

Sprachproduktionstheorien (4).

**Regelung der Sprachproduktion durch afferente
Information**

Welche Arten von sensorischer Information sind beim Sprechen relevant?

1. Auditorische Information

2. Somatosensorische Information

Taktile Afferenzen (Zungenoberfläche, harter Gaumen)

Dehnung der Haut (Gesichtsbereich)

Muskelafferenzen (Länge, Änderung der Länge; vgl.

Muskelspindeln in den Folien zur Neurophysiologie);

allerdings vermutlich nicht überall vorhanden in der

Sprechmuskulatur

(3. Visuelle Information)

Teil 1. Auditorische Afferenzen

Erste Überlegungen:

Einerseits:

Auch nach langer Gehörlosigkeit (ev. nach Jahrzehnten) ist Sprache von Gehörlosen meistens noch gut verständlich.

Ist auditorisches Feedback also gar nicht so wichtig?

Andererseits

Sprecher reagieren ziemlich schnell, wenn sie plötzlich lautem Lärm ausgesetzt sind: **Lombard-Effekt** (Lane & Tranel, 1971)

Sie werden lauter, um für den Gesprächspartner bei ungünstigen Umweltbedingungen verständlich zu bleiben.

Dieses Verhalten muss natürlich über auditorisches Feedback gesteuert sein.

Wie schnell ist eigentlich die Reaktion auf plötzliche Änderungen im auditorischen Feedback?

Sprecher hört sich über Kopfhörer

Äußerung eines Dauerlauts (/ a::::: /)

Plötzliche Änderung durch den Computer des Signals an den Kopfhörer in Hinblick auf z.B. Lautstärke, Grundfrequenz, auch Formanten

Reaktion nach ca. 100-200ms

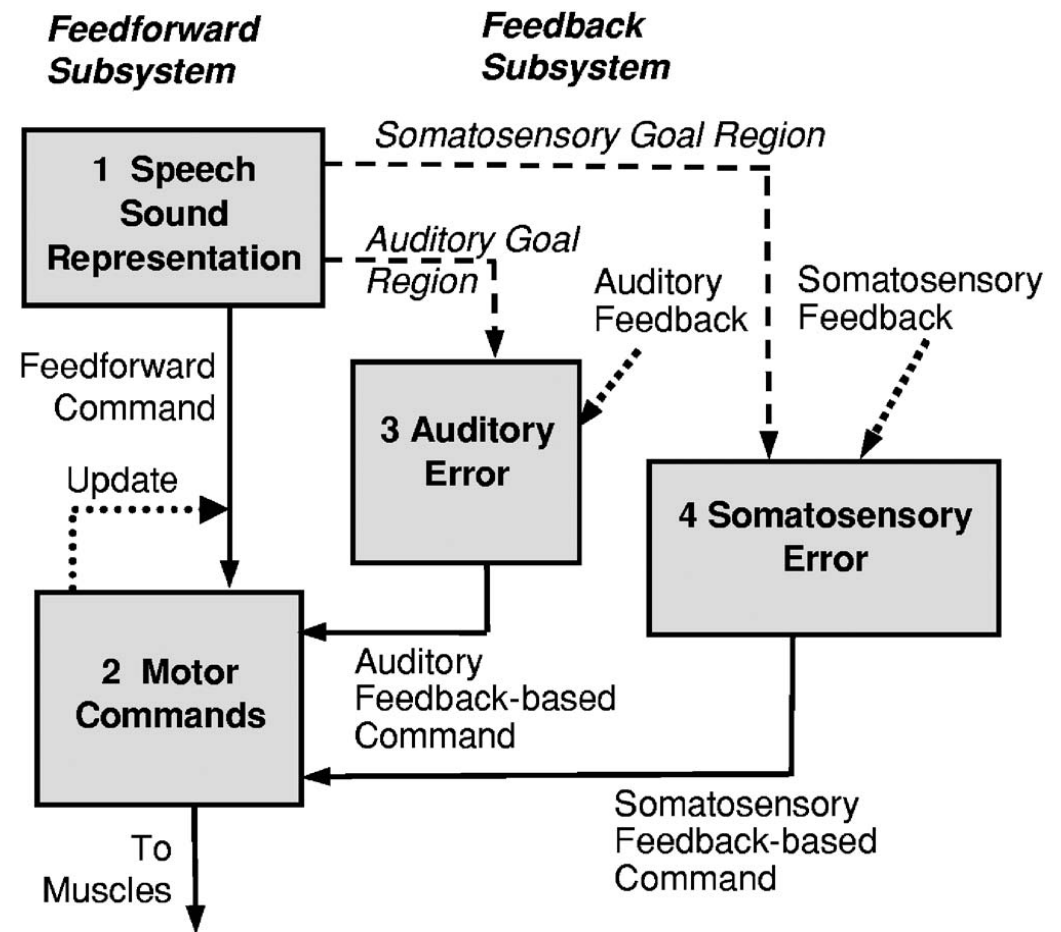
Wahrscheinlich schnell genug, um Lautstärke, Intonation, Deutlichkeit der Artikulation '*online*' (während der Äußerung) an Änderungen der Umweltbedingungen anzupassen.

Aber zu langsam, um einzelne Laute als '*closed-loop*'-Rückkopplungsschleife zu steuern.

closed-loop: Während der Bewegung wird Augenblick für Augenblick die sensorische Information mit dem für das Erreichen des Ziels erwarteten Zustand verglichen
===> Fehlersignal ===> Korrektur der Bewegung.
Wiederholung bis das Ziel erreicht wird.

Auf dem Hintergrund der bisherigen Überlegungen: Zweigleisiger Aufbau des Sprachproduktionssystems ist naheliegend.

Hier im DIVA-Modell:
Feedforward subsystem
vs.
Feedback subsystem



aus Lane et al, 2007

Welche Rolle spielt das Feedback-System bei der Artikulation einzelner Laute?

Selbstverständlich unbedingt erforderlich beim Erwerb der Lautsprache.

Aber nachher?

Tuning der Feedforward-Befehle
(‘Update’-Pfeil beim Kästchen ‘motor commands’)

‘Update’ heißt **nicht**, dass die gerade ablaufende Bewegung anhand des Feedbacks korrigiert wird (closed-loop), sondern dass die gespeicherten Befehle angepasst werden, und erst beim nächsten Versuch, den gleichen Laut zu artikulieren, wirksam werden.

Warum ist ein Tuning-Mechanismus nützlich?

Warum ist ein Tuning-Mechanismus nützlich?

Anpassung an körperliche Änderungen (Wachstum, Verletzungen)

Anpassung an die Umweltbedingungen

Sprechen in neuen Situationen (Pfeife im Mund)

Anmerkungen und Beispiele zum Ablauf des Tunings:

(1)

Gehörlosigkeit hat bei Kindern gravierende Auswirkungen auf das Sprechen als bei Erwachsenen, (auch wenn Sprachentwicklung zum Zeitpunkt der Ertaubung weit fortgeschritten ist), weil die körperliche Entwicklung noch nicht abgeschlossen ist.

(2) Fallbeispiel aus Perkell et al. (2000)

Erwachsene Patientin wird durch die Entfernung eines Tumors im Bereich des Hörnervs gehörlos.

Lautbildung (gemessen am spektralen Abstand zwischen /s/ und /ʃ/) ändert sich nicht.

Nach 70 Wochen weitere OP (als Folge der Tumorchirurgie) im Bereich des **motorischen** Zungennervs (N. Hypoglossus) erforderlich. ==> leichte Schwächung der Zungenmuskulatur links.

/s/-/ʃ/-Kontrast bricht innerhalb wenigen Wochen komplett zusammen.

==> ohne Gehör, keine Möglichkeit auf Änderungen der Artikulationsorgane zu reagieren.

(3) Untersuchungen von Perkell et al. bei Versorgung von erwachsenen Gehörlosen mit Cochlea-Implantat (z.B Lane et al. 2007)

Typischer Ablauf:

Vor der Versorgung mit CI ist die Artikulation oft etwas schlechter als bei Normalhörenden.

Unmittelbar (bis einigen Wochen) nach der Versorgung wird die Sprachproduktion oft sogar noch etwas schlechter, auch wenn die Patienten bereits besser verstehen als vorher.

Interpretation (anhand des DIVA-Blockschaltbilds):

Die gespeicherten auditorischen Repräsentationen (Kästchen 1) passen überhaupt nicht zu den neuen auditorischen Afferenzen.

Deswegen großer 'auditory error' (Kästchen 3), der aber zu keinen sinnvollen motorischen Befehlen und zu keinem sinnvollen Update führt.

===> zweistufiger Lernvorgang erforderlich

zuerst Neuaufbau der Repräsentationen in Kästchen 1

erst dann können die feedforward-Befehle angepasst und optimiert werden.

Nach einem Jahr typischerweise deutlich verbesserte Artikulation.

===> die angepassten feedforward-Befehle sind jetzt stabil, d.h sie funktionieren auch ohne auditorisches Feedback ganz gut (z.B durch Ausschalten des CI-Prozessors)

(4) Formantperturbationsexperimente

Typischer Ablauf:

Proband spricht einen mittleren Vokal wie engl. “bet” (sehr oft ...)

Auditorisches Feedback über Kopfhörer

In Echtzeit (ca. 20ms Verzögerung) werden die Formanten analysiert, und die Äußerung wird mit neuen Formantwerten resynthetisiert.

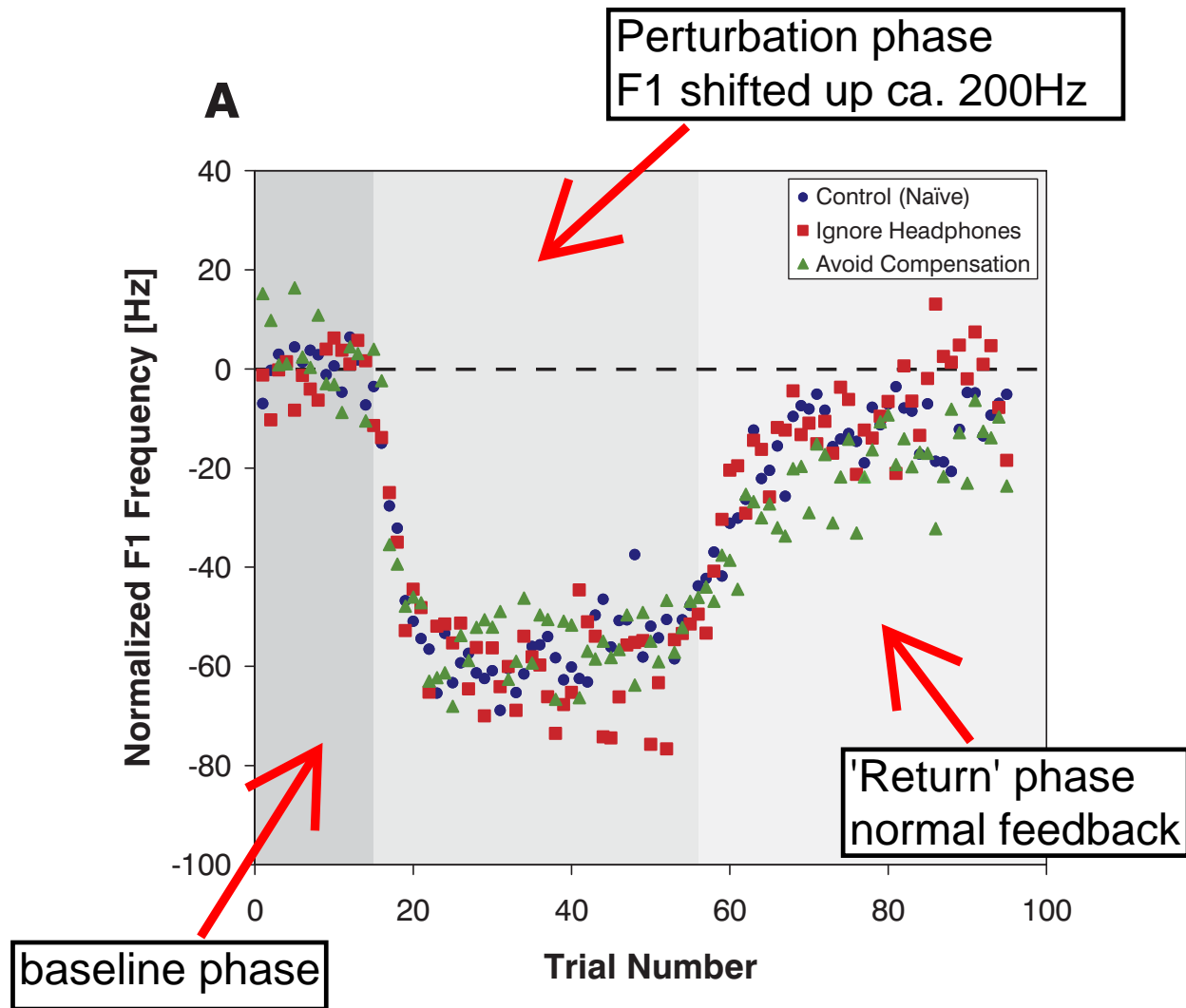
Z.B bei Verschiebung von F1 nach oben

Proband spricht “bet” und hört dabei “bat”

Um diesen Effekt zu kompensieren, d.h um “bet” zu hören, müsste der Proband (in etwa) “bit” sprechen



Typische Ergebnisse; aus Munhall et al. 2009



Die meisten Probanden kompensieren, aber fast nie zu 100%

Hier ca. 30% (Perturbation ca. 200Hz nach oben; Sprecher kompensieren mit ca. 60Hz nach unten)

(vermutlich wegen Konkurrenz mit somatosensorischem Feedback: perfekte auditorische Kompensation würde zu einem großen somatosensorischen Fehlersignal führen; vgl. DIVA-Blockschaltbild)

Die Kompensation setzt meistens ziemlich schnell ein (hier nach ca. 10 Trials)

Nach Beendigung der Perturbation ('return phase') bleibt eine (kleinere) Verschiebung der Formanten vorhanden: **Adaptation**

d.h. Proband spricht mit normalem Feedback etwas anders als am Anfang des Experiments

====> bereits nach einigen Minuten ist eine Anpassung der feedforward-Befehle nachweisbar

Das besondere am Experiment von Munhall et al.:

Kompensation (und Adaptation) findet statt, auch wenn Probanden über die Art der Perturbation informiert werden, und gebeten werden, nicht zu kompensieren!

Interpretation:

Kompensation/Verarbeitung der auditorischen Afferenzen ist beim normalen Sprechen im Hintergrund immer aktiv - nicht nur in speziellen Laborsituationen.

Sprechen ist immer adaptativ.

Manchmal langsamer: körperliche Änderungen, Verletzungen

Manchmal schneller: Anpassung an wechselnde Umweltbedingungen

Gibt es einen Zusammenhang zwischen Präzision in Produktion und Perzeption (weitere Experimentreihe von Perkell et al.)?

Beispiel für diese Untersuchungen:

Perkell, J., Guenther, F. , Lane, H., Matthies, M., Stockmann, E., Tiede, M., Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *Journal of the Acoustical Society of America*, 116, 2338-2344.

Contrast in production estimated from EMA and formant data.

EMA example: '***cod***' vs. '***cud***'

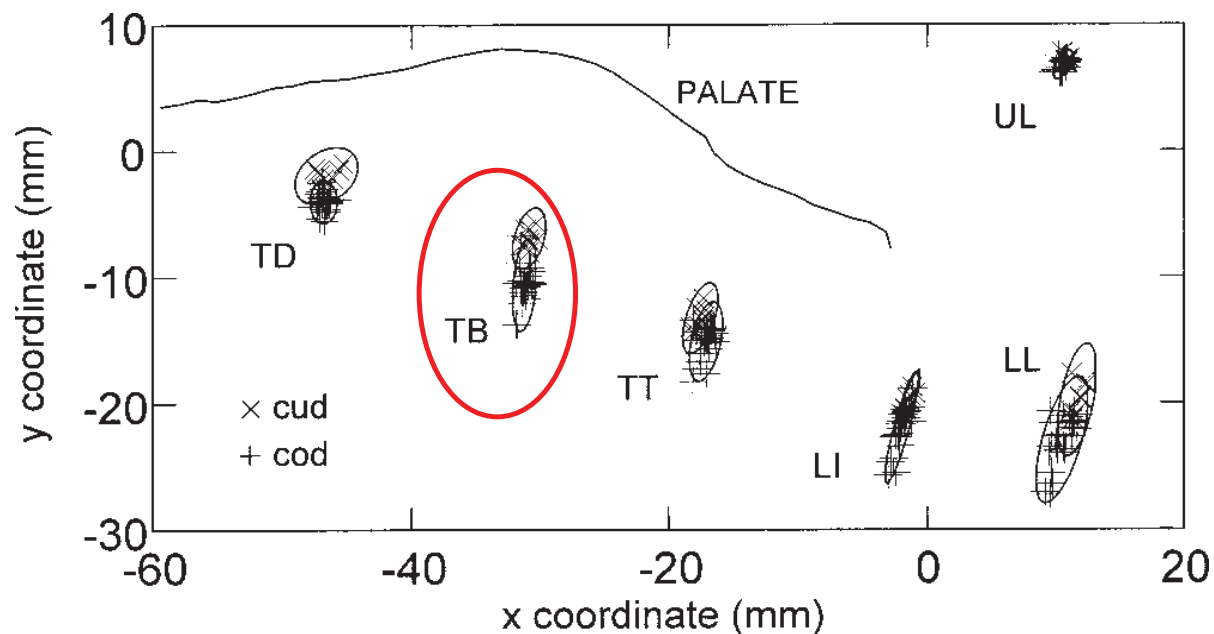
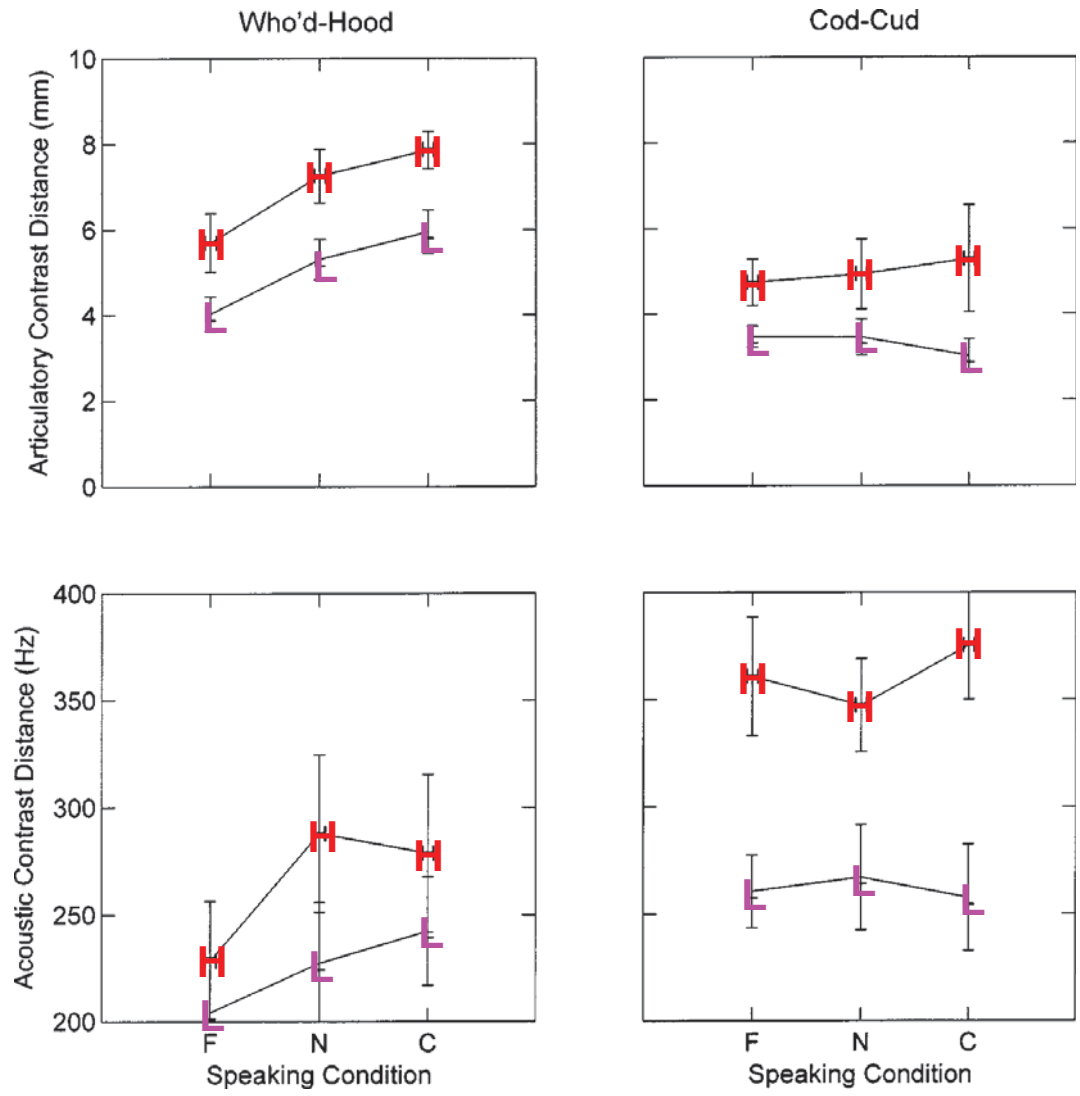


FIG. 1. EMMA transducer coil locations during all the productions of tokens containing the /ɑ/ in *cod* (+) and the /ʌ/ in *cud* (×) by one female subject in the normal condition. The transducer coils are located at points on the tongue dorsum (TD), tongue blade (TB), tongue tip (TT), lower incisor, (LI), lower lip (LL), and upper lip (UL). The midsagittal palatal contour (up to about the dento-alveolar junction on the right) is shown for reference.

Perceptual acuity estimated from ABX discrimination test
(7-step continua from *'cod'* to *'cud'* and *'who'd'* to *'hood'*)

Based on results, subjects divided into high discriminators and low discriminators



Contrast distance in production for **High** and **Low** discriminators in perception

FIG. 4. Articulatory contrast distance (for tongue body position—upper panel) and acoustic contrast distance (for separation in the formant plane—lower panel) as a function of the three speaking conditions. The left-hand panel gives results for *who'd-hood*, the right for *cod-cud*. Findings for high discriminators (labeled “H”) and low discriminators (“L”) are plotted separately. Error bars are one standard error about the mean.

Speaking condition:
 F=Fast
 N=Normal
 C=Clear

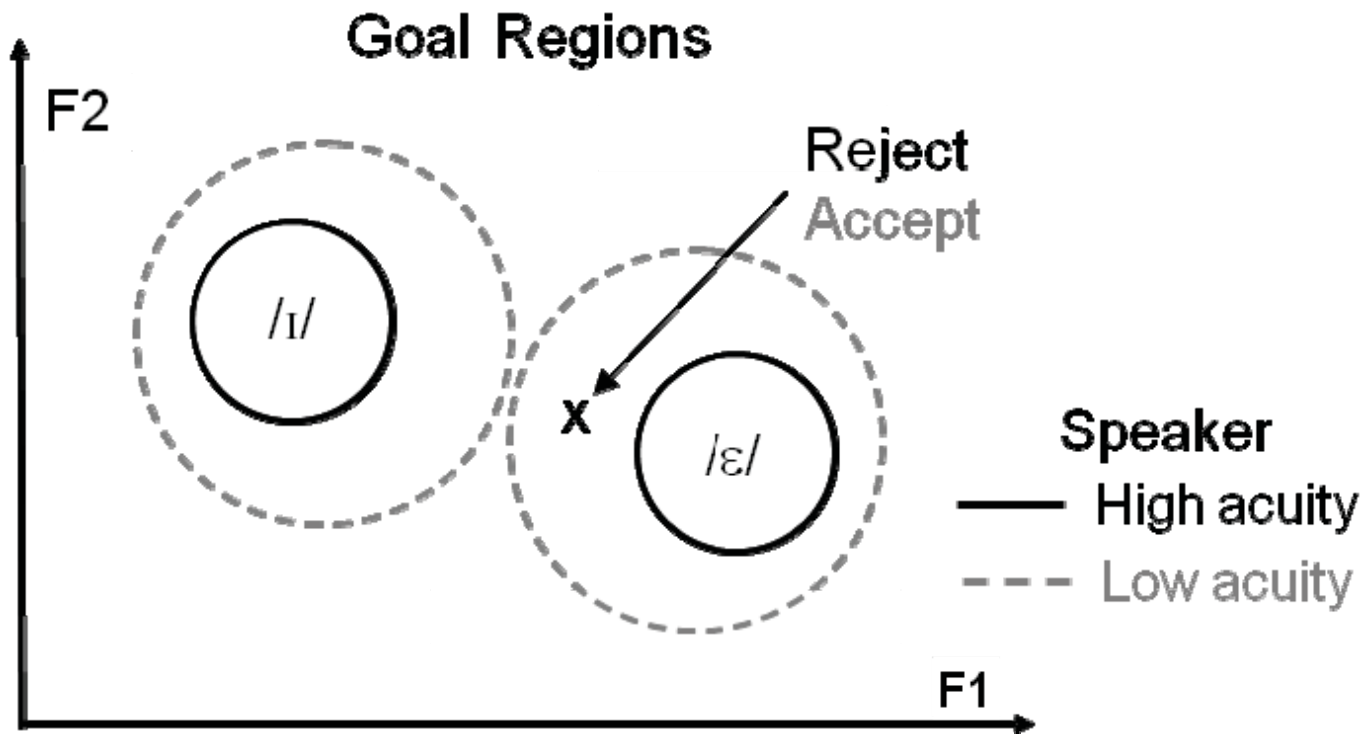


Figure 4. *Schematic illustration of goal regions of a high-acuity speaker and a low-acuity speaker for the vowels /i/ and /ε/ in F1 x F2 space.*

aus Perkell et al., 2008

Demo

Delayed Auditory Feedback (DAF)

Maximale störende Wirkung (bei normalen Probanden) bei Verzögerungen von ca. 100-200 ms.

Viele Untersuchungen von DAF, weil “fluency” von Stotterern oft **besser** bei DAF.

(Lee, 1950; Zimmermann et al., 1988; Stuart et al., 2002)

Teil 2. Somatosensorische Afferenzen

Wie kann man zeigen, ob diese Informationsquellen beim Sprechen eine Rolle spielen?

- (1) K.O.-Verfahren: Anästhesie. Beispiele später ('bite-block'-Experimente)
- (2) Zusammenhang mit Sprechleistungen untersuchen
- (3) Manipulation der Information

Zusammenhang mit Sprechleistungen:

Untersuchung von Ghosh et al.

Ghosh, S., Matthies, M., Maas, E., Hanson, A., Tiede, M., Ménard, L., Guenther, F., Lane, H., Perkell, J. (2010). An investigation of the relation between sibilant production and somatosensory and auditory acuity, *Journal of the Acoustical Society of America*, 128(5), 3079-3087.

Gibt es einen Zusammenhang zwischen Genauigkeit bei der Frikativartikulation und Diskriminationsfähigkeit für taktile Reize?

Vorgehensweise ähnlich wie bei Untersuchungen von Perkell et al. zur Vokalartikulation und Diskriminationsfähigkeit im akustischen Bereich.

Somatosensorische Diskriminationsfähigkeit: Probanden müssen Rillenorientierung mit der Zunge oder Lippe erkennen

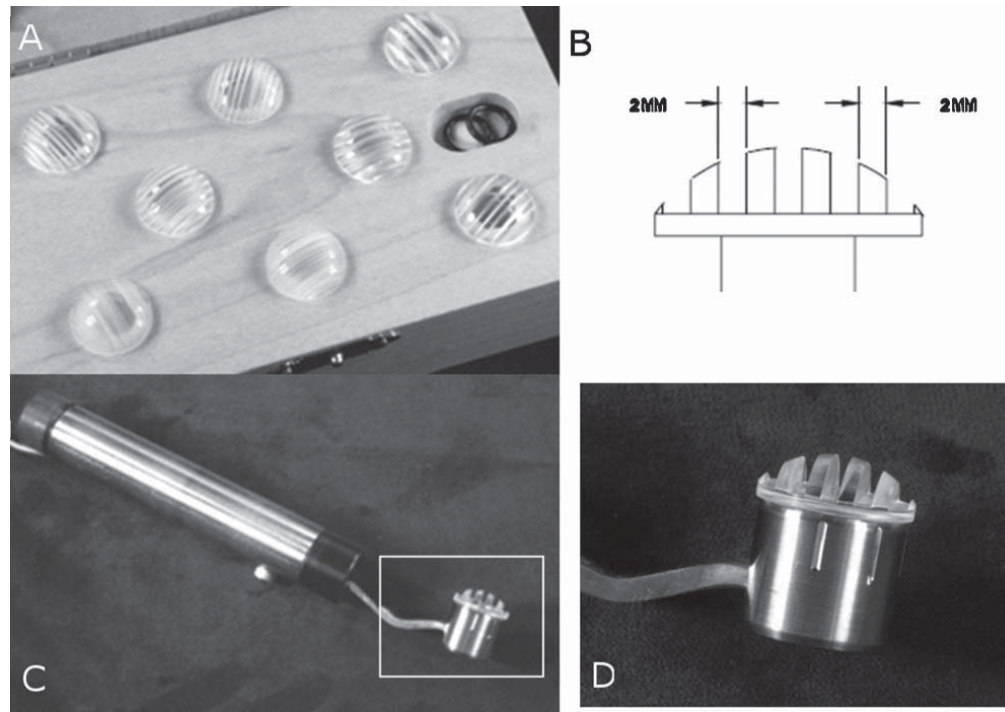


FIG. 1. JVP Domes and custom probe. (a) The set of 8 domes used in the study. (b) Grid spacing on one dome. (c) Custom holder used to apply pressure. The strain gauges are inside the handle and are not visible. (d) Magnified image of the region marked by the white box in (c).

Es wird die kleinste Rillenbreite ermittelt, bei der die Rillenorientierung noch richtig erkannt wird.

Akustische Diskriminationsfähigkeit wurde auch gemessen:

Kontinuum zwischen /s/ und /ʃ/

Als Merkmal für die artikulatorische Präzision:

Abstand zwischen /s/ und /ʃ/ im Spektrum gemessen ('contrast distance')

Ergebnisse

Somatosensorische und akustische Diskriminationsfähigkeit korrelieren ähnlich stark mit 'contrast distance'.

Zwar nicht besonders stark ($r = \text{ca. } 0.4$) aber statistisch signifikant.

Probanden eingeteilt in 2 Gruppen (über- vs. unterdurchschnittlich) jeweils bezüglich auditorischer und somatosensorischer Diskriminationsleistung. Beste Produktionsgenauigkeit (contrast distance) bei Probanden mit überdurchschnittlicher Leistung in **beiden** sensorischen Modalitäten = group **b**

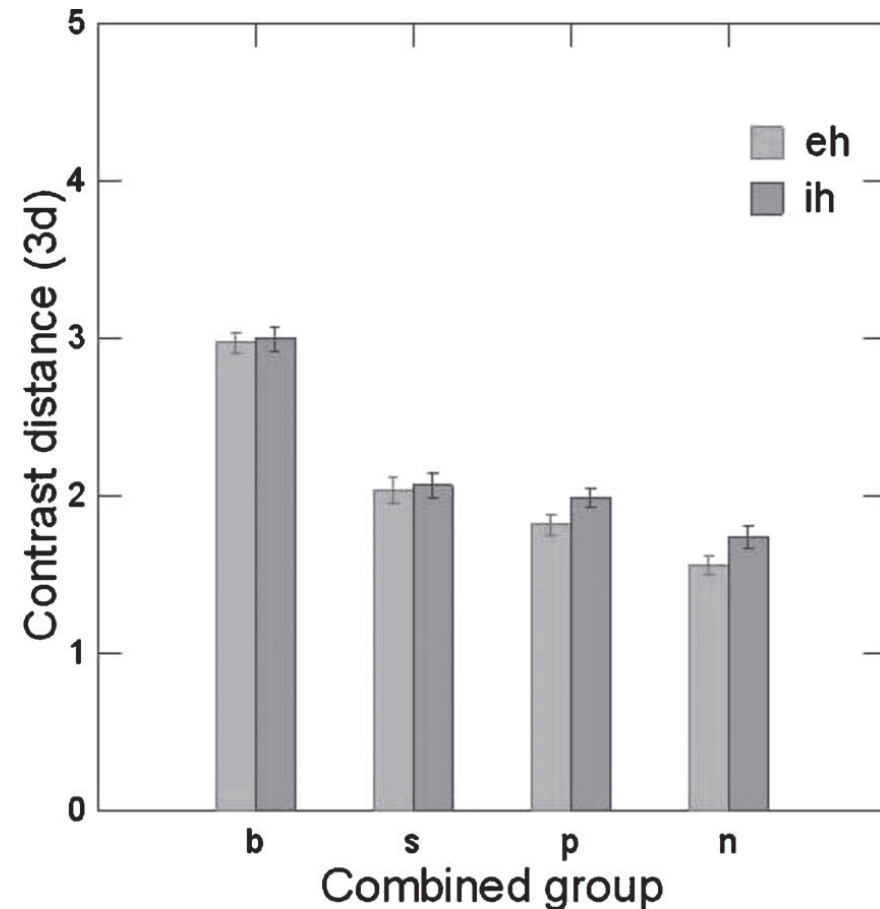


FIG. 5. Differences in 3-D contrast distance for said/shed and sid/shid as a function of group based on acuity and vowels ('eh' as in said, 'ih' as in sid; Groups-b: SomGrp=1, AudGrp=1; s: SomGrp=1, AudGrp=0; p: SomGrp=0, AudGrp=1; n: SomGrp=0, AudGrp=0; 1 indicates higher than median, and 0 indicates lower than median).

Schlechteste Produktionswerte (group **n**) bei unterdurchschnittlicher Leistung in beiden sensorischen Modalitäten

Weiterer Hinweis für den Zusammenhang zwischen somatosensorischer und artikulatorischer Präzision:

Untersuchungen von Wohlert (1996).

Ältere Probanden zeigen schlechtere Diskrimination als junge beim Rillentest.

Und auch schlechtere Präzision bei Lippenbewegungen

Aber:

Nachteil dieses taktilen Tests: Nichtsprachlich

Beispiel für einen Versuch, die Wirkung somatosensorischer Mechanismen bei phonetisch relevanteren Aufgaben nachzuweisen:

Loucks TM, De Nil LF (2006) Anomalous sensorimotor integration in adults who stutter: a tendon vibration study. *Neuroscience Letters* 402(1-2): 195-200.

Ein Vibrationsreiz wird verwendet, um die Muskelrezeptoren 'auszutricksen'.

Probanden trainieren eine bestimmte Kieferöffnungsbewegung (z.B. 18mm) mit visueller Kontrolle.

In der Testphase machen sie die Öffnungsbewegung ohne visuelle Kontrolle. Der Vibrationsreiz wird einem Muskel appliziert, der der Bewegung **antagonistisch** ist: Hier dem Masseter.

Die Muskelrezeptoren 'denken', dass der Muskel gedehnter ist als eigentlich der Fall ist.

====> wenn die Muskelafferenzen intakt sind, ist die Bewegungsamplitude kleiner in der Testphase als in der Trainingsphase.

Noch näher an das Sprechen heran: Manipulation von Kieferbewegungen

Noch direktere Hinweise, dass Sprecher eine sehr genaue Vorstellung der somatosensorischen Information haben, die durch Sprechbewegungen ausgelöst werden.

Nasir, S., Ostry, D. (2008). "Speech motor learning in profoundly deaf adults". *Nature Neuroscience*, 11(10), 1217-1222.

Tremblay, S., Shiller, D., Ostry, D. (2003). "Somatosensory basis of speech production", *Nature*, Volume 423, Issue 6942, 866-869

Mittels eines "Roboters" wird der Kiefer während Öffnungsbewegungen für Silben wie "saw", "say", "sass", "sane" in Protrusionsrichtung perturbiert.



Perturbation of jaw movements using a robotic arm. Image courtesy of David Ostry, McGill University, Montreal, Canada.

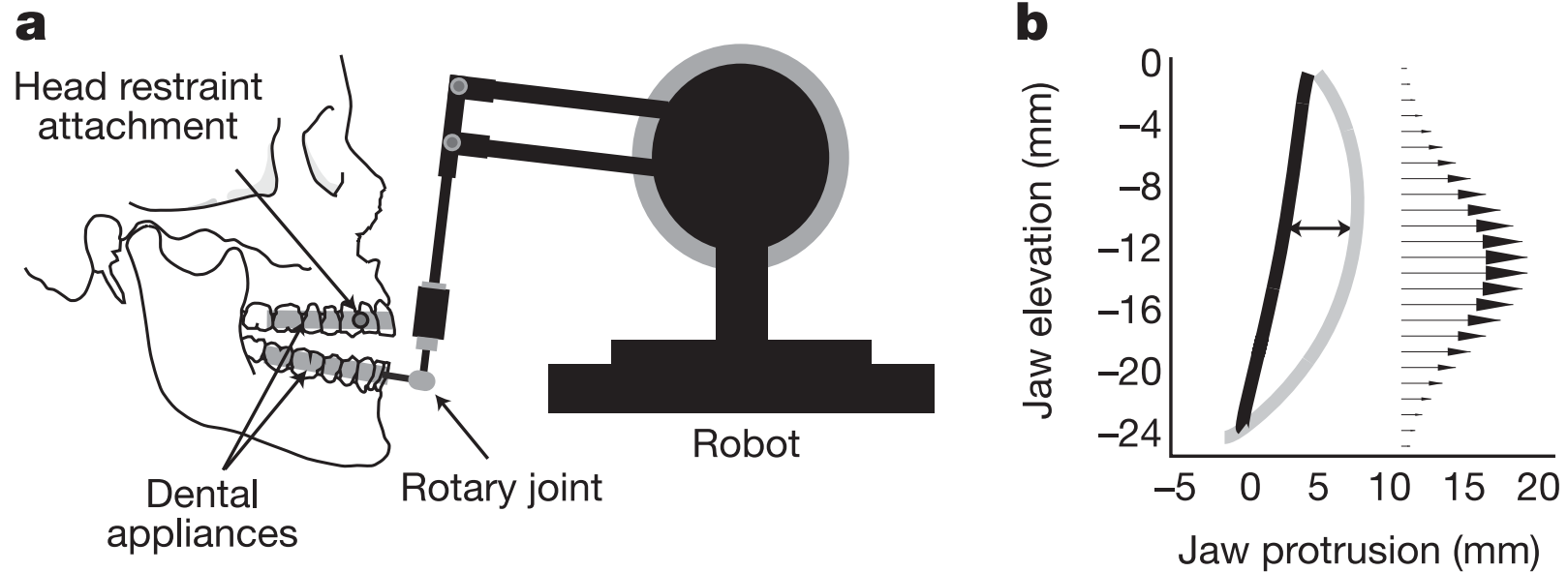


Figure 1 Experimental set-up and representative data. **a**, Diagram showing subject attached to the robotic device. **b**, Jaw opening movement with the force field off (black) and on initial exposure to the field (grey). Vectors depict the magnitude and direction of force applied by the robot over the course of the movement. The double-headed arrow shows the maximum horizontal deviation between null-field and force-field movements that served as a performance index.

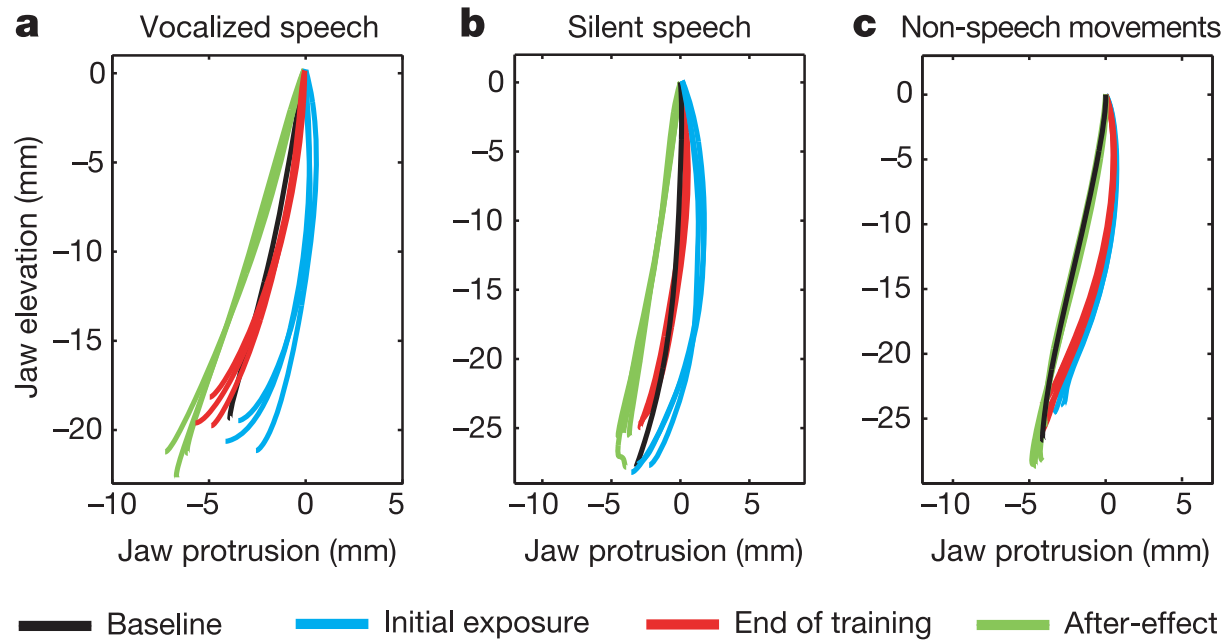


Figure 2 Sagittal plane jaw motion paths. Data were acquired during the baseline condition (black trace), on initial exposure to the force field (blue), at the end of training (red), and following unexpected removal of the field (green). The figure shows individual trials for single subjects. **a**, During vocalized speech, adaptation to the force field and a subsequent after-effect are observed. **b**, During silent speech, the pattern of adaptation and after-effect observed in vocalized speech are unaltered by removal of acoustic feedback. **c**, Matched non-speech movements show neither adaptation nor an after-effect.

Die Kompensation für die Perturbation lässt sich nicht durch auditorische Information erklären:

Diese Art von Perturbation hat fast keine Auswirkung auf die Formantwerte der Vokale

Die gehörlosen Probanden können sowieso nichts hören
Sprecher müssen sehr genaue Erwartungen darüber haben, zu welchen somatosensorischen Empfindungen ihre Sprechbewegungen führen werden.

Mögliche Erklärung, warum das Sprechen sich bei lang andauernder Gehörlosigkeit nur langsam verschlechtert.

Nicht nur auditorische, auch somatosensorische Information kann das Erlernen/die Anpassung von Sprechbewegungen steuern.

Ähnliche nichtsprachliche Bewegungen sind offensichtlich somatosensorisch nicht so genau spezifiziert: hier findet keine Kompensation statt (vgl. oben Fig. 2c).

Because the 'robot' experiments are so much fun:

Evidence that manipulating the articulators influences *perception*:

Ito T, Tiede M, Ostry DJ (2009)
Somatosensory function in speech perception, Proceedings of the National Academy of Sciences, 106: 1245-1248

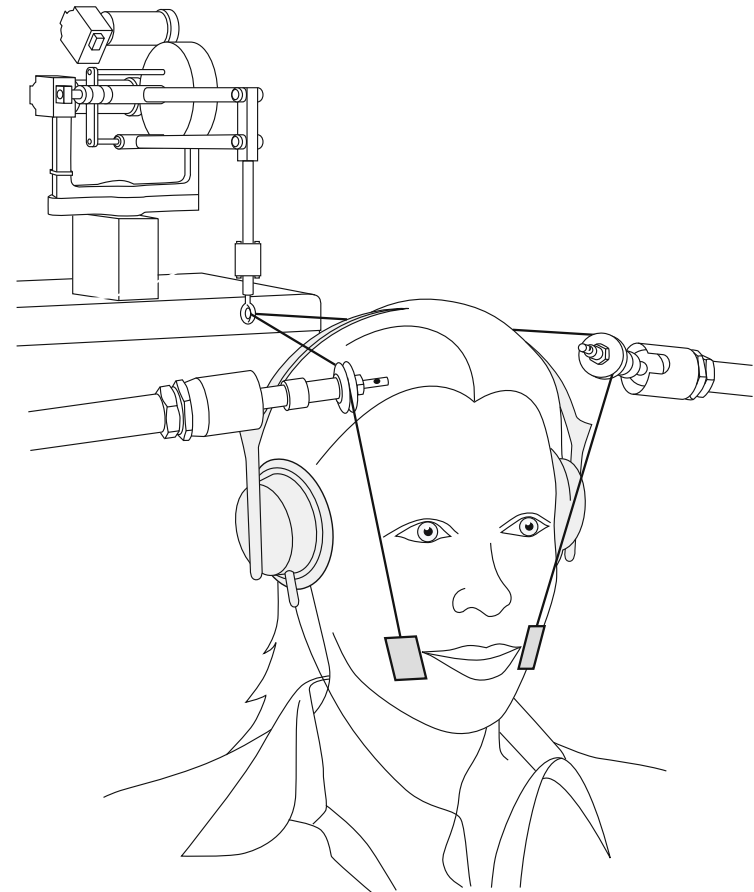


Fig. 1. Experimental setup for the delivery of skin stretch perturbations.

In einem Perzeptionsexperiment hören Probanden Stimuli aus einem Kontinuum von “head” nach “had”.

Während dem Vorspielen des Stimulus zieht der Roboter an der Gesichtshaut: Ähnliche Verformung der Haut, wie beim Sprechen dieser Vokale.

Bei Zug nach oben verschiebt sich die Kategoriengrenze zum tieferen Vokal hin (Richtung “had”): d.h. mehr Stimuli klingen wie der höhere Vokal (“head”).

Bei Zug nach unten umgekehrt.

Bei Zug nach hinten keine Auswirkung (diese Hautbewegung kommt bei der Artikulation dieser Vokale nicht vor)

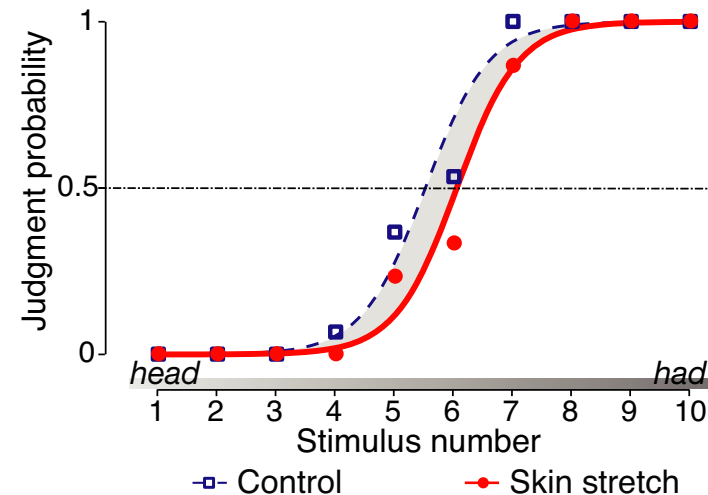


Fig. 2. Representative changes to speech perception with a 3-Hz sinusoidal pattern of upward skin stretch. The blue squares are for judgments without skin stretch. The red circles show judgments that occur in conjunction with skin stretch. Squares and circles show the judgment probability for each audio stimulus. The two solid lines show the estimated psychometric functions.

Literatur (sofern oben nicht schon angegeben)

Lane, H., and Tranel, B. (1971). "The Lombard sign and the role of hearing in speech," *J. Speech Hear. Res.* 14, 677–709.

Lane, H., Matthies, M., Guenther, F., Denny, M., Perkell, J., Stockmann, Tiede, M., Vick, J., Zandipour, M. (2007). Effects of short- and long-term changes in auditory feedback on vowel and sibilant contrasts. *Journal of Speech, Language and Hearing Research*, 50, 913-927.

Lee, B. (1950). "Effects of delayed speech feedback". *JASA* 22(6), 824-826.

Munhall, K.G., E. N. MacDonald, S. K. Byrne, and I. Johnsrude (2009). "Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate". *J. Acoust. Soc. Am.* 125, 384-390.

Perkell, Joseph S.; Guenther, Frank H.; Lane, Harlan; Matthies, Melanie L.; Perrier, Pascal; Vick, Jennell; Wilhelms-Tricarico, Reiner; Zandipour, Majid (2000). "A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss", *J. of Phonetics*, 28(3), 233-272.

Perkell, J., Lane, H., Ghosh, S., Matthies, M., Tiede, M., Guenther, F., Ménard, L. (2008). *Mechanisms of Vowel Production: Auditory Goals and Speaker Acuity.*

Proc. 8th International Seminar on Speech Production, 29-32.

Stuart, A. et al. (2002). "Effect of delayed auditory feedback on normal speakers at two speech rates". JASA 111(5), 2237-2241.

Zimmermann, G. et al. (1988). "The association between acoustic and articulatory events in a delayed auditory feedback paradigm". J. Phonetics, 16, 437-451.

Wohlert, A. (1996). Tactile perception of spatial stimuli on the lip surface by young and older adults, Journal of Speech, Language and Hearing Research, 39, 1191-1198.