

Hirnhälfte an der Schrift-zu-Laut-Konversion beteiligt ist (*Indefrey, Gruber, Brown, Hagoort, Posse, Kleinschmidt*).

*Semantische Koordination zwischen Sprache und spontanen ikonischen Gesten: eine sprachvergleichende Untersuchung*

Wenn Menschen sprechen, bewegen sie nicht nur ihre Sprechwerkzeuge, sondern sie bewegen auch ihre Hände und andere Körperteile, um Gesten zu erzeugen. Gesten können nach ihren semiotischen Eigenschaften in mehrere Klassen eingeteilt werden. Bei einem bekannten Typ von Gesten ist die Beziehung zwischen Form und Inhalt durch Konvention festgelegt; ein Beispiel ist das OK-Zeichen, bei dem Daumen und Zeigefinger einen Ring bilden. Eine andere Gestenklasse weist keine konventionalisierte Form - Inhaltsrelation auf. Hier wird situativ und momentan durch ikonische Abbildung des Referenten repräsentiert. Diesen Typ nennt man ‚spontane ikonische Gesten‘; er steht im Zentrum dieses Berichts.

Es gibt Evidenz dafür, daß spontane ikonische Gesten nicht nur aufgeführt werden, um dem Adressaten die Botschaft visuell zu präsentieren, sondern daß sie ein integraler Bestandteil von Sprachproduktionsprozessen sind. So ist die Information, die durch Sprache und Gesten übermittelt wird, normalerweise nicht streng komplementär, sondern sie überlappt. Darüberhinaus ist bekannt, daß spontane ikonische Gesten auch dann produziert werden, wenn den Sprechern bewußt ist, daß der Adressat die Gesten nicht sehen kann (z. B. am Telefon). Es stellt sich die Frage, in welcher Weise Sprachproduktion und die Erzeugung spontaner Gesten interagieren. Die hier dokumentierte Studie versucht, diese

Frage durch eine sprachvergleichende Untersuchung der Informationskoordination zwischen Sprache und Gesten zu beantworten.

In der Literatur werden drei konkurrierende Auffassungen über die Beziehung zwischen Gestik und Sprachproduktion vertreten. Die erste dieser Auffassungen ist die ‚Freie-Bildinhalte-Hypothese‘ (‚Free Imagery Hypothesis‘), nach der Gesten auf der Grundlage von bildhaften Vorstellungen erzeugt werden, die aus dem Langzeitgedächtnis von Ereignissen stammen oder aus bestimmten Gedankenprozessen. Gesten werden nach dieser Vorstellung ‚vorsprachlich‘ produziert, d. h. unabhängig von sprachlichen Repräsentationen. Aus dieser Hypothese läßt sich die Voraussage ableiten, daß die in Gesten kodierte Information nicht davon beeinflusst sein sollte, wie diese Information sprachlich ausgedrückt werden kann.

Die zweite Auffassung ist die ‚Lexikalische-Semantik-Hypothese‘ (‚Lexical Semantics Hypothesis‘), nach der Gesten von semantischen Merkmalen lexikalischer Einheiten der zu produzierenden Rede generiert werden. Mit anderen Worten, zunächst wird ein geeigneter Eintrag in einem semantisch strukturierten mentalen Lexikon abgerufen, und daraufhin werden ikonische Gesten auf der Grundlage von semantischen Merkmalen dieses Lexikoneintrags generiert. Nach dieser Hypothese sollten Gesten keine Information systematisch kodieren, die nicht auch sprachlich kodiert wird.

Die dritte Hypothese über den Zusammenhang von Gestik und Sprachproduktion ist die ‚Schnittstellenhypothese‘ (‚Interface Hypothesis‘). Nach dieser Auffassung werden Gesten auf der Grundlage der Schnittstellenrepräsentation zwischen Sprache und räumlich-motorischem Denken generiert. Diese Schnittstellenre-

präsentation ist eine spatio-motorische Repräsentation, die den jeweiligen Möglichkeiten der einzelsprachlichen Enkodierung optimal angepaßt ist. Der Mechanismus, der diesen Informationsverpackungsprozeß steuert, wird ‚Schnittstellenmechanismus‘ (‚Interface Mechanism‘) genannt. Der Schnittstellenmechanismus leistet die konzeptuelle Planung der Rede mittels zwei einander entgegengesetzten Prinzipien, die die Form der Schnittstellenrepräsentation bestimmen. Das erste dieser beiden Prinzipien besteht darin, die zu versprachlichende Information so weit als möglich den Bedingungen des einzelsprachlichen Codes anzupassen: bestimmte Aspekte der zu versprachlichenden Information werden ausgewählt, auf einer anderen Abstraktionsebene rekonstruiert, in kleinere Einheiten aufgeteilt und neu gruppiert. Das entgegengesetzte Prinzip besteht in der maximalen Bewahrung der nicht-sprachlichen Eigenschaften der zu übermittelnden Information. Entsprechend sagt die Schnittstellenhypothese voraus, daß die den Gesten zugrundeliegenden Bildinhalte *gleichzeitig* bestimmt werden durch 1) einen sprachlichen Ausdruck, der leicht zugänglich ist und kompakt genug, um innerhalb einer Planungseinheit des Sprachproduktionsprozesses formuliert und artikuliert zu werden (entgegen der Freie-Bildinhalte-Hypothese), und 2) die Eigenschaften der zu übermittelnden spatio-motorischen Information (entgegen der Lexikalische-Semantik-Hypothese).

In der hier vorgestellten Studie werden die Voraussagen dieser drei Hypothesen im Sprachvergleich überprüft. Als Testfall dienen dabei bestimmte Stimulusereignisse, die in verschiedenen Sprachen unterschiedlich sprachlich kodiert werden. Bei den zu vergleichenden Sprachen handelt es sich um das amerikanische

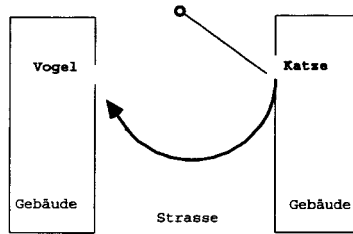
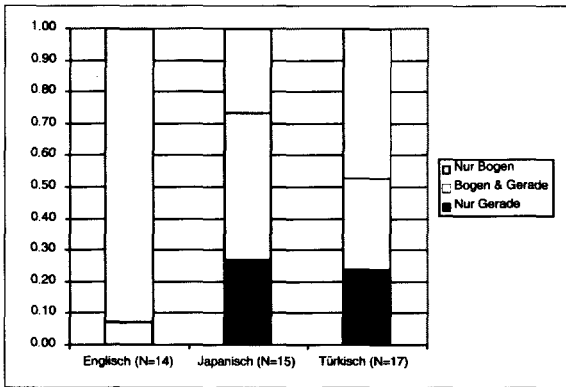


Abb. 2: Schematische Darstellung der Schwingszene im Stimulus.

Englisch, das Türkische und das Japanische. Erwachsenen Sprechern dieser Sprachen wurde der gleiche Zeichentrickfilm gezeigt. Sie wurden aufgefordert, die Geschichte für eine Person nachzuerzählen, die den Film nicht gesehen hatte. In dem Zeichentrickfilm schwingt eine Katze an einem Seil von einem Gebäude zu einem anderen, um einen Vogel zu fangen (Abb. 2). Im Türkischen und Japanischen wird diese Szene auf andere Weise sprachlich kodiert als im amerikanischen Englisch. Bei der Auswertung wird die gestische und sprachliche Repräsentation dieser Szene in den drei Sprachen analysiert.

In den Ausdrucksmitteln des Türkischen und Japanischen existiert kein einzelner sprachlicher Ausdruck, der die bogenförmige Bahn des Bewegungseignisses beschreibt. Die japanischen und die türkischen Versuchspersonen waren ausnahmslos nicht in der Lage, diese Trajektorform in ihren Beschreibungen der Szene sprachlich zu kodieren. Im Englischen dagegen steht das gebräuchliche Wort *swing* ‚schwingen, schaukeln‘ zur Verfügung, um diese Trajektorform von Bewegungsereignissen zu bezeichnen. 94% der Englischsprecher benutzten dieses Wort in ihren Beschreibungen.

Gestisch kann die ‚Schwing‘-Szene auf zwei unterschiedliche Weisen repräsentiert werden: als reine Ortsveränderung, in Abstraktion von der bogenförmigen Bahn



**Abb. 3:** Procentsatz des Informanten mit den drei Gebrauchsmustern von Bogen- und Geradengesten bei der Beschreibung der Schwingenszene.

(eine „Geradengeste“, es wird gestisch eine gerade Linie beschrieben) oder als Ortsveränderung in Kombination mit dem Bogentrajektor (eine „Bogengeste“). Die gestische Beschreibung der Schwing-Szene variiert parallel zu der Variation in der einzelsprachlichen Beschreibung. **Abbildung 3** zeigt den Procentsatz der Sprecher in jeder der drei Sprachen, die ausschließlich Bogengesten verwenden, den Procentsatz von Sprechern, die ausschließlich geradlinige Gesten verwenden, sowie den Procentsatz von Versuchspersonen, die eine Kombination der beiden Typen von Gesten produzieren. Sprecher des Türkischen und des Japanischen produzieren geradlinige Gesten mit größerer Wahrscheinlichkeit in ihren Beschreibungen der Schwing-Szene als Sprecher des Englischen. Die Mehrzahl der englischsprachigen Versuchspersonen produziert ausschließlich Bogengesten. Dieses Ergebnis widerspricht der

Freie-Bildinhalte-Hypothese, nach der gestische Repräsentationen nicht durch die sprachlichen Kodierungsmöglichkeiten beeinflusst werden.

Darüber hinaus zeigt sich, daß Gesten, die die Schwing-Szene beschreiben, systematisch Informationen kodieren, die niemals sprachlich ausgedrückt werden. Vom Standpunkt des Betrachters aus findet die bogenförmige Bewegung in der Szene von links nach rechts statt. Diese Bewegungsausrichtung wird systematisch in den Gesten reproduziert, unabhängig davon, ob eine Bogengeste oder eine geradlinige Geste ausgeführt wird. Die **Tabellen 1** und **2** geben die Procentsätze der Gestentokens über alle Versuchspersonen summiert an, die unter die drei relevanten Kategorien der Kodierung der horizontalen Bewegungsrichtung fallen.

Die Mehrzahl der Gestentokens kodiert die laterale Orientierung der Schwingenszene so, wie sie von der Versuchsperson gesehen wird (d. h. linksgerichtet), und nur sehr wenige Gestentokens haben die entgegengesetzte Orientierung. Frühere Untersuchungen, in denen Gesten mit demselben Stimulus und derselben Methode elizitiert wurden, bestätigen eine generell konsistente gestische Abbildung der Links-Rechts-Orientierung der Stimulusereignisse. Es kann daher geschlossen werden, daß die Schwingereignisgesten unabhängig von der Trajektorform und dem Sprachtyp systematisch die laterale Orientierung der Schwingenszene ko-

**Tab. 1:** Seitlicher Bias bei den Tokens für Bogengesten („links“ und „rechts“ ist abhängig vom Standpunkt des Sprechers. N = Gestentokens).

	Links-Bias	Rechts-Bias	Nur weg vom Körper
Türkisch (N = 20)	85%	0%	15%
Japanisch (N = 23)	74%	0%	26%
Englisch (N = 22)	77%	0%	23%

	Links-Bias	Rechts-Bias	Nur weg vom Körper
Türkisch (N = 10)	60%	0%	40%
Japanisch (N = 13)	62%	8%	31%
Englisch (N = 4)	75%	0%	25%

**Tab. 2:** Seitlicher Bias bei den Belegen für Geradengesten („links“ und „rechts“ ist abhängig vom Standpunkt des Sprechers. N = Gestentokens).

dieren, die niemals versprachlicht wird. Dieses Resultat ist in Übereinstimmung mit der Schnittstellenhypothese, nicht jedoch mit der Lexikalische-Semantik-Hypothese, nach der Gesten nur diejenige Information kodieren, die auch sprachlich vermittelt wird.

Zusammenfassend läßt sich zeigen, daß der gestische Ausdruck der Schwingszene sowohl systematische linguistische Variation aufweist als auch Ähnlichkeiten zeigt, so wie von der Schnittstellenhypothese vorausgesagt. Die einzelsprachliche Variation in der gestischen Repräsentation der Schwingszene folgt dem gleichen Muster wie die Variation in der

sprachlichen Kodierung der Ereignisinformation. Gleichzeitig sind in den Schwingszenengesten systematisch räumliche Informationen repräsentiert, die sprachlich nicht kodiert werden. Die laterale Orientierung der Schwingszenengesten kodiert in allen drei Sprachen systematisch die Ausrichtung der Schwingbewegung nach links, unabhängig von der Kodierung des Bogens. Diese Resultate sprechen dafür, daß Gesten durch den Schnittstellenmechanismus zwischen räumlichem Denken und Sprechen erzeugt werden. Mit diesem Mechanismus wird räumliche Information selektiert, so daß sie einfach verbalisiert werden kann (*Kita, Özyürek*).