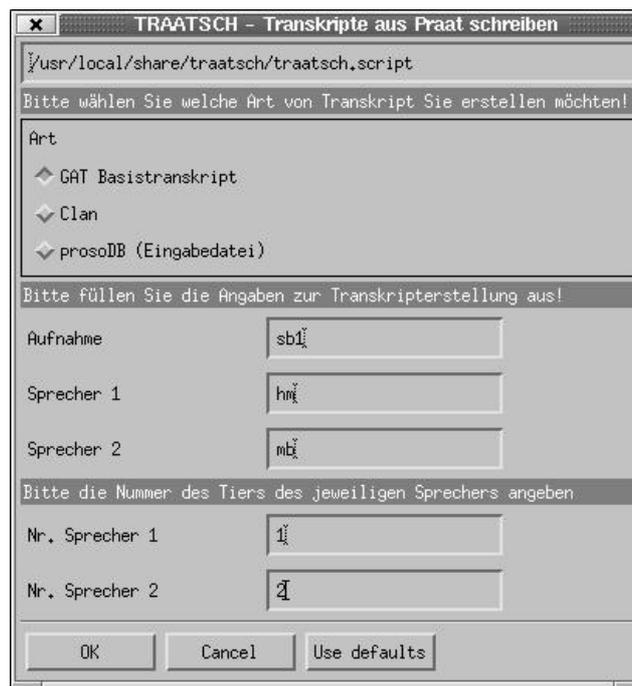


## Traatsch - Ein Werkzeug zur Konvertierung von Transkripten aus Praat

Frank Kügler

**Kurzbeschreibung:** *Traatsch* ist ein Werkzeug zur Konvertierung von Transkripten aus *Praat*. Ein in *Praat* erstelltes Transkript einer Sounddatei kann wahlweise in eine Textdatei nach GAT-Konventionen, in eine *Clan*-Datei oder in eine Eingabedatei für die Prosodiedatenbank *prosoDB* konvertiert werden. *Praat* ermöglicht durch die Transkripterstellung eine Verknüpfung von Text und Sound, die für weitere phonetisch akustische Analysen in *Praat* notwendig ist. Mithilfe von *Traatsch* wird der Transkripttext in ein für die Gesprächsanalyse notwendiges Transkript nach GAT-Konventionen konvertiert. *Traatsch* ist kostenlos verfügbar.

**Bezugsquelle:** *Traatsch* ist frei verfügbar und im Internet unter <http://talk.to/traatsch> zum Herunterladen bereitgestellt. Es gibt zwei Versionen: (1) für Linux und (2) für Windows.



## 1. Einführung

In diesem Beitrag wird vorgeschlagen, die Transkription gesprochener Sprache im Phonetikprogramm *Praat*<sup>1</sup> durchzuführen. Dieses Vorgehen bietet zwei entscheidende Vorteile:

(1) Durch die Transkription in *Praat* entsteht neben dem eigentlichen Transkripttext gleichzeitig eine zeitliche Verknüpfung von Text und Sound. Das heißt, es entstehen Intervalle, also einzelne Phrasen oder Turns, die die Grundlage für weitere phonetisch akustische Analysen in *Praat* bilden. Außerdem können mithilfe der Transkription sehr genaue Dauermessungen in *Praat* durchgeführt werden, so zum Beispiel bei Sprecherüberlappungen.

(2) Da für die Arbeit mit gesprochener Sprache ein lesbares Transkript unerlässlich ist, wurde *Traatsch* entwickelt. *Traatsch* konvertiert das Transkript, in *Praat* aus Intervallen bestehend, in eine Textdatei nach GAT-Konventionen (Selling et al. 1998).

Das Gesprächsanalytische Transkriptionssystem (GAT) wurde gewählt, weil es sich seit einigen Jahren im deutschsprachigen Raum als Standard für die Transkription gesprochener Sprache etabliert hat. Die Idee, die dem GAT zugrunde liegt, ist eine weitgehende Theorienunabhängigkeit, um allen Linguisten den Gebrauch des Systems zu ermöglichen. GAT beschreibt ein allgemeines Layout eines Transkriptes und die zu transkribierenden Kategorien, wobei es versucht einen Mindeststandard anzubieten, der in jedem Textverarbeitungsprogramm verwendet werden kann.

Es ist unbestritten, dass das Transkript ein zentrales Datum in der Analyse der gesprochenen Sprache darstellt, worauf Schwitalla (1997:15) verweist:

Dennoch können wir gesprochene Sprache als einen Forschungsgegenstand nicht anders als durch seinen einzigen Gegenpart, das Geschriebene (und dessen Varianten: das Gedruckte, das Getippte, das auf dem Bildschirm Erscheinende), in den Blick bekommen.

Es ermöglicht das Identifizieren von Turns, Sprecherwechseln, Handlungs- und Erzählsträngen, Nebensequenzen, kohäsiven Beziehungen, konditionellen Relevanzen etc., bietet einen Überblick über das Gespräch und ist dauerhaft vorhanden, so dass einzelne Sequenzen oder Abschnitte beliebig lange und wiederholt zur Analyse herangezogen werden können (vgl. Deppermann 1999).

Jedoch gibt es zahlreiche Ansätze und unterschiedliche Methoden der Transkription (für einen Überblick siehe Ehlich und Switalla 1976, Edwards und Lampert 1993). Neben der Transkription mit Kassettenrecorder und Textverarbeitungsprogramm kommen vorhandene Programme wie *Clan*<sup>2</sup>, *Transcriber*<sup>3</sup> oder *TuT*<sup>4</sup> für die Transkripterstellung zum Einsatz, wobei keines dieser Programme die Möglichkeit der akustischen Analyse der Sprachdaten vorsieht. Insbesondere im Vergleich zur Transkription im Textverarbeitungsprogramm, wo keinerlei zeitliche Verknüpfung zwischen Text und Sound vorliegt, und im Vergleich zur Transkription in *Clan*, wo eine zeitliche Verknüpfung nicht gleichzeitig, sondern erst nach einer erfolgten Transkription eingefügt werden kann, ist die Transkrip-

<sup>1</sup> *Praat* ist kostenlos unter <http://www.praat.org> erhältlich.

<sup>2</sup> *Clan* ist kostenlos unter <http://childes.psy.cmu.edu/html/clan.html> erhältlich.

<sup>3</sup> *Transcriber* ist kostenlos unter <http://morph ldc.upenn.edu/mirror/Transcriber> erhältlich.

<sup>4</sup> *TuT* ist kostenlos unter <http://www.ton-und-text.de> erhältlich.

tion in *Praat* vorteilhafter: Ein bestehendes Transkript im Wordformat ist wegen der fehlenden zeitlichen Verankerung des Textes im Sound nach *Praat* unmöglich zu konvertieren. Weitergehende akustische Analysen der Sprachdaten sind in so einem Fall nur mit einer erneuten Transkription in einem Schallanalyseprogramm möglich. Eine 'doppelte' Transkription in *Clan*, Texteingabe und spätere Verknüpfung des Textes mit dem Sound, ist vermeidbar. Schließlich verursachen fehlende Hilfsmittel zur Konvertierung der Transkriptdaten aus *Transcriber* oder *TuT* auch eine erneute Segmentierung der Sprachdaten, um akustische Analysen durchzuführen.

Diese doppelte Transkription zu vermeiden, hilft *Traatsch*: Wird direkt im Schallanalyseprogramm *Praat* transkribiert, kann dieser Text in ein Transkript nach GAT-Konventionen konvertiert werden, und die entstandene Segmentierung bildet gleichzeitig die Grundlage für jede weitere phonetisch akustische Analyse.

## 2. Transkription in *Praat*

*Praat* ist ein Schallanalyseprogramm, das vor allem für phonetische Zwecke genutzt wird. Gerade in der phonetischen Forschung werden Audiodaten verwendet, die mit Text verknüpft sind. Für die Analyse ist es notwendig, eine sehr genaue Verknüpfung von Sound und Text (Laute, Silben, Phrasen) zu haben. Der Beginn und das Ende einer Einheit, beispielsweise einer Silbe, markieren auf diese Weise ein Intervall, das die Grundlage jeder weiteren Analyse darstellt. So können zum Beispiel Silbendauern ermittelt werden. Deshalb werden Audiodaten vor einer Analyse zunächst mit entsprechendem Text verknüpft. Diese Vorgehensweise wird hier für die Gesprächstranskription vorgeschlagen, weil auf diese Weise neben dem Transkript gleichzeitig die Verknüpfung von Text und Sound entsteht, die für jede weitere phonetisch akustische Analyse notwendig ist. Die Intervalle bestehen in diesem Fall aus Phrasen oder Turns.

Vor der Transkription muss die Aufnahme in digitalisierter Form als Sounddatei auf dem Computer (entweder auf Festplatte oder auf CD) vorliegen. Die Digitalisierung kann in jedem verfügbaren Soundeditorprogramm oder in *Praat* selber durchgeführt werden. Hervorzuheben ist, dass *Praat* in der Lage ist, sehr unterschiedliche Soundformate mit unterschiedlichen Samplingsfrequenzen zu lesen. Dabei darf die Datei eine maximale Größe von zwei Gigabyte nicht übersteigen, was drei Stunden Stereo-CD-Qualität entspricht oder zwölf Stunden Mono-Qualität bei einer 16bit-Auflösung und 22,5 kHz Samplingsrate.

### 2.1 Vorgehensweise

Eine vorhandene Sounddatei wird in *Praat* über das Menü *Read/Open long sound file...* geladen. Die geöffnete Datei erscheint in der Liste der *Praat Objects*.

Nach dem Laden des *LongSounds* wird eine Datei generiert, in die das Transkript geschrieben wird. Dies geschieht über die Auswahl *To TextGrid...*, und in der Liste der *Praat Objects* erscheint eine sogenannte TextGrid-Datei mit dem Namen des LongSounds. Diese Datei muss vor Beendigung des Programms über das Menü *Write/Write to text file...* gespeichert werden, da *Praat* über keine automatische Speicherfunktion verfügt. Besteht bereits eine TextGrid-Datei zu einem

LongSound, wird diese über das Menü *Read/Read from file...* geladen ohne eine weitere TextGrid-Datei zu generieren.

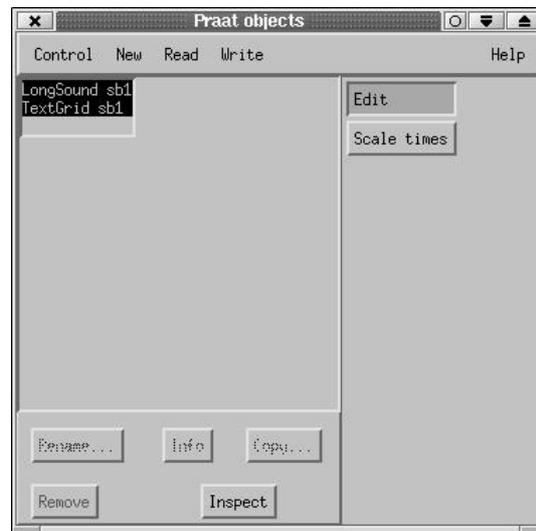


Abbildung 1. Aktivierung des Editiermodus.

Um zum Transkriptionsmodus zu gelangen, wird sowohl der LongSound als auch das TextGrid markiert, woraufhin eine Auswahl *Edit* erscheint (siehe Abbildung 1). Durch Anklicken dieser Auswahl erscheint ein Fenster, bestehend aus dem Oszillogramm der geöffneten Audiodatei und mehreren, sich darunter befindenden Zeilen, sogenannte 'tiers' (siehe Abbildung 2), in die Text eingefügt werden kann. Diese Zeilen sind mit dem LongSound zeitlich verknüpft. Hier wird in sogenannten Intervallen eine Phrase markiert, was einer Zeile oder einem Turn in GAT entspricht (siehe Abbildung 2).

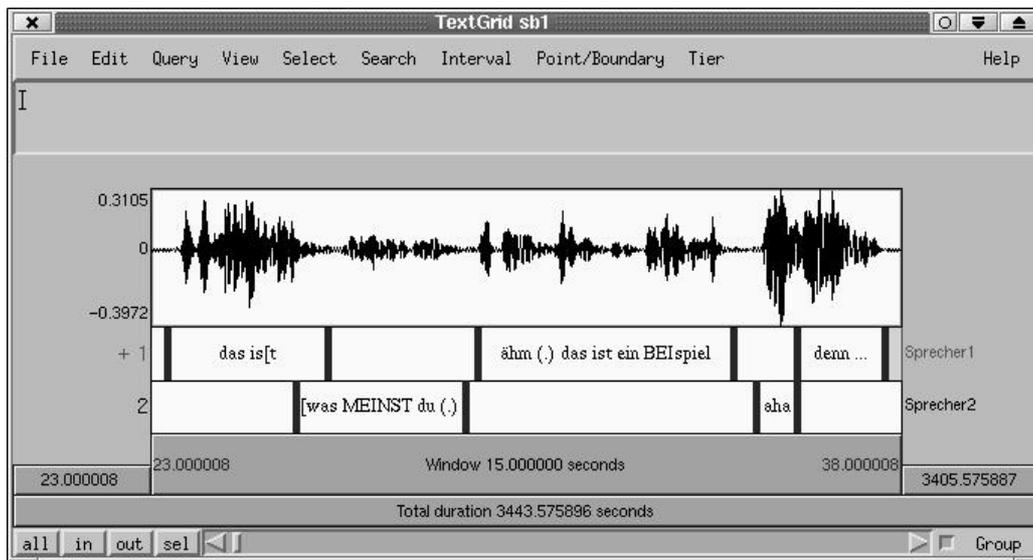


Abbildung 2. Sound verknüpft mit Transkription im TextGrid.

Ein Intervall wird erstellt, indem eine beliebige Zeitspanne im Sprachsignal markiert wird. Die Markierung befindet sich automatisch im Textfenster und muss dort durch Anklicken bestätigt werden. Daraufhin kann der Text in das Intervall

eingefügt werden. Die Grenzen der Intervalle lassen sich durch Anklicken mit festgehaltener Maustaste verschieben. Dadurch kann eine Feinabstimmung des Intervalls in Bezug auf den Sound vorgenommen werden.

Durch Anklicken einzelner Markierungen oder des unteren Balken *Window* (siehe Abbildung 2) wird der markierte Bereich beziehungsweise der Sound, der im Fenster abgebildet ist, abgespielt. Je nach Länge der Markierung erhält man so die Möglichkeit sehr fein und genau zu transkribieren.

Bei der Transkription von Gesprächen ist für jeden Sprecher eine eigene Zeile vorgesehen. In Abbildung 2 handelt es sich um zwei Zeilen, jeweils eine für Sprecher 1 und Sprecher 2. Dies ist eine notwendige Bedingung, damit *Traatsch* ein Transkript erstellen kann. Außerdem wären Sprecherüberlappungen in nur einer Zeile für alle Sprecher nicht zu transkribieren. Neben dem eigentlichen Text müssen weitere Elemente transkribiert werden, die im Basistranskript vorhanden sein sollen, wie beispielsweise Pausenmarkierungen und Intonationsmarkierungen.

Mithilfe der segmentierten Gesprächsdaten können in *Praat* alle erdenklichen Messungen durchgeführt werden, wie zum Beispiel Analysen der Intonation oder Messungen der Pausenlänge oder Sprecherüberlappungen. Außerdem können mit dieser Art der Transkription auditiv festgestellte Merkmale apparativ mithilfe akustischer Analysen verifiziert und vor allem anschaulich dargestellt werden, da *Praat* über einen bedienungsfreundlichen Modus zur Graphikerstellung verfügt.

### 3. Traatsch

#### 3.1 Technische Voraussetzungen

*Traatsch* ist auf der Basis sogenannter 'Open Source Software' entwickelt, das heißt, die benötigten Softwarekomponenten stehen Anwendern kostenlos im Internet zur Verfügung. Das Programmpaket *Traatsch* läuft auf den Betriebssystemen Windows<sup>5</sup> und Linux. Für die Arbeit mit *Traatsch* benötigt man *Praat* und *Perl*. *Praat* steht unter <http://www.praat.org> kostenlos zur Verfügung. Unter Linux wird *Perl* standardmäßig installiert. Für Windows steht ein kostenloses Paket unter <http://www.activestate.com> zur Verfügung. Hierbei ist zu beachten, dass das *MSI-package* von *ActivePerl* gewählt wird.

#### 3.2 Installation

*Traatsch* basiert auf einem *Praatskript* und muss in *Praat* installiert werden: Über das Menü *Controll/Open script...* wird eine Eingabemaske geöffnet, in der das Praatskript *traatsch.script* geladen werden muss. In dem Skriptfenster wird das Skript über das Menü *File/Add to fixed menu...* in die Menüleiste der *Praat Objects* unter dem Menüpunkt *New* mit dem Kommando *Traatsch...* eingebunden.

---

<sup>5</sup> Windows Systeme ab Windows 95, wobei ein Mindestspeicherplatz von 35 MB auf der Festplatte zur Verfügung stehen muss.

### 3.3 Funktion

Als Ausgangspunkt für die Konvertierung benötigt *Traatsch* eine TextGrid-Datei aus *Praat*. Diese muss so organisiert sein, dass für jeden Sprecher eine Zeile existiert (vergleiche Abbildung 2). Daraufhin kann *Traatsch* über das Menü *New/Traatsch...* aufgerufen werden. In einer Eingabemaske müssen folgende Angaben eingetragen werden (siehe Abbildung 3): Zunächst muss sich für die Art der Konvertierung entscheiden werden. Entweder soll eine Textdatei nach GAT-Konventionen, eine *Clan*-Datei oder eine Eingabedatei für die Prosodiedatenbank *prosoDB* (Gilles 2001) erzeugt werden. Danach muss der Aufnahmename eingegeben werden. Dieser entspricht dem Dateinamen der TextGrid-Datei. Es folgen die Sprecherkürzel (hier: Sprecher hm / Sprecher mb) und die entsprechenden Zeilennummern (tier-Nummer) der einzelnen Sprecher aus dem TextGrid (hier: Zeile 1 / Zeile 2). Schließlich wird *Traatsch* durch Anklicken des *OK*-Knopfes gestartet.

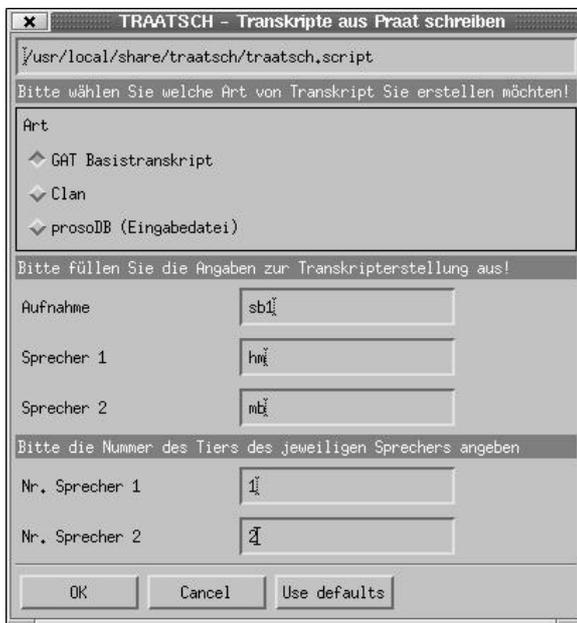


Abbildung 3. *Traatsch* Eingabemaske.



Abbildung 4. Information über das Ergebnis.

Als Ergebnis erhält man eine Information, unter welchem Namen und in welchem Ordner die Ergebnisdatei gespeichert worden ist (siehe Abbildung 4).

### 3.4 Vorhandene Konvertierungsmodule

*Traatsch* besteht bisher aus drei Modulen, die die Transkripttexte der TextGrids in drei verschiedene Ausgaben konvertieren (siehe Tabelle 1, erste Zeile). Zunächst kann ein GAT-Basistranskript in Form einer Textdatei, die mit jedem beliebigen Textverarbeitungsprogramm bearbeitet werden kann, erstellt werden. Dieses Basistranskript erfüllt die layouttechnischen Anforderungen des GAT (vgl. Selting et al. 1998). Alle weiteren Aspekte des Basistranskriptes müssen, wie bei jeder anderen Transkription, bereits im TextGrid vorhanden sein. Insbesondere handelt es sich um die prosodischen Elemente wie Pausen und Intonationsmarkierungen

(vergleiche Abbildung 2). Für zukünftige Versionen ist geplant, die Pausen-Transkription automatisch durchzuführen, sofern im TextGrid eine Extra-Zeile (tier) angelegt wird, in der Intervalle gesetzt werden, die die jeweilige Pause umfassen.<sup>6</sup>

Außerdem kann eine *Clan*-Datei erzeugt werden. Diese enthält alle Konventionen eines *Clan*-Dokumentes und liefert bereits eine Verknüpfung mit dem Sound.

Schließlich kann das TextGrid-Transkript in eine Eingabedatei für die Prosodiedatenbank *prosoDB* konvertiert werden. Diese Eingabedatei wird mit einem Update-Befehl in die Datenbank eingelesen.

*Tabelle 1: Vorhandene und geplante Konvertierungsmodule von Traatsch. Vorhandene Module sind in der ersten Zeile aufgeführt, geplante Module sind in den Zeilen 2,3 und 4 aufgeführt. \* = technisch nicht möglich, da eine Zeitverknüpfung fehlt.*

TRAATSCH	
INPUT	OUTPUT
vorhandene Module: <i>name</i> .TextGrid ( <i>Praat</i> )	Textdatei nach GAT (Basistranskript) <i>Clan</i> -Datei (mit Soundverknüpfung) <i>prosoDB</i> (Eingabedatei)
geplante Module: <i>name</i> .txt (z.B. aus Word)	<i>Clan</i> -Datei (ohne Soundverknüpfung) <i>prosoDB</i> (Eingabedatei) * <i>Praat</i> (TextGrid)
geplante Module: <i>name</i> .ca ( <i>Clan</i> -Datei mit Soundverknüpfung)	Textdatei nach GAT (Basistranskript) <i>prosoDB</i> (Eingabedatei) <i>Praat</i> (TextGrid)
geplante Module: <i>name</i> .ca ( <i>Clan</i> -Datei ohne Soundverknüpfung)	Textdatei nach GAT (Basistranskript) <i>prosoDB</i> (Eingabedatei) * <i>Praat</i> (TextGrid)

### 3.5 Geplante Module

Neben den bestehenden drei Modulen ist es geplant weitere Module zur Transkriptkonvertierung zu entwickeln (siehe Tabelle 1, Zeile zwei bis vier).

Ein Modul soll zum Beispiel eine Textdatei (zum Beispiel ein Transkript aus Microsoft Word) in eine *Clan*-Datei umwandeln. Diese hätte jedoch keine Soundverknüpfung, da im Textverarbeitungsprogramm keine Angaben über die zeitliche

<sup>6</sup> Ein relativ zuverlässiges Pausenmarkierungsskript ist bereits in der Testphase, so dass die Transkription der Pausen im TextGrid automatisch durchgeführt werden kann.

Verknüpfung zur Sounddatei vorliegen. Deshalb ist es auch unmöglich, bereits bestehende Transkripte im Wordformat nach *Praat* zu konvertieren.

Weiterhin sind Module geplant, die vorhandene *Clan*-Dokumente in TextGrid-Dateien konvertieren, in eine Textdatei nach GAT-Konventionen oder in eine Eingabedatei für *prosoDB*. Für die Konvertierung in ein TextGrid muss jedoch in *Clan* eine Soundverknüpfung vorhanden sein.

## Literatur

- Deppermann, Arnulf (1999): Gespräche analysieren. Opladen: Leske und Budrich.
- Edwards, Jane A. / Lampert, Martin D. (eds) (1993): Talking Data. Transcription and Coding in Discourse Research. Hillsdale: Lawrence Erlbaum.
- Ehlich, Konrad / Switalla, Bernd (1976): Transkriptionssysteme - Eine exemplarische Übersicht. In: Studium Linguistik 2, 78-105.
- Gilles, Peter (2001). *prosoDB*: Eine multimediale Datenbankumgebung für konversationelle und prosodische Analysen. In: Gesprächsforschung 2, 75-89.
- Schwitalla, Johannes (1997): Gesprochenes Deutsch. Eine Einführung. Berlin: Erich Schmidt.
- Selting, Margret / Auer, Peter / Barden, Birgit / Bergmann, Jörg / Couper-Kuhlen, Elizabeth / Günthner, Susanne / Meier, Christoph / Quasthoff, Uta / Schlobinski, Peter / Uhmman, Susanne (1998): Gesprächsanalytisches Transkriptionssystem (GAT). In: Linguistische Berichte 173, 91-122.

Frank Kügler  
Universität Potsdam  
Institut für Linguistik  
Postfach 601553  
14415 Potsdam  
kuegler@rz.uni-potsdam.de

Veröffentlicht am 12.2.2002

© Copyright by GESPRÄCHSFORSCHUNG. Alle Rechte vorbehalten.