

3 Perzeptive Phonetik (VII)

3.5 Sprachwahrnehmung

Inhalt

- Einführung
- Das Problem der Nichtsegmentierbarkeit, der fehlenden Invarianz und der Variabilität des akustischen Sprachsignals
- Die Motortheorie der Sprachwahrnehmung
- Kategoriale Wahrnehmung und „Enkodiertheit“
- Selektive Adaptation
- REA: „right ear advantage“
- Duale auditive Wahrnehmung und Fusionen
- Intermodale / Heteromodale Sprachwahrnehmung: Der McGurk-Effekt <--
- Diskussion der Theorien / Modelle der Sprachwahrnehmung aufgrund der vorgestellten experimentellen Ergebnisse

- Neurophysiologie der Sprachwahrnehmung
- Periphere Hörstörungen und zentrale Störungen der Sprachwahrnehmung
- Die nichtlinearen Relationen zwischen Artikulation, Akustik und Wahrnehmung:
Quantaltheorie und kategoriale Wahrnehmung
- Anhänge

3.5.8 Intermodale / Heteromodale Sprachwahrnehmung: Der McGurk-Effekt

Experiment von McGurk et al. (1976):

- optisch: Mundbild (Nahaufnahme eines Sprechers) [gaga]
- akustisch: [baba] mit guter Synchronisation der konsonantischen Verschlusslösung
- Wahrnehmung: die Versuchsperson „hört(!)“ /dada/ [GS 10 13](#)

Ergebnisse:

- Die auditive Wahrnehmung wird durch den visuellen Stimulus beeinflusst
- Das Wahrnehmungsergebnis ist *substanziell* weder akustisch noch optisch vorhanden

Interpretation:

- Da auditive und optische Integration erst auf hoher Ebene stattfindet, ist dies ein Indiz für einen phonetischen Sprachverarbeitungsmechanismus (Erkennung von distinktiven Merkmalen des Artikulationsortes) auch auf hoher Ebene
- Gibt es eine modalitätsfreie Ebene der Sprachwahrnehmung, auf der distinktive Merkmale spezifiziert werden (modalitätsfreie phonetisch-phonologische Ebene)? (siehe unten)

Idee einer modalitätsfreien phonetisch-phonologischen Ebene der Sprachverarbeitung

These 1:

Es gibt eine phonetisch-phonologische Ebene der Sprachwahrnehmung, die oberhalb der auditiven und über der visuellen Verarbeitung liegt.

- Auf dieser Ebene werden Phoneme aus (modalitätsspezifisch abgeleiteten) distinktiven Merkmalen zusammengesetzt.

Siehe dazu speziell die Effekte zur *audiovisuellen Integration* wie z.B.

Experiment: [ba] (optisch) + [ka] (akustisch) → [pa]

Modell: „bilabial“ (optisch) + Ort „stimmlos“ (akustisch) → [pa]

- Auf dieser Ebene werden auch distinktive Merkmale (modalitätsfrei) spezifiziert

Siehe McGurk-Effekt (oben)

Experiment: [ga] (optisch) + [ba] (akustische) → [da]

→ *heteromodale oder modalitätsfreie Ebene*

These 2:

Die Artikulation stellt einen übergeordneten Modus der Sprachwahrnehmung dar

Also:

- Ich sehe Artikulationsbewegungen

Das leuchtet unmittelbar ein: Bewegungen der Lippen, des Unterkiefers,;

siehe auch McGurk-Effekt

- Ich „höre“ Artikulationsbewegungen

Das leuchtet vielleicht nicht unmittelbar ein; ist aber wissenschaftlicher Fakt:

z.B. Spracherwerb:

Lallphase: Aufstellen der artikulatorisch-akustischen Beziehungen

Imitationsphase: Nachahmen von akustischen Reizen z.B. der Mutter ist jetzt möglich

Die These „Ich höre Artikulationsbewegungen“ ist letztlich die zentrale These der Motortheorie der Sprachwahrnehmung:

- Erinnerung an Enkodiertheit: Auf der akustischen Ebene gibt es kein invariantes Merkmal für den Artikulationsort
- Genau deshalb ist es gut, distinktive Merkmale nicht direkt akustisch sondern eher (übergeordnet) artikulatorisch zu definieren

Weitere Thesen der Motortheorie in diesem Zusammenhang:

- Sprachproduktion und Sprachwahrnehmung liegen (auf kortikaler Ebene) eng beieinander;
- Wir nutzen neuronale Prozesse der Sprachproduktion bei der Sprachwahrnehmung

Weitere Überlegungen:

- Phonetische Transkription ist detaillierte Fixierung einer Äußerung in artikulatorischen (nicht in akustischen!) Kategorien.

McGurk-Effekt mit nichtsprachlichen Stimuli

Experiment von Saldana und Rosenblum (1993)

- Tonerzeugung: Cello gezupft (visuell) + Cello gestrichen (akustisch)
- Ergebnis: Die Probanden tendierten zur Wahrnehmung des Tonerzeugungsmodus, den sie gesehen haben: Cello gezupft
- Das Experiment funktioniert auch bei Umkehrung der Tonerzeugungsmodi

Achtung: dies ist nicht vergleichbar mit dem McGurk-Effekt, da

- das Wahrnehmungsergebnis substanziell immer als Stimulus vorhanden war
- die Probanden immer eher zum visuellen Stimulus tendierten

Selektive Adaptation mittels McGurk-Stimuli

Erinnerung:

- Wenn die Klassifikation eines akustischen Kontinuums zur Variation der Artikulationsstellen aufgrund eines vorher dargebotenen Stimulus an den Artikulationsort dieses Stimulus „selektiv adaptiert“, liegt die Vermutung nahe, dass es ein spezielles Neuron zur Wahrnehmung des distinktiven Merkmals „Artikulationsort“ gibt.
- Da selektive Adaptation nur bei akustisch-auditiv bzw. spektral ähnlichen Adaptoren funktioniert, kann die These von speziellen Neuronen zur Wahrnehmung des artikulationsortes auf hoher Ebene verworfen werden.

(Erinnerung: Der zur Adaptation benutzte Stimulus muss nicht aus der Menge der Stimuli des Artikulationsstellenkontinuums kommen, muss aber spektral ähnlich sein: Adaptation mittels [di] an das /bε dε gε/-Kontinuum klappt nicht, wohl mit [da])

Vermutung:

Dementsprechend wird eine selektive Adaptation mittels dichotischer (und damit auf akustisch-auditiver Ebene noch mischbarer) Reize funktionieren, mittels McGurk-Stimuli aber nicht.

Experimente:

- Experiment 1 (Pompino-Marschall 1980):
Selektive Adaptation mittels des dichotischen Stimulus: [ba](Ohr1) + [ga](Ohr2)
→ Selektive Adaptation an [da]
- Experiment 2 (Roberts und Summerfield 1981):
Selektive Adaptation mittels des McGurk-Stimulus: [ba](akus) + [ga](visuell)
→ Selektive Adaptation an [ba]

Erinnerung: Selektive Adaptation passiert immer nur an der Phonemgrenze des Artikulationsortes des *adaptierten Stimulus*: [PM_086](#) (siehe die 3 Zeilen unter b)

Selektive Adaptation bei dichotischen Stimuli funktioniert, weil die Stimuli sich noch auf akustisch-auditiver Ebene im Wahrnehmungsapparat mischen.

Eine Mischung der Stimuli auf einer unimodalen Ebene (auditiv oder visuell) passiert bei McGurk-Stimuli nicht. Hier entsteht der Wahrnehmungseindruck von /d/ modalitätsfrei erst auf höher Ebene.

Schlussfolgerungen:

- Der Effekt der heteromodalen Fusion (McGurk-Effekt) kommt auf einer hohen nicht-auditiven Verarbeitungsebene zustande
Das hier wahrgenommene [da] adaptiert nicht! Grund: Zur Adaptation wird ein akustisch-auditiv (ein auditiv unimodal substanzielles) /d/ benötigt.
- McGurk-Effekt zeigt: Da es audio-visuelle Fusion gibt, kommt auch auf dieser Ebene erst die Spezifikation distinktiver Merkmale (zumindest für den Artikulationsort) zustande.
- Die fehlende selektive Adaptation an McGurk-Stimuli zeigt: Auf dieser hohen Ebene der Verarbeitung gibt es keine spezifischen Merkmalsdetektoren (keine „Hardware“) für das distinktive Merkmal „Artikulationsort“.
- Frühere Experimente zur Selektiven Adaptation mittels rein auditiver Stimuli zeigen aber, dass es auf akustisch-auditiver Ebene, also auf niedriger Ebene im Wahrnehmungsapparat, auf der noch der auditive Eindruck des akustischen Signals vorliegt, spezifische Merkmalsdetektoren zumindest für Merkmale wie ansteigende / abfallende Formanttransienten gibt.

Also:

Die Existenz des McGurk-Effektes belegt, dass es eine weitere Verarbeitungsebene für distinktive Merkmale gibt, die oberhalb der auditiven Verarbeitung liegt, aber wohl nicht mittels spezialisierter Neuronen (spezialisierter Detektoren, „Hardware“) geschieht, sondern mittels komplexer Verarbeitungsschritte (algorithmische Verarbeitung, „Software“)

3.5.9 Diskussion der Theorien / Modelle zur Sprachwahrnehmung aufgrund der vorgestellten experimentellen Ergebnisse

Einige Fragen wurden noch nicht beantwortet:

Gibt es einen speziellen Mechanismus zur Sprachwahrnehmung?

- Ja. Hinweis darauf geben die Experimente zur *dualen Wahrnehmung* (Wahrnehmung von Sprachlauten und einem nichtsprachlichen Zirpen, siehe oben)
Aber: duale Wahrnehmung kann auch bei nichtsprachlichen Reizen auftreten (Fowler und Rosenblum 1990)
- Ja. Siehe Experimente zur kategorialen Wahrnehmung und die Theorie der Enkodiertheit: Kategoriale Wahrnehmung bzw. das Phänomen der Enkodiertheit ist spezifisch für Sprache.
Aber: kategoriale Wahrnehmung tritt auch bei nichtsprachlichen Geräuschen auf (Summen, Musik) (Miller et al. 1976, Pisoni 1977, Zatorre et al. 1979)
Gegenargument: Summen und Musik können bedingt informationstragende Signale sein.

Kategoriale Wahrnehmung kann auch bei Tieren nachgewiesen werden (Affen, Chinchillas, Wachteln), also bei nicht sprechenden Lebewesen (May et al. 1989, Kuhl 1986, Kluender et al. 1987)

Gegengegenargument: Tiere verarbeiten auch Information (kategoriale Reize), auch ohne Sprechen zu können!

Weiter: Gibt es einen speziellen Mechanismus zur Sprachwahrnehmung?

- Ja. Siehe McGurk-Effekt: Es gibt eine hohe Ebene der Fusion, die rein sprachlichen Input voraussetzt (sonst würde die Signaldiskrepanz wahrgenommen)

Aber: McGurk-Effekt wurde (in abgeschwächter Form) auch für Instrumentalklänge nachgewiesen (Rosenblum 1993)

Gegengegenargument: Erkennung von Streichen und Zupfen eines Cellos ist eine spezielle Art kategorialer Wahrnehmung; deshalb tritt der Effekt auf, aber schwächer.

Der Effekt ist kein „vollständiger“ McGurk-Effekt, da das wahrgenommene Ergebnis (Zupfen bzw. Streichen) ja den substanziellen (auf einem Kanal vorliegenden) Basisreizen entsprechen. (McGurk: /da/ lag substanziell in keinem Kanal vor)

Gibt es spezielle linguistische Merkmalsdetektoren?

- Ja. Siehe Experimente zur kategorialen Wahrnehmung:

Es existieren spezialisierte Merkmalsdetektoren z.B. für die Wahrnehmung von „Stimmhaftigkeit“, für die Wahrnehmung von „labial“, „alveolar“ und „dorsal“, siehe die Experimente zur selektiven Adaptation mit auditiven Stimuli

Aber siehe oben: Kategoriale Wahrnehmung tritt auch bei nicht sprachlichen Geräuschen und bei Tieren auf. (siehe aber auch die Gegengegenargumente)

Es kommen Merkmalsdetektoren auf unterschiedlichen Ebenen vor:

- Die Ergebnisse zur selektiven Adaptation mit McGurk-Stimuli zeigen, dass die Detektion distinktiver Merkmale auf auditiver Ebene stattfindet
- McGurk-Effekt zeigt, dass eine „Korrektur“ / „Beeinflussung“ der Wahrnehmung distinktiver Merkmale aber auch auf einer nachfolgenden heteromodalen Ebene erfolgen kann.
- Andere (evtl. einfacherer herleitbare) distinktive Merkmale werden evtl. auf noch anderen Ebenen der neuronalen Verarbeitung spezifiziert.

Weitere Theorien und Modelle zur Sprachwahrnehmung

Bisher wurde die Motortheorie der Sprachwahrnehmung nur ansatzweise besprochen.

Diese Theorie soll jetzt vor dem Hintergrund der beschriebenen Experimente noch einmal geschlossen dargestellt werden.

Und es gibt weitere bekannte Modelle und Theorien, die vorgestellt werden sollen:

- Die Motortheorie nach Liberman et al. (1963 und 1985)
- Die passive Theorie der Sprachwahrnehmung nach Pisoni et al. (1975) (als Gegenpol zur Motortheorie)
- Die Kohorten-Theorie von Marslen-Wilson et al. (1980)
- Das TRACE-Modell nach McClelland et al. (1986) (aufbauend auf Kohorten)

Die Motortheorie nach Liberman et al. (1963 und 1985)

Hauptthema dieser Theorie: Erklärung des Problems der Lauterkennung (Erkennung distinktiver Merkmale) aus kontinuierlichen akustischen Sprachsignalen:

- Sprachwahrnehmung nutzt einen speziellen (hochkomplexen) nur für Sprachsignale geschaffenen *Modul zur Dekodierung der distinktiven Merkmale aus dem akustischen Signal*
 - > Es gibt einen eigenen Modul für Sprachwahrnehmung (im Unterschied zur Musik- und Geräuschwahrnehmung)
- Distinktive Merkmale sind im Sprachsignal komplex „kodiert“ (verschlüsselt)

- Die Dekodierung gelingt nur über „aktiven Nachvollzug“ des Gehörten
Anders ausgedrückt: Wir rufen zur Entschlüsselung des akustischen Signals unsere für die Phonemrealisierungen gelernten artikulatorischen Muster und die zugehörigen zentral parallel abgelegten perzeptiven Muster ab.
→ Wir nutzen unser Wissen zum Zusammenhang zwischen Artikulation und Perzeption. Dieses Wissen existiert (zumindest für hochfrequente Silben einer Sprache) in Form von *zentral abgelegten artikulatorisch-perzeptiven Plänen*.
Die neben den artikulatorischen Mustern existierenden perzeptiven Muster werden ebenfalls bei der (sensomotorischen) Produktion von Sprache zur Planung und Steuerung der Artikulation benötigt (siehe unten: Modelle der Sprachproduktion)
- Diese Fakten münden in das einfache Zitat: *Wir hören Sprechbewegungen*.

- In der Formulierung der Motortheorie von Liberman und Mattingly (1985) werden die *artikulatorischen Gesten als die zentralen invarianten Einheiten der Sprachproduktion und –wahrnehmung* herausgestellt.

Von Browman und Goldstein wurde eine phonologische Theorie zur artikulatorischen Geste (siehe klinische Phonetik I), die „gestische Phonologie“ entwickelt. Die Geste kann danach auch als phonologische (d.h. als bedeutungsunterscheide) Basiseinheit angesehen werden.

Gesten sind somit sowohl phonologisch als auch phonetisch. Phonetisch sind sie modalitätsübergreifend sowohl artikulatorisch als auch perzeptiv. Das Ziel einer Geste ist immer perzeptiv definiert: Realisierung der entsprechenden Bedeutungsunterscheidung, des entsprechenden distinktiven Merkmals.

- Ein Hauptproblem der Theorien zur Sprachwahrnehmung ist die Erklärung der Invarianz. Bsp.: /di/ vs. /du/: man hört ein (gleiches) /d/ obwohl einmal eine ansteigende und einmal eine abfallende F2-Transition vorliegt ([GS 10 05](#))

Invarianz (Enkodiertheit) entsteht einfach aufgrund der zeitlichen Koproduktion von artikulatorischen Gesten. Invarianz stellt somit gar kein Problem mehr da, wenn das eigentliche „Perzept“ bei Sprachwahrnehmung nicht die Formanttransition sondern direkt die artikulatorische Geste ist.

- Der Unterschied von sprachlichen und nichtsprachlichen akustischen Signalen ist die Enkodiertheit (siehe oben). Enkodiertheit folgt (wie Invarianz) letztlich aus der parallelen Produktion von Gesten (d.h. aus Koartikulation)

- Die Sprachwahrnehmung funktioniert aus Sicht der Motortheorie also über kortikale Aktivierung der (gestischen) Sprachproduktion, d.h. über die kortikale Aktivierung der artikulatorischen Planung und Programmierung (siehe unten: Modelle der Sprachproduktion).
- Der Prozess der Erkennung von artikulatorischen Gesten passiert in der Motortheorie auch nicht durch (planlosen) Vergleich des akustisch-auditiven Musters mit allen möglichen abgelegten perzeptiv-artikulatorischen Mustern von Gesten.

Vielmehr wird mit Eingang des akustisch-auditiven Musters direkt die Generierung nur einer bestimmten Auswahl von Geste angeregt (z.B. nur konsonantische labiale Gesten) und nach und nach weiter spezifiziert (Verschluss oder kritische Enge)

Zusammenfassung: Module der Sprachwahrnehmung nach der Motortheorie

- *periphere (spektrale) Analyse* des Sprachsignals → akustisch-auditive Muster (ähnlich einem Sonagramm) im auditiven sensorischen Kurzzeitspeicher

Dieser Modul ist nicht sprachspezifisch

- *zentraler Modul der Sprachwahrnehmung*: Generierung von Gesten durch Abgleich der akustisch-auditiven Muster mit den zentral abgelegten auditiven *Mustern der perzeptiv-artikulatorischen Pläne der Gesten* für die (sensomotorische) Produktion

Hier also enge Verbindung: *Sprachwahrnehmung <-> Sprachproduktion*

Die Sprachwahrnehmung bedingt die kortikale Aktivierung zur Planung artikulatorischer Abläufe, d.h. die kortikale Aktivierung von Sprechgesten.

- weitere zentrale phonologisch-lexikalisch-linguistische Verarbeitung
(Diese Ebene wird von der Motortheorie nicht weiter problematisiert)

Die passive Theorie der Sprachwahrnehmung nach Pisoni und Sawush (1975)

Wesentliche These: Sprachverarbeitung ohne Aktivierung der Artikulation

Die akustisch-auditiven Muster werden mit einer Menge von im Wahrnehmungsmodul abgelegten Mustern für distinktive Merkmale, Phoneme, Wörter etc. verglichen

Aufbau des Modells: [PN_70](#) (Peturson Neppert)

- erste auditive (spektrale) Analyse --> sensorischer Kurzzeitspeicher (als Teil des allgemeinen Kurzzeitgedächtnisses)
- Vergleich dieses Musters mit im Langzeitgedächtnis abgelegten Strukturen:
template-matching

Dieser Vergleich geschieht in mehreren Stufen:

- distinktive Merkmale
- Phoneme
- Silben / Wörter
- Semantische Analyse von Satzteilen

Dabei werden die analysierten distinktiven Merkmale im (modalitätsfreien) Kurzzeitgedächtnis abgelegt und dienen der Ermittlung der Phoneme, der Silben, der Wörter ...

[PN_70](#)

Die Kohorten-Theorie von Marslen-Wilson und Tyler (1980)

- Ausgangspunkt: Wichtiges Merkmal von Sprachwahrnehmung ist die Schnelligkeit: Sprachwahrnehmung passiert bis hin zur semantischen Interpretation in Echtzeit
- Diese Geschwindigkeit kann nur erreicht werden, wenn *auf jeder Verarbeitungsebene* (Laute, Silben, Wörter, Sätze) kurz nach Beginn der Wahrnehmung der Äußerung parallel mehrere potentiell richtige Kandidaten aktiviert werden

- Beispiel: Ebene Wort: Wenn der Wortanfang (die ersten Konsonanten auf akustisch-auditiver Ebene erkannt wurden, werden alle Wörter mit der Anfangskombination dieser Laute aktiviert (-> Wort-Kohorten) [PN_71](#)

Gleichzeitig werden alle anderen Wörter des Lexikons (also die meisten) gehemmt und verzögern damit nicht den weiteren Erkennungsprozess für das aktuelle Wort

Diese Einschränkung der möglichen Lösungen führt dann auch vor weiterer detaillierter phonetischer Lautanalyse zur richtigen Lösung und es kann mit der Verarbeitung des nächsten Wortes weitergehen

- Die Einflüsse z.B. des Lexikons auf die Sprachwahrnehmung sind bekannt: Wir nehmen wegen der Aktivierung des Lexikons oft real auftretenden phonetische Reduktionsprozesse gar nicht wahr (z.B.: Zahlwort „Fünfunddreißig“)

Das TRACE-Modell nach McClelland und Elman (1986)

- Umfassendes Modell der Sprachwahrnehmung
- Das Modell schließt die Kohorten-Theorie mit ein

Drei Hauptebenen: [PN_72](#)

- Ebene der distinktiven Merkmale (feature level)
- Phonem-Ebene
- Wort-Ebene

Horizontal: die Zeitebene [PN_72](#)

- jedes Kästchen entspricht einer Zeiteinheit
- Die Zeiteinheiten werden von Ebene zu Ebene länger
- Das Kurzzeitgedächtnis / der Speicher wird durch die horizontale Länge des Modells repräsentiert
- Die Signale laufen als Spur (trace) mit der Zeit durch den Speicher

Vertikal: Anordnung aller Merkmale / Phoneme / Wörter [PN_72](#)

Neuronale Verschaltung:

- Zwischen den Ebenen Merkmal / Phonem / Wort sind die Verschaltungen exhibierend (auf- und absteigend):
z.B. werden alle Phoneme angeregt, wenn die zugehörigen Merkmale angeregt sind
- Innerhalb einer Ebene sind die Verschaltungen inhibierend:
z.B. wird das Merkmal [-kons] gehemmt, wenn [+vocalic] angeregt wurde
- Die neuronale Verschaltung ist zeitinvariant
(-> Gilt in senkrechter Richtung in Abb. [PN_72](#))

Ablauf der Sprachwahrnehmung im TRACE-Modell: [PN_72](#)

- Input-Stufe: Die Extraktion distinktiver Merkmale aus dem akustischen Signal
- Die Merkmale (Phoneme / Wörter) können zeitlich überlappend und mit kontinuierlicher Stärke aktiviert werden
- (Überlappend) gleichzeitig angeregte Merkmale bilden ein Merkmalsbündel und erregen das zugehörige Phonem
- Phonemfolgen erregen dann das Zielwort

Eigenschaften des TRACE-Modells: [PN_72](#)

- Modellierung von Kontexteffekten: eine Wortaktivierung wirkt auf die Phonemaktivierung zurück
- Robustheit gegen Störgeräusche: Wort- und Phonemaktivierung wirkt auf die Merkmalsaktivierung zurück
- Die Nichtsegmentierbarkeit und die zeitliche Überlappung akustischer Lautmerkmale wird auf die Ebene der Merkmale übernommen und ist charakteristisch für dieses Modell

Dagegen stören diese Eigenschaften des Sprachsignals oftmals technische Spracherkennungssysteme.

Das Trace-Modell ist sehr umfangreich und es werden eine Menge Effekte zur kategorialen Wahrnehmung und zum Einfluss des Kontextes simuliert.

→ weiteres siehe:

McClelland JL, Elman JL (1986) The TRACE model of speech perception. *Cognitive Psychology* 18: 1-86

Diskussion

Im großen ganzen stehen alle oben erläuterten Modelle nicht in Widerspruch zueinander sondern betonen unterschiedliche Aspekte der Sprachproduktion

- Motortheorie: Die Verarbeitung von akustisch-auditiven Mustern zu distinktiven Merkmalen (z.B. Artikulationsort bei Konsonanten)
- Passive Theorien:
 - Die Strukturierung des Wahrnehmungsprozesses durch hierarchische Ebenen: distinktive Merkmale, Phoneme, Silben, Wörter etc.
 - die Einbeziehung von auditivem Kurzzeitspeicher und nichtmodalem Kurz- und Langzeitgedächtnis
- Kohorten-Theorie / Trace-Modell:
 - Wichtigkeit der Parallelverarbeitung bei der Sprachwahrnehmung
 - Wichtigkeit des mentalen Lexikons bei der Sprachwahrnehmung

3.5.10 Neurophysiologie der Sprachwahrnehmung:

Oben: Nur Neuroanatomie und Neurophysiologie der auditiven Wahrnehmung im Allgemeinen.

Jetzt speziell: einige Aspekte der neuronalen zentralen Verarbeitung zur Sprachwahrnehmung

Überblick: Sprachwahrnehmung: [PN_65](#)

Eingehender Schall am Ohr

→ Gehör → Basilarmembran → Rezeptorzellen

→ periphere und zentrale Hörbahn

→ Endstrecke der zentralen Hörbahn: primärer Hörkortex [PN_63](#)
(hier als Hörzentrum bezeichnet)

→ auditorischer Assoziationskortex

→ Wernicke-Areal: sprachspezifische Verarbeitung

Kortikale Verarbeitung

Ganz wichtig ist die Trennung zwischen primären Kortexarealen und anderen:

- primäre kortikale Areale sind *modalitätsspezifisch* (Wahrnehmung: sehen, hören, somatosensorisch; Willkürbewegungen) aber nicht spezifisch hinsichtlich der Sprache

[PN_63](#)

(Anmerkung: primärer somatosensorischer Kortex liegt hinter der Zentralfurche im Parietallappen parallel zum primären Motorkortex)

Diese Areale liegen alle bilateral vor.

Also:

- primärer auditorischer Kortex verarbeitet alle akustischen Signale
- primärer Motorkortex erzeugt alle Bewegungen, nicht nur Sprechbewegungen

- Während Gehör, Hörbahn und primäre sensorische Areale wohl weitgehend unspezifisch arbeiten, gibt es im Kortex dann weitere sprachspezifische Areale: [PN_66](#)

Diese Areale liegen unilateral vor (Sprache: meist linksseitig)

- Wernicke-Areal für die sensorische Sprachverarbeitung

Hier wird das mentale Lexikon vermutet

Störung dieses Areals führt zu Problemen bei der Sprachwahrnehmung von lexikalischen Inhalten aber auch beim Benennen von Gegenständen

- Broca-Areal für die motorische Sprachverarbeitung

Initiierung des Sprachproduktionsprozesses; Initiierung und Erstellung von motorischen Plänen für geplante Äußerungen

Störungen in diesem Gebiet führen zu Störungen der Expression und der Artikulation

Das Broca-Zentrum und seine enge Verbindung zum Wernicke Zentrum (Fasciculus Arcuatus = Assoziationsbündel im Gehirn [PN_66](#)) ist aus Sicht der Motortheorie der Sprachwahrnehmung wichtig.

- Weitere kortikale Gebiete, wichtig für die Sprachwahrnehmung:
 - sekundäres motorisches Zentrum = prämotorischer Kortex [PN_66](#)
 - primärer Motorkortex: liegt eingerahmt von Broca-Areal und Prämotor-Area [PN_63](#)
 - Wichtig für die sensomotorische Produktion: primärer Motorkortex und primäres somatosensorisches Areal liegen eng benachbart um die Zentralfurche herum (mit paralleler Struktur der Körperoberflächen)
 - Wernicke-Areal liegt eingerahmt von Broca-Areal und sekundärem Sehzentrum (das primäre Sehzentrum liegt dahinter) [PN_66](#)
- Hat das Bedeutung für die heteromodale Verarbeitung bei der Sprachwahrnehmung (McGurk-Effekt) ?

- Verbindungen zwischen den kortikalen Arealen heißen: Assoziationsbündel oder Assoziationsareale

Wichtig:

- Fasciculus Arcuatus [PN_66](#)
als Verbindung zwischen Broca- und Wernicke-Arealen
- Gyrus Angularis: [PN_66](#)
ist eine Verbindung nicht nur zwischen sekundärem visuellem Areal und Wernicke-Areal sondern auch eine Verbindung zu motorischen auditiven und somatosensorischen Arealen

3.5.11 Periphere Hörstörungen und zentrale Störungen der Sprachwahrnehmung

Beispiele für periphere Hörstörungen

siehe Skript → Birger Kollmeier

Ein Beispiele für zentrale Hörstörungen:

Reine Worttaubheit

Ist eine neurologische *Leitungsstörung*: Störung der neuralen Verbindungen zwischen Arealen

Leitsymptome:

- Nichtverstehen gesprochener Sprache
- Fähigkeiten des Sprechens, Lesens und Schreibens intakt (Unterschied zur Aphasie)
- Töne und nicht-verbale Stimuli werden richtig erkannt und zugeordnet (verarbeitet)

Pathoneuroanatomie:

- bilaterale Läsion des auditorischen Assoziationskortex
- primärer Hörkortex sowie Wernicke-Region sind aber intakt

Das Wernicke-Areal wird durch diese Läsion von akustischen Afferenzen abgeschnitten

Also afferenter Verlauf:

primärer Hörkortex → auditorischer Assoziationskortex → Wernicke-Areal ...

dagegen: Kortikale Taubheit

Leitsymptome:

- Taubheit in Bezug auf alle akustischen Reize

Pathoneuroanatomie:

- bilaterale Läsion des primären Hörkortex (und wahrscheinlich auch seiner Umgebung)
- aber intaktes Gehör und intakte periphere Hörbahn

3.5.12 Die nichtlinearen Relationen zwischen Artikulation, Akustik und Wahrnehmung: Quantaltheorie und kategoriale Wahrnehmung

Erinnerung an die Kommunikationskette:

Sprecher		Übertragung		Hörer
Artikulation	->	akustisches Signal	->	Wahrnehmung

Was ist „erfolgreiche Kommunikation“?

- Erfolgreiche Übertragung von Information
- Wahrnehmung einer Lautkette bedeutungsunterscheidender Einheiten (Phoneme)
- → Fortlaufende Unterscheidung und Identifikation von Lauten bzw. distinktiven Merkmalen (auf der Stufe unterhalb der Laute) (Artikulationsort, -art, ...)
- → fortlaufende *Kategorisierung*

Dagegen:

- Das artikulatorische und akustische Signal ist *kontinuierlich* skaliert.
 - artikulatorische Ebene: z.B. die Variablen: Zungenhöhe, Zungenposition, ...
 - akustische Ebene: z.B. die kontinuierlichen Verläufe der Formanten im akustischen Signal
- Wie kommen ich im „Sprachwahrnehmungsapparat“ von kontinuierlichen Signalen zur Kategorisierung?
- Wie komme ich zur Einsortierung der *kontinuierlichen Signale* zu *diskreten Kategorien* ?

Beispiele für diskrete Kategorien:

- Artikulationsort: labial, alveolar, ...
- Artikulationsart, Plosiv, Frikativ,
- Vokaltyp hoch /i/, tief /a/...

Zentrale Frage:

Wie komme ich innerhalb der Kommunikationskette (Artikulation -> akustisches Signal -> Wahrnehmung) von *kontinuierlichen Artikulationsbewegungen* zu *diskrete Einheiten*?

Antwort:

Es gibt zwei Mechanismen (einen physikalischen und einen biologischen) die dabei behilflich sind:

- *Artikulatorisch-akustische Quantalsprünge* bei der Entstehung des akustischen Signals aus Artikulationsbewegungen, (Quantaltheorie nach KN Stevens 1972 und 1985)
- Das Phänomen der *kategorialen Wahrnehmung*

Artikulatorisch-akustische Quantalsprünge: die Quantaltheorie nach KN Stevens

Es gibt Effekte beim Übergang von der Artikulation zum akustischen Signal, die „quantaler Natur“ sind:

- Die Kardinalvokale: z.B. /a/ und /i/ zeigen quantalen Charakter
- Die Artikulationsarten: z.B. /Plosiv/ und /Frikativ/ zeigen quantalen Charakter
- Artikulationsorte zeigen quantalen Charakter

Anmerkung:

Die Schrägstriche stehen hier (in Einklang zur Phonologie) für diskrete Kategorien.

Der quantale Charakter der Vokale /a/ und /i/

- /a/ und /i/ können durch zwei gleichlange Röhren, ein enges und ein weites Röhren approximiert werden (Zweiröhrenmodell)
- Bei kontinuierlicher Änderung der Länge der Röhren ergibt sich genau für den Fall der Äquidistanz ein
 - Minimum für F1 und Maximum für F2 für /i/ Abb. -> Tafelbild
 - Maximum für F1 und Minimum für F2 für /a/ Abb.

Also: *Benachbarte artikulatorische Zustände* führen im Bereich der Extrema (Miniam, Maxima) nur zu *geringen* (verschwindenden) akustischen Effekten (geringe Änderungen der Formantfrequenzen), in größerer Entfernung aber zu *großen* akustischen Effekten

Also:

- Einerseits: artikulatorisches Kontinuum
- Andererseits: akustisch finden wir Unterschiede in der Steigung der Formanttransienten
- → nichtlineare Abbildung der Artikulation in den akustischen Raum

Dennoch:

Der Übergang von der Artikulation zur Akustik bleibt bei Vokalen kontinuierlich

Dies ist bei Konsonanten nicht mehr der Fall:

Der quantale Charakter der Artikulationsarten /Plosiv/ und /Frikativ/

Die kontinuierliche Verringerung einer Engebildung im Ansatzrohr führt zu zwei quantalen Effekten:

- Ab einer bestimmten Querschnittsfläche der Engebildung entstehen (plötzlich) Turbulenzen durch die
 - starke Bündelung des Luftstroms
 - und Erhöhung der Strömungsgeschwindigkeit in der Engstelle
- Bei Vollverschluss endet das akustische Signal abrupt
 - Beispiel: Bilabialer Verschluss: Die Stärke des Zusammenpressens kann variieren; der akustische Effekt bleibt gleich: kein Signal
 - Bei schon geringer Öffnung des Verschlusses liegt sofort ein akustisches Signal vor

Abbildung: → Tafelbild

Stärke des Reibegeräusches als Funktion Fläche der oralen Engebildung

Man erkennt drei diskrete Bereiche:

- Bereich 1: beinahe Vollverschluss: kein Reibegeräusch
- Bereich 2: kritische Enge: maximales Reibegeräusch
- Bereich 3: vokalische Enge / Weite: kein Reibegeräusch

Also:

- einerseits: artikulatorisches Kontinuum
- andererseits: 3 diskrete Bereiche durch definierte aerodynamisch/akustische Effekte
- Also: → Nichtlineare Abbildung der Artikulation in den akustischen Raum

Noch stärker als *Nichtlinearitäten* sind *Singularitäten*:

→ Hier fallen darüber hinaus auch mehrere artikulatorische Zustände zu einem akustischen Zustand zusammen

Der quantale Charakter des Artikulationsorts

Die kontinuierliche Variation des Artikulationsortes zeigt quantale Effekte:

- Artikulation: kontinuierliche Variierung einer konsonantischen Enge von vorne nach hinten oder umgekehrt
 - Akustik: die Resonanzfrequenzen des vorderen und hinteren Resonanzraumes kreuzen sich (→ Abb: Tafelbild)
 - hier entstehen „Fokalfpunkte“: Die Änderung zweier Formantfrequenzen wird Null.
 - → ähnlicher Effekt wie bei den Vokalen, nur quantitativ ausgeprägter.
- Nichtlineare Abbildung von der Artikulation in den akustischen Raum

Kategoriale Wahrnehmung

Erinnerung an die Experimente zur kategorialen Wahrnehmung:

- Ausgangspunkt: ein akustisches Stimulus-Kontinuum (x-Achse)
- Ergebnis: eine spezielle Wahrnehmungskurve [GS_10_11](#) [GS_10_12](#)
Die y-Achse: Skalierung unterschiedlicher (benachbarter) Wahrnehmungszustände

Also: → ein nichtlinearer Übergang von Akustik zur Wahrnehmung:

- Mehre (ungleiche) akustische Stimuli werden perzeptiv nicht mehr unterschieden
- .. fallen im Raum der Wahrnehmung zu einem Zustand zusammen:

Abbildung:

Zustände:

Aku11 -> *

Aku12 -> * Perz 1 (z.B. /da/)

Aku13 -> *

Ebenso andere Zustände:

Aku21 -> *

Aku22 -> * Perz 2 (z.B. /ta/)

Aku23 -> *

Abschluss: Zur Relation Artikulation – Akustik – Wahrnehmung

Wir haben gesehen:

- Der Übergang Artikulation zur Akustik ist nichtlinear: Abb. Stev -> Folie
- Der Übergang Akustik zur Wahrnehmung ist nichtlinear: [GS_10_12](#)

Also:

- Wir haben eine Verringerung der Dimensionen:
- Wir haben ein „Mapping“ / ein Abbildung mehrerer artikulatorischer Zustände auf einen akustischen Zustand und wiederum mehrerer akustischer Zustände auf einen perzeptiven Zustand

Diese „Flaschenhalsstruktur“ ermöglicht letztlich die zur Informationsübertragung nötige Kategorisierung (siehe Skizze)

Skizze:

Art111	->	*					
Art112	->		*	Aku11	->	*	
Art113	->	*					
						*	
Art131	->	*					
Art132	->		*	Aku13	->		* Perz1
Art133	->	*					
						*	
Art151	->	*					
Art152	->		*	Aku15	->	*	
Art153	->	*					
Art311	->	*					
Art312	->		*	Aku31	->	*	
Art313	->	*					
						*	
Art331	->	*					
Art332	->		*	Aku33	->		* Perz3
Art333	->	*					
						*	
Art351	->	*					
Art352	->		*	Aku35	->	*	
Art353	->	*					

Also: Kategorisierung hier als Zusammenfassung vieler artikulatorischer (bzw. akustischer) Zustände zu jeweils einem Wahrnehmungszustand (Perz).