

Registrierung von Augenbewegungen bei Studien zur Sprachproduktion

Dobel, Christian E.
Meyer, Antje S.
Levelt, Willem J.M.

Max-Planck-Institut für Psycholinguistik, Postbus 310, 6500 AH Nijmegen,
Niederlande

Einleitung

Die Messung von Augenbewegungen gewinnt in den letzten Jahren immer stärkere Bedeutung bei der Untersuchung komplexer kognitiver Prozesse (ein Überblick findet sich z.B. bei Rayner, 1998). Mit dieser Methode lassen sich Blickbewegungen mit einer Genauigkeit von Millisekunden in der zeitlichen und mit ca. 0.5 Grad in der räumlichen Domäne registrieren vor, während und nachdem eine Versuchsperson eine overt Handlung wie z.B. einen Tastendruck oder Sprechen vollzieht. Dabei ist bemerkenswert, dass das Zusammenspiel von Augenbewegungen und visueller Aufmerksamkeit relativ gut verstanden wird. Wenn Probanden Szenen betrachten oder Objekte manipulieren, ohne dass sie spezifische Instruktionen bzgl. ihre Augenbewegungen erhalten haben, folgen Augenbewegungen einem Wechsel der Aufmerksamkeit (Rayer & Pollatsek, 1992; Tanenhaus et al., 1995). Es ist zwar möglich die Konzentrationsrichtung ohne eine Blickbewegung zu ändern (z.B. Posner, 1980), aber Studien jüngerer Datums konnten zeigen, dass der umgekehrte Fall nicht vorkommt. Einer Blickfolge von einem Objekt auf ein anderes geht zwingend eine Änderung der Aufmerksamkeit voraus (Deubel & Schneider, 1996; Irwin & Gordon, 1998). Man kann also schliessen, dass eine Person, die eine Augenbewegung gemacht hat, auch ihre Konzentrationsrichtung geändert hat.

Im Rahmen der Sprachproduktion gibt es eine Reihe von Studien jüngerer Datums, die diese Methode verwandt haben. In Studien, bei denen Probanden Objektpaare benennen sollten fanden Meyer und Mitarbeiter (Meyer et al., 1998; Meyer & van der Meulen, 2000), dass Versuchspersonen das erst genannte Objekt so lange fixierten und beachteten, bis sie den phonologischen Code für das entsprechende Objekt aktiviert hatten. Dieses Ergebnis wurde bestätigt in einer weiteren Untersuchung, bei der Probanden kurze (Artikel + Substantiv) und lange (Artikel + zwei Adjektive + Substantiv) Substantivsätze produzieren sollten. Es zeigte sich, dass der Wechsel der Fixation zeitlich abgestimmt war auf den Beginn der Äußerung des jeweils letzten Wortes. Dementsprechend zeigte sich auch, dass Objekte kürzer und weniger häufig fixiert werden, wenn sich Sprecher auf diese mit einem Pronomen bezogen (van der Meulen et al., 2001). Diese Resultate fanden ihre Bestätigung ebenfalls, wenn Probanden komplexe Szenen mit Aktiv-, und Passivsätzen beschrieben (Griffin & Bock, 2000): Elemente eines Bildes wurden weniger als eine Sekunde vor dem Benennen fixiert. Des Weiteren zeigte sich in eine klare Trennung zwischen einer Fixationsphase vor und nach Sprachbeginn, wobei die Verteilung der Fixationszeiten die Reihenfolge des Benennens im Satz vorwegnahm, unabhängig von der verwendeten Struktur.

Auf dem Hintergrund dieser Untersuchungen wurde in der vorliegenden Arbeit untersucht, wie Augenbewegungen der Produktion komplexer Sätze (Sätze mit präpositionaler Struktur oder mit direktem und indirektem Objekt) vorausgehen und diese begleiten. Dabei wurde ein Design verwendet, das seine Anwendung häufig bei Studien zu syntaktischem Priming findet

(z.B. Bock, 1986). Syntaktisches Priming beschreibt die Tendenz eine syntaktische Struktur zu verwenden, wenn diese gerade vom Sprecher selbst gebraucht wurde oder wenn er diese gehört hatte. Diese Tendenz ist mittlerweile vielfach repliziert worden, sowohl mit unterschiedlichen Designs (ein Überblick findet sich z.B. bei Pickering & Branigan, 1999) als auch in mehreren Sprachen (für Holländisch: Hartsuiker & Kolk, 1998).

Methoden

Zwölf Studenten der Universität Nimwegen nahmen gegen Bezahlung an zwei Sitzungen teil. Alle sprachen niederländisch als Muttersprache.

Das Bildmaterial bestand aus 28 Dativ Bildern mit den Bildelementen Agent, Objekt und Rezipient, die im Holländischen entweder mit einer präpositionalen Struktur (PP, z.B. „Der Cowboy gibt einen Hut an den Clown,“) oder einer Struktur mit zwei Objekten (DO, z.B. „Der Cowboy gibt dem Clown einen Hut,“) beschrieben werden konnten. Die Bilder wurden in einem Vortest auf ihre Grösse überprüft, so dass sich die Fixationen auf die Elemente gut unterscheiden liessen (visueller Winkel horizontal: ca. 18 Grad; vertikal 16 Grad). Diesen Target-Bildern gingen Prime-Sätze voraus, die bedinungsabhängig ebenfalls eine der beiden Strukturen hatten. 28 zusätzliche Bilder und Sätze dienten als Füller. Diese konnten am besten mit transitiven, intransitiven oder lokativen Sätzen beschrieben werden, bzw. hatten diese Struktur. Zwischen einer Prime-Target Kombination wurden im Durchschnitt sechs Fülleritems dargeboten. Ein Trial bestand aus der Darbietung eines Fixationskreuzes am oberen, mittleren Bildschirmrand für 850ms. Danach verschwand diese Kreuz und wurde ersetzt durch ein Fragezeichen oder eine Null. Zeitgleich erschien eines der Items. Für die Versuchspersonen wurde die Aufgabe als Gedächtnisaufgabe deklariert, wo sie abhängig vom Zeichen („?,“ oder „0,“) reagieren sollten oder nicht. Bei einem Fragezeichen war ihre Aufgabe per Tastendruck anzugeben, ob dieses Item schon einmal vorgekommen ist. Sie wurden angehalten zur Verbesserung der Gedächtnisleistung den dargebotenen Satz laut zu lesen, bzw. das Bild zu beschreiben.

Jeder Proband nahm an zwei Sitzungen teil, die im Abstand von ca. drei Wochen durchgeführt wurden. Bei der ersten Sitzung war der Agent der Target Bilder immer links positioniert, bei der zweiten Sitzung rechts (Agent-links vs. Agent-rechts Bedingung).

Die Äusserungen der Versuchspersonen wurde mit einem DAT Rekorder aufgezeichnet. Die Messung vom Satzbeginn und Äusserungen innerhalb eines Satzes wurde mit der xwaves/ESPS Software manuell gemessen. Augenbewegungen wurden mit einem SMI EyeLink-Hispeed 2D head-mounted eye tracking System registriert (SensoMotoric Instruments GmbH, Teltow, Deutschland). Augenbewegungen wurden mit Hilfe graphischer Software ausgewertet, wobei Fixationen auf ein Element des Bildes zusammengefasst wurden, wenn diese nicht durch Fixieren eines anderen Elements unterbrochen wurden.

Das Verhältnis von Fixation und Sprachbeginn wurden sprachrelevante Fixationen, d.h. die letzte Fixation auf eine Element vor der Benennung desselben, und die „Auge-Sprach-, Spanne bestimmt. Dies ist die mittlere Zeit zwischen sprachrelevanter Fixation auf eine Bildelement und der Beginn für die Benennung desselben (einschl. direktem oder indirektem Artikel).

Ergebnisse

Bei der Analyse der Äusserungen der Probanden ergab sich eine signifikante Interaktion zwischen Lokalisation des Agenten und verwendeter Struktur ($F(1,11)=5.7$; $p,.05$). In der Agent-links Bedingung wurden häufiger präpositional Strukturen verwendet, wohingegen das Verhältnis in der Agent-rechts Bedingung ausgeglichen war. Das Priming hatte keinen

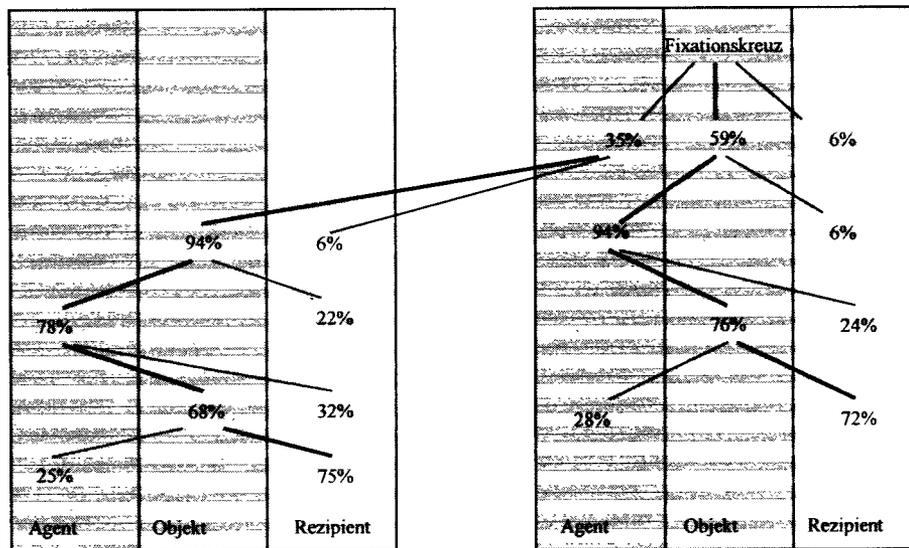
signifikanten Effekt auf die Wahl der syntaktischen Struktur, lag jedoch in der richtigen Richtung.

Es zeigte sich eine Tendenz in der Agent-rechts Bedingung Sätze zu produzieren, die mit einer Beschreibung des linken Bildelements begannen ($t=-1.9$; $df=11$; $p<.1$).

In dieser Bedingung waren die Probanden signifikant weniger flüssig, d.h. sie zögerten häufiger innerhalb eines Satzes ($t=-2.3$; $df=11$; $p<.05$).

Die Analyse der Blickfolgebewegungen ergab, dass die Probanden im Allgemeinen zürst das Objekt betrachteten, dann den Agenten, wieder das Objekt und schliesslich den Rezipienten. Auch in den Fällen wo die Versuchspersonen zuerst den Agenten fixierten, betrachteten sie dann das Objekt und fuhren fort wie in der Mehrzahl, nämlich Agent, Objekt Rezipient (als Beispiel für Agent-links siehe Figur 1).

Figur 1:
Fixationsreihenfolge für Agent links



Bei der Darstellung der Fixationshäufigkeit auf ein Objekt in Abhängigkeit von der Zeit wurde dieses Ergebnis bestätigt und es zeigte sich ein Trend für unterschiedliche Fixationshäufigkeiten abhängig der produzierten Satzstruktur (als Beispiel fuer Agent-links PP Saezte siehe Figur 2).

Dies wurde genauer untersucht durch die sprachabhängigen Fixationen. In der weit über dem Zufall liegenden Mehrheit der Fälle war die Reihenfolge der sprachrelevanten Fixationen Agent, Objekt, Rezipient bei PP Sätzen und Agent, Rezipient, Objekt bei DO Sätzen. Dieser Effekt war für beide Bedingungen hochsignifikant (Agent-links: $F(2,16)= 64$; $p<.01$; Agent-rechts: $F(2,14)= 16$; $p<.01$), wobei posthoc Analysen zeigten, dass der Effekt von der umgekehrten Reihenfolge der syntaktischen Strukturen von Objekt und Rezipient herrührte (siehe Tabelle 1).

Die gemittelte „Auge-Sprach,“ Spanne betrug für die benannten Objekte zwischen 400 und 700 msec, d.h. Probanden begannen zwischen 400 und 700 msec ein Element zu fixieren bevor sie anfangen es innerhalb eines Satzes zu benennen (siehe Tabelle 2).

Tabelle 1: Beginn sprachrelevanter Fixationen (Mittelwerte und Standardabweichungen)

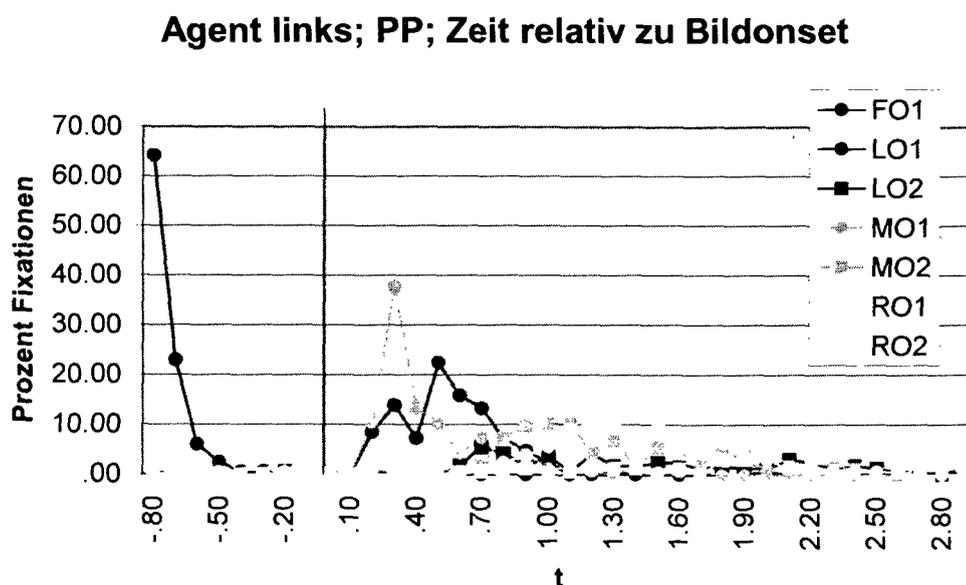
		Agent-links	Agent-rechts
PP	Agent	660±199	791±168
	Objekt	1294±524	1120±400
	Rezipient	2117±679	1830±376
DO	Agent	629±175	798±215
	Objekt	1693±440	1507±533
	Rezipient	1443±310	1320±500

Tabelle 2: „Auge-Sprach,, Spanne (Mittelwerte und Standardabweichung)

	Agent links			Agent rechts		
	Agent	Objekt	Rezipient	Agent	Objekt	Rezipient
PP	645±416	703±335	704±305	465±231	632±252	665±253
DO	574±311	427±281	616±307	506±242	516±272	646±258

Die Analyse der Fixationen während Benennungsbeginn zeigte über Satzstruktur und benanntem Objekt hinweg, dass eine starke Tendenz besteht, bereits das Objekt zu fixieren, welches als nächstes benannt wird. Dieser Effekt war weniger stark ausgeprägt bei DO Konstruktionen in der Agent-links Bedingung. Der Vergleich der „Auge-Sprach,, Spannen zwischen den beiden Konstruktionen in dieser Bedingung erbrachte ebenfalls einen hochsignifikanten Effekt mit kürzeren Zeiten bei DO Konstruktionen ($F(1,8)=21$; $p<.01$).

Figur 2: Prozent Fixationen relativ zu Bildonset (Schritte von 100 msec). LO bezeichnet Agent, MO Objekt, RO Rezipient; erste und zweite Fixation durch Nummerierung gekennzeichnet



Diskussion

In der vorliegenden Arbeit produzierten Probanden im Rahmen eines Priming Paradigmas anhand von dargebotenen Bildern syntaktisch komplexe Sätze während ihre Augenbewegungen gemessen wurden. Die Auszählung der bevorzugten Satzstrukturen in Abhängigkeit vom Priming ergab zwar einen Effekt in der erwünschten Richtung, der jedoch klein und nicht signifikant war. Dies scheint im Widerspruch zur existierenden Literatur zu stehen. Dabei muss jedoch berücksichtigt werden, dass diese Studien mindestens viermal grössere Probandengruppen verwendet haben, so dass die statistische Power wesentlich höher war. Des Weiteren berichteten bestehende Studien keinen Einfluss der Positionierung des Agenten auf die Wahl der Konstruktion, was in der vorliegenden Studie nur bei der Agent-rechts Bedingung der Fall war. Wir vermuten, dass dies an der Grösse des Bildmaterials lag, da die existierenden Studien wesentlich kleineres Bildmaterial verwendeten. Wenn das Bild in der bevorzugten bzw. geprimten Orientierung dargeboten wurde, war es für die Probanden am ökonomischsten eine Beschreibung von links nach rechts durchzuführen, was mit einer präpositionalen Konstruktion übereinstimmen würde. Dazu muss angemerkt werden, dass die Repräsentation der visuellen Welt erheblich schematischer und abstrakter vorliegt, als es die subjektive Erfahrung vermuten lässt (Intraub, 1997). Nur durch die aufmerksamkeitsgesteuerte Ausrichtung der Fovea auf ein Bildelement werden Details repräsentiert, wenn die Notwendigkeit dafür besteht. Eine Reihe von jüngeren Untersuchungen legen nahe, dass Probanden bei visuo-motorischen Aufgaben nur einen kleinen Teil der verfügbaren visuellen Information repräsentieren und bei einer Fixation nur ein Aspekt (z.B. Farbe oder Position) repräsentiert wird (Hayhoe, 2000).

Wenn das Bildmaterial nicht in der geprimten oder bevorzugten Weise vorliegt (Agent-rechts), scheinen sprachliche Gewohnheiten einen stärkeren Einfluss auszuüben, so dass beide Konstruktionen gleich häufig vorkommen. Dass die Agent-links Bedingung bevorzugt ist scheint darin bestätigt zu werden, dass die Sprache hier flüssiger ist und auch in der Agent-rechts Bedingung ein Trend besteht die Bildbeschreibung mit links zu beginnen.

Bei der Analyse der Augenbewegungen war ein überraschendes Ergebnis, dass in der Mehrzahl der Fälle das Objekt zuerst fixiert wurde, bzw. dass es im Falle eines anderen zuerst fixierten Elements das nächst fixierte war. Dies lässt vermuten, dass dadurch auf eine Information zugegriffen wird, die für die aktuelle Aufgabe, das Bilden eines Satzes, wichtig ist. An dieser Stelle kann noch nicht widerlegt werden, dass es sich dabei nicht um ein Artefakt handelt (z.B. vertikale Sakkaden werden gegenüber diagonalen bevorzugt). Es sprechen jedoch ein paar Gründe dagegen. Wie schon erwähnt wurde das Objekt dann als nächstes fixiert, wenn zuvor ein anderes Element betrachtet wurde und danach nochmal sprachbegleitend. Es muss also nicht unbedingt als erstes fixiert werden, aber vermutlich vor Satzbeginn. Des Weiteren zeigten Studien mit Textmaterial, dass Probanden durchaus diagonale Sakkaden durchführen können, um zum Beginn eines Satzes zu kommen. Ähnlich dazu wurde in der nicht-sprachlichen Kontrollbedingung bei Griffin und Bock (2000) festgestellt, dass Versuchspersonen das intendierte Bildelement ohne Suchbewegungen fixieren können. Als Arbeitshypothese für künftige Experimente vermuten wir, dass die Probanden von der Region, wo sich das Objekt befindet, ableiten, welche Handlung vollzogen wird. Obwohl der Rezipient vor Sprachbeginn nicht fixiert wurde, lässt sich daraus nicht schliessen, dass die Probanden überhaupt keine Information über dieses Bildelement hatten. Zum einen ergibt sich das aus dem Kontext, der durch die Handlung gegeben wird. Zum anderen zeigten Untersuchungen mit Paradigmen der visuellen Suche, dass im präattentiven Zustand die Identität von Objekten zwar nicht festliegt, aber Informationen über Eigenschaften wie Grösse, Farbe, Neigung etc. liegen vor (Wolfe, 1998), so dass bestimmte Objekte wahrscheinlicher werden als andere.

Die Analyse der sprachbegleitenden Fixationen konnten die früheren Studien von Meyer und Mitarbeitern (1999; 2000, Griffin & Bock (2000) und van der Meulen (2001) bestätigen. Elemente eines Bildes werden vor der Benennung fixiert, was auf eine inkrementelle Verarbeitungsweise hindeutet. In der vorliegenden Studie ergab sich daraus eine umgekehrte Betrachtungsreihenfolge auf Objekt und Rezipient bei PP und DO Konstruktionen. Auf diese Weise lässt sich anhand der Blickfolgebewegungen die geäußerte Satzstruktur erschliessen. Die Auswertung der „Auge-Sprach,, Spanne konnte ebenfalls existierende Ergebnisse replizieren (Meyer et al., 1998; Meyer et al., 2000) und zeigt, dass die Zeitspanne zwischen Fixation und Benennung zwischen 400 und 700 msec beträgt, wobei zu Beginn der Äusserung bereits das nächst zu benennende Bildelement fixiert wurde. Basierend auf den vorhandenen Studien bekräftigt das die Annahme, dass Bilder bis zum Zugriff auf den phonologischen Code fixiert werden und bei Beginn der Artikulation das nächste Element betrachtet wird. Es lässt sich vermuten, dass diese Verarbeitungsweise dazu dient mögliche Fehlerquellen bis zum erfolgreichen Zugriff auf die Phonologie eines Wortes auszuschliessen. Eine Ausnahme davon scheinen bei der Agent-links Bedingung DO Sätze zu bilden. Hier war die „Auge-Sprach,, Spanne kürzer und Probanden tendierten dazu ein Objekt weiter zu fixieren während sie die Benennung begannen. Es lässt sich nur vermuten, dass die Probanden hier eine andere Strategie verwendeten, da bei dieser Bedingung präpositional Konstruktionen den Normalfall und alles andere eine Ausnahmen darstellte. Das Fixieren auf das Bildelement bis zum Sprachbeginn kann eine Strategie sein, um Fehleranfälligkeit bis einschliesslich zur Artikulation möglichst gering zu halten. Da dadurch ein Fixationswechsel auf ein anderes Element entfällt, scheint es naheliegend, dass auch die „Auge-Sprach,, Spanne kürzer ist. Die vorliegende Arbeit zeigte wie die Registrierung von Augenbewegungen dazu dienen kann die Interaktion von visueller Aufmerksamkeit und linguistischen Prozessen bei der Produktion syntaktisch komplexer Konstruktionen zu untersuchen.

Literatur

- Bock, J.K. (1986). Meaning, sound, and syntax: Lexical priming in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 575-586.
- Deubel, H., & Schneider, W.X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, 36, 1827-1837.
- Griffin, Z.M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 4, 274-279.
- Hartsuiker, R., & Kolk, H.H.J. (1998). Syntactic persistence in Dutch. *Language and Speech*, 41, 143-184.
- Hayhoe, M. (2000). Vision using routines: a functional account of vision. *Visual Cognition*, 7, 43-64.
- Intraub, H. (1997). The representation of visual scenes. *Trends in Cognitive Sciences*, 6, 217-222.
- Irwin, D.E., & Gordon, R.D. (1998). Eye movements, attention and trans-saccadic memory. *Visual Cognition*, 5, 127-155.

- Meyer, A.S., & van der Meulen, F.F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin & Review*, 7, 314-319.
- Meyer, A.S., Sleiderink, A.M., & Levelt, W.J.M. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition*, 66, B25-B33.
- Pickering, M.J., & Branigan, H.P. (1999). Syntactic priming in language production. *Trends in cognitive sciences*, 4, 136-141.
- Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rayner, K., & Pollatsek, A. (1992). Eye movements and scene perception. *Canadian Journal of Psychology*, 46, 342-376.
- Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M., & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.
- Van der Meulen, F.F., Meyer, A.S., Levelt, W.J.M. (2001). Eye movements during the production of nouns and pronouns. *Memory and Cognition*, in press.
- Wolfe J.M. (1998). Visual Search. In: H. Pashler (Hrsg.), *Attention* (pp. 13-73). East Sussex, GB.: Psychology Press.