermauern, wurde ein Programm erstellt, welches für jeden Kanal je zwei Wortklassen von einander subtrahiert und dann den Mittelwert für Zeitfenster von 10ms Länge berechnet. Sodann werden diese Werte für alle Kanäle quadriert, aus ihrer Summe sodann die Quadratwurzel gezogen, die anschließend durch die Anzahl der Kanäle geteilt wird (RMS, root mean square). Man erhält auf diese Weise einen Graphen, der auf 10ms genau die Abweichung zweier Wortklassen im Verlaufe der Zeit darstellt. Legt man nun die 6 (2 aus 4, Paarvergleich aus Anzahl der Bedingungen) möglichen Graphen übereinander, ergibt sich für die Wortklassen folgendes Bild:

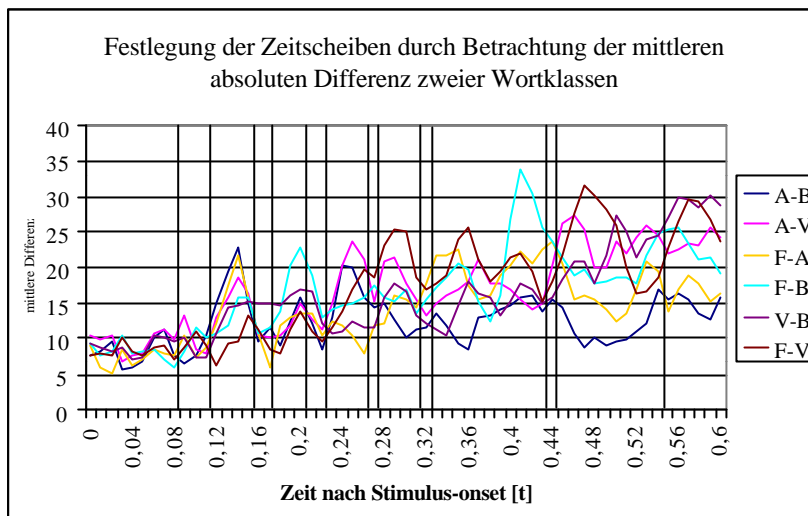


ABBILDUNG. DIE ZEITSCHLEIBEN FÜR DIE WORTUNTERSCHIEDE

	Intervall
I0	90-120ms
I1	120-160ms
I2	180-210ms
I3	230-270ms
I4	280-320ms
I5	330-440ms
I6	440-550ms

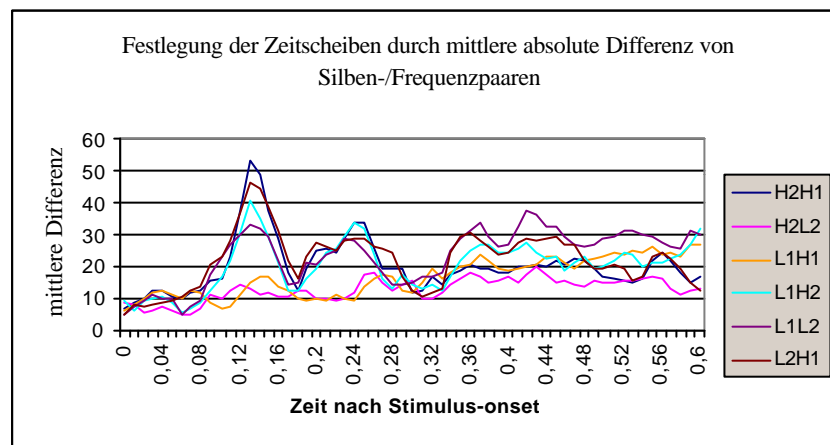


ABBILDUNG. DIE ZEITSCHLEIBEN FÜR DIE SILBEN/FREQUENZ-VERGLEICHE

	Intervall
I0	90-120ms
I1	120-170ms
I01	90-170ms
I2	180-220ms
I3	240-290ms
I23	180-290ms
I4	320-370ms
I5	370-470ms
I6	500-800ms
I7	600-700ms
I8	240-260ms
I9	270-290ms

Wie zu sehen, orientierten wir uns bei der Festlegung der Zeitscheiben am zeitlichen Verlauf der PRMSD (Paarweise RMS Differenz).

Die Zeitscheiben für den Silben/Frequenzvergleich wurden außerdem noch in etwa denen der Wortklassen angepaßt, was zu Folge hatte, daß einige Scheiben als Zusammenfassungen nochmals auftreten (z.B. i01 = i0+i1) bzw. einige Zeitbereiche doppelt analysiert wurden.

7.7. Suche nach den Foci

Es wurde ein Programm geschrieben, welches für die Differenz zweier Stimuli jeweils die fünf negativen und die fünf positiven größten Abweichungen ermittelt und in sortierter Form ausgibt. Danach wurde in einer Karte (flache Darstellung der räumlichen Verteilung der Sensoren, sog. Sensor-Layout-Map) eingezeichnet, wo die Kanäle der größten Abweichung sind und dies über alle Wortklassen hinweg. Die Kanäle wurden dann per Hand in kleine 5er Gruppen eingeteilt, die verschiedene Vergleiche zusammenfaßten. Diese Gruppen (zumeist 6 oder 7) waren dann die Foci, die für weitere Analysen herangezogen wurden. Es gab für jede Zeitscheibe eigene Gruppen. Als Beispiel hier die Gruppen (Foci) für den Zeitabschnitt i0 (90-120ms) bei den Wortklassenvergleichen:

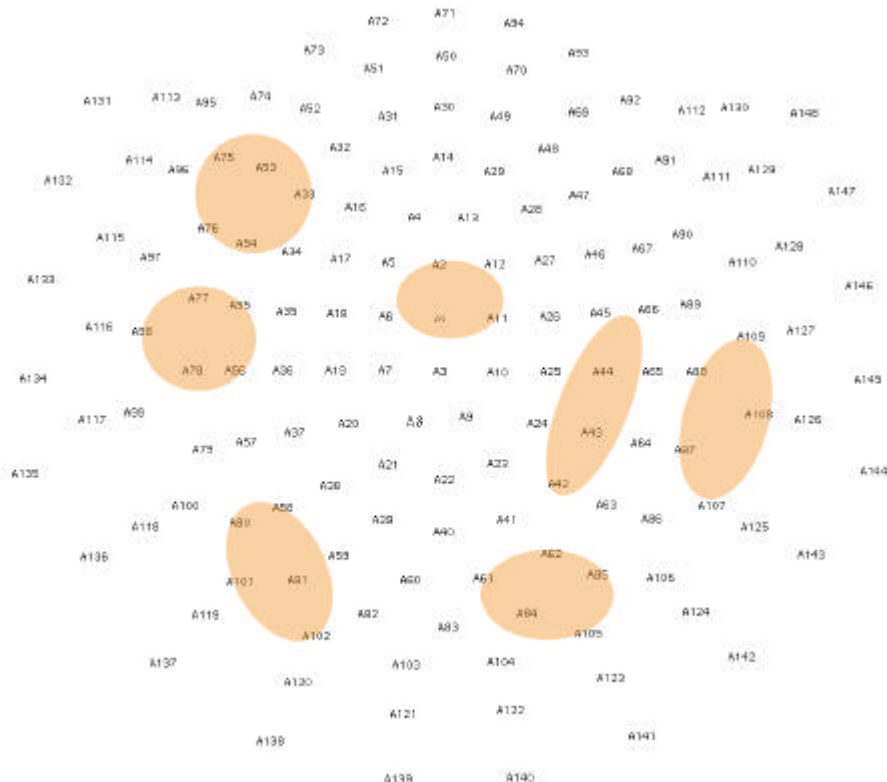


ABBILDUNG. FOCI FÜR DAS ZEITFENSTER 90-120MS FÜR WORTKLASSENVERGLEICHE

7.8. Berechnung der Foci-Daten

Für eine globale, focus-unabhängige Untersuchung der Wortklassen- und Silben-/Frequenz-Unterschiede wurde für alle einzelnen Tage und alle einzelnen Kategorien die sogenannte Global Field Power pro Zeitfenster berechnet. Dazu wurden die Mittelwerte pro Kanal und Zeitabschnitt berechnet, quadriert, aufaddiert, durch ihre Anzahl geteilt und schließlich aus dem Ergebnis die Quadratwurzel gezogen. Dies entspricht einem RMS (root mean square) über die Zeit eines Intervalls.

Die gleiche Berechnung wurde dann auch für die Foci gemacht, mit dem Unterschied, daß hier lediglich jeweils die fünf Focus-Kanäle zur Berechnung herangezogen wurden.

8. Die Statistik.

Es ist die Hypothese zu testen, welchen Einfluß die Silbenzahl und die Worthäufigkeit auf die neuromagnetische Antwort hat und ob und wo sich die Wortkategorien unterscheiden. Für die Untersuchung wurden Zeitbereiche ermittelt, an welchen die Divergenz im Signal am größten ist. Um zunächst generelle Unterschiede bei den Wörtern zu untersuchen, wurde die Global Field Power für jede Zeitscheibe berechnet und einer statistischen Analyse (repeated measurement MANOVA) unterzogen. Dann wurde für spezifischere Untersuchungen der Suchraum eingeschränkt, indem für diese Zeitbereiche die Orte gesucht wurden, an welchen sich die Wörter am stärksten unterscheiden (Foci).

Die Strategie bei beiden Analysen (Silben-/Frequenzpaare, Wortklassen) war gleich: Die Daten wurden von grober zu feiner Unterteilung untersucht. Eine sehr grobe Untersuchung geschah durch die Betrachtung der Global Field Power jeder Zeitscheibe, wo also keine Unterteilung des Kortex in Foci oder Areale vorgenommen, sondern das ganze Feld des

Kopfes untersucht wurde. Die nächst feinere Untersuchung war dann die Betrachtung aller Foci einer Zeitscheibe, i.e. gesamthaft für alle Wortklassen- bzw. Silben-/Frequenzunterschiede. Danach wurden noch feinere Untersuchungen gemacht: eine, die die Foci mit einer Signalstärke über 100fT betrachtet und eine, die genau diejenigen Foci untersucht, die für den jeweiligen Vergleich zweier Kategorien den größten Unterschied zeigten. Damit ergaben sich insgesamt vier Analysestufen:

- Global Field Power
- Alle Foci
- Foci mit Signalstärke $> 100fT$
- Vergleichsspezifische Foci

Im Folgenden werden zwei dieser vier Stufen näher erläutert, da sie als Methoden empirisch besser abgesichert sind. Die Begründung zur Heranziehung der Global Field Power ist sehr einfach: Ändert sich der globale, energetische Zustand (in Form von magnetischen Feldstärken) des Gehirns bei Präsentation von verschieden beschaffenen Wörtern? Dies ist eine Frage, die unabhängig von der Frage nach der Lokalisation gestellt werden kann.

Als zweite, spezifischere Untersuchung soll die Statistik für diejenigen Foci vorgestellt werden, die über 100fT liegen. Ein solches Vorgehen läßt sich durch das höhere Signal-Rausch-Verhältnis dieser Foci begründen, denn Foci mit geringeren Amplituden sind stärker verrauscht und besitzen somit weniger Aussagekraft.

8.1. Statistik für Silben-/Frequenzpaare

8.1.1. Global Field Power

Die Untersuchung der Global Field Power zeigte für die Zeitscheiben i0-i2 (90-220ms) jeweils einen signifikanten Haupteffekt zwischen einsilbigen und zweisilbigen Wörtern ($F(1,10) > 24.0$, $p < 0.0006$). Desweiteren wiesen die Fenster i5-i7 (370-800ms: $F(1,10) > 11.2$, $p < 0.007$) wie auch die Zeitbereiche von 180 bis 290ms (i23, i8: $F(1,10) > 15.4$, $p < 0.003$) diesen Haupteffekt vor.

Während sich also die Hirnantworten auf verschiedene Silbenzahl in den Fenstern i0 wie auch i1 unterschieden, trat der Unterschied zwischen den Wortfrequenzen erst in der Zeitscheibe i1 (120-160ms) signifikant ($F(1,10) = 8.3$, $p = 0.016$) auf. Dies bedeutet, daß ein Silbeneffekt einsetzte, bevor ein Frequenzeffekt zu beobachten ist. Der signifikante Haupteffekt zwischen den Frequenzen tritt auch in den Zeitscheiben i01 ($F(1,10) = 9.7$, $p = 0.01$) und i8 ($F(1,10) = 5.9$, $p = 0.036$) auf.

Eine signifikante Interaktion der beiden Faktoren ist bis zum Zeitpunkt 240ms nicht zu beobachten, dann jedoch im Fenster i8 (240-260ms: $F(1,10) = 6.5$, $p = 0.03$) vorhanden. Diese Interaktion ist erst später wieder signifikant (600-800ms, i7: $F(1,10) = 5.2$, $p = 0.04$). Geplante Vergleiche der Interaktionen zeigen, daß ein Frequenzeffekt für die zweisilbigen Wörter (i.e. H2 vs. L2) nur in den Fenstern 240-260ms (i8: $F(1,10) = 13.6$, $p = 0.004$) und 370-470ms (i5: $F(1,10) = 7.5$, $p = 0.02$) vorhanden ist, die einsilbigen Wörter sich aber später, im Zeitraum von 500-800ms (i6, i7: $F(1,10) > 5.7$, $p < 0.038$) hinsichtlich der Frequenz unterscheiden.

Übersicht (HF Haupteffekt, I Interaktion):

	HF:syl		HF:freq		I:syl x freq	
	F	p	F	p	F	p
i0: 90-120ms	34.3	.0001		n.s.		n.s.
i1: 120-170ms	507.4	.000001	8.4	.015		n.s.
i01: 90-170ms	169.4	.000001	9.7	.01		n.s.
i2: 180-220ms	24.0	.0006		n.s.		n.s.
i3: 240-290ms		n.s.		n.s.		n.s.
i23: 180-290ms	17.2	.001		n.s.		n.s.
i4: 320-370ms		n.s.		n.s.		n.s.
i5: 370-470ms	11.2	.007		n.s.		n.s.
i6: 500-800ms	29.2	.0003		n.s.		n.s.
i7: 600-700ms	34.0	.0002		n.s.	5.3	.04
i8: 240-260ms	15.4	.003	5.9	.04	6.5	0.03
i9: 270-290ms		n.s.		n.s.		n.s.

8.1.2. *Foci > 100fT*

Die Analyse der hochamplitudigen Foci untermauern die im vorigen Abschnitt vorgestellten Ergebnisse. Ausschließlich der Silbeneffekt tritt in der Zeitscheibe i0 (links und rechts parietal, rechts zentro-okzipital: $F(1,10) = 39.4$, $p = 0.0001$) auf, wohingegen die Zeitscheibe i1 eine signifikante Interaktion zwischen Silben und Frequenzen ($F(1,10) = 20.4$, $p = 0.001$), sowie eine signifikante Dreifachinteraktion zwischen Silben, Frequenzen und den Foci (links parietal, rechts zentro-okzipital: $F(1,10) = 5.6$, $p = 0.039$) aufweist:

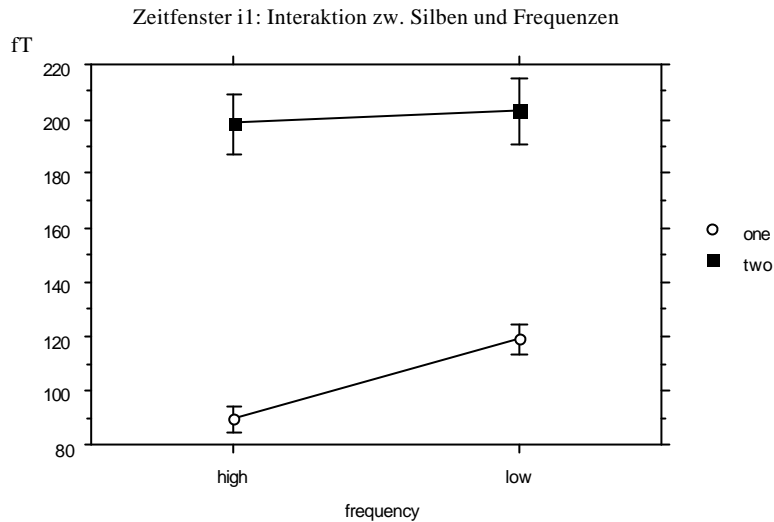


ABBILDUNG. INTERAKTION ZW. SILBENZAHL UND FREQUENZEN IN I1 (120-170MS).

Der Haupteffekt beim Faktor Silben ist sehr stabil ($F(1,10) > 16.7$, $p < 0.002$, $\eta^2 = 1.0$) und unterbleibt lediglich in den Zeitbereichen 320-370ms (i4) und 270-290ms (i9); i23 besaß keine Foci mit Feldstärken über 100fT. Der schwächer ausgeprägte Frequenzeffekt zeigt sich nur in den Intervallen von i1, i01 (90-170ms: $F(1,10) > 16.7$, $p < 0.002$) und i6-i9 (500-800ms, 240-260ms, 270-290ms, 370-470ms: $F(1,10) > 5.1$, $p < 0.047$) und verteilt sich damit unzusammenhängend über den untersuchten Zeitraum.

Die signifikante Interaktion der beiden Faktoren beschränkt sich auf die Bereiche i1, i01, i5, i6, i7 ($F(1,10) > 8.0$, $p < 0.01$). Post-hoc Tests zeigen dabei, daß im Zeitfenster i1 alle Vergleiche bis auf den Frequenzeffekt bei Zweisilbern (hochfrequente Zweisilber vs. niederfrequente Zweisilber) signifikant werden (Haupteffekte: $F(1,10) > 46.1$, $p < 0.00004$). Im Gegensatz dazu weist das Fenster i01 (90-170ms) den Frequenzeffekt sowohl bei Ein- als auch bei Zweisilbern als signifikant aus (Vergleiche hochfreq. einsilb. vs. niederfreq. einsilb.: $F(1,10) = 28.2$, $p = 0.0003$ und hochfreq. zweisilb. vs. niederfreq. zweisilb.: $F(1,10) = 19.9$, $p = 0.001$). Bis auf den Silbeneffekt bei den hochfrequenten

Wörtern (hochfreq. einsilb. vs. hochfreq. zweisilb.: $F(1,10) = 42.6$, $p = 0.00006$) werden jedoch keine weiteren Vergleiche signifikant.

Post-hoc Tests für das Zeitfenster i5 (370-470ms) zeigen, daß die niederfrequenten Zweisilber sich von allen anderen Kombinationen signifikant ($F(1,10) > 7.5$, $p < 0.02$) unterscheiden, während diese sich ihrerseits bis auf den Vergleich niederfreq. einsilb. vs. hochfreq. zweisilb. (Haupteffekt: $F(1,10) = 10.0$, $p = 0.01$) sich nicht unterscheiden. Eine andere Kombination, nämlich die niederfrequenten Einsilber unterscheiden sich in den darauffolgenden Zeitscheiben i6 und i7 (500-800ms) von allen anderen ($F(1,10) > 25.2$, $p < 0.0007$). Diese wiederum unterscheiden sich untereinander nicht.

Übersicht (i23: keine Foci > 100fT):

	HE: syl		HE: freq		I: syl x freq		I: syl x freq x foc	
	F	p	F	p	F	p	F	p
i0: 90-120ms	39,4	0.0001		n.s.		n.s.		n.s.
i1: 120-170ms	388,7	0.0001	29,6	0.0003	20,4	0.001	5,6	0,04
i01: 90-170ms	355,0	0.0001	16,6	0.002	21,4	0.0009		n.s.
i2: 180-220ms	16,7	0.002		n.s.		n.s.		n.s.
i3: 240-290ms	33,4	0.0002		n.s.		n.s.		n.s.
i23: 180-290ms		--		--		--		--
i4: 320-370ms		n.s.		n.s.		n.s.		n.s.
i5: 370-470ms	15,5	0.003		n.s.	7,6	0.019		n.s.
i6: 500-800ms	17,5	0.002	15,6	0.003	17,9	0.002		n.s.
i7: 600-700ms	35,4	0.0001	11,3	0.007	8,1	0.017		n.s.
i8: 240-260ms	40,6	0.0001	21,6	0.0009		n.s.		n.s.
i9: 270-290ms		n.s.	5,1	0.047		n.s.		n.s.

8.2. Statistik für Wortklassenvergleiche

8.2.1. Global Field Power

Die Analyse der Wortklassenunterschiede in Maß der Global Field Power zeigt keinen signifikanten Unterschied in der Zeitscheibe i0 (90-120ms), jedoch in der Zeitscheibe i1 (120-160ms: $F(3,30) = 9.3$, $p = 0.0007$, $(= 0.77)$). Im Zeitbereich von 180-270ms (i2, i3) bleibt ein solcher Unterschied aus und wird erst wieder bei i4 (280-320ms: $F(3,30) = 4.3$, $p = 0.032$, $(= 0.62)$) signifikant. Das Zeitfenster um 400ms (i5, 330-440ms) zeigt keine

signifikanten Unterschiede, wohingegen das späte Fenster um 500ms (i6, 440-550ms: $F(3,30) = 15.3$, $p = 0.0001$, $\eta^2 = 0.64$) die Unterscheidung wieder signifikant werden läßt.

Zusammenfassend sind in der Global Field Power Analyse lediglich die Unterschiede in den Intervallen i1, i4 und i6 signifikant.

8.2.2. *Foci > 100fT*

Eine Greenhouse-Geisser-korrigierte Analyse der drei stärksten Foci (links und rechts parieto-okzipital, rechts temporal) im Zeitbereich i0 wies einen signifikanten Haupteffekt ($F(3,30) = 5.5$, $p = 0.005$, $\eta^2 = 0.93$) auf:

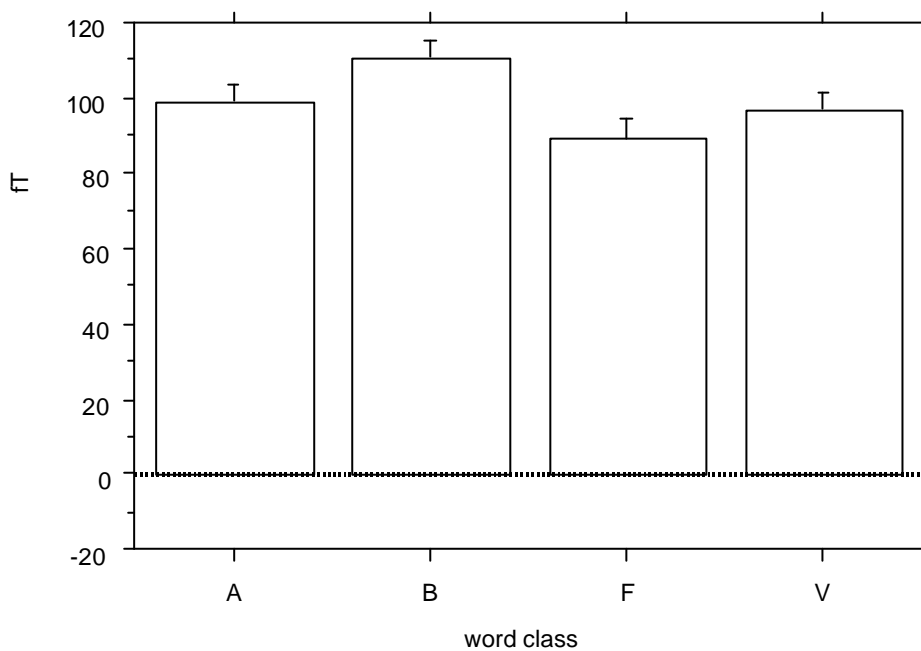


ABBILDUNG. HAUPTEFFEKT DER WORTKLASSEN IM INTERVALL I0.

Geplante Vergleiche zeigten signifikante Unterschiede zwischen der Wortklasse Both (multimodal) und allen anderen Wortklassen:

action vs. both: $F(1,10) = 6.2, p = 0.03$

function vs. both: $F(1,10) = 15.3, p = 0.002$

visual vs. both: $F(1,10) = 6.8, p = 0.02$

Die geplanten Vergleiche zwischen den anderen Wortklassen waren nicht signifikant.

Demnach divergierte die multimodale Klasse both im Zeitbereich i0 von allen übrigen Klassen, wogegen diese ihrerseits sich nicht unterschieden. Dies spricht für einen fundamentalen biomagnetischen Unterschied der multimodalen Wörter zu allen anderen Klassen.

Für die Zeitscheibe II (120-160ms) ergaben sich noch höher signifikante Unterschiede zwischen allen Wortklassen:

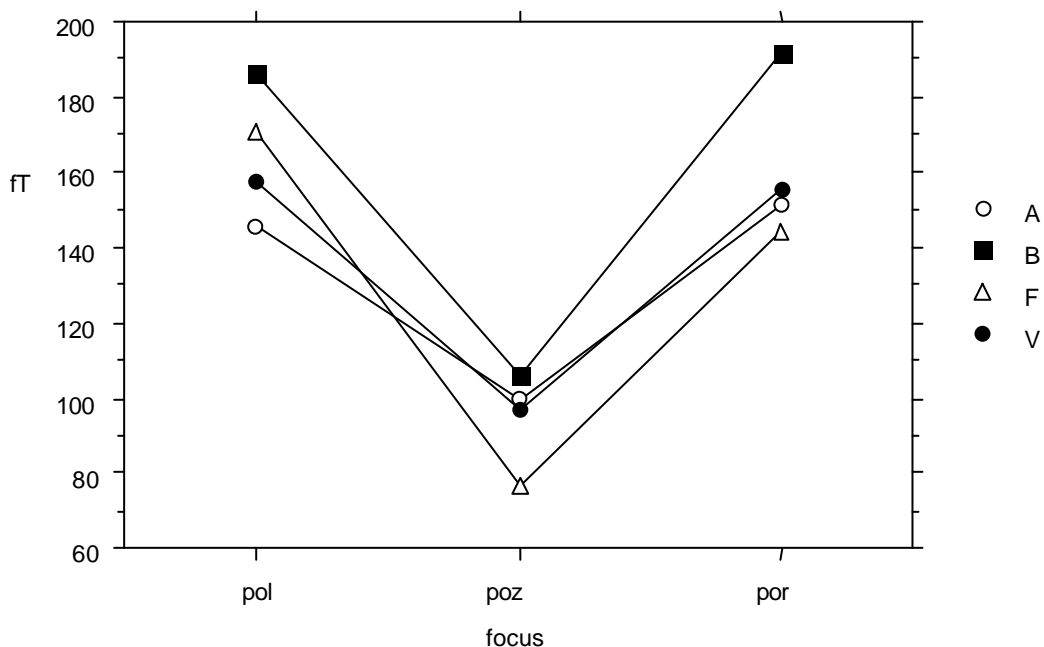


ABBILDUNG. SIGNIFIKANTE INTERAKTION ZWISCHEN WORTKLASSEN UND FOCI IM ZEITBEREICH II (120-160MS, POL PARIETO-OKZIPITAL-LINKS, POZ -ZENTRAL, POR - RECHTS)

Die Statistik zeigt hier eine signifikante Interaktion zwischen den Faktoren Wortklasse und Focus ($F(6, 60) = 4.039, p = 0.0216, \eta^2 = 0.431$). Das deutliche Abheben der Klasse Both im Zeitbereich i0 wurde hier weiter ausgebaut. Die Klasse Both zeigte in geplanten Vergleichen folgende signifikanten Effekte im links parieto-okzipitalen Focus (pol). Es zeigte sich aber, daß nun auch die anderen Vergleiche über die Foci (siehe oben) signifikant wurden, so daß sich nun alle Wortklassen voneinander unterschieden:

Im Zeitbereich i2 (180-210ms) zeigten die Faktoren Wortklasse und Focus wieder eine signifikante Interaktion ($F(12, 120) = 4.6, p = 0.0049, \eta^2 = 0.313$). Dabei waren die 5 Foci links zentral (f2), links parieto-okzipital (f3, f4, f5) und rechts temporal (f6) angesiedelt.

Ein signifikanter Haupteffekt war bei den Wortklassen im Intervall i4 (280-320ms: $F(3,30) = 10.3, p = 0.0003, \eta^2 = 0.8$) links und rechts zentral zu beobachten.

I5 und i6 (330-440ms, 440-550ms: $F(3,30) > 3.2, p < 0.008, \eta^2 > 0.58$) zeigten signifikante Interaktionen auf, die die Klassen Both und Visual in den Foci (links temporal, links und rechts zentral) von den beiden anderen abtrennen:

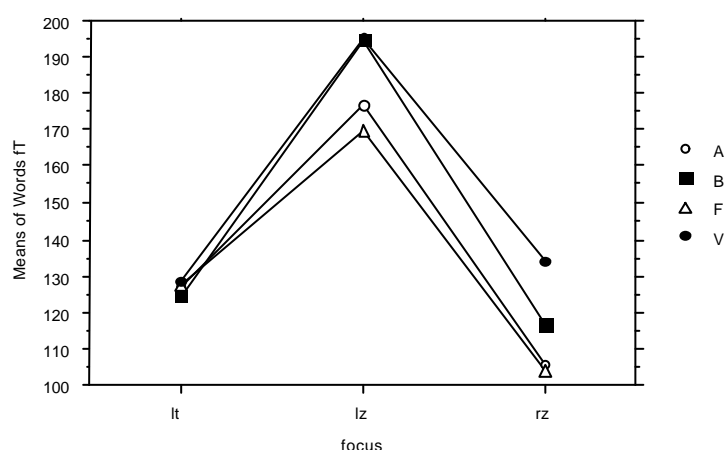


ABBILDUNG. INTERAKTION DER WORTKLASSEN IM ZEITFENSTER I5 (330-440MS, LT LINKS-TEMPORAL, LZ LINKS-ZENTRAL, RZ RECHTS-ZENTRAL).

8.3. Weitere Ergebnisse

Ob sich die im Fragebogen erhobenen psychologischen Daten in den biomagnetischen Hirnantworten widerspiegeln, ist eine interessante Frage. Wenn die kognitiven Einschätzungen der Wörter mit ihren neurobiologischen Strukturen und deren Antwortcharakteristika übereinstimmen, müßte sich eine Korrelation zwischen den beiden ergeben. Diese Antworten werden durch die kortikal motivierte Theorie vorhergesagt: Ein Funktionswort sollte als lokal beschränktestes Wort eine kleine Hirnantwort verursachen, und weiterhin sollte ein Wort, welches eine visuelle oder eine motorische Assoziation besitzt, wegen größeren Netzen eine größere Antwort erzeugen und schließlich sollte ein Wort, welches beide Assoziationen aufweist, die größte Hirnantwort liefern.

Die Modalitätsspezifika haben sich bereits in den Fragebogenwerten gezeigt. Im Kapitel 4.2.3. wurden die Unterschiede zwischen den Wortklassen durch einen Vergleich der Bewertungen "visuell" und "motorisch" gezeigt. Wenn jene Bewertungen addiert werden, sollte sich das hypostasierte Bild ergeben. Der andere Faktor der Korrelation wurde durch Addition von zwei signalstarken Kanälen berechnet. Die Berechnung geschah durch die Mittelung der Werte im Zeitbereich i0 (90-120ms).

Eine Analyse der beiden Faktoren ergab hochsignifikante Korrelationen mit einem $R = 0.796$ für den Fragebogen über zehn Befragte und einem $R = 0.766$ für den Fragebogen, den die Versuchsperson selbst ausfüllte.

Die Regressionsplots für die beiden Korrelationen bestätigen die Hypothese:

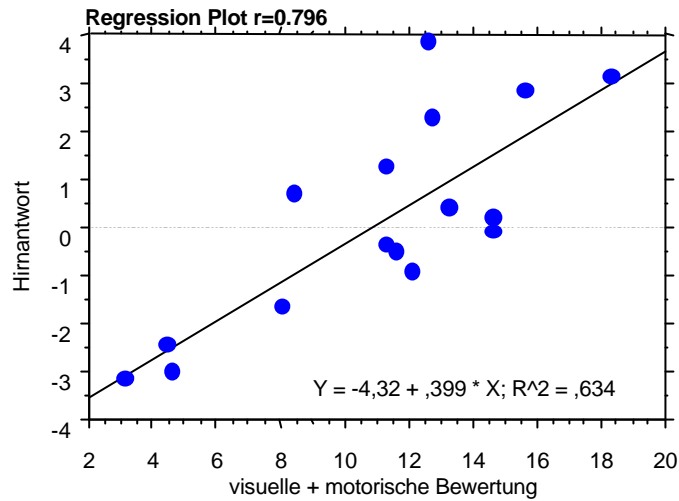


ABBILDUNG. KORRELATION DER HIRNANTWORT AUF DIE WÖRTER UND DER BEWERTUNG DURCH 10 PERSONEN.

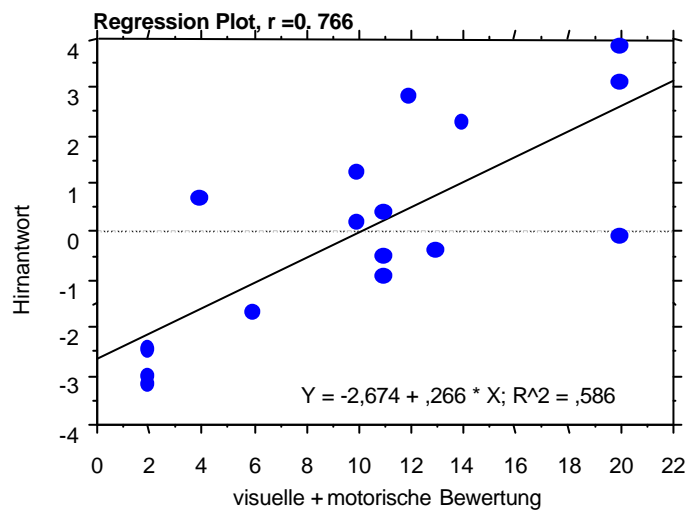


ABBILDUNG. KORRELATION DER HIRNANTWORT AUF DIE WÖRTER UND DER BEWERTUNG DURCH DIE VERSUCHSPERSON SELBST

9. Die Diskussion.

Die Ergebnisse des letzten Kapitels lassen nun eine Diskussion der Hypothesen aus Kapitel 3 zu. Was wissen wir über Worthäufigkeiten, -längen im Hinblick auf die Unterscheidung von Funktions- und Inhaltswörtern (Zipf 1935)?

* Es ist bekannt, daß Wortlänge und -häufigkeit stark korrelieren (z.B. Osterhout et al. 1997, King & Kutas 1998): Je höher die Frequenz eines Wortes, desto wahrscheinlicher ist es, daß es sehr kurz ist.

* Die häufigsten Wörter sind zumeist Funktionswörter (Francis & Kucera, 1982).

Einige EEG-Studien haben ergeben, daß Wortlänge und Frequenz einen Einfluß auf die Latenz einer negativen Komponente, die auf die P2 folgt, haben (Osterhout et al. 1997), andere jedoch, daß ausschließlich die Frequenz einen Effekt hat (King & Kutas 1998).

Neville et al. fanden 1992 heraus, daß Funktionswörter eine stärkere links anteriore Aktivität um 280ms nach Stimulus-Onset (N280) aufweisen, während Inhaltswörter eine nicht-lateralisierte Hirnantwort mit Maximum um 350-400ms (N400) zeigten. Diese N400 war zwar auch bei Funktionswörtern vorhanden, jedoch stärker lokalisiert und deutlich schwächer ausgeprägt. Die Interpretation dieser Ergebnisse ist allerdings diskutabel, da in dieser Untersuchung äußerst wenig auf Wortlänge, -frequenz, Abstraktheit, Kontextabhängigkeit, Referenzierung oder grammatische Rolle geachtet wurde. Es ist daher laut Osterhout et al. 1997 schwierig, die vorgefundenen Effekte den Eigenschaften der Wortklassen zuzuordnen. Eine weitere Problematik besteht im Versuchsaufbau selbst: Die Wörter wurden innerhalb eines Satzkontextes präsentiert. Es ist daher möglich, daß der Kontext und die syntaktische Verarbeitung die Hirnantworten auf die einzelnen Wörter verändert haben.

Osterhout et al. selbst haben daher 1997 zwei Experimente gemacht, die einander sehr ähnlich waren. In einem Experiment wurden den Versuchspersonen Wörter eines zusammenhängenden Textes einzeln gezeigt und in einem weiteren wurden die gleichen Wörter einzeln in zufälliger Reihenfolge präsentiert. In beiden Fällen ergab sich dabei, daß die Latenz der bereits angesprochenen negativen Komponente mit der Wortlänge steigt und mit der Wortfrequenz fällt. Dieser Latenzeffekt ist unabhängig von der Wortklasse. Es zeigte sich auch, daß die späte positive Komponente (late positive component, LPC), welche bei Inhaltswörtern auftritt, bei nicht zusammenhängenden Wortfolgen kleiner wird. Auch die korrespondierende Negativierung für Funktionswörter über dem Zeitraum 400-700ms fällt deutlich geringer in der zufälligen Wortreihenfolge aus. Das Ergebnis von Neville et al. 1992, bei welchem sich eine starke, links anteriore Negativierung für Funktionswörter bei 280ms zeigt, konnte im Großen und Ganzen nicht bestätigt werden, lediglich Artikel zeigten diesen Effekt und auch nur in der normalen Wortreihenfolge.

Dies kann jedoch auch an dem großen Frequenzspektrum oder der verschiedenen Länge der Wörter liegen, welches unter Umständen die feinen Effekte durch Überlagerung der verschiedenen Latenzen zunichte macht. Dennoch kann aus diesen Studien gefolgert werden, daß oftmals fälschlich Wortklassenunterschiede als Ursache angenommen worden waren, obwohl die Frequenzverteilung in den Wortklassen den Effekt verursachte. Um diesen Frequenzeffekt auszuschalten, haben King & Kutas 1998 die Rohdaten 4Hz-hochpassgefiltert. Die genaue Motivation dieses Verfahrens bleibt jedoch in der Publikation unklar. Wenn man aber annimmt, daß eine abgesicherte Rechtfertigung für dieses Verfahren besteht, dann ist immer noch unverständlich, weshalb eine Untersuchung der Amplituden im Bereich 100ms nach Stimulus-Onset, die offensichtlich die beiden Wortklassen trotz Korrektur zu unterscheiden scheint, unterbleibt. King & Kutas folgern weiterhin aus einer Korrelation von 0.34 zwischen Wortlänge und Frequenz abhängiger

Negativität (frequency dependent negativity, FSN), welche zusammen mit der Wortfrequenz 22% der Varianz erklärt, daß die Wortlänge keine gute Vorhersage über die Latenz macht. Auch wird kein Einfluß der Wortklasse auf die Latenz oder die kortikale Distribution der Hirnantwort gefunden. Aber auch hier ist wiederum eine Untersuchung gemacht worden, welche die Wörter in einen Satzkontext stellt und damit die genaue Analyse der Gründe schwierig macht.

Im Gegensatz zu den vorgestellten Studien gehen andere Gruppen den umgekehrten Weg. Um eine mögliche Konfundierung der Ergebnisse durch Wortlänge oder Frequenz a priori auszuschließen, werden einzelne Wörter über die Wortklassen hinweg so gut wie möglich auf einander abgestimmt und nicht in einer sinnvollen Reihenfolge präsentiert. Garnsey hat dieses Vorgehen 1985 zuerst vorgebracht, welches dann auch von Pulvermüller et al. 1995 (Frequenz und Längenübereinstimmung) und Nobre & McCarthy 1994 (Längenübereinstimmung) benutzt und verfeinert wurde. Garnsey 1985 zeigte durch PCA (principle component analysis), daß sich ein Unterschied zwischen den Wortklassen bei 400ms finden läßt. Dennoch war ein solcher Unterschied bei Standardanalysen nicht signifikant. Hier läßt sich argumentieren, daß trotz der Abstimmung der Wörter in den Aspekten Frequenz und Länge eine "Verwaschung" der Effekte stattfand, denn die Silbenzahl betrug 1-4 und die Frequenz variierte zwischen 3 und 1300 pro Million Wörter Text.

Die Studie von Pulvermüller et al. 1995 zeigte bei genauerer Abstimmung der Stimuli eine Unterscheidung der hoch- und niederfrequenten Wörter bei 160ms und Inhaltswörter zusätzlich eine stärkere Negativierung in der rechten Hemisphäre (d.h. eine geringere Lateralisation) aufwiesen.

In der vorliegenden Studie, die auch den letzteren Weg beschreitet, wurden viele dieser Faktoren berücksichtigt. Alle Wortklassen wurden so genau wie möglich aufeinander

abgestimmt. Die Frequenz der Wörter wurde auf hoch- und niederfrequent beschränkt, so daß sich keine Überlagerung der verschiedenen Frequenzen ergeben sollte, wenn man die einzelnen Konditionen untersucht.

Zur Analyse der Wortklasseneffekte wurden also Silbenzahl und Frequenz konstant gehalten, um einen solchen konfundierenden Effekt auszuschalten, und zu untersuchen, ob dennoch ein Wortklasseneffekt auftritt. Auch wurde trotz geringer Stimuluszahl bei den Funktionswörtern darauf geachtet, daß diese Kategorie möglichst viele verschiedene grammatikalische Kategorien enthält (Präposition, Konjunktion, Pronomen).

Zunächst einmal existiert auch in dieser Studie ein Wortlängeneffekt, und er äußert sich darin, daß bereits schon etwa 100ms nach Präsentation eines Wortes seine Silbenzahl die Stärke der magnetischen Hirnantwort bestimmt. Wie auch bei der Osterhout-Gruppe zeigt sich, daß mit steigender Silbenzahl höhere evozierte Potentiale auftreten. Nicht bestätigt werden kann jedoch, daß die höhere Silbenzahl zu einer längeren Verarbeitungsdauer führt, was sich in einer größeren Dauer der Wellenform widerspiegeln würde. Das kortikal motivierte Modell erklärt dieses Phänomen durch eine größere Anzahl von Cell Assemblies auf der Ebene der Silbenrepräsentation, die zünden und damit mehr meßbare Ionenströme erzeugen. Physikalisch gesehen würde die zusätzliche Lichtenergie stärkeren Input verursachen, welcher über mehr EPSPs zu höheren magnetischen Amplituden führt.

Auch ein Frequenzeffekt wurde vorgefunden, er zeigt sich ca. 30-60ms später und ist für einsilbige Wörter ausgeprägter als für zweisilbige. Keine der obengenannten Studien hat bisher über eine zeitliche Anordnung der beiden Effekte berichtet. Jedoch zeigt sich in Übereinstimmung mit Osterhout et al. 1997, daß die Hirnantwort mit steigender Worthäufigkeit geringer ausfällt. Der hier vorgefundene Effekt drückt sich allerdings schon deutlich früher aus, denn die Osterhout-Gruppe fand ihn erst nach 200ms.

Da dieser Frequenzeffekt erst nach dem Silbeneffekt einsetzt, ist anzunehmen, daß die frequenzabhängige Antwort etwas mit dem Wort als Ganzes (nachdem also das Wort komplett eingelesen wurde) zu tun hat, oder umgekehrt, daß die silbenabhängige Antwort auch physikalisch erklärt werden kann. Die Zahl der Silben oder die geometrischen Bestandteile des Wortes läßt sich messen, auch wenn keine semantische Information zu dem Wort vorhanden ist - ganz im Gegensatz zur Wortfrequenz, die man nur ermitteln kann, wenn man das ganze Wort erkannt hat und eine wie auch immer geartete Antwort produzieren kann. Diese Antwort würde sich beispielsweise aus modularistischer Sicht in einem Tabellen-lookup (ein Nachsehen in einem tabellarisch aufgebauten Lexikon, Hogaboam & Perfetti 1975) ausdrücken, in einer kortikal motivierten Theorie ist die Signalstärke von der Kopplungsstärke eines kortikalen Netzwerks (welches das Wort repräsentiert) abhängig. Werden also häufig die Silben eines Wortes zusammen aktiviert, ergibt sich eine stärkere Verschaltung der Silben-Assemblies, was eine erhöhte Grundaktivität der Assembly zur Folge hat. Damit kann die Aktivierung eines Wortes leichter stattfinden, wenn es häufig aktiviert wurde (hohe Frequenz), da aufgrund der hohen Koppelungsstärke und der vorhandenen Grundaktivität der Schwellenwert leichter erreicht wird. Damit kann sich eine geringere Amplitude in der Messung ergeben, weil ja dabei ausschließlich die zusätzliche Aktivität gemessen wird. Der Wortfrequenz-Effekt kann sich also erst nach der Präsentation aller Silben einstellen, da dann erst die Verbindungen zwischen den Silben zum Tragen kommen können. Dies muß jedoch nicht notwendigerweise in einer zeitlichen Reihenfolge geschehen, da ein- und zweisilbige Wörter parallel auf der Retina abgebildet werden können und somit in der Lage sind, die korrespondierenden Cell Assemblies gleichzeitig zu aktivieren. (Es könnte also sein, daß eine Latenz erst dann einsetzt, wenn die parallelen Aufnahmekapazitäten erschöpft sind

und eine Art Zwischenspeicher benutzt werden muß. Dies würde sich bei mehrsilbigen Wörtern ergeben.)

Die von King & Kutas 1998 und Osterhout et al. 1997 gefundene hohe Korrelation zwischen Wortlänge und Wortfrequenz hat in der vorliegenden Studie keine direkte Entsprechung in der Hirnantwort, denn dann müßten sich die jeweiligen Unterschiede gleichzeitig ergeben. Hinweise auf schnellere Aktivationsprozesse nach hochfrequenten Wörtern ergaben sich in den Untersuchungen der Foci mit mehr als 100fT um den Bereich 280-335ms. Der Unterschied war jedoch nicht signifikant. Dieser Befund bestätigte sich auch in der Global Field Power Analyse. Ein Zusammenhang mit der FSN (frequency sensitive negativity) der King & Kutas (1998) Studie ist möglich, der Unterschied ergibt sich in der vorliegenden Studie deutlich früher und es ist hier keine Konfundierung (durch Länge und Frequenz) gegeben.

Abgesehen von den Silben- und Frequenzeigenschaften der Wörter ergaben sich auch Erkenntnisse über die Verarbeitung der Wortklassen: Die biomagnetischen Korrelate von verschiedenen Wortklassen sind verschieden in ihrer kortikalen Verteilung und in der magnetischen Ausprägung.

Es zeigt sich beispielsweise, daß sich die Funktionswörter 120-160ms nach Präsentation von allen anderen an einem parieto-zentralen Focus unterscheiden, wohingegen die Klasse Both sich von allen anderen rechts-parieto-okzipital unterscheidet. Dies deutet auf eine differenzielle Lokalisation hin, kann aber in dieser Form noch nicht für die kortikale Lokalisation gelten, da Magnetwellen im Gegensatz zum EEG nicht direkt den Ort Aktivität preis geben. Eine genauere Untersuchung der Lokalisationen für die Wortkategorien hätte den Rahmen dieser Arbeit gesprengt, wird jedoch außerhalb der Diplomarbeit mittels Minimum Norm oder einfacheren Mehr-Dipolmodellen erfolgen.

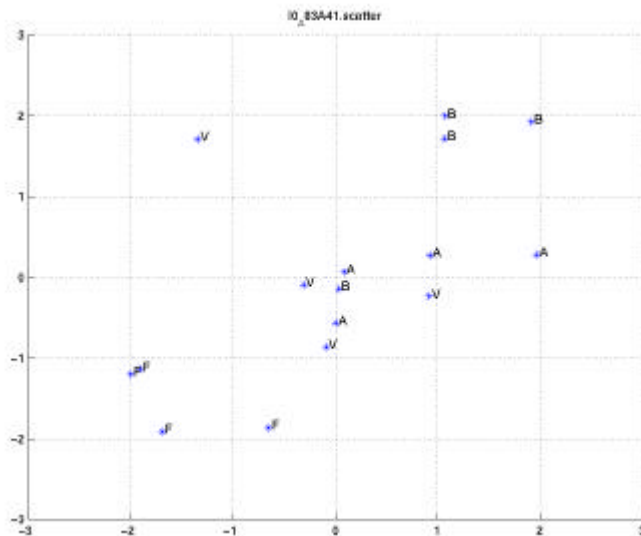
Genauerer lässt sich hingegen über die zeitlichen und quantitativen Unterschiede in den Wortklassen sagen. Die Hirnantworten auf Wörter der selben lexikalischen Kategorie (Nomen) unterscheiden sich sehr früh (90-120ms nach Stimulus-Onset), wenn sie Wortklassen von verschiedenem Modalitätsgrad angehören. Die Kategorie Both, welche aus Nomen besteht und sowohl visuelle als auch motorische Assoziationen vorweist, hebt sich früh von der Kategorie Visual ab, gleichwohl auch diese ausschließlich aus Nomen besteht, jedoch nur visuelle Assoziationen hervorruft.

Offensichtlich geschieht dies bevor die lexikalische Unterscheidung nach etwa 120ms einsetzt. Dies ist für eine Theorie, die Lexika definiert (z.B. Ellis & Young 1988), nicht einfach zu erklären, da die Kategorie Both im gleichen Lexikon wie die Kategorie Visual verarbeitet werden müsste. Nach dieser Theorie würde zunächst die eingehende visuelle Information in einen Code (Zuordnung zu grammatischen Kategorien) übersetzt werden, der dann der Semantikverarbeitung weitergereicht wird. Dort würde dann die Unterscheidung zwischen der Kategorie Both und der Kategorie Visual vorgenommen. Daher sollte nach einer solchen Theorie diese Unterscheidung frühestens nach der Unterscheidung der lexikalischen Klassen stattfinden. Die Unterscheidung aller lexikalischen Klassen geschieht hier jedoch in der Zeit von 120-160ms. Es scheint also die umgekehrte Reihenfolge aufzutreten: Ein semantischer Unterschied, dann die lexikalische Unterscheidung. Um eine Erklärung in diesem Modell zu schaffen, müsste die Semantikstufe also vor die Lexikonstufe gestellt werden.

Eine kortikal motivierte Theorie hatte die Vorhersage getroffen, daß die Hirnantwort abhängig von der Größe der repräsentierenden Netze ist und damit, daß ein Wort, welches Assoziationen in mehreren Modalitäten umfaßt, auch stärkere Signale liefert. Zur Untersuchung dieser Vorhersage wurden zwei signalstarke Kanäle (links und rechts

parietal) aus den Foci der Zeitscheibe i0 ausgewählt und ihr Mittel über den Zeitraum 90-120ms berechnet. Die Werte wurden in z-scores umgerechnet, um den Einfluß von Wortlänge und Worthäufigkeit zu minimieren und in einem Scatterplot gegeneinander abgetragen. Dies geschah für alle sechzehn Wörter:

ABBILDUNG. SCATTERPLOT FÜR 2 KANÄLE, RECHTS UND LINKS PARIETAL IM



FENSTER I0.

Wie zu sehen, sind die Feldstärken für die Funktionswörter die geringsten und die multimodalen Wörter bilden die Fraktion mit den höchsten Feldstärken. Dies scheint die obengenannte Hypothese sogar bereits 100ms nach Stimulus-onset zu bestärken.

Auch der im vorigen Kapitel vorgestellte Regressionsplot, welcher zusätzlich zu dieser Erkenntnis noch zeigt, daß kognitive Einschätzung der Wörter stark mit der biomagnetischen Antwort korreliert, untermauert die Hypothese.

Nun könnte aber der Vorwand eingebracht werden, daß das, was gemessen wurde nicht der Stärke der Assoziationen entspricht, sondern dem Grad an Imageability (Kounios & Holcomb 1994) oder umgekehrt, dem Grad an Abstraktheit. Wenn dies zuträfe, dann sollten die Bewertungen im Fragebogenteil für die Klasse Both den geringsten Wert an Abstraktheit liefern. Was sich aber zeigt ist, daß sich die Klasse Both in der Bewertung

nicht von den Klassen Visual und Action unterscheidet. Damit kann der gemessene Effekt nicht mit dem Abstraktheitsgrad begründet werden.

Der Befund von Osterhout et al. 1997, daß Funktionswörter eine Negativität im Bereich 400-700ms aufzeigen und Inhaltswörter eine späte positive Komponente, ließe sich in diesem Kontext umformulieren in die Frage, ob sich die Wortklassen im Zeitbereich 400-700ms unterscheiden. Obwohl dieser Effekt bei zufälliger Reihenfolge der Wörter deutlich kleiner ausfiel, ist er hier sowohl in der Global Field Power Analyse vorhanden, wie auch in der Analyse der stärksten Foci.

Warum gibt es einen semantischen Effekt auf die hirmagnetische Antwort, bevor die Worthäufigkeit ins Spiel kommt?

Beide Effekte sollten doch nach der kortikal motivierten Theorie erst einsetzen, wenn das Wort vollständig präsent ist. Als Erklärungsversuch könnte man zwei Aspekte der Cell Assemblies betrachten: die Kopplungsstärke und die Größe des Netzes.

Während die Kopplungsstärke durch die Häufigkeit der Präsentation des Wortes bestimmt wird, ist die kortikale Distribution und die Größe des Netzes von den Modalitäten der Bedeutung abhängig. Es könnte also sein, daß Aktivität, die sich in einem großen Netz (multimodal, Both) ausbreitet, noch unterschwellig (d.h. die Cell Assembly zündet noch nicht) aber bereits meßbar ist. Erst bei vollständiger Zündung zeigt sich dann der Frequenzeffekt.

Abschließend können zwei Befunde formuliert werden:

- Die biomagnetische Unterscheidung von Wörtern aufgrund ihrer Silbenzahl (100ms) geschieht vor ihrer Unterscheidung aufgrund ihrer Worthäufigkeiten (150ms).
- Die Stärke der Hirnantwort auf Wörter korreliert mit ihrer Assoziationsstärke bereits nach 100ms.

Die Frage nach der Relevanz einer Einzelfallstudie für wissenschaftliche Erkenntnisgewinnung kann positiv beantwortet werden. Die gewonnenen Erkenntnisse zeigen, daß mindestens eine Person Hirneigenschaften hat, welche diese Phänomene produzieren. In der neuropsychologischen Literatur hat es sich etabliert, Einzelfälle zu reportieren, da sie Auskunft darüber geben

10. Das Netz.

10.1. Einleitung

Teuvo Kohonen (1982) hat eine Architektur für ein Neuronales Netz vorgeschlagen, das bestimmte Eigenschaften des Kortex modellieren soll. Ausgangspunkt war für ihn die Beobachtung, daß benachbarte Stellen auf der Haut eines Menschen benachbarte Stellen in seinem Kortex aktivieren: die Somatotopie. Der Überbegriff für dieses Konzept ist die Topologieerhaltung, die in verschiedenen Modalitäten der Sensorik wie auch der Motorik auftritt. Für die vorliegende Arbeit von besonderem Interesse ist die tonotope Organisation der akustischen Kortexes. Im primären auditorischen Kortex sind beispielsweise ähnliche Frequenzen in Nachbarschaft repräsentiert.

Hiervon inspiriert, konstruierte Kohonen (1984) ein Netz, welches auf verschiedene Frequenzbänder der akustischen, menschlichen Sprache trainiert wurde. Die Vorgehensweise ist auch hier von großer Relevanz. Frequenzen von 200-5000Hz wurden in 15 einzelne Bänder eingeteilt. Die Energie dieser Bänder wurde für Zeitfenster der Länge 25,6ms berechnet. Dann wurde das Fenster um 10ms verschoben und erneut wurde die Energie der Bänder berechnet. Der Vorgang wird fortgeführt bis zum Ende des Signals - völlig unabhängig von Wort- oder Lautgrenzen. Man kann die Energie eines Bands als

sogenanntes Feature betrachten, ein Merkmal, dessen Vorhandensein oder seine Ausprägung eine Eigenschaft des zugrundeliegenden Lautes ist. Kombinationen von Features erhöhen die Sicherheit für eine Zuordnung eines Signalstücks zu einem Laut und werden Feature-Vektoren genannt.

Der Prozeß der Feature-Generierung erzeugt durch die zeitliche Überlappung der Fenster Daten, die sich von Zeitpunkt zu Zeitpunkt nur graduell unterscheiden, denn die redundante Informationsdarbietung modelliert einen Kontext. Es können somit später auch Formantenverläufe, wie sie bei Diphthongen auftreten, repräsentiert werden.

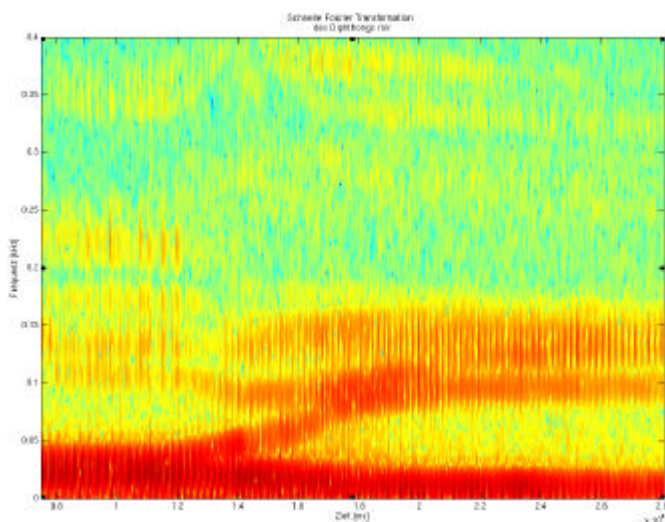
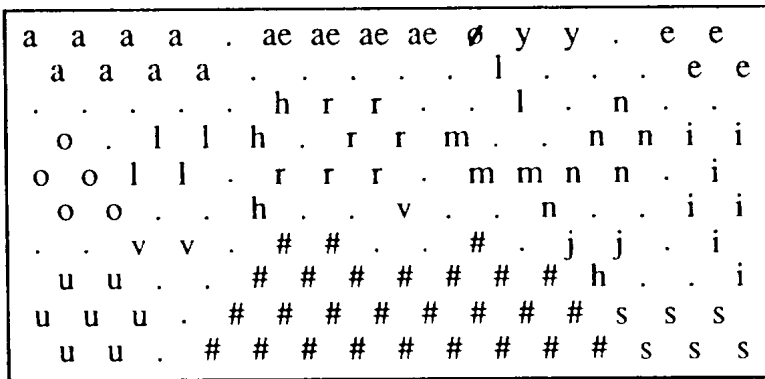


ABBILDUNG. FORMANTENVERLAUF DES DIPHTHONGS /OI/

Kohonen benutzte zur Repräsentation der 21 finnischen Phoneme ein 2 dimensionales Netz mit 10x15 Neuronen, was zur Folge hat, daß jedes Phonem in verschiedenen, sich ähnelnden Formen gespeichert werden kann. Ein Neuron in diesem Netz steht für einen Punkt in einem n-dimensionalen Raum, wobei n die Anzahl der Eingangswerte (Features), also hier der 15 Bänder ist. Punkte, die in diesem 15-dimensionalen Raum benachbart sind, sind also ähnliche Lautstrukturen (oder allgemeiner: ähnliche Featurestrukturen) und werden in der zwei-dimensionalen Karte in Nachbarschaft gespeichert.



= /k, p, t/

ABBILDUNG: KOHONENKARTE DER FINNISCHEN LAUTE

Die Architektur des Netzes kann angegeben werden in der Anzahl der Neuronen (10x15=150) und in der Stärke der Dimensionsreduktion pro Präsentation (15:2). Diese beiden Faktoren sind selbstverständlich von einander abhängig, denn wenn die Zahl der Dimensionen des Eingangsvektors sehr hoch ist und die Dimensionszahl der Karte klein, dann muß mehr Speicherplatz (Anzahl der Neuronen) investiert werden, um die adäquate Abbildung zu gewährleisten - stets vorausgesetzt, daß jede weitere Dimension zusätzliche Information beiträgt und daß das Training optimal ist. Weitere Dimensionen im Eingangsvektor könnten jedoch auch Redundanz bergen oder das Signal-Rausch-Verhältnis verschlechtern, welches Auswirkungen auf die Erkennungsrate hat.

Mit Hilfe dieser Überlegungen konnte eine erste Schätzung einer Architektur für die anstehende Aufgabe gemacht werden. In Kohonens Fall erstreckte sich die Repräsentation des Phonems /a/ über acht Neuronen in der Karte. Wenn die Daten des Signals nacheinander präsentiert und im Verlauf alle diese acht Neuronen aktiv werden, dann kann das Netz eine Signaldauer von rund 100ms (25,6ms + 7 * 10ms = 95,6ms) repräsentieren. Diese Signaldauer steht dann für Signalabschnitte, die sich nicht wiederholen, was bedeutet, daß jedes Neuron einen relativ distinkten Signalabschnitt modelliert. Würde sich

die Formantstruktur im Verlauf mehrerer Zeitabschnitte nicht ändern, so würde immer das gleiche Neuron aktiv.

Wenn man für die Repräsentation der Wörter den Zeitraum von 100ms-800ms nach Stimulus-Onset für notwendig erachtet, die Zeitfenster mit der Länge 50ms und einer Verschiebung von 30ms festlegt, so erhält man 22 Fenster für das Signal. Wird jedem Fenster ein Neuron zugeordnet, so kann würde man mindestens $22 \cdot 16 = 352$ Neuronen benötigen, um alle Wörter zu repräsentieren. Es wurde also als vorläufige Annahme ein Netz von 400 Neuronen (20x20) konstruiert. Generell ist es lediglich eine Frage des Rechenaufwandes ein Netz größer zu machen, wenn die Varianz in den Daten klein ist. Es verbessert die Erkennungsrate nicht. Wenn jedoch die Varianz groß ist und der Speicherplatz im Netz zu hoch, dann kann es sein, daß das Netz sich zu sehr auf die Trainingsdaten spezialisiert und dann schlechte Erkennungsraten auf dem Testdatensatz zeigt:

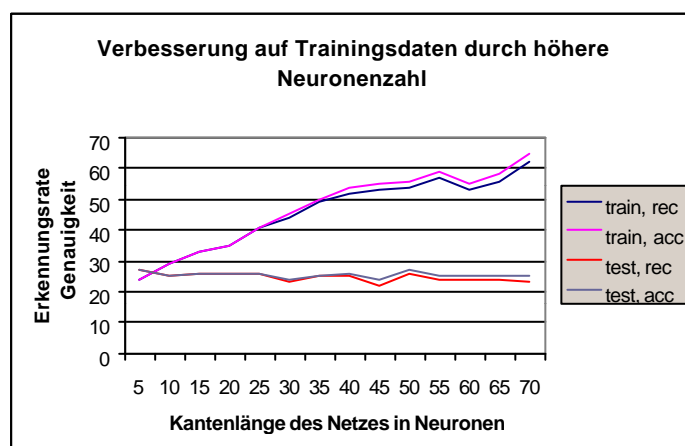


ABBILDUNG. DIE SPEICHERKAPAZITÄT DES NETZES ERHÖHT DIE LEISTUNG AUF TRAININGSDATEN

Es kann bei großer Varianz von Vorteil sein, das Netz in der Speicherkapazität zu begrenzen, um es zur Abstraktion zu zwingen. Dies würde die Leistung auf Trainingsdaten

verschlechtern, jedoch die Leistung auf Testdaten verbessern. Experimente hierzu wurden gemacht und werden weiter unten ausführlicher vorgestellt.

10.2. Daten.

Die Daten für das Neuronale Netz wurden zunächst sehr einfach generiert. Es wurden sieben Foci aus dem statistischen Teil dieser Arbeit genommen, von denen ja bereits bekannt war, daß sie Unterschiede in den Wörtern sehr gut einkreisen. Zusätzlich wurde auch ein zentro-zentraler Focus (7) gewählt:

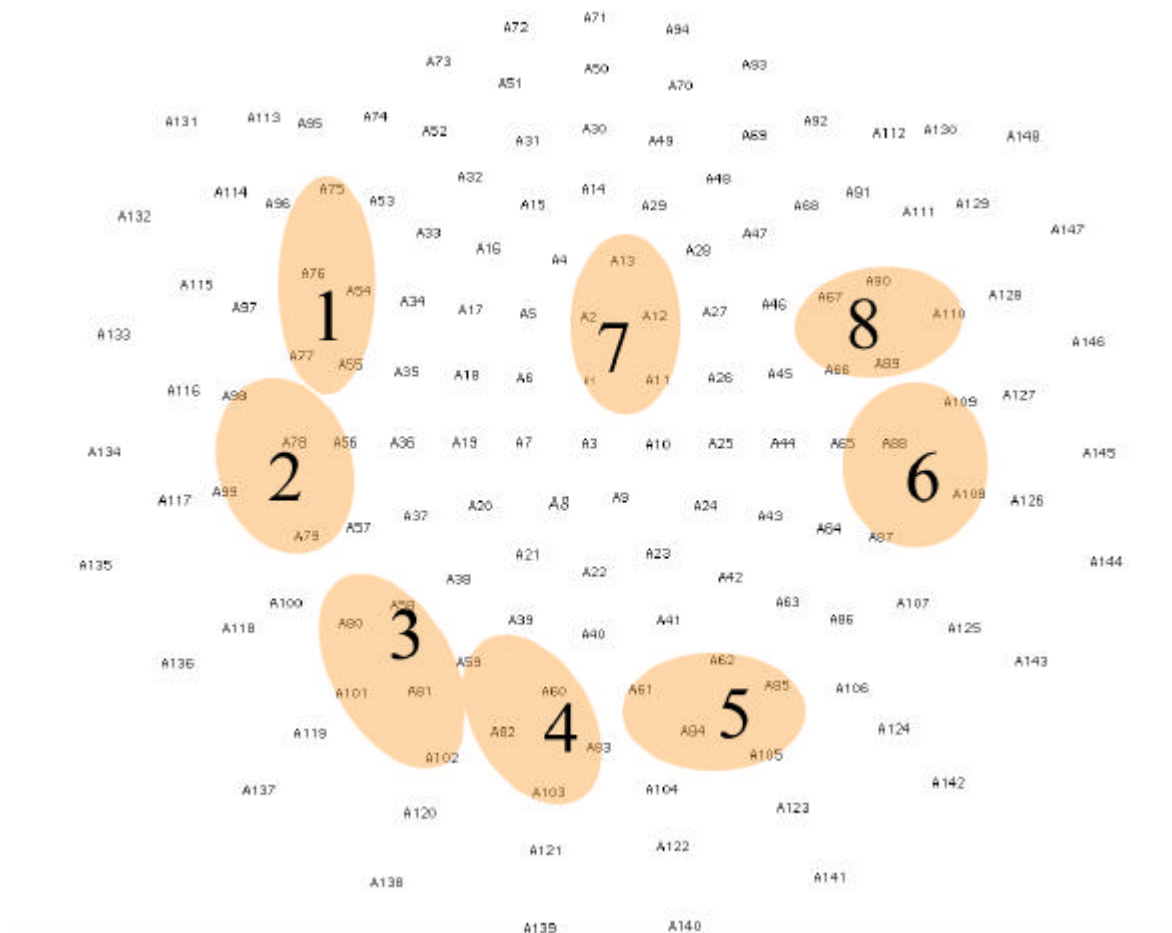


ABBILDUNG. DIE FOCI, DIE ZUM TRAINIEREN DES NEURONALEN NETZES BENUTZT WURDEN.

Für diese 8 Foci wurden für alle obengenannten 22 Zeitfenster der RMS berechnet und lieferten somit die Eingangsdaten für das Netz. Die Datenmenge wurde im Verhältnis 2:1 in Trainings- und Testdaten aufgeteilt.

10.3. Training.

Sodann wurde ein universelles Netz geschrieben, welches mit diesen Daten trainiert werden konnte. Ein Kohonen-Netz arbeitet nach der nearest-neighbour-Methode, was bedeutet, daß dasjenige Neuron am meisten lernen darf, welches als Punkt im 8-dimensionalen Raum dem Mustervektor (dem momentanen Trainingsbeispiel) am nächsten ist. Dazu muß dieses Neuron ermittelt werden, was über die euklidische Distanz berechnet wird. Ist dieses Neuron gefunden, wird die Distanz zum Mustervektor über eine Lernrate ϵ verringert - das Neuron lernt. Um der Topologie-Erhaltung Rechnung zu tragen, müssen die Nachbarneuronen in der zweidimensionalen Karte zu einem gewissen Maß mitlernen. Dies bedeutet, daß auch sie ihre Distanz zum Mustervektor verringern, jedoch nicht im gleichen Ausmaß wie der nearest neighbour. Die Rate, mit der Nachbarneuronen mitlernen, wird im Folgenden σ genannt. Die beiden Faktoren ϵ und σ definieren eine Gaussglocke:

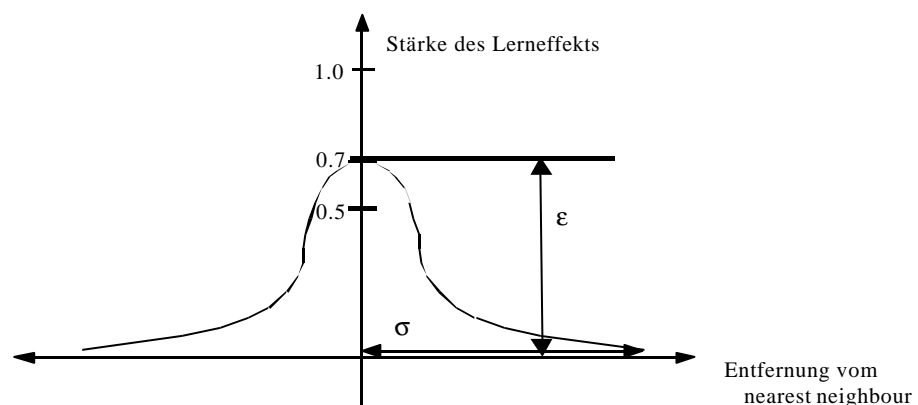


ABBILDUNG. GAUSSGLOCKE DER LERNFUNKTION.

Die Gaussglocke beschreibt das Maß, in wie weit sich das Netz den präsentierten Daten anpaßt und ob es die Daten eher speichern soll (großes ε , kleines σ) oder ob es die Daten abstrahieren soll, in dem viele Beispiele dazu benutzt werden, viele Neuronen zu verändern (kleines ε , großes σ). Wenn ε und σ beide groß sind, wird eine Abstraktion mit wenigen Beispielen vorgenommen bzw. die abstrakte Repräsentation wird von Beispiel zu Beispiel stark geändert. Dennoch ist diese Kombination in frühen Stadien sinnvoll, da das Netz im Grundzustand mit Zufallswerten belegt ist und daher keine Nachbarschaftsbeziehung besteht. So sind also beide Werte am Anfang des Trainings relativ hoch, um grobe Strukturen auszubilden und am Ende relativ klein, um die Daten detailliert zu repräsentieren.

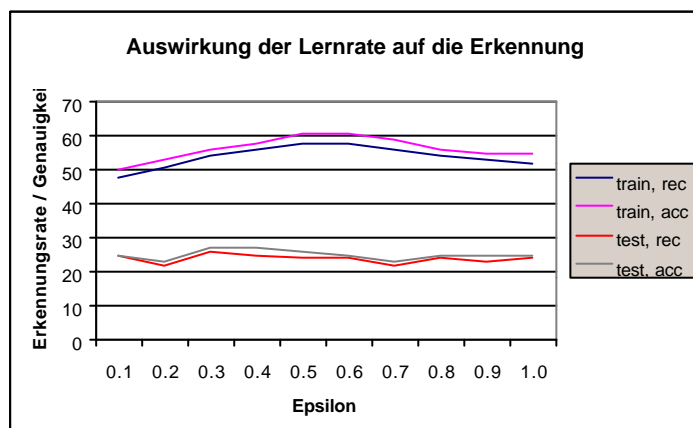


ABBILDUNG. DIE BESTE LEISTUNG AUF TESTDATEN WIRD BEI $\varepsilon=0.3$ ERZIELT, AUF TRAININGSDATEN JEDOCH BEI $\varepsilon=0.6$.

10.4. Generierung der Karte.

Nach Abschluß des Trainings beginnt ein weiterer Durchgang durch die Trainingsdaten, der nun dazu dient, für jedes Neuron dasjenige Wort (analog: diejenige Wortklasse) zu finden, das es am besten repräsentiert. Jedes Neuron bekommt für 16 Zähler für die einzelnen Wörter und 4 zusätzlich für die Wortklassen. Dann wird für jeden Datenpunkt

der nearest neighbour errechnet und der Zähler dieses Neurons für das präsentierte Wort wird um 1 erhöht.

Man erhält am Ende bei jedem Neuron für alle Wörter eine Zahl, wie oft das Neuron am besten dem Wort entsprach. Das Programm gibt sodann eine Karte aus, die für die Wortklassen jeweils das Wort mit der höchsten Belegung anzeigt.

AAB++AABFA
BF'BBF+_ABB
FBBBAV+VA+
BAFAAAA+VB
B_ABBV+VVB
AVBFABFAAB
FFABFVBA+A
B_A+VBAAB+
F_FVVABFVF
AAFBBAVABF

ABBILDUNG. EINE KOHONENKARTE (10X10) FÜR DIE WORTKLASSEN

10.5. Der Erkennungsvorgang.

Um eine Erkennung der Daten vorzunehmen, muß ermittelt werden, welchem Wort die 22 Datenpunkte entsprechen könnten. Hierzu werden die 22 Datenpunkte dem Netz präsentiert und die Häufigkeiten der einzelnen Wörter (gewonnen aus der Häufigkeit des jeweiligen nearest neighbour) aufaddiert. Das Wort mit dem höchsten Wert wird dann als "erkannt" ausgegeben. Im Anschluß wird dann noch ermittelt, ob das Erkannte dem Präsentierten entspricht, was als hit gewertet wird. Mit den Werten total (wie oft kam das Präsentierte vor = hits + misses) und swacc (wie oft gab das Netz das Wort als erkannt aus, unabhängig davon, ob richtig oder nicht = hits + false positives) lassen sich Erkennungsrate (recognition rate) und Genauigkeit (accuracy) berechnen:

Erkennungsrate = hit / total (Erwartungswerte: 4 Wortklassen 25%, 16 Wörter 6,25%)

Genauigkeit = hit / swacc (Erwartungswerte: 4 Wortklassen 25%, 16 Wörter 6,25%)

10.6. Erstes Ergebnis.

Ein erster Lauf des Netzes (20x20 Neuronen, alle 8 Foci, 22 Zeitfenster = 100-780ms) über die Daten lieferte folgendes Ergebnis:

```
BFFB+F_FA+_+BFAFVBF+
V_____+F_A_____+V_+B
____+__+____+____V_B+
B+_____F+VB_+_++_VV
F_+++_B+_++VA++_V++V
B+_AA_F_____+++_V
F+_F_+BB_++F___A_VV
BB_____+_++B
_A+_+_B+_+_+____+VV
B+BVBAB+_+___ABA+_VB
BVA++AA_V+____+____+A
BV_BB_F____+B_+_+++A
B+F+ABF____+B_+_+_+_
B_F_A++_A_A+_VV_B
F++_A_+_BAVAVA__+
+++AB_+++++_VFB+BF_A
B_AF+_+_B_____A_A++AV
+_AB+B++++_A_+++_V
V_B_+_+____+A+F__B
AAAB___F+FBVBABA+F_A
```

ABBILDUNG. DAS ERSTE NETZ.

```
computing scores
... for training data
rates for single words:
train: L1   F: 21% (20%) A: 9% (19%) B: 42% (14%) V: 13% (24%)
train: H1   F: 49% (9%) A: 41% (27%) B: 11% (20%) V: 13% (33%)
train: L2   F: 9% (20%) A: 8% (19%) B: 10% (31%) V: 12% (26%)
train: H2   F: 7% (23%) A: 24% (18%) B: 5% (21%) V: 6% (28%)
train: rates for classes: F: 46% (30%) A: 39% (40%) B: 35% (42%) V: 21% (46%)
train: overall classes: 35% (37%) overall words: 17% (22%)

... for testing data
rates for single words:
test:  L1   F: 1% (2%) A: 1% (3%) B: 28% (8%) V: 5% (11%)
test:  H1   F: 33% (7%) A: 13% (6%) B: 1% (4%) V: 3% (10%)
test:  L2   F: 4% (7%) A: 0% (0%) B: 1% (5%) V: 2% (5%)
test:  H2   F: 0% (0%) A: 13% (7%) B: 4% (7%) V: 0% (0%)
test: rates for classes: F: 38% (27%) A: 30% (26%) B: 18% (24%) V: 12% (26%)
test: overall classes: 25% (26%) overall words: 6% (5%)
```

10.7. Diskussion.

Wie zu sehen, sind die Ergebnisse relativ enttäuschend und am Zufallsniveau. Die Erkennungsraten für die Funktionswörter sind die höchsten, während die visuellen Nomen am schlechtesten Abschneiden.

Es ist auch keine Anordnung der Wörter in Bereichen zu beobachten, so wie es bei der phonetischen Karte bei Kohonen der Fall war. Dies könnte in einer relativ starken Ähnlichkeit der Hirnantworten in ihrer Grobstruktur begründet sein. Im Gegensatz zu akustischen Lauten, die in einer Fouriertransformation sehr starke Unterschiede vorweisen, scheinen Wörter - zumindest bei einer Berechnung der RMS eines Focus - sich stark zu ähneln. Damit ergibt sich ein Problem der Frequenzauflösung. Kohonens Aufteilung in Frequenzbänder der Fouriertransformation ließ machte dem Netz eine detailliertere Repräsentation der Frequenzeigenschaften der Laute möglich. Diese Möglichkeit hatte das vorliegende Netz nicht.

Es ist auch möglich, daß bei dieser RMS-Analyse Werte aus verschiedenen Zeitbereichen gleich ausfallen und daher zusammen gespeichert werden. Diese Verschiedenheit in der Zeit könnte jedoch ein Charakteristikum für die Wörter oder Wortklassen darstellen: Das Netz modelliert vielleicht nicht genug Kontext.

10.8. Weitere Architekturen.

Daher wurde versuchsweise eine interne Uhr eingeführt, um eine Zeitmodellierung zu realisieren. Diese Uhr gestaltete sich in einem weiteren Feature Vektor, welcher für jedes präsentierte Fenster einer Wortableitung erhöht wurde. Damit wurden also alle anderen Werte immer zusammen mit einem Zeitpunkt nach Stimulus-onset gelernt.

Um das Problem der Frequenzauflösung zu adressieren, wurde das Zeitfenster auf 16ms gekürzt und jeweils um die halbe Fensterlänge über das Signal geschoben. Wenn jeweils 16ms eines Signals präsentiert werden, und die sinusförmige Welle eine Dauer von 32ms hat, so wird im besten Falle genau eine positive oder negative Flanke des Signals als RMS berechnet:

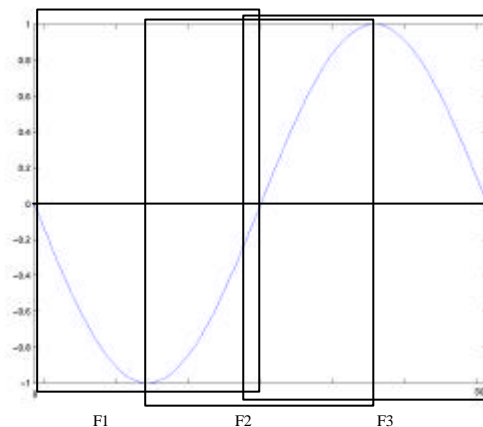


ABBILDUNG.VERSCHOBENE FENSTER ERFASSEN GAMMAWELLEN VERSCHIEDEN GUT

Wellen von 30ms und mehr bezeichnet man als Gamma-Wellen und Forschungen von Lutzenberger et al. (1994) haben ergeben, daß einige Charakteristika der Wortverarbeitung im Gamma-Band geschehen.

Darüber hinaus hat ein Screening der Ableitungen von den einzelnen Wörtern ergeben, daß sie sich in einen späten Zeitbereich links zentral und links parietal stark unterscheiden:

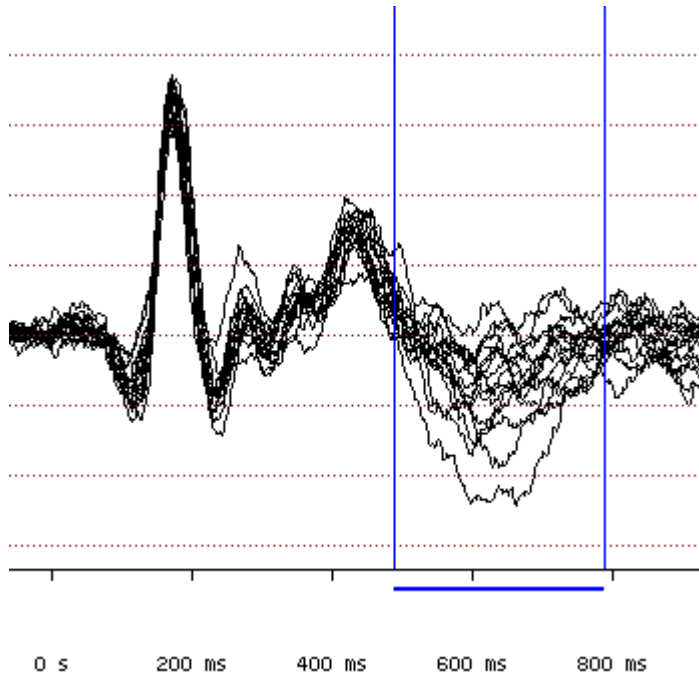


ABBILDUNG. LINKS PARIETAL UNTERSCHIEDEN SICH DIE EINZELNEN WÖRTER SPÄT.

Damit ergaben sich einige mögliche neue Architekturen von welchen folgende weiter erforscht wurden:

- interne Uhr auf ursprünglichen Daten
- ursprüngliche Daten (22 Fenster: 100-780ms, Länge 50ms, Shift 30ms)
- ursprüngliche Daten als Gamma (48 Fenster: 100-500ms, Länge 16ms, Shift 8ms)
- späte Daten, neue Foci (28 Fenster: 490-790ms, Länge 30ms, Shift 10ms)
- späte Daten als Gamma (40 Fenster: 490-790ms, Länge 15ms, Shift 7.5ms)

10.9. Endergebnisse.

Unter den vielen Trainings- und Testläufen, die liefen sollen im Folgenden die besten Ergebnisse vorgestellt werden (Versuche mit der internen Uhr steigerten die Ergebnisse nicht):

```

ursprüngliche Daten:
bester Wortklassenerkennung:
test:  overall classes: 26% (28%)  overall words:  6% ( 8%)

bester Einzelworterkennung:
test:  overall classes: 24% (25%)  overall words:  8% ( 8%)

ursprüngliche Daten als Gamma:
bester Wortklassenerkennung:
test:  overall classes: 29% (29%)  overall words:  6% (13%)

bester Einzelworterkennung:
test:  overall classes: 26% (26%)  overall words:  7% (12%)
(test:  overall classes: 26% (26%)  overall words:  7% (21%))

späte Daten:
bester Wortklassenerkennung und bester Einzelworterkennung:
test:  overall classes: 29% (30%)  overall words:  8% ( 9%)

späte Daten als Gamma:
bester Wortklassenerkennung:
test:  overall classes: 29% (29%)  overall words:  6% (10%)

bester Einzelworterkennung:
test:  overall classes: 26% (26%)  overall words:  7% (12%)

```

10.10. Diskussion.

Obwohl eine genaue statistische Absicherung bislang nicht stattfand, scheinen die Erkennungsraten über dem Erwartungswert zu liegen. Es zeigt sich dabei, daß ein Erkennung, welcher ein Fenster benutzt, das eine halbe 30Hz Schwingung lang ist, am besten abschneidet. Die Gamma-Erkennung sind daher anscheinend besser auf die Worterkennungsaufgabe zu geschneitten. Umgekehrt scheinen die Erkennung mit der größeren Fensterlänge besser für die Wortklassenerkennung zu funktionieren. Dabei zeigt sich der Erkennung über den späten Zeitbereich als leicht überlegen.

Die meisten Ergebnisse sind von den Netzen mehrfach in verschiedenen Lernkonfigurationen erbracht worden, so daß man sagen kann, daß nicht zufällig eine Netzstruktur gefunden wurde, die die Testdaten am besten erkennt. Das einzige Ergebnis,

welches nur einmal auftrat, ist in Klammern angegeben. Dieses daher vermutlich ein Produkt einer zufällig sehr guten Konfiguration des Netzes.

Auffällig ist auch, daß die Erkennungsraten für die Wortklassen zwar deutlich über 25% liegen, aber da vier Wörter einer Klasse angehören, ist die durchschnittliche Erkennungsrate eines Vertreter einer Klasse nur ein Viertel so hoch. Damit ergibt sich bei einer Wortklassenerkennungsrate von 29% eine Rate von 7,25% pro Vertreter. Die Erkennung der einzelnen Wörter ist aber deutlich besser. Dies kann mit einer hohen Interklassenvarianz erklärt werden. Die Vertreter einer Klasse sind oft mit den Vertretern einer anderen Klasse verwechselt worden und daher ist die Erkennungsrate für Wortklassen relativ niedrig. Wie im Statistikeil dieser Arbeit zu erfahren, sind die Unterschiede einsilbigen und zweisilbigen Wörtern sehr stark, daher ist es möglich, daß zweisilbige Wörter (oder einsilbige) einer Klasse mit zweisilbigen (bzw. einsilbigen) einer anderen Klasse verwechselt werden. In der früheren Diskussion (Kapitel 10.7) wurde dies auch Grund für das nicht sehr ausgeprägte Zusammenhängen der Repräsentationen der Wortklassen angesehen. Daher müßte in folgenden Arbeiten eine Repräsentation für die Werte gefunden werden, die Wortklassen unspezifische Eigenschaften nicht in der Karte abbildet.

Weitere Erkennen sollten sich auch den andeutungsweisen Vorteil der Gamma-Erkennen zu nutze machen und die Rohdaten mit einem Bandpassfilter bearbeiten, der ausschließlich das Gamma-Band in den Daten erhält. Weitere Möglichkeiten wären in Algorithmen zu sehen, die zunächst eine grobe Lokalisierung vornehmen und die Daten dann erst zu Erkennung weiterreichen. Hier wäre das Minimum Norm Verfahren eine interessante Perspektive. Die Daten könnten auch einer principle component Analyse unterzogen werden, um einen genaueren Einblick auf die Struktur der Daten zu ermöglichen und so die Erkennungsrate erheblich zu steigern.

11. Die Literatur.

- Petersen & Fietz, 1993: The processing of single words studied with positron emission tomography. *Annu. Rev. Neurosci.*, 16 509-530
- Pulvermüller, F., Lutzenberger, W., Birbaumer, N., 1995: Electrocortical distinction of vocabulary types. *Electroencephalography and clinical Neurophysiology*
- Pulvermüller, F., Mohr, B., 1996: The Concept of Transcortical Cell Assemblies: a Key to the Understanding of Cortical Lateralization and Interhemispheric Interaction. *Neuroscience and Biobehavioral Reviews*, Vol. 20, No. 4, pp. 557-566.
- Pulvermüller, F., Preissl, H., 1994: Explaining aphasia in neurobiological terms. *J. Neuroling.* 8: pp. 75-81
- W. Lutzenberger, F. Pulvermüller, N. Birbaumer: „Words and pseudowords elicit distinct patterns of 30-Hz EEG responses in humans“ *Neuroscience Letters* 1994, 176 pp. 115-118
- Friederici, D., Pfeifer, E., Hahne, A., 1993: „Event-related brain potentials during natural speech processing: effects of semantics, morphological and syntactic violations“ *Cognitive Brain Research*, 1 pp.183-192
- Osterhout, L., Berwick, M., McKinnon, R. 1997: „Brain potentials elicited by words: word length and frequency predict the latency of an early negativity“ *Biological Psychology*, 46 pp. 143-168
- Kutas, M. & Hillyard, S.A., 1980: „Reading senseless sentences: brain potentials reflect semantic incongruity“ *Science*, 207 pp. 203-205
- Kutas, M. & Hillyard, S.A., 1983: „Event-related brain potentials to grammatical errors and semantic anomalies.“ *Mem. Cogn.*, 11 539-550
- Caplan, D. 1992: *Language: Structure, processing, and disorders*. Cambridge, MA: MIT Press.
- Fodor, J. A. 1983: *The modularity of mind*. Cambridge, MA: MIT Press.
- Frazier, L. 1990: Exploring the architecture of the language processing system. In G.T.M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 409-433). Cambridge, MA: MIT Press.
- Morton, J. 1979: Facilitation in word recognition: Experiments causing change in logogen model. In P.A. Kolers, M.E. Wrolstad & M. Bouma (Eds.), *Processing of visible language* (pp. 259-268). New York: Plenum.
- Ellis, R. & Young, A.W., 1988: *Human cognitive neuropsychology*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Petersen, S.E., Fox, P.T., Posner, M.I., Mintun, M.E. & Raichle, J. 1989: Positron emission tomographic studies of the processing of single words.
- Monsell, S. 1987: On the relation between lexical input and output pathways for speech. In A. Allport, D. Mackay, W. Prinz, & E. Sheerer (Eds.), *Language perception and production: Shared mechanisms in listening, speaking, reading, and writing*. (pp. 273-311). London: Academic Press.
- Glucksberg, S., Kreuz, R.J. & Rho, S.H., 1986: Context can constrain lexical access: Implications for models of language comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 323-335.
- Hogaboam, T.W. & Perfetti, C.A., 1975: Lexical ambiguity and sentence comprehension: the common sense effect. *Journal of Verbal Learning and Verbal Behaviour*, 14, 265-275.
- Onifer, W. & Swinney D.A., 1981: Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory and Cognition*, 9, 225-236.
- Shannon, C. E., 1948: A mathematical theory of communication. *Bell System Technical Journal*, vol. 27, pp. 379-423, 623-656.
- Kohonen, T., 1982: Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, vol. 43, pp. 59-69.
- Kohonen, T., Makisra, K., Saramaki, T., 1984: Phonotopic maps – insightful representation of phonological features for speech recognition. In: *Proc. of the Seventh Int. Conf. on Pattern Recognition*, Montreal, Canada, pp. 182-185.
- Kohonen, T., 1988: *Self-Organization and Associative Memory*. 2nd Edition, Springer Verlag, Berlin.
- Zipf, G. K., 1935: *Psycho-Biology of Language*. Houghton Mifflin, Boston.
- King, J. W., Kutas, M., 1998: Neural Plasticity in the dynamics of human visual word recognition, *Neuroscience Letters*, pp. 61-64.

Neville, H.J., Mills, D.L., Lawson, D.S., 1992: Fractionating language: different neural subsystems with different sensitive periods, *Cerebral Cortex*, 2 (3), pp. 244-258.

Francis, W. N., Kucera, H. 1982: *Computational Analysis of English Usage: Lexicon and Grammar*. . Houghton Mifflin, Boston.

Garrett, M. F., 1982: A perspective on research in language production. In J. Mehler, E.C.T. Walker and M.F. Garrett (Eds.): *Perspectives on mental representation*. Hillsdale, NJ: Erlbaum.

12. Das Abbildungsverzeichnis.

ABBILDUNG. EIN KLEINES, MODULARES SPRACHSYSTEM.....	9
ABBILDUNG: DAS ELLIS & YOUNG MODELL VON 1988.....	12
ABBILDUNG. EIN FORMALES NEURON (E EINGANGSWERT, W GEWICHT, A AUSGANGSWERT, F AUSGABEFUNKTION).....	13
ABBILDUNG. AKTIVATIONEN FÜR VERSCHIEDENE WORTKLASSEN (F FUNKTIONSWÖRTER, A VERBEN MIT MOTORISCHER ASSOZIATION, V NOMEN MIT VISUELLER ASSOZIATION, B NOMEN MIT VISUELLER UND MOTORISCHER ASSOZIATION).....	17
ABBILDUNG. CELL ASSEMBLIES ALS REPRÄSENTATIONEN FÜR HOCH- UND NIEDERFREQUENTE WÖRTER. (DUNKEL: GRUNDAKTIVATION, HELL: ZUSÄTZLICHE AKTIVATION BEI ZÜNDUNG).....	20
ABBILDUNG. EINSCHÄTZUNGEN, MOTORISCH.....	27
ABBILDUNG. EINSCHÄTZUNGEN, VISUELL.....	28
ABBILDUNG. EINSCHÄTZUNGEN, ABSTRAKT.....	28
ABBILDUNG. DIE BEWERTUNGEN VON VISUELL UND MOTORISCH FÜR DIE WORTKLASSEN.....	29
ABBILDUNG. WORTKLASSENUNTERSCHIEDE: BOTH (1. V. U.) UND FUNCTION (2. V. U.) HEBEN SICH FRÜHEITIG VON VISUAL (3. V. U.) UND ACTION (4. V. U.) AB.....	34
ABBILDUNG. WORTKLASSENUNTERSCHIEDE: ACTION (1. V. U.) HEBT SICH VON DEN DREI ANDEREN KATEGORIEN AB.....	35
ABBILDUNG. SILBEN-/FREQUENZUNTERSCHIEDE: ZUNÄCHST SPALTEN SICH DIE SILBEN AUF, DANN DIE FREQUENZEN. (V. U. N. O.: H1, L1, L2, H2).....	35
ABBILDUNG. VERSCHIEDENE MITTELUNGEN V.L.N.R.: EIN RUN, EINE SESSION, ALLE.....	37
ABBILDUNG. DIE BEIDEN ÄUSSEREN RINGE WURDE WEGEN ZU HOHER ARTEFAKTRATE VON WEITEREN AUSWERTUNGEN AUSGESCHLOSSEN.....	39
ABBILDUNG. DIE ZEITSCHLEIBEN FÜR DIE WORTUNTERSCHIEDE.....	41
ABBILDUNG. DIE ZEITSCHLEIBEN FÜR DIE SILBEN/FREQUENZ-VERGLEICHE.....	41
ABBILDUNG. FOCI FÜR DAS ZEITFENSTER 90-120MS FÜR WORTKLASSENVERGLEICHE.....	42
ABBILDUNG. INTERAKTION ZW. SILBENZAHL UND FREQUENZEN IN I1 (120-170MS)....	47
ABBILDUNG. HAUPTEFFEKT DER WORTKLASSEN IM INTERVALL I0.....	49
ABBILDUNG. SIGNIFIKANTE INTERAKTION ZWISCHEN WORTKLASSEN UND FOCI IM ZEITBEREICH I1 (120-160MS, POL PARIETO-OKZIPITAL-LINKS, POZ -ZENTRAL, POR -RECHTS).....	50

ABBILDUNG. INTERAKTION DER WORTKLASSEN IM ZEITFENSTER 15 (330-440MS, LT LINKS-TEMPORAL, LZ LINKS-ZENTRAL, RZ RECHTS-ZENTRAL).....	51
ABBILDUNG. KORRELATION DER HIRNANTWORT AUF DIE WÖRTER UND DER BEWERTUNG DURCH 10 PERSONEN	53
ABBILDUNG. KORRELATION DER HIRNANTWORT AUF DIE WÖRTER UND DER BEWERTUNG DURCH DIE VERSUCHSPERSON SELBST	53
ABBILDUNG. SCATTERPLOT FÜR 2 KANÄLE, RECHTS UND LINKS PARIETAL IM FENSTER 10.....	61
ABBILDUNG. FORMANTENVERLAUF DES DIPHTHONGS /OI/.....	64
ABBILDUNG: KOHONENKARTE DER FINNISCHEN LAUTE.....	65
ABBILDUNG. DIE SPEICHERKAPAZITÄT DES NETZES ERHÖHT DIE LEISTUNG AUF TRAININGSDATEN.....	66
ABBILDUNG. DIE FOCI, DIE ZUM TRAINIEREN DES NEURONALEN NETZES BENUTZT WURDEN.....	67
ABBILDUNG. DIE BESTE LEISTUNG AUF TESTDATEN WIRD BEI ($\epsilon=0.3$ ERZIELT, AUF TRAININGSDATEN JEDOCH BEI ($\epsilon=0.6$	69
ABBILDUNG. EINE KOHONENKARTE (10X10) FÜR DIE WORTKLASSEN.....	70
ABBILDUNG. DAS ERSTE NETZ.....	71
ABBILDUNG. VERSCHOBENE FENSTER ERFASSEN GAMMAWELLEN VERSCHIEDEN GUT	73
ABBILDUNG. LINKS PARIETAL UNTERSCHIEDEN SICH DIE EINZELNEN WÖRTER SPÄT.	74