

B. M. Jansma¹
J. Möller²
A. Rodriguez-Fornells³
T. F. Münte²

Elektrophysiologie der Sprachproduktion

Electrophysiology of Language (Speech) Production

Zusammenfassung

Einleitung: Bei der Sprachproduktion müssen semantische, syntaktische und phonologische Information über das zu produzierende Wort abgerufen werden. Die genaue zeitliche Abfolge der einzelnen Stadien ist Gegenstand der in dieser Übersichtsarbeit zusammengestellten Studien. **Methodik:** Bei gesunden Probanden wurden ereigniskorrelierte Hirnpotenziale in einer impliziten Bildbenennungsaufgabe registriert. In jedem Einzelexperiment wurden zwei verschiedene Informationen (z.B. Semantik und Phonologie) gegeneinander getestet. Als Marker des Zeitpunktes, an dem ein bestimmter Informationstyp zur Verfügung steht, diente die N200-Komponente. **Ergebnisse:** Es zeigte sich folgende Abfolge bei der Bildbenennung: Semantik früher als Syntax und Phonologie, Syntax früher als Phonologie. In den zum Vergleich durchgeführten Untersuchungen zum auditorischen Wortverstehen ergab sich hingegen: Phonologie früher als Semantik, Semantik früher als Syntax. **Diskussion:** Ereigniskorrelierte Hirnpotenziale gestatten Aussagen über die zeitliche Verfügbarkeit von Informationen in Sprachproduktionsparadigmata. Diese Befunde können benutzt werden, um die Voraussagen von Sprachverarbeitungsmodellen zu testen.

Abstract

Purpose: Language (speech) production involves semantic, syntactic and phonological information that has to be retrieved for the word to be produced. The studies reviewed in this paper concern the temporal order in which each information becomes available. **Method:** Event-related brain potentials were recorded in young healthy subjects in an implicit picture naming task. In every experiment two types of information (e.g. semantic information and phonology) were tested against each other. The N200 NoGo component served as a time marker for the availability of the studied type of information. **Results:** The following time course was found during picture naming: semantic information preceded syntax, while syntax preceded phonology. By contrast, two parallel experiments on auditory word comprehension showed the following pattern: phonology earlier than semantic information, whereas semantics preceded syntax. **Discussion:** Event-related brain potentials permit a temporally precise delineation of information availability in language/speech production tasks. These findings can be used to test predictions of language/speech processing models.

Key words

Event-related potentials · language/speech production · N200 · semantics · syntax · phonology

Institutsangaben

¹ Faculteit der Psychologie, Neurocognitie, Universiteit Maastricht, Maastricht, Niederlande

² Abteilung Neuropsychologie, Otto-von-Guericke-Universität Magdeburg

³ Fakultät für Psychologie, Universität Barcelona, Spanien

Korrespondenzadresse

Prof. Dr. Thomas F. Münte · Abteilung Neuropsychologie · Otto-von-Guericke-Universität · Universitätsplatz 2 · 39106 Magdeburg · E-mail: Thomas.muente@med.uni-magdeburg.de

Bibliografie

Klin Neurophysiol 2004; 35: 61 – 68 © Georg Thieme Verlag KG Stuttgart · New York

DOI 10.1055/s-2003-814954

ISSN 1434-0275

Einleitung

Der menschliche Sprechakt ist ein hoch komplexer und immens schneller Prozess. Ein normaler Sprecher kann ohne Schwierigkeiten mehr als 100 Wörter pro Minute mit hoher Präzision und einer vergleichsweise geringen Fehlerrate (ca. 1 auf 1000 Wörter) produzieren [1, 2]. Aphasische Syndrome mit ihren charakteristischen Fehlermustern (semantische, phonematische Paraphasien, Agrammatismus, Paragrammatismus) legen nahe, dass die Sprachproduktion in verschiedene Stadien gegliedert werden muss, die jeweils mit unterschiedlichen Informationen (u. a. semantisch, syntaktisch, phonologisch) befasst sind. Für das Verständnis von pathologischen Sprachproduktionsmustern ist jedoch die Kenntnis der Verhältnisse beim normalen, gesunden Sprecher notwendig. Aufgrund der hohen zeitlichen Präzision und Dynamik der Sprechplanung und Ausführung sind hierfür Verfahren mit einer entsprechend hohen zeitlichen Auflösung, also insbesondere elektrophysiologische Verfahren, notwendig.

Im Vergleich zur elektrophysiologischen Forschung im Bereich des Sprachverstehens, die in den vergangenen 25 Jahren eine Reihe von Komponenten der ereigniskorrelierten Potenziale herausgearbeitet hat, die für semantische (N400), syntaktische (P600, LAN), phonologische (MMN) und prosodische (CPS) sensitiv bzw. spezifisch sind (Übersichten in [3, 4]), steckt die Elektrophysiologie der Sprachproduktion noch in den Kinderschuhen.

Dies liegt sicher daran, dass elektrophysiologische Untersuchungen von Sprachproduktion mithilfe von ereigniskorrelierten Hirnpotenzialen (EKP) mit einer Reihe von Schwierigkeiten konfrontiert sind. Beim natürlichen Sprechen sind die Äußerungen der Sprecher sehr schwer zu kontrollieren und zu standardisieren. Eine freie Sprechäußerung liefert z. B. in der Regel keine geeigneten Triggerpunkte, auf die die einzelnen EEG-Segmente für eine Mittlung synchronisiert werden könnten. In der psychologischen Sprachproduktionsforschung hat sich daher das Bildbenennen als Paradigma durchgesetzt, weil es hierbei möglich ist, die Sprachäußerungen der Probanden weitgehend zu kontrollieren und durch Zusatzinformationen (z. B. über akustisch präsentierte Wörter) Interferenzeffekte oder Fazilitationseffekte zu erzeugen (siehe z. B. [5]).

Ferner ist zu berücksichtigen, dass die Vokalisationen durchaus zu erheblichen Artefakten bei der EEG-Messung führen können.

Wir haben daher, wie zuvor van Turrenout et al. [6, 7], einen Versuchsaufbau gewählt, der als implizites Bildbenennen bezeichnet werden kann und bei dem eine Vokalisation des Probanden nicht notwendig ist. In solchen Experimenten wird den Probanden auf einem Bildschirm ein Bild präsentiert, welches „im Sinn“ benannt werden muss. Der Proband hat anhand des Bildes bzw. anhand des Namens des Abgebildeten zwei Entscheidungen zu treffen, von denen die eine bestimmt, ob der Proband einen Knopf drücken muss („Go“-Antwort) oder nicht („NoGo“-Antwort), die andere die Antworthand (rechts oder links) determiniert. Beispiele für Stimuli finden sich in Abb. 1.

Der N200-Effekt

Eine Reihe von elektrophysiologischen Studien haben die EKP-Korrelate von NoGo-Stimuli untersucht, d. h. von solchen Stimuli,

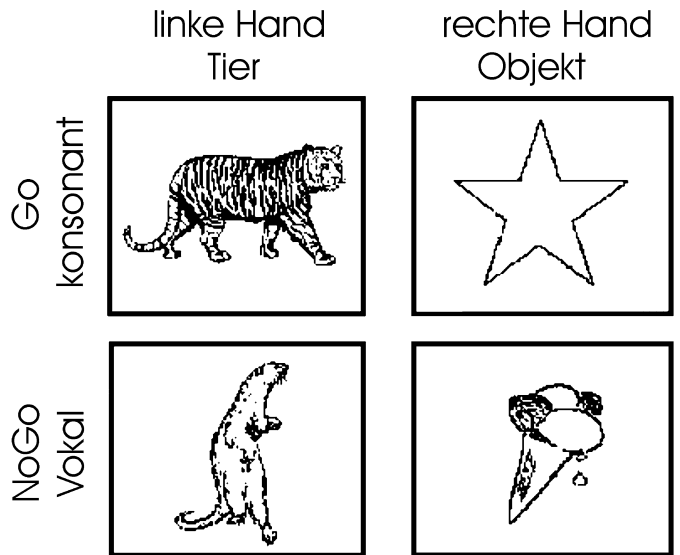


Abb. 1 Illustration des Stimulusmaterials. Die Go/NoGo-Entscheidung wird in diesem Falle durch phonologische Information (Vokal/Konsonant), die Handentscheidung durch die semantische Information (Tier/Objekt) determiniert. Durch Permutation der Entscheidungsreihenfolge ergeben sich 8 verschiedene Bedingungen (s. Tab. 1).

für die eine motorische Antwort unterdrückt werden muss. Relativ zu den Go-Versuchen findet sich für die NoGo-Stimuli eine Negativierung von 1–4 μ V-Amplitude, die mit aufgabenabhängiger Latenz über frontozentralen Regionen registriert werden kann [8–12]. Dieser so genannte frontale N200-Effekt steht mit Inhibitionsprozessen in Zusammenhang. Dies ergibt sich aus Versuchen an Affen, bei denen invasive Ableitungen aus dem präfrontalen Kortex für NoGo-Stimuli Aktivität im dorsalen Teil des Sulcus principalis gezeigt haben [13]. Darüber hinaus ist es gelungen, durch gezielte Stimulation des Sulcus principalis während Versuchen, die einen Go-Response erfordern hätten, diesen zu unterdrücken. Schließlich wurde durch funktionell-bildgebende Studien ein Netzwerk von Hirnarealen herausgearbeitet, welches den dorsolateralen und inferioren Frontallappen einschließt [14, 16]. All diese Befunde legen nahe, dass die N200 in NoGo-Versuchen die bei diesen erforderliche Inhibition reflektiert. Die N200 kann herangezogen werden, um den Zeitpunkt zu determinieren, zu dem die für die Inhibition erforderliche Information verfügbar ist [17]. Dabei zeigt die Gipfellatenz des N200-Effektes den Zeitpunkt an, an dem genug Information für die Inhibition zur Verfügung steht. Die Anfangslatenz spiegelt den Zeitpunkt wider, an dem erstmals Information zur Verfügung steht, die Go- und NoGo-Versuche voneinander unterscheidet.

In unseren eigenen Studien fanden wir, dass der N200-Effekt für praktisch jede Versuchsperson erhältlich ist mit einer großen inter- und intraindividuellen Konstanz.

Wir benutzten den N200-Effekt zur Bestimmung der zeitlichen Verfügbarkeit von semantisch/konzeptueller, syntaktischer und phonologischer Information während der impliziten Benennung von Bildern. Um ein komplettes Bild zu erstellen, wurden Entscheidungen entlang dieser Informationen in den Versuchen jeweils paarweise kombiniert. Darüber hinaus wurde in zwei Ex-

perimenten der Zeitverlauf beim Bildbenennen mit dem beim Hören verglichen.

Generelle Versuchsbedingungen

In jedes Experiment gingen jeweils zwischen 12 und 20 gesunde junge Versuchspersonen, die jeweils Deutsch als Muttersprache hatten und über die Hypothesen des Experimentes nicht informiert waren, ein. Es wurden für jedes Experiment 120 Strichzeichnungen ausgewählt, wobei die spezifischen Stimuli von der Art der Information abhingen, die abgeprüft wurde. So wurden beispielsweise in Studie 1 60 Bilder von Tieren und 60 Bilder von unbelebten Objekten benutzt. In jeder der beiden Kategorien begannen die entsprechenden Namen zu 50% mit einem Vokal und 50% mit einem Konsonant. 20 der Stimuli wurden für die Übungsaufgaben verwendet, die übrigen 100 für das eigentliche Experiment. Von diesen gehörten jeweils 25 in die Tier/Konsonant-, Tier/Vokal-, Objekt/Konsonant- und Objekt/Vokal-Kategorien. Die Wortfrequenz und Länge der Namen in den Kategorien wurden jeweils mithilfe der Datenbasis CELEX [18] angeglichen.

Für jedes Experiment ergaben sich acht verschiedene Instruktionen (siehe Beispiel in Tab. 1). Die Probanden wurden einzeln getestet. Zunächst wurden ihnen die Stimuli mit dem jeweils zugehörigen Namen gezeigt. Sodann wurden die Bilder ohne Namen auf einem Bildschirm gezeigt, wobei die Probanden den Namen laut aussprechen sollten. So wurde erreicht, dass die Probanden bei Wörtern mit mehreren möglichen Namen (Kröte/Frosch) den gewünschten Namen produzierten. Während der eigentlichen Datenaufzeichnung mussten die Probanden dann die oben geschilderte Aufgabe durchführen und durften den Namen des Bildes nicht laut aussprechen. Die Reihenfolge der acht Instruktionen war über die Probanden ausbalanciert.

Jeder einzelne Versuch begann mit der Präsentation eines Fixationskreuzes in der Mitte des Bildschirms, dann erschien das Bild für 1500 ms, welches wiederum vom Fixationskreuz abgelöst wurde. Die Reaktionszeiten wurden registriert. Das EEG wurde von 26–32 Kanälen in typischer Weise aufgezeichnet und die ereigniskorrelierten Potenziale für die einzelnen Signalkategorien gemittelt.

Studie 1: Semantik und Phonologie beim impliziten Bildbenennen

In dieser ersten Studie wurden Entscheidungen bezüglich der semantischen (Tier/Objekt) und der Phonologie (Konsonant/Vokal) gefordert. Die Reaktionszeiten für die beiden Bedingungen waren sehr ähnlich (Semantik = go/nogo 1053, Phonologie = go/nogo 1097 ms, n.s.).

Wie bereits oben beschrieben, gehen wir davon aus, dass eine erhöhte Negativierung im Sinne eines N200-Effektes in NoGo-Versuchen den Moment anzeigt, an dem genügend Information zur Verfügung steht, um eine motorische Antwort zu unterdrücken. Somit sind die Latenzcharakteristika des N200-Effektes in diesem Versuch von besonderer Bedeutung. Diese können am besten in den Differenzwellen, die durch Subtraktion der Go-Versuch-EKP von den NoGo-EKP erhalten werden, sichtbar gemacht werden. Die Abb. 2 illustriert das Gruppenmittel der beiden Bedingungen sowie die entsprechenden Differenzwellen. Hierbei wird deutlich, dass beide Antwortbedingungen (Semantik = Go/

Tab. 1 Illustration der acht verschiedenen Instruktionen am Beispiel des Stimulus „Feder“

	8 verschiedene Instruktionen 1–4 Phonologie = Go/NoGo 5–8 Semantik = Go/NoGo	erforderliche Antworten für Stimulus „Feder“ (Objekt, Konsonant)
1	drücke links für Tier, drücke rechts für Objekt, drücke nur, wenn erster Buchstabe Konsonant	Semantik rechts, Phonologie GO
2	drücke rechts für Tier, drücke links für Objekt, drücke nur, wenn erster Buchstabe Konsonant	Semantik links, Phonologie GO
3	drücke links für Tier, drücke rechts für Objekt, drücke nur, wenn erster Buchstabe Vokal	Semantik rechts, Phonologie NOGO
4	drücke rechts für Tier, drücke links für Objekt, drücke nur, wenn erster Buchstabe Vokal	Semantik links, Phonologie NOGO
5	drücke rechts für Konsonant, drücke links für Vokal, drück nur wenn Objekt	Phonologie rechts, Semantik GO
6	drücke rechts für Vokal, drücke links für Konsonant, drück nur wenn Objekt	Phonologie links, Semantik GO
7	drücke rechts für Konsonant, drücke links für Vokal, drück nur wenn Tier	Phonologie rechts, Semantik NOGO
8	drücke links für Konsonant, drücke rechts für Vokal, drück nur wenn Tier	Phonologie links, Semantik NOGO

NoGo; Phonologie = Go/NoGo) mit einem N200-Effekt assoziiert sind. Allerdings zeigen die Differenzpotenziale, dass die Latenzen in den beiden Bedingungen deutlich unterschiedlich sind. Wenn die Go/NoGo-Entscheidung durch laut Sprachproduktionsmodellen „frühe“ Informationen determiniert wird, trat der N200-Effekt mit einer früheren Gipfelatenz auf (384 ms) als in der Phonologie = Go/NoGo-Bedingung (473 ms). Die Latenzdifferenz war signifikant ($p < 0,01$). Auch die Anfangslatenz, definiert über den Zeitpunkt an dem die Amplitude des Differenzpotenzials 10% ihrer maximalen Amplitude erreicht hatte, war unterschiedlich (Semantik 206 ms, Phonologie 325 ms). Somit kann von einer Differenz der zeitlichen Verfügbarkeit von semantischer und phonologischer Information während der Sprachproduktion von 89 ms (Gipfelatenz) bis 119 ms (Anfangslatenz) ausgegangen werden. Dies passt gut zu den Vorhersagen von Sprachproduktionsmodellen.

Studie 2: Semantik und Phonologie: Vergleich von impliziter Bildbenennung und auditorischer Wortverarbeitung

Dieses Experiment verfolgte zwei Ziele. Zum einen sollten die Latenzwerte der Studie 1 repliziert werden, zum anderen sollte der Versuchsaufbau für die Domäne des Wortverstehens adaptiert werden. Ebenso wie das Sprechen ist auch das Verstehen ein komplexer Prozess, bei dem mit hoher Geschwindigkeit die verschiedenen Informationsebenen bearbeitet werden müssen. Dabei erscheint es vernünftig anzunehmen, dass beim Verstehen von Sprache, das heißt beim Übersetzen von Schallwellen in Bedeutung, phonologische Informationen früher zur Verfügung stehen sollte als konzeptuelle/semantische Information [19].

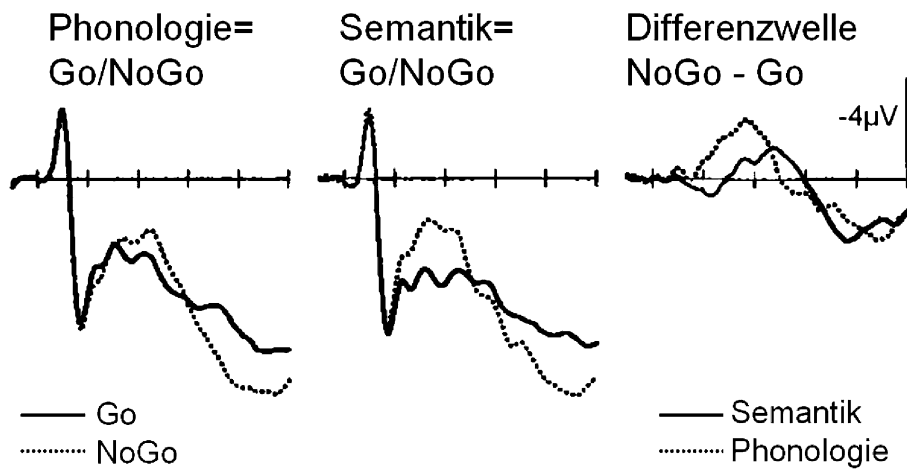


Abb. 2 Studie 1. Gruppenmittelpotenziale von einer frontalen Mittellinienelektrode (nahe Fz). NoGo-Versuche sind mit einer größeren Negativität im Zeitbereich 200–450 ms in der Semantik = Go/NoGo-Bedingung und 300–600 ms in der Phonologie = Go/NoGo-Bedingung assoziiert. Dieser N200-Effekt kann durch die Bildung von Differenzwellen (rechte Spalte) isoliert werden (Negativ nach oben; 200 ms pro Zeiteinheit).

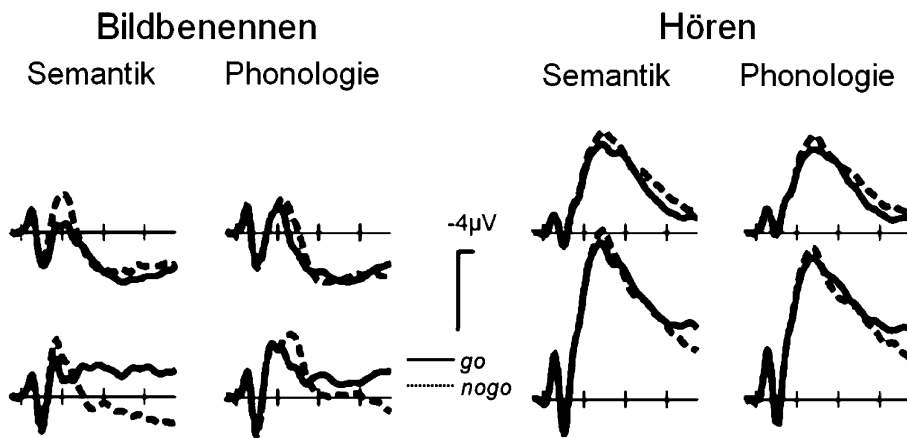


Abb. 3 Studie 2. Gruppenmittelpotenziale für die Bildbenennungsaufgabe und die auditorische Aufgabe. Die linken Spalten illustrieren jeweils die Semantik = Go/NoGo-Bedingung, die rechten Spalten geben die Daten der Phonologie = Go/NoGo-Bedingung wieder. Die obere Zeile zeigt die Potenziale einer frontopolen Elektrode, die untere Zeile die einer frontalen Elektrode (Negativ nach oben; 400 ms pro Zeiteinheit).

Während also Sprachproduktionsmodelle ein zeitliches Primat semantischer Informationen annehmen, ist für das Verstehen eine umgekehrte Verarbeitungsreihenfolge anzunehmen. Somit sollte in einer auditorischen Verstehensaufgabe der N200-Effekt für phonologische Information demjenigen für semantische Information vorausgehen.

Dabei gilt es jedoch zu berücksichtigen, dass sich Sprachproduktionsmodelle und Sprachverstehensmodelle in Bezug auf ihre Architektur bedeutsam unterscheiden. Erstere nehmen gewöhnlich eine serielle Abarbeitung der verschiedenen Informationen an, während letztere häufig eine parallele Architektur aufweisen [20,21]. Auf die Details soll hier aber nicht eingegangen werden (siehe hierzu [22]).

Während die Bildbenennungsaufgabe wie in Studie 1 angelegt war, wurden für die Sprachverstehensaufgabe die Namen der Bilder von einer weiblichen Sprecherin gesprochen, digitalisiert und mit den gleichen zeitlichen Vorgaben wie die Bilder in einem Szenario arrangiert. Die Probanden nahmen an zwei Sitzungen teil, während derer sie entweder die Bildbenennung oder die Wortverstehensaufgabe durchführen mussten, für die es jeweils wieder acht verschiedene Instruktionen gab.

Die mittleren Reaktionszeiten in der Bildbenennungsaufgabe waren 1006 für die Phonologie = Go/NoGo-Bedingung und 973 ms für die Semantik = Go/NoGo-Bedingung. In der Sprachverstehensaufgabe waren die Reaktionszeiten 1090 ms für Pho-

nologie = Go/NoGo und 1116 für Semantik = Go/NoGo. Bei der statistischen Analyse mithilfe der Varianzanalyse führte dieses Muster zu einem Haupteffekt für den Aufgabenfaktor ($p < 0,001$) sowie, kritisch für den vorliegenden Versuch, zu einer Wechselwirkung zwischen den Faktoren Aufgabe (Bildbenennen vs. Verstehen) und Bedingung (Semantik vs. Phonologie). Dies bedeutet, dass beim Bildbenennen semantische Entscheidungen schneller waren, beim Sprachverstehen hingegen phonologische Entscheidungen.

Die entsprechenden elektrophysiologischen Daten sind in Abb. 3 dargestellt. Beide Aufgaben führen demnach zu N200-Effekten in der NoGo-Bedingung. Dieser Effekt ist sehr unterschiedlichen EKP überlagert, was durch die unterschiedlichen Stimulationsmodalitäten erklärt ist. Die zeitliche Relation der verschiedenen N200-NoGo-Effekte geht aus Abb. 4 hervor. Es zeigt sich hierbei, dass der durch semantische Informationen bedingte N200-Effekt beim Bildbenennen dem phonologischen vorausgeht (390 ms gegenüber 563 ms). Dies stellt eine Replikation der Studie 1 dar. Bei der Sprachverstehensaufgabe verhält es sich genau umgekehrt (591 ms gegenüber 499 ms). Dieses Muster führte statistisch bei der Analyse der Gipfelatenzen zu einem Haupteffekt des Faktors Aufgabe (Bildbenennen vs. Verstehen, $p < 0,003$) und einer Wechselwirkung zwischen Aufgabe und Bedingung ($p < 0,0001$). Diese Wechselwirkung erklärt sich dadurch, dass die Gipfelatenz des phonologischen N200-Effektes in der Bildbenennungsaufgabe um 173 ms gegenüber der des semantischen verzögert war, wohingegen sie in der Verstehensaufgabe 92 ms früher als

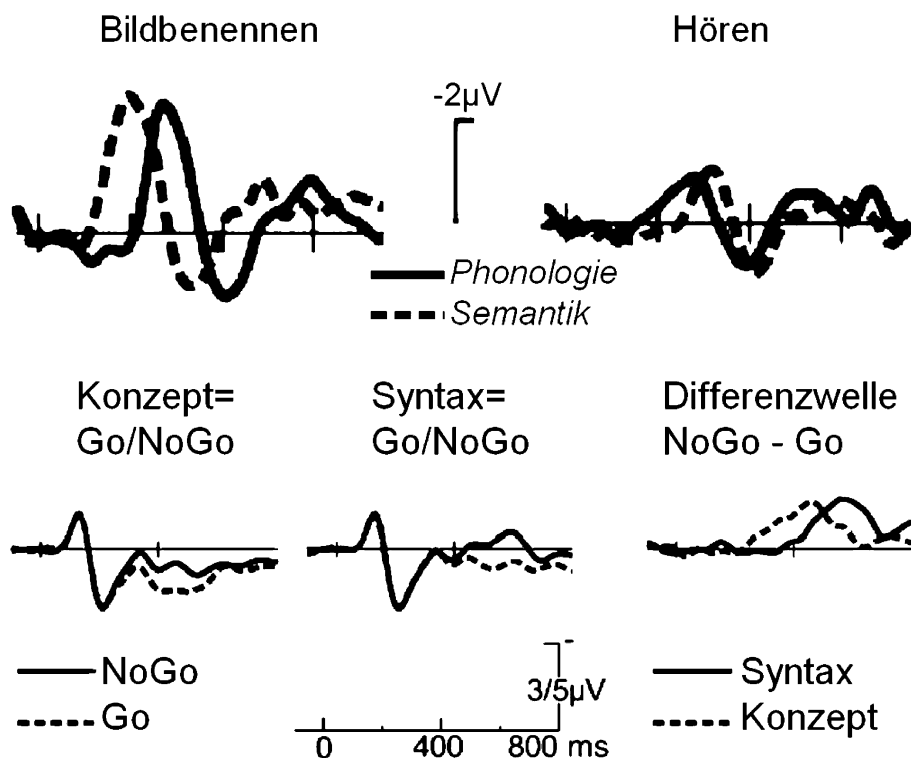


Abb. 4 Studie 2. Differenzwellen (NoGo minus Go) für eine frontale Mittellinienelektrode. In der Bildbenennungsaufgabe ist die Gipfellatenz des N200-Effektes früher, wenn die Go/NoGo-Entscheidung anhand semantischer Information gefällt werden muss. In der auditorischen Aufgabe ist die Reihenfolge umgekehrt (Negativ nach oben; 400 ms pro Zeiteinheit).

Abb. 5 Studie 3. Gruppenmittelpotenziale für eine frontale Mittellinienelektrode. Sowohl konzeptuelle Information als auch syntaktische Information ist mit einem N200-Effekt assoziiert. Die Differenzwellen (rechts) zeigen, dass der N200-Effekt eine spätere Anfangs- und Gipfellatenz in der Bedingung Syntax = Go/NoGo hat (Kalibrierung: 5 µV für die Original-EKP, 3 µV für die Differenzwellen).

die des semantischen Effektes lag. Auffällig ist die deutliche Modalitätsabhängigkeit der Amplitude des N200-Effektes, die sich auch in früheren, nicht linguistischen Experimenten von Falkenstein und Kollegen [23,24] gefunden hat. Unterstellt man, wie bisher in dieser Arbeit getan, dass die N200 die Inhibition der motorischen Antwort anzeigt, ist eine solche Antwort nicht zu erwarten. Gleichwohl sprechen die Daten von Falkenstein und Kollegen für die Inhibitionsinterpretation und für die funktionelle Äquivalenz der N200-Effekte in der visuellen und auditorischen Modalität.

Als Hauptresultat von Studie 2 kann somit festgehalten werden, dass die erwartete zeitliche Umkehr der semantischen und phonologischen Effekte in der Tat erhalten wurde. Andererseits zeigt sich aber auch, dass die zeitliche Trennung der Verarbeitungsstufen beim Sprachverstehen deutlich geringer ist als beim Bildbenennen und eine größere Überlappung der beiden N200-Effekte beim Sprachverstehen gefunden werden kann. Es stellt sich daher die Frage, ob diese Differenz Unterschiede in der funktionellen Organisation von Sprachproduktion und -verstehen widerspiegelt. Für eine ausführlichere Diskussion dieser Frage siehe Jansma et al. [22].

Studie 3: Konzeptuelle Verarbeitung und Syntax beim Bildbenennen

In einem weiteren Experiment untersuchten wir die relative zeitliche Verarbeitung von konzeptueller und syntaktischer Information beim Bildbenennen. Erstere wurde über eine Gewichtsentscheidung („Wiegt das abgebildete Objekt mehr oder weniger als 500 g?“), letztere über eine syntaktische Entscheidung (maskulines oder feminines Geschlecht?) realisiert. Die sich bei dieser Versuchsanordnung ergebenden Befunde sind in Abb. 5 dargestellt.

Die mittleren Reaktionszeiten lagen bei 1001 ms für die konzeptuelle Go/NoGo-Entscheidung und bei 1032 ms für die syntaktische Entscheidung. Diese Differenz war nicht signifikant. Im Gegensatz hierzu zeigten die ereigniskorrelierten Potenziale einerseits einen typischen N200-Effekt für beide Entscheidungen und darüber hinaus eine signifikante Latenzdifferenz ($p < 0,0015$) zugunsten der konzeptuellen Entscheidung (477 ms gegenüber 550 ms). Somit kann ein zeitliches Primat der konzeptuellen Information über die syntaktische Information geschlossen werden.

Studie 4: Semantik und Syntax: Vergleich zwischen impliziter Bildbenennung und auditorischem Wortverstehen

In diesem Experiment wurden zeitliche Verhältnisse der Verarbeitung semantischer und syntaktischer Informationen für die Bildbenennung und das auditorische Wortverstehen verglichen. Die semantische Verarbeitungsstufe wurde über eine Tier/Objekt-Klassifikation, die syntaktische Stufe über das grammatische Geschlecht abgeprüft. Dabei waren die experimentellen Materialien ähnlich denen der Studie 1. Die sich ergebenden ereigniskorrelierten Potenziale sind in Abb. 6 gezeigt. Hierbei ergibt sich, dass erwartungsgemäß die semantische Go/NoGo-Entscheidung beim Bildbenennen zu einem deutlich früheren (Gipfellatenz 384 ms) N200-Effekt führt als die syntaktische Entscheidung (464 ms, $p < 0,001$). Überraschenderweise ergab sich allerdings auch in der Sprachverstehensaufgabe eine zeitliche Präzedenz der semantischen Entscheidung (Semantik = Go/NoGo 451 ms, Syntax = Go/NoGo 520 ms, $p < 0,001$).

Die Daten der Bildbenennungsstudie replizieren somit Studie 3, die allerdings eine komplexe konzeptuelle Entscheidung erfordert hatte. Die unerwartete Reihenfolge in der auditorischen Aufgabe fand kürzlich Bestätigung durch eine ähnlich angelegte Studie von Müller u. Hagoort [25], die jedoch geschriebene statt

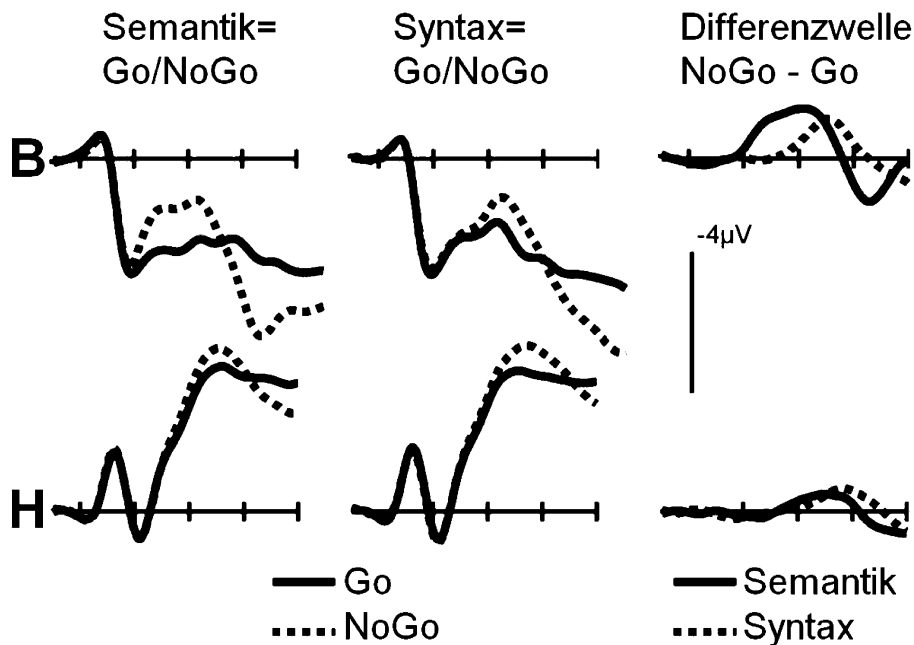


Abb. 6 Studie 4. Gruppenmittelpotenziale für frontale Mittellinienelektrode links und korrespondierende Differenzwellen. B Bildbenennexperiment: Hier hat der N200-Effekt für semantische Information eine frühere Latenz als derjenige für die syntaktische Information. Sowohl konzeptuelle Information als auch syntaktische Information. H Höraufgabe: Die relative zeitliche Abfolge des N200-Effektes für semantische und syntaktische Information ist der Bildbenennaufgabe vergleichbar (Negativ nach oben, 200 ms pro Zeitmarke).

gesprochene Wörter benutzen. Auch hier wurde ein früherer N200-Gipfel für die semantische Information gefunden. Dies widerspricht den Voraussagen von Sprachverarbeitungsmodellen, z. B. dem einflussreichen WEAVER⁺⁺-Modell von Roelofs [26]. In Computersimulationen konnte allerdings dieser Widerspruch inzwischen aufgelöst werden: Die semantische Information steht schneller zur Verfügung, weil sich diese in einem reich verzweigten Netzwerk ausbreiten und verstärken kann, selbst dann, wenn sie später angewählt wird als syntaktische (Geschlecht) Information, für die nur ein rudimentäres Netzwerk zur Verfügung steht.

Weitere Entwicklungen

Will man Sprachproduktionsprozesse verstehen, ist eine Beschränkung auf die Produktion von einzelnen Wörtern wie beim impliziten Bildbenennen nicht hinreichend. Wir haben daher – in Zusammenarbeit mit Van Nice u. Dietrich, Berlin – unsere Technik in einer ersten Studie auf die Satzproduktion ausgeweitet. In diesem Experiment wurden den Probanden Bilder präsentiert, auf denen etwas oder jemand („Agent“) etwas mit etwas oder jemand („Patient“) tat. Die Belebtheit von Agent und Patient wurde systematisch manipuliert, sodass vier Kombinationen (belebt/belebt; belebt/unbelebt; unbelebt/belebt; unbelebt/unbelebt) entstanden.

Die entscheidende Frage in diesem Experiment war, ob die Belebtheitsinformation für den Agenten (der gewöhnlich das Subjekt eines Satzes darstellt und daher am Anfang des Satzes zu finden ist) vor der Belebtheitsinformation für den Patienten (Objekt und Satzende) zur Verfügung steht. In analoger Weise zu Tab. 1 wurden in diesem Experiment die Belebtheitsentscheidungen für den Agenten und Patienten eines Bildes den Hand- (rechts/links) und Go/NoGo-Entscheidungen zugeordnet, sodass sich acht verschiedene Instruktionen ergaben. In diesem Versuch fanden sich wiederum systematische Zeitdifferenzen für den N200-Effekt: Die Agent = Go/NoGo-Entscheidung zeigte einen Gipfel bei 400 ms, die Patient = Go/NoGo-Entscheidung hingegen bei 580 ms, sodass unsere Voraussagen erfüllt wurden.

In einer weiteren Serie von Experimenten haben wir die Go/NoGo-Technik eingesetzt, um Interferenzen zwischen Sprachen bei zweisprachigen Probanden zu untersuchen [27].

Diskussion

Die vorliegende Serie von Experimenten liefert Daten über den Zeitverlauf der Verfügbarkeit von semantischen, syntaktischen und phonologischen Informationen während des Bildbenennens. In der Tab. 2 sind diese, basierend auf den Daten zur Gipfellatenz des N200-Effektes, zusammengefasst. Hieraus ergeben sich substantielle Zeitunterschiede. Die konservativste Interpretation dieser Daten ist, dass die entsprechenden Informationen zu verschiedenen Zeitpunkten zur Verfügung stehen. Trotz der Tatsache, dass die Daten in einem recht artefiziellen Paradigma erhoben wurden, möchten wir gleichwohl argumentieren, dass die hier erhaltene zeitliche Reihenfolge eine Konsequenz der Architektur des Sprachproduktionssystems ist und somit auch Aussagen über diese Architektur auf dem Boden der vorliegenden Daten erlaubt sind. Das erhaltene Befundmuster erscheint kompatibel mit Modellen, die eine strikt serielle Architektur der Verarbeitungsschritte vorsehen, bei denen also ein nachgeordneter Informationstyp erst dann abgearbeitet wird, wenn der vorher gelegene Informationstyp vollständig analysiert ist. Die Befunde sich jedoch auch mit so genannten Kaskadenmodellen vereinbar, bei denen zwar eine zeitliche Staffelung der Informationen (semantisch/konzeptuell → syntaktisch → phonologisch) vorgesehen ist, die Verarbeitung in den einzelnen Schritten aber z. T. parallel auf der Basis von Teilergebnissen der vorgeschalteten Stufe erfolgt [1,28,29]. Nicht gestützt werden hingegen Modelle, die eine komplett parallele Verarbeitung der verschiedenen Informationstypen annehmen würden. Eine weitere Frage betrifft, ob die vorliegenden Befunde eher mit nicht dekompositionalen oder mit dekompositionalen Modellen der Wortrepräsentation vereinbar sind. Erstere nehmen eine Wortspeicherung „im Ganzen“ an, wobei die semantischen Eigenschaften (wie „Pflanzenfresser“, „rot“, „schuppig“) über „links“ zu verwandten Knoten

Tab. 2 Latenzunterschiede für die Gipfellatenz des N200-Effektes. Die Zahlen in den eckigen Klammern beziehen sich auf die Nummer der Studie, wie sie in diesem Manuskript verwendet wurde. Ein Minuszeichen indiziert, dass die Gipfellatenz des Effektes, der durch die Zeilenlegende bezeichnet wird, vor dem Effekt lag, der durch die Spaltenlegende bezeichnet wird

Bildbenennung	syntaktisch	phonologisch
semantisch/konzeptuell	- 90 ms [3]	- 100 - 170 ms [1, 2]
syntaktisch	-	- 40 ms [4]
auditorische Aufgabe	syntaktisch	phonologisch
semantisch/konzeptuell	- 70 ms [4]	+ 90 ms [2]
syntaktisch	-	nicht untersucht

abgerufen werden [30]. In dekompositionalen Modellen hingegen ist die Selektion eines Konzeptes äquivalent mit der Aktivierung seiner semantischen Eigenschaften [31].

Eine ausführliche Diskussion der sprachpsychologischen Implikationen der vorliegenden Befunde und eine Behandlung kritischer Aspekte des vorliegenden Paradigmas sprengt den Rahmen dieses Artikels. Daher verweisen wir den interessierten Leser auf eine ausführlichere Abhandlung [22].

Abschließend möchten wir festhalten, dass die elektrophysiologische Untersuchung von Sprachproduktionsprozessen nicht nur möglich ist, sondern anderweitig nicht oder kaum erhältliche Informationen über den Zeitverlauf von Informationsverarbeitungsschritten zu liefern vermag. Die Elektrophysiologie stellt daher eine notwendige Ergänzung zu funktionell kernspintomographischen Untersuchungen dar, die über die anatomischen Grundlagen der Sprachproduktion Auskunft geben können [32]. In den nächsten Jahren gilt es, die vorliegende Datenbasis durch realistischere Experimente zu untermauern und vor allem den vorhandenen Ansatz auf Probleme der Satzproduktion auszuweiten.

Anmerkungen

Die Daten wurden im Rahmen eines im DFG-Schwerpunktprogramms „Sprachproduktion“ geförderten Projekts erhoben (MU 1311/7-2, T. F. Münte). Weitere Unterstützung erfolgte durch die niederländische Wissenschaftsorganisation NWO (B. M. Jansma) und die spanische Regierung (Ramon Y Cajal Programm, A. Rodriguez-Fornells). An einzelnen Experimenten wirkten mit (alphabetische Nennung): Kati Van Nice, Kolja Schiltz, Marta Kutas, Mart Bles, Niels Schiller, Rainer Dietrich, Wanda Zaake. Die hier vorgestellten Daten sind im Original publiziert in [33–36]. Bernadette Jansma publizierte bis 2002 unter ihrem Geburtsnamen Schmitt.

Literatur

- Levelt WJM. Speaking. From Intention to Articulation. Cambridge: MIT Press, 1989
- Leuninger H. Reden ist Schweigen, Silber ist Gold. Gesammelte Vorträge. Zürich: Ammann, 1993
- Münte TF, Urbach TP, Düzel E, Kutas M. Event-related brain potentials in the study of human cognition and neuropsychology. In: Boller F, Grafman J, Rizzolatti G (eds): Handbook of Neuropsychology. Amsterdam: Elsevier, 2000: 139–235
- Kutas M, Federmeier KD, Coulson S, King JW, Münte TF. Language. In: Cacioppo JT, Tassinary LG (eds): Handbook of Psychophysiology (2nd ed.). New York: Cambridge University Press, 2000: 576–601
- Levelt WJM, Schriefers H, Vorberg D, Meyer AS, Pechmann T, Havinga J. The time course of lexical access in speech production: A study of picture naming. Psychol Rev 1991; 98: 122–142
- Turenout M Van, Hagoort P, Brown CM. Electrophysiological evidence on the time course of semantic and phonological processes in speech production. J Exp Psychol Learn Mem Cogn 1997; 23: 787–806
- Turenout M Van, Hagoort P, Brown CM. Brain activity during speaking: From syntax to phonology in 40 milliseconds. Science 1998; 280: 572–574
- Gemba H, Sasaki K. Potential related to no-go reaction of go/no-go hand movement task with color discrimination in human. Neurosci Lett 1989; 101: 262–268
- Kok A. Effects of degradation of visual stimuli on components of the event-related potential (ERP) in Go/NoGo reaction tasks. Biol Psychol 1986; 23: 21–38
- Pfefferbaum A, Ford J, Weller B, Kopell B. ERPs to response production and inhibition. Electroenceph Clin Neurophys 1985; 60: 423–434
- Sasaki K, Gemba H, Nambu A, Matsuzaki R. No-go activity in the frontal association cortex of human subjects. Neurosci Res 1993; 18: 249–252
- Simson R, Vaughan G, Ritter W. The scalp topography of potentials in auditory and visual go/nogo tasks. Electroenceph Clin Neurophys 1977; 43: 864–875
- Sasaki K, Gemba H, Tsujimoto T. Suppression of visually initiated hand movement by stimulation of the prefrontal cortex in the monkey. Brain Res 1989; 495: 100–107
- Konishi S, Nakajima K, Uchida Y, Kikyo H, Kameyama M, Miyashita Y. Common inhibitory mechanism in human inferior prefrontal cortex revealed by event-related functional MRI. Brain 1999; 122: 981–991
- Garavan H, Ross TJ, Stein EA. Right hemispheric dominance of inhibitory control: an event-related functional MRI study. Proc Nat Acad Sci USA 1999; 96: 8301–8306
- Menon V, Adelman NE, White CD, Glover GH, Reiss AR. Error-related activation during a go/nogo response inhibition task. Human Brain Mapping 2001; 12: 131–143
- Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. Nature 1996; 381: 520–522
- Baayen RH, Piepenbrock R, Rijn H Van. The CELEX Lexical Database. Philadelphia: Linguistic Data Consortium, University of Pennsylvania, 1993
- Cutler A, Clifton C. Comprehending spoken language: a blueprint of the listener. In: Brown CM, Hagoort P (eds): The neurocognition of language. Oxford: Oxford University Press, 2000: 132–166
- Elman JL, McClelland JL. Speech perception as a cognitive process: the interactive activation model. In: Lass N (ed): Speech and Language, vol. 10. New York: Academic Press, 1984: 337–374
- Marslen-Wilson W, Tyler LK. The temporal structure of spoken language understanding. Cognition 1980; 8: 1–71
- Jansma BM, Rodriguez-Fornells A, Möller J, Münte TF. Electrophysiological studies of speech production. In: Pechmann T, Habel C (eds): Multidisciplinary approaches to language production. Berlin: Mouton De Gruyter, 2004, in press
- Falkenstein MJ, Hoormann J, Hohnsbein J. ERP components in Go/no-go tasks and their relation to inhibition. Acta Psychol 1999; 101: 267–291
- Falkenstein MJ, Koshlykova NA, Kiroj VN, Hoormann J, Hohnsbein J. Later ERP components in visual and auditory Go/noGo tasks. Electroenceph Clin Neurophys 1995; 96: 36–43
- Müller O, Hagoort P. Access to lexical information in comprehension – semantics before syntax. Nijmegen: Poster presented at the TUBS meeting, 2001

- ²⁶ Roelofs A. A spreading-activation theory of lemma retrieval in speaking. *Cognition* 1992; 42: 107–142
- ²⁷ Rodriguez-Fornells A, Lugt A van der, Rotte M, Britti B, Heinze HJ, Münte TF. Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. *J Cogn Neurosci*, 2004 im Druck
- ²⁸ Dell GS. A spreading-activation model of retrieval in sentence production. *Psychol Rev* 1986; 93: 283–321
- ²⁹ Dell GS, O'Seaghdha PG. Stages of lexical access in language production. *Cognition* 1992; 42: 287–314
- ³⁰ Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. *Behav Brain Sci* 1999; 22: 1–75
- ³¹ Bierwisch M, Schreuder R. From lexical concepts to lexical items. *Cognition* 1992; 42: 23–60
- ³² Indefrey P, Levelt WJM. The neural correlates of language production. In: Gazzaniga M (ed): *The New Cognitive Neurosciences*. Cambridge, MA: The MIT Press, 2000: 845–865
- ³³ Schmitt BM, Münte TF, Kutas M. Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiol* 2000; 37: 473–484
- ³⁴ Schmitt BM, Schiltz K, Zaake W, Kutas M, Münte TF. An electrophysiological analysis of the time course of conceptual and syntactic encoding during tacit picture naming. *J Cogn Neurosci* 2001; 13: 510–522
- ³⁵ Schmitt BM, Rodriguez-Fornells A, Kutas M, Münte TF. Electrophysiological estimates of semantic and syntactic information access during tacit picture naming and listening to words. *Neurosci Res* 2001; 41: 293–298
- ³⁶ Rodriguez-Fornells A, Schmitt BM, Kutas M, Münte TF. Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia* 2002; 40: 778–787