

Sprachproduktionstheorien (1)

‘Articulatory Phonology’ und ‘Task Dynamics’

Gesture als Kernbegriff der Phonetik

Besonders geprägt durch Forscher der Haskins Labs (u.a. Browman & Goldstein, Fowler)

“By vocal tract (or phonetic) gestures, we mean coordinated actions of vocal tract articulators that achieve some linguistic goal. It is important to note that according to this definition, gestures are not movements of individual articulators. For example, a bilabial consonant such as /b/ corresponds to one gesture—the complete closure of the lips—achieved by the coordination of the movements of three articulators: the jaw and the two lips”

(Galantucci et al., 2006, p. 377)

Jede phonologische Theorie geht von einer begrenzten Anzahl von Grundbausteinen aus, die zu einer sehr großen Anzahl größerer linguistischer Einheiten kombiniert werden können.

In der articulatory phonology sind die Grundbausteine (“Atome”) eben *gestures*.

Charakteristisch für diesen Ansatz: Der Versuch, die Grenzen zwischen Phonologie und Phonetik aufzuheben.

gestures sind demnach sowohl *units of combination and contrast* (“phonologisch”) als auch *units of action* (“phonetisch”)

“units of combination (atoms) are constriction actions of vocal organs” (Goldstein & Fowler)

Die Grundstruktur des Modells ergibt sich aus den 5 Organen, die (relativ) unabhängig voneinander Verengungen bilden können:

Lippen

Zungenspitze

Zungendorsum

Velum

Glottis

(ev. noch Zungenwurzel (tongue root) als 6. Organ)

====> Darstellung auf getrennten “*tiers*”

“As abstract, discrete, dynamic linguistic units, the gestures are invariant across different contexts. Yet, because the gestures are also inherently spatio-temporal, it is possible for them to overlap in time. Such overlapping activation of several invariant gestures results in context-varying articulatory trajectories, when the gestures involve the same articulators” (Browman & Goldstein, 1990, p. 342).

Um besser eingrenzen zu können, was hier mit *abstract, discrete, dynamic* gemeint ist:

Einordnung des Gesture-Begriffs in einem Gesamtmodell der Sprachproduktion.

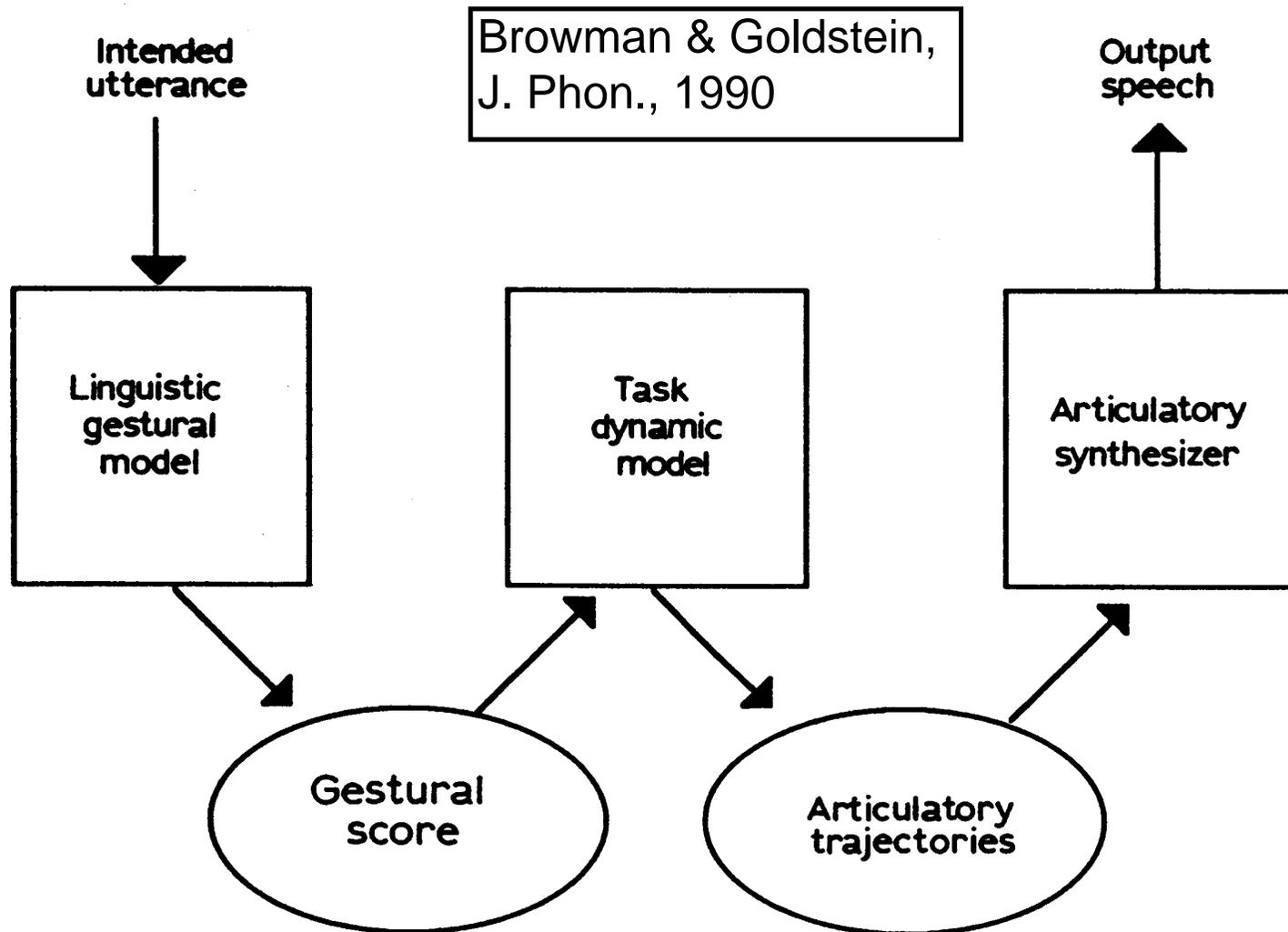


Figure 1. Gestural computational model.

Für das Verständnis des Modells ist folgende sehr sinnvolle Unterscheidung wichtig: ***tract variables vs. articulators***

	tract variable	articulators involved
LP	lip protrusion	upper & lower lips, jaw
LA	lip aperture	upper & lower lips, jaw
TTCL	tongue tip constrict location	tongue tip, tongue body, jaw
TTCD	tongue tip constrict degree	tongue tip, tongue body, jaw
TBCL	tongue body constrict location	tongue body, jaw
TBCD	tongue body constrict degree	tongue body, jaw
VEL	velic aperture	velum
GLO	glottal aperture	glottis

The figure contains two diagrams of the vocal tract. The left diagram is a sagittal cross-section with arrows indicating the locations of tract variables: LP (lip protrusion) at the front, LA (lip aperture) at the lip opening, TTCL (tongue tip constrict location) and TTCD (tongue tip constrict degree) at the tongue tip, TBCL (tongue body constrict location) and TBCD (tongue body constrict degree) at the tongue body, VEL (velic aperture) at the velum, and GLO (glottal aperture) at the glottis. The right diagram is a similar cross-section with '+' markers indicating the articulators involved: upper lip, lower lip, jaw, tongue tip, tongue body center, velum, and glottis.

Figure 2. Tract variables and the articulators involved.

Aus den 5 “Basisorganen” ergeben sich 8 *tract variables*,
v.a. durch die paarweise Anordnung der Zungenvariablen in
‘*constriction location*’ und ‘*constriction degree*’

Die meisten *tract variables* werden durch mehrere *articulators*
umgesetzt.

Der gleiche *articulator* (‘*jaw*’, ‘*tongue body*’) kann bei mehreren *tract variables* vertreten sein.

Gestures sind in Hinblick auf die *tract variables*, **nicht** in Hinblick auf
die *articulators* spezifiziert (vgl. oben Zitat von Galantucci et al.)

Warum das sinnvoll, ist sehen wir weiter unten.

In einfachen Fällen, z.B. *gestures* wie 'lip closure', 'glottal opening', werden die *gestures* durch eine einzige *tract variable* umgesetzt.

Bei allen oralen Gesten werden aber 2 *tract variables* erforderlich sein.

Beispiel: hinteres [a] wird durch die *gesture* 'pharyngeal narrowing' implementiert. Dies erfordert eine Spezifikation für die 2 *tract variables* 'tongue body constriction location' und 'tongue body constriction degree'.

Näheres zu den Bausteinen *linguistic gestural model*, *gestural score*, *task dynamic model*

Annahme:

Ein Wort ist im Lexikon durch eine *gestural structure* spezifiziert, die angibt, welche *gestures* aktiviert werden müssen, und wie sie miteinander gekoppelt sind. Aufgrund dieser Information über die Koordinierung der Gesten, zusammen mit Information über die dynamischen Eigenschaften der einzelnen Gesten, berechnet das *linguistic gestural model*, wann die Gesten aktiviert sind und hält dies in einer '*gestural score*' fest.

In einer '*gestural score*' wird also sichtbar, welche *gestures* sich zeitlich überlappen.

Gestural structure (oben) und gestural score (unten) für “pen”

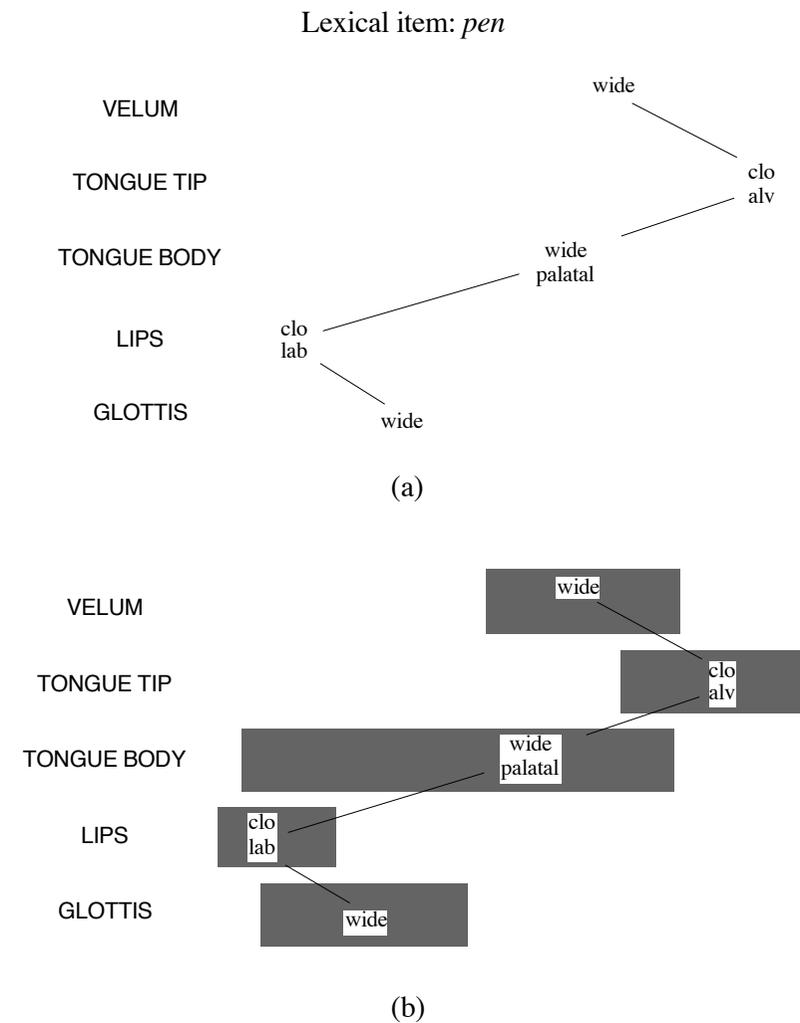


Figure 1: (a) Gestural structure for the lexical item *pen*. Lines connect gestures whose coordination is specified lexically in the computational model. (b) Gestural score. Horizontal extent of boxes represents temporal activation intervals. After Browman & Goldstein, 1995.

In der *gestural score* stecken auch die weiteren dynamischen Parameter für die mit der *gesture* verbundenen *tract variables*, z.B. Zielwerte für *tongue tip constriction location/degree* für einen alveolaren Verschluss.

Die *gestural score* bildet den Input zum *task dynamic model*. Dies berechnet den tatsächlichen zeitlichen Verlauf aller *tract variables*.

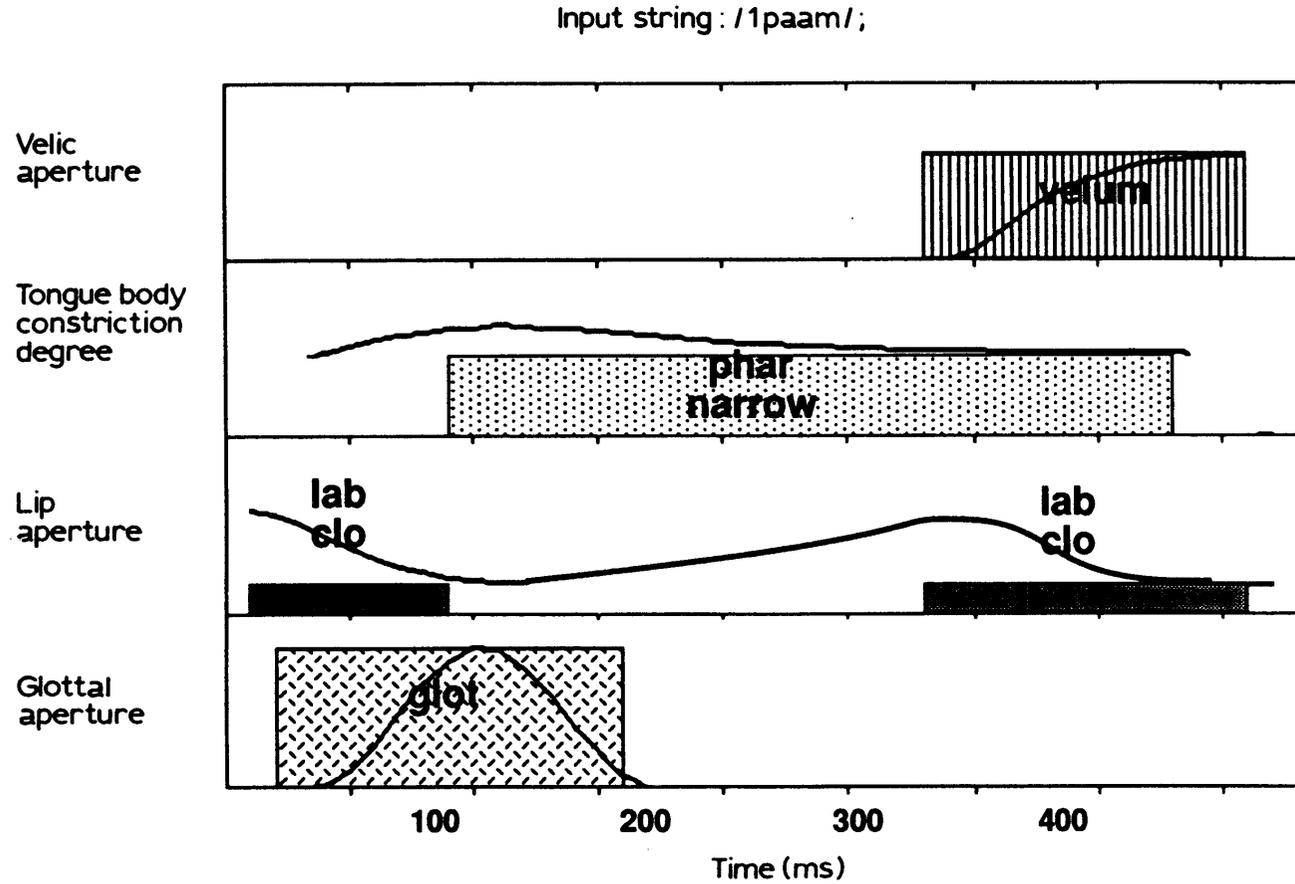


Figure 3. Gestural score and generated tract variable motions for “palm” (pronounced [pam]). The input is specified in ARPAbet, so IPA /pam/ = ARPAbet input string /paam/. The boxes indicate gestural activation, and the curves the generated tract variable movements. Within each panel, the height of the box indicates the targeted degree of opening (aperture) for the relevant constriction: the higher the box (or curve) the greater the amount of opening.

Dynamische Systeme

Das dynamische Verhalten einer *tract variable* wird durch ein gedämpftes Masse-Federsystem modelliert

- allerdings in etwas vereinfachter/'abgespeckter' Form (s.u.)

Allgemeine Formulierung eines solchen Systems:

$$m \cdot a + b \cdot v + k \cdot (x - x_0) = 0$$

m = Masse

b = Dämpfung

k = Steifigkeit

a = Beschleunigung

v = Geschwindigkeit

x = aktuelle Position der *tract variable*

x_0 = Ruheposition der *tract variable*

entspricht dem Bewegungsziel, z.B. in Hinblick auf
constriction location oder *constriction degree*

Spezielle Annahmen:

Masse = 1

Es geht um die Modellierung abstrakter Gesten, nicht physikalischer Artikulatoren

Steifigkeit und Dämpfung

In einem ungedämpften System regelt die Steifigkeit ('Federkonstante') die Frequenz mit der die Masse um die Ruheposition oszilliert.

Um normale, nichtzyklische Bewegungen zu modellieren, nimmt man hier 'kritische Dämpfung' an:

Die tract variable erreicht die Ruheposition ohne Überschießen.

Dann bestimmt die Steifigkeit, wie schnell die tract variable die Ruheposition erreicht.

Es hat sich gezeigt, dass man weitgehend mit zwei Werten für Steifigkeit auskommen kann: einem Wert für Konsonanten (schnelle Bewegungen), einem Wert für Vokale (langsame Bewegungen) (ev. noch mit einem dazwischen liegenden Wert für Glides ('Halbvokale'))
Deswegen sind oben in den *gestural scores* die Aktivierungsintervalle für die vokalischen Gesten länger.

Demo

Warum sind dynamische Systeme für ein Sprachproduktionsmodell attraktiv?

Gesten können diskret sein, d.h sie enthalten eine dynamische Spezifikation (v.a. für Ruheposition und Steifigkeit), die gleich bleibt, so lange die Geste aktiv ist.

Anders ausgedrückt: die Geste enthält keine Spezifikation, die den zeitlichen Ablauf der Artikulationsbewegungen Punkt für Punkt festlegt.

Aus der dynamischen Spezifikation lässt sich aber eine vollständige Bewegungsbahn in Raum und Zeit generieren.

Dabei ist Ziel der Bewegung (Ruheposition) und wie schnell das Ziel erreicht werden soll (Steifigkeit) vorgegeben.

Der genaue Weg dahin hängt vom Ausgangszustand des Systems ab, kann also beträchtlich variieren, muss aber nicht vom Sprecher Schritt für Schritt geplant werden.

Auch wenn eine ablaufende Bewegung gestört wird, wird der gleiche Endpunkt erreicht (mehr zu Perturbationen unten).

Der dynamische Ansatz verbindet also die diskrete Welt der Kategorien (Linguistik) mit der realen Welt kontinuierlicher Artikulationsbewegungen.

Was leistet das Kästchen *Task-dynamic model*?

Erste Antwort oben: generiert den zeitlichen Ablauf der *Tract variables* aufgrund der dynamischen Spezifikation der *gesture*.

Dabei müssen 'Konkurrenzsituationen' der folgenden Art aufgelöst werden:

Gesten können sich überlappen.

z.B. Gesten für Anlautkonsonant und Vokal werden in etwa zeitgleich aktiviert

Konkretes Beispiel: Silbe /pa/. Konsonant wird durch *tract variable lip aperture*, Vokal durch *tongue body constriction location/degree* realisiert.

Allen diesen *tract variables* ist der Artikulator 'jaw' zugeordnet.

====> Konkurrenz unter den Gesten um die Kontrolle dieses Artikulators.

Das *task dynamic model* berechnet die Bewegungsbahnen der Artikulatoren, die diese Konkurrenz berücksichtigen, aber gleichzeitig die in den *tract variables* linguistisch definierten Ziele ('*tasks*') einhalten.

Die Einteilung in *tract variables* vs. *articulators* bietet hier aber sehr viel Flexibilität:

Beispiel für die *tract variable lip aperture*:

Unendlich viele Kombinationen von Oberlippe, Unterlippe, Kiefer, um ein linguistisch definiertes Ziel wie “bilabialer Verschluss” zu erreichen.

Wird bereits während der Bildung von /p/ in /pa/ der Kiefer vom Vokal nach unten gezogen, kann der bilabiale Verschluss dadurch erreicht werden, dass die Unterlippe etwas mehr macht.

Solche ‘sets’ von Artikulatoren, die zusammen mit der übergeordneten Geste gemeinsam aktiviert werden, werden *coordinative structures* genannt.

In Anlehnung an Konzepte, die für andere motorische Systeme entwickelt wurden.

Durch dieses Konstruktionsprinzip ist das System nicht nur flexibel, sondern auch in der Lage das Problem der vielen **Freiheitsgrade** im Sprechapparat zu bewältigen.

Viele Artikulatoren, noch mehr Muskeln (auch für kurze Äußerungen ca. 70 Muskeln erforderlich).

Sprechen wäre gar nicht möglich, wenn alle diesen Möglichkeiten zu jedem Zeitpunkt im Detail geplant und gesteuert werden müsste.

Analogie Flugzeugsteuerung (Fowler et al., 1980, pp. 393-394):

Beschränkung der mechanischen Freiheitsgrade

Eine *coordinative structure* beinhaltet also nicht nur einen Artikulatorverbund, sondern auch die Beziehungen ('constraints') der Artikulatoren untereinander, die zum Erreichen des jeweiligen Ziels sinnvoll sind.

Wie *coordinative structures* neurophysiologisch zustandekommen, ist noch etwas mysteriös.

Aber durchaus plausibel, dass dieses Organisationsprinzip tatsächlich existiert

====> Perturbationsexperimente (Stichwort "Bite-block"; näheres am Semesterende)

Zwischenfazit Articulatory Phonology

Linguistisch relevante Vokaltraktverengungen als Grundbausteine der Sprachproduktion (und -Perzeption)

Gegensatz zu auditorisch orientierten Theorien wie DIVA
(Thema der Sitzung “Sprachproduktionstheorien(3): Phonetische Ziele”)

Durch die dynamische Spezifizierung haben Gesten auch eine inhärente zeitliche Strukturierung

Gegensatz zu traditionellen phonologischen Merkmalen

Auch durch die Orientierung an dynamischen Modellen eine sehr starke Berücksichtigung der Gesetzmäßigkeiten, wie Sprechbewegungen miteinander koordiniert werden.

Wie werden die “Atome” zu “Molekülen” verbunden?

====> nächste Stunde: zeitliche Strukturierung der Silbe

Literatur

- Browman, C. & Goldstein, L. (1990). *Tiers in articulatory phonology with some implications for casual speech*. In J. Kingston & M. Beckman, M. (eds.), *Papers in Laboratory Phonology I*. Cambridge University Press, Cambridge: 341-367.
- Browman, C. & Goldstein, L. (1990). *Gestural specification using dynamically-defined articulatory structures*. *J. Phonetics* 18, 299-320.
- Fowler, C. A., Rubin, P., Remez, R. E., & Turvey, M. T. (1980). *Implications for speech production of a general theory of action*. In B. Butterworth (Ed.), *Language production* (pp. 373-420). New York: Academic Press.
- Galantucci, B., Fowler, C. & Turvey, M. (2006). *The motor theory of speech perception reviewed*. *Psychon. Bull. Rev.* 13: 361-377.
- Pouplier, M. (2011). *The atoms of phonological representations*. In M. van Oostendorp, K. Rice, B. Hume, C. Ewen (eds.), *The Blackwell Companion to Phonology*. Wiley-Blackwell.