

Messung und Analyse von Mund- und Nasensignalen bei der Sprachproduktion

Manon Bettinelli, Arild Lacroix

Institut für Angewandte Physik, Goethe-Universität Frankfurt am Main

Email: {Bettinelli, Lacroix}@iap.uni-frankfurt.de

Einleitung

Das Sprachsignal entsteht im allgemeinen durch die Überlagerung des Mund- und Nasensignals. Bei Aufnahmen von natürlicher Sprache liegen diese beiden Signale gemischt vor. Zur Untersuchung der Sprachproduktion bei nasalierten Lauten sind Kenntnisse über die Stellung und Bewegung des Velums unverzichtbar. Mittels eines geeigneten Messaufbaues wird die Abstrahlung von Lippen und Nasenlöchern akustisch separiert. Mit den getrennt aufgezeichneten synchronen Mund- und Nasensignalen kann die Ankopplung des Nasaltrakts analysiert und somit indirekt die Velumbewegung beobachtet werden. Bei der Auswahl des Sprachmaterials werden insbesondere Nasalvokale und Vokale in Kombination mit Nasalen und Explosiven berücksichtigt.

Versuchsaufbau

Signaltrennung

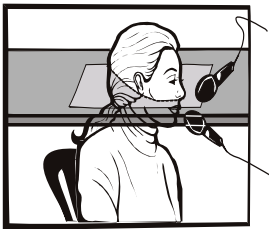


Abbildung 1: Proband im Messraum

Die Signaltrennung wird in einem Messraum mit den Ausmaßen $1.71 \times 1.71 \times 2.20$ (in m) realisiert. Die Wände des Messraums sind gegeneinander abgeschrägt, um stehende Wellen zu vermeiden. Die Trennung erfolgt über eine horizontal angeordnete Dämmplatte mit einer Aussparung für den Kopf. Die Aussparung wird

schalldicht abgeschlossen mit Hilfe eines flexiblen Kragens. Durch die Dämmplatte wird der Meßraum in zwei Teilräume aufgeteilt.

Aufnahme

Für die Aufnahme der Signale werden zwei Kondensator-Mikrophone verwendet, die unmittelbar vor dem Mund und der Nase positioniert sind. Der Proband sitzt aufrecht während der Messungen, um eine natürliche Position zu gewährleisten. Alle Messungen erfolgen mit einer Abtastfrequenz von 48 kHz und einer Auflösung von 16 bit je Kanal.

Auswahl des Sprachmaterials

Die Auswahl der Sprachäußerungen orientiert sich an der Aktivität des Velums, wobei die Ankopplung des Nasaltrakts an den Vokaltrakt von Interesse ist. Dies bedeutet, dass vor allem Nasale und Nasalvokale in Kombinationen mit anderen Lauten für die Sprachäußerungen ausgesucht werden. Zwei Beispiele mögen dies erläutern:

- Im Parlament findet die Ernennung zum Präsidenten statt
- Ich liebe Majoran und Premium Weine.

Bei der Messung eines Wortes, in dem möglichst viele der genannten Laute auftreten, kann der zeitliche Verlauf der Nasalisierung bestimmt werden, wodurch Kenntnisse über die Ankopplung

des Nasal- an den Vokaltrakt erlangt werden und dadurch die Velumbewegung ermittelt werden kann.

Bei der Artikulation der Nasale /n/, /m/ und /N/ kommt es zu einem Verschluss der Mundhöhle. Der Schall geht bei abgesenktem Velum vom Rachen- in den Nasenraum und wird von der Nasenöffnung abgestrahlt; dabei wirkt die verschlossene Mundhöhle als angekoppelter Resonator.

Bei den Vokalen breitet sich der Schall durch die Rachen- und Mundhöhle aus, wobei für deren Unterscheidung die Zungenstellung, sowie die Lippenöffnung maßgeblich ist. Befinden sich die Vokale in der unmittelbaren Umgebung der Nasale, so nasalisieren die Vokale. Dies bedeutet, dass sich der Schall auch durch den Nasenraum ausbreitet und das Sprachspektrum durch den Nasaltrakt mitbestimmt wird. Ein markantes Beispiel in diesem Zusammenhang ist die Lautkette [majon], bei der die Vokale nasaliert werden.

Nasalierte Vokale, auch als Nasalvokale bezeichnet, treten vorwiegend in französischen und französischstämmigen Äußerungen auf. Bei den Nasalvokalen wird der Schall gleichzeitig von den Öffnungen des Mund- und Nasenraums abgestrahlt.

Im Zusammenhang mit der Velumstellung sind die Explosive interessant wobei /b/, /d/, /g/ stimmhaft und /t/, /p/, /k/ stimmlos sind. Der Luftstrom wird durch den Verschluss von Nasal- und Vokalraum aufgestaut. Dabei verschließt das angehobene Velum den Nasaltrakt. Nachdem sich ein genügend hoher Druck aufgebaut hat, wird der Verschluss schlagartig gelöst. Lautkombinationen mit Nasalen und Explosiven verursachen Bewegungen des Velums.

Für die Aufnahmen wurden je zwei männliche und weibliche Sprecher im Alter von 25-32 Jahren ausgesucht. Je einer jeden Geschlechtes ist deutsch- bzw. französischsprachig, um die Authentizität der Landessprache zu gewährleisten.

Auswertung der Messungen

Velumbewegung

Nachdem die getrennten Mund- und Nasensignale aufgezeichnet sind, wird ein Hamming-Fenster auf 960 Abtastwerte angewendet und die Bestimmung der Signalleistungen von Mund- und Nasensignal mit einer Schrittweite von 120 Werten vorgenommen. Diese Kurzzeitsignalleistungen werden benutzt, um die Nasaltrakt-Ankopplung in ihrem zeitlichen Verlauf zu analysieren [1]. Aus der Nasaltrakt-Ankopplung kann die Velumbewegung ermittelt werden.

Bei der Auswertung der Äußerung "Amseln" ergibt sich der in Abbildung 2 dargestellte Verlauf. Die Lautfolge [zɛl] befindet sich zwischen den Nasalen [m] und [n]. Eine deutliche Nasalisierung kann an den Lauten [ɛ] und [l] beobachtet werden. Es lassen sich weitere Beispiele für Nasalierungen infolge Koartikulation in der deutschen Sprache finden.

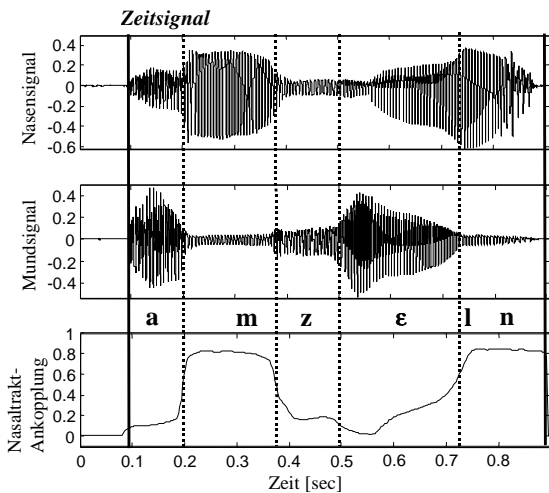


Abbildung 2: Nasaltrakt-Ankopplung während der Äußerung „amseln“

Frequenzspektren von nasalierten Vokalen

Die bei der Sprachproduktion auftretenden Resonanzen und Antiresonanzen des Sprechtraktes sind in den Spektren des Mund- und Nasensignals wiedergegeben. Diese Kenntnis erlaubt einen genaueren Einblick in die realen Verhältnisse im Nasal- und Vokaltrakt bei der Sprachbildung. Die in Abbildung 3 gezeigten Frequenzspektren des Mund- und Nasensignals werden aus einem Signalabschnitt der Äußerung "Mammut" des nasalierten Vokals /ã/ berechnet. Dafür werden 15 Grundperioden verwendet; dargestellt ist der niederfrequente Bereich.

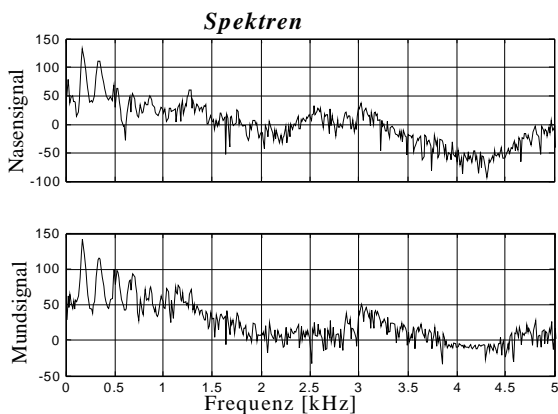


Abbildung 3: Frequenzspektren von Mund- und Nasensignal des nasalierten Vokals /ã/ aus dem Wort "Mammut" (männl. 29 Jahre)

Die getrennt aufgezeichneten Mund- und Nasensignale können auch zur Schätzung der Modellparameter eines verzweigten Rohrmodells verwendet werden [2].

Sprachcharakteristiken

Die untersuchten Nasalvokale in der deutschen Sprache lassen sich mit französischen Nasalvokalen vergleichen. Für die Analyse wurde das französischstämmige Wort "Restaurant" ausgewählt, das in die deutsche Sprache übernommen wurde.

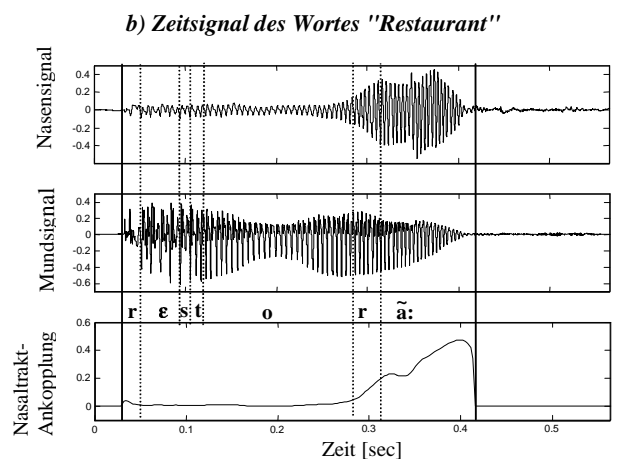
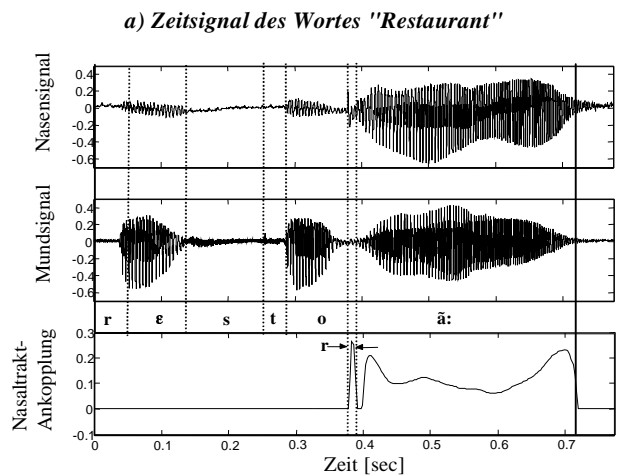


Abbildung 4: Nasaltrakt-Ankopplung während der Äußerung "Restaurant"

- (a) in der deutschen Aussprache (weibl. 28 Jahre)
- (b) in der französischen Aussprache (weibl. 32 Jahre)

In den Abbildungen 4a und 4b zeigen sich die unterschiedlichen Charakteristiken der deutschen und französischen Aussprache des Wortes "Restaurant". Im Vergleich zur deutschen Aussprache ist im Falle der französischen Interpretation ein fließender Verlauf der Nasaltrakt-Ankopplung zu beobachten.

Zusammenfassung

Die bisher vorliegenden Analysen deutscher Sprachäußerungen zeigen, dass Vokale in der Nachbarschaft von Nasalen zumindest teilweise nasaliert sind. Neben der Untersuchung der Velumbewegung werden die Mund- und Nasensignale auch spektral analysiert, um Aussagen über die Resonanzen und Antiresonanzen des Sprechtraktes zu ermöglichen.

Referenzen:

- [1] Bettinelli, M.; Schnell, K.; Lacroix, A.: „Separate Messung und Analyse von Mund- und Nasensignalen bei natürlicher Sprache“, Studientexte zur Sprachkommunikation: Band 24, 13. Konferenz ESSV-2002, Dresden 2002, S.237-244.
- [2] Schnell, K.; Lacroix, A.: „Parameter Estimation from Speech Signals for Tube Models“, Joint Meeting: ASA/EAA/DEGA, Berlin 1999, CD-ROM; contr. paper 2pSCa4.