EIN DRUCKSENSOR FÜR (LABIALE) PLOSIVE

Wolfgang Wokurek

Universität Stuttgart
wolfgang.wokurek@ims.uni-stuttgart.de

Kurzfassung: Wann erfolgt der Verschluss bei Plosivlauten ([p, t, k, b, d, g]), wie lange dauert er und wie hoch ist der Luftdruck währenddessen. Zur Beantwortung dieser Detailfragen wird eine Messanordnung aus einfachen elektronischen Komponenten entwickelt.

1 Einführung

Simultan zur Sprachaufnahme mit dem Mikrofon wird ein dünner Schlauch in den Mundwinkel gelegt und ein wenig in den Mund gesteckt, so dass er während jedes Verschlusslauts den Luftdruck im Mundraum nach außen zum Drucksensor bringt. Das Messsignal des Drucksensors soll simultan mit dem Sprachsignal aufgezeichnet werden. Das kann grundsätzlich über einen beliebigen Strom von Messwerten über eine der digitalen oder analogen Schnittstellen eines Rechners erfolgen. Da in unserem Labor das Mikrofonsignal, das Elektroglottogramm und Beschleunigungssensorsignale über eine mehrkanalige Soundkarte oder ein Mischpult aufgezeichnet werden, wird das Drucksignal als Schwingungsamplitude moduliert und dieser Ton in einem weiteren Kanal aufgenommen. So sind die Druckverläufe durch die Messung der Einhüllenden dieser Schwingung jederzeit abrufbar, und es kann kein Zweifel über den Zeitpunkt im Artikulationsverlauf entstehen.

Doch was kann der Zweck solcher Messungen sein? Zunächst ist die Druckmessung im geschlossenen Mund (bei offenen Stimmlippen) das einfachste Mittel, um den Luftdruck in der Lunge zu messen (Rothenberg [1],Rothenberg [2]). Im Sprechverlauf ist so z.B. dessen Veränderung (Abnahme) quantifizierbar. Auch eine Beziehung zur Betonung ist zu erwarten. Weiters hat nicht jeder Verschlusslaut eine Transiente (burst) im Mikrofonsignal. Mit der Druckmessung kann dieses phonetische Artikulationsdetail auf Verschlussdauer und -stärke bezogen, untersucht werden. Schließlich kann die Messung bei der Dokumentation von Sprechstörungen (z.B. bei Parkinson) nützen.

1.1 Druck

Zur Orientierung, welche Druckmessbereiche in Frage kommen, dienten dem Autor die Druckmessgeräte PG-20E und PG-100E von Glottal.com. Diese verwenden cm-H2O als Druckeinheit. 1 cm-H2O entspricht knapp 100 Pa (exakt 98.0665Pa). Die Gerätebezeichnung enthält den Skalenendwert des höchsten Messbereichs. Jedes Gerät hat eine zweite Einstellung mit dem halben Skalenendwert. Der größte Skalenendwert von 100 cm-H2O entspricht 10kPa und der kleinste von 10 cm-H2O entspricht 1 kPa.

2 Aufnahmen

Aufgenommen wurden gelesene Silben [pi, pa, pu, psi, psa, psu, pfi, pfa, pfu] eines im Aufnahmeraum sitzenden Sprechers. Verwendet werden piezoresistive Drucksensoren. Es wurden

Sensoren verschiedener Empfindlichkeit erprobt. Die ersten Sensoren wurden mt einem statischen bilabialen Verschluss ausgewählt. Dabei war das Ziel, den Sensor voraussichtlich nicht zu übersteuern. Praktisch kann man aber alle erprobten Sensoren übesteuern. Also wurden starke, aber realistisch scheinende Drucke beim bilabialen Verschluss zu realisieren. Insbesondere war der Sprecher der ersten Aufnahmesitzung dem Autor nicht bekannt und der gewählte Sensor war zu unempfindlich und lieferte unbrauchbare Signale, die sich nur selten aus der Ruhelage bewegten. Die wenigen Ausschläge waren aber stufenförmig. Zwar wurde der Analogausgang dieses Sensors verwendet, offenbar erfolgte aber eine interne digitale Verarbeitung.

Daher wurden Sensoren ohne interner digitaler Zwischenstufe ausgewählt (Freescale Semiconductor MPX5100, MPX5050, MPX5010). Die Bauteilbezeichnung enthält den Maximaldruck 100 kPa, 50 kPa und 10 kPa. Der 100 kPa Sensor war in der Lage, bei den isoliert produzierten p-Lauten nicht zu übersteuern und wurde deshalb zunächst für die Aufnahmen verwendet. Bei der kontinuierlichen Sprechweise lieferte dieser Sensor aber zu kleine Amplituden. Umgekehrt war der 10 kPa Sensor meist nahe dem Vollausschlag oder ging in Sättigung. Für den Sprecher (MP) dieser bislang letzten Sitzung war der 50 kPa Sensor die richtige Wahl.

Nach den bisherigen Erfahrungen sind alle 3 Messbereiche sinnvoll und werden beim nächsten Umbau parallel am selben Schlauch verwendet. Möglicherweise kann nach Aufnahmen mit vielen verschiedenen Sprecherinnen und Sprechern der 100 kPa Sensor wegfallen.

Anfangs war der Wunsch, alle Plosive [p, t, k, b, d, g] mit dieser Anordnung zu vermessen. Die ersten Versuche stellten aber klar, dass die Messung scheitert, sobald die Zunge an der Verschlussbildung beteiligt ist. Zur Druckmessung muss der Schlauch nämlich hinter den Verschluss ragen. Einzelne Messungen gelingen manchmal; aber rasch verschiebt die Zunge vor, bei oder nach dem Verschluss den Schlauch wieder aus einer für die Messung geeigneten Lage. Ein Schlauch ohne weitere Befestigung eignet sich also nur für die Laute p und b.

3 Signalverarbeitung

Der Sensor hat zwei Flansche und liefert eine zum Differenzdruck proportionale Spannung. Im derzeitigen Aufbau bleibt der negative Flansch offen am Umgebungsdruck und der positive Flansch empfängt durch einen Kunststoffschlauch den intraoralen Druck des Sprechers. Der gemessene Druck ist also der den umgebenden Luftdruck übersteigende intraoralen Druck. Derzeit ist keine Messung und Aufzeichnung des umgebenden Luftdrucks vorgesehen, kann aber z.B. bei Versuchen zur Artikulation in extremen Bedingungen ergänzt werden.

Nach dem Datenblatt misst der Sensor den Differenzdruck, ist aber nur für positive Druckdifferenzen spezifiziert. Es ist aber nicht ausgeschlossen, dass bei der Artiklulation der intraorale Druck kurzfristig unter den Luftdruck fällt, z.B. bei der Einatmung oder durch eine Zungenverformung oder -bewegung, die während des Verschlusses das eingeschossene Luftvolumen vergrößert. Weitere Versuche werden klären, ob der Sensor für geringe Unterdrucke linear bleibt oder der negative Flansch mit einem Unterdruck versorgt werden muss. Auch die Ergänzung durch einen umgedrehten zweiten Sensor wäre möglich, wenn die dauerhafte Bereitstellung des Unterdrucks zu aufwändig ist.

Das Drucksignal soll gemeinsam mit dem Sprachsignal und anderen Sensorsignalen mit einer mehrkanalige Soundkarte oder einem Mischpult und einem Rechner aufgezeichnet werden. Der Frequenzbereich dieser analogen Eingänge ist insbesondere nach unten im 1-20 Hz Bereich begrenzt und das Drucksignal hat wesentliche Signalanteile ab 0 Hz. Deshalb wird das Drucksignal als Schwingungsamplitude einer Tonfrequenz moduliert, mit dem Ziel einer einfachen Rückgewinnung durch Messung der Einhüllenden dieser Schwingung.

Da derzeit mit mindestens 48 kHz aufgezeichnet wird, steht der ganze Hörbereich zur Verfügung. Verwendet wird eine sinusförmige 10 kHz-Schwingung, da deren Periode 0.1ms dauert

und das eine ausreichend feine Untergrenze für zeitliche Auflösung des Druckverlaufs ist.

Die Grundidee der Berechnung der Einhüllenden ist der Betrag des analytischen Signals. Da das Drucksignal aber symmetrisch um 10 kHz liegt, wird nur ein komplexer Bandpass von 7-13 kHz als FIR-Filter implementiert und der Absolutbetrag dessen Ausgangssignals ausgegeben. Diese Vorgangsweise liefert ein Drucksignal von 0-3 kHz. Schnellere Druckschwankungen und Störungen sind geglättet. Insbesondere sollte der Druckanstieg vor der Verschlusslösung, während eines Frikativs und während der Vokalartikulation enthalten sein.

4 Ergebnisse

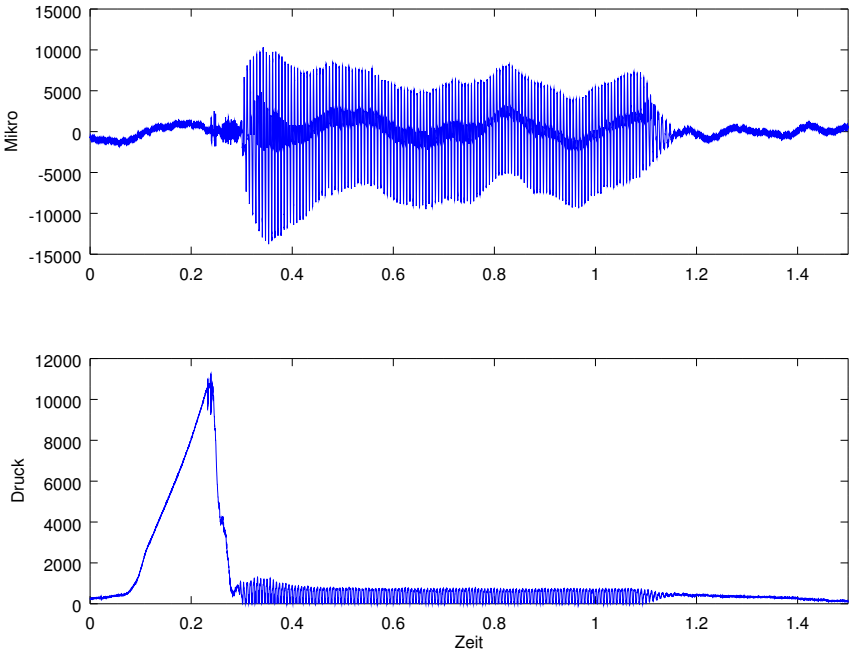

Abbildung 1 – Äußerung [pi], Zeit in Sekunden

Die Abbildungen 1, 2, 3 und 4 zeigen das Mikrofonsignal und das (demodulierte) Drucksignal des 10 kPa Sensors der Silben [pi], [psi], [psa] und [pfi]. Der Druck steigt gleichmäßig an, d.h. folgt etwa einer Geraden und dauert etwas weniger als 200 ms. Die Vokalqualtät hat auf Verschlussbildung und -lösung keinen Einfluss (Abbildungen 2 und 3).

Die Abbildungen 1, 2 und 3 zeigen kein Anzeichen von Sättigung. Erst Abbildung 4 erweckt den Verdacht, dass knapp über der Amplitude 12000 das Sensor-Verstärker-Modulator-System dem möglicherweise vorhanden Druckanstieg nicht mehr weiter folgen kann, also sättigt. Als Konsequenz wurden die Aufnahmen mit dem 50 kPa Sensor wiederholt, worauf alle Druckamplituden 5000 nicht weit überstiegen und nie eine Sättigung sichtbar war.

Wegen der Sättigung bei 12000 des 10 kPa-Sensors lassen sich Druckamplituden ohne weitere Kalibrierung zumindest schätzen, nämlich 12000 $\sim$ 10 kPa $\sim$ 100 cm-H2O.

Der 50kPa Sensor würde ja ebenfalls bei etwa 12000 sättigen und die Drucke der Aufnahme

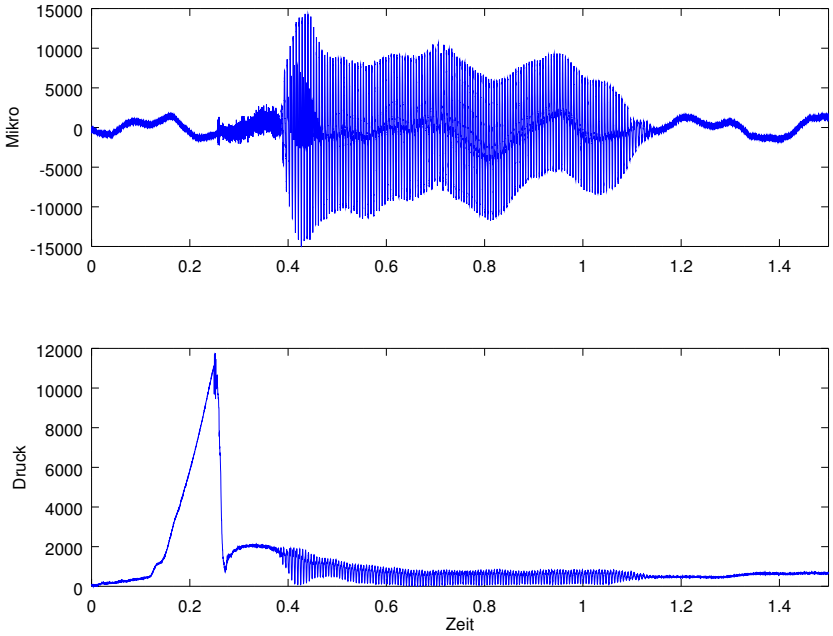

Abbildung 2 – Äußerung [psi], Zeit in Sekunden

bleiben unter 5000/12000*50 kPa $\sim$ 20 kPa oder 200 cm-H2O.

Die Zeitkonstante des Sensors (*response time*) ist im Datenblatt mit 1 ms angegeben. D.h., dass Druckschwankungen ab etwa 1 kHz zu rasch für den Sensor erfolgen. Umgekehrt sollten Druckschwankungen bis etwa 500Hz erfasst werden. Das sollte für den ersten Vokalformanten von Erwachsenen ausreichen. Abbildung 5 zeigt etwa 15 ms des aus dem [iː] der Silbe [psi]. Mikrofon- und Drucksignal zeigen beide die Sprachgrundperiode von etwa 6 ms und die Periode des ersten Formanten von etwa 2,5 ms. Das entspricht einer Sprachgrundfrequenz $F_0 \sim 170Hz$ und einer Formantfrequenz $F_1 \sim 400Hz$.

Abbildung 6 zeigt etwa 15 ms des aus dem [aː] der Silbe [psa]. Mikrofon- und Drucksignal zeigen beide die Sprachgrundperiode von etwas mehr als 6 ms und die Periode des ersten Formanten von etwas weniger als 2 ms. Das entspricht einer Sprachgrundfrequenz $F_0 \sim 160Hz$ und einer Formantfrequenz $F_1 > 500Hz$.

5 Zusammenfassung

Die Verwendung von piezoresistiven Drucksensoren zur Messung des intraoralen Druckverlaufs bei bilabialen Plosiven wurde vorgestellt. Ein Sensor mit einem Maximaldruck von 50 kPa erwies sich als ausreichend für den aufgenommenen Sprecher. Die Bandbreite des Drucksensors ist so hoch, dass neben dem langsam veränderlichen Druckverlauf der Verschlussphase auch die Sprachgrundfrequenz und der erste Formant im Druckverlauf sichtbar sind. Erste Messkurven wurden vorgestellt. Der Druckaufbau vor der Verschlusslösung des [p] dauert etwas weniger als 200 ms und der intraoralen Druck blieb unter 20 kPa d.h. 200 cm-H2O.

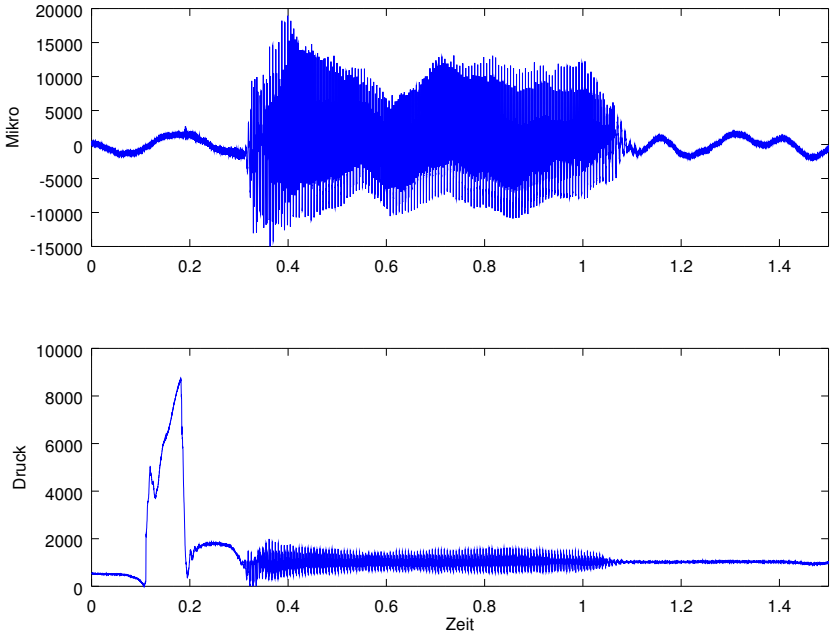

Abbildung 3 – Äußerung [psa], Zeit in Sekunden

Literatur

[1] ROTHENBERG, M.: *Rethinking the interpolation method for estimating subglottal pressure*. In *10th International Conference Advances in Quantitative Laryngology*, S. 111. 2013.

[2] ROTHENBERG, M.: *A new inverse-filtering technique for deriving the glottal air flow waveform during voicing*. *The Journal of the Acoustical Society of America*, 53(6), S. 1632–1645, 1973.

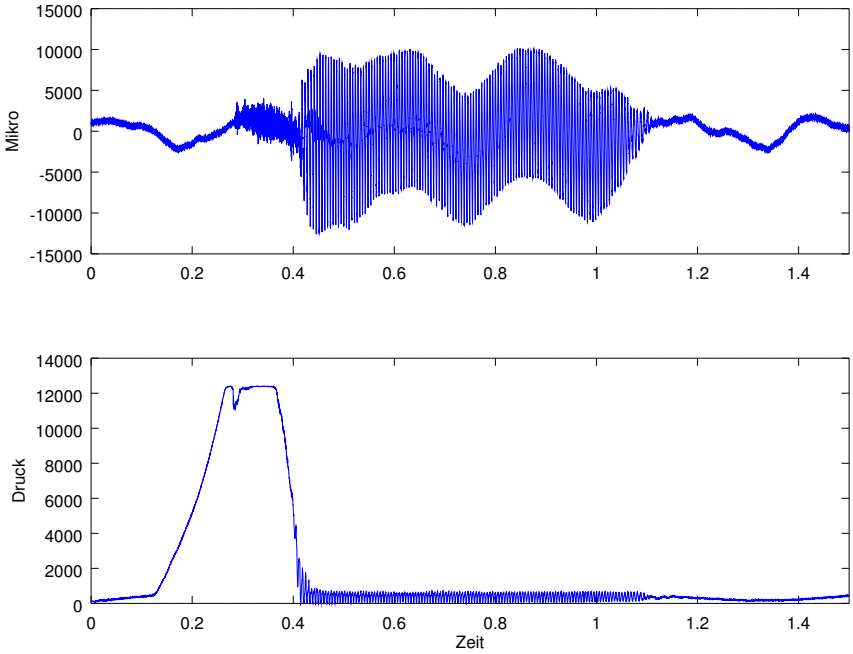

Abbildung 4 – Äußerung [pfi], Zeit in Sekunden

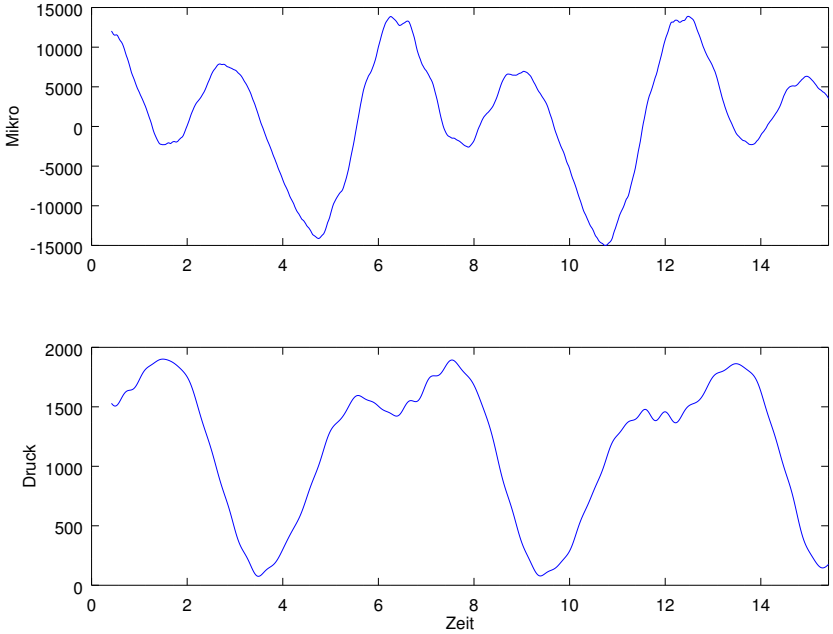

Abbildung 5 – Äußerung [psi] (Signalausschnitt aus Abbildung 2), Zeit in Millisekunden

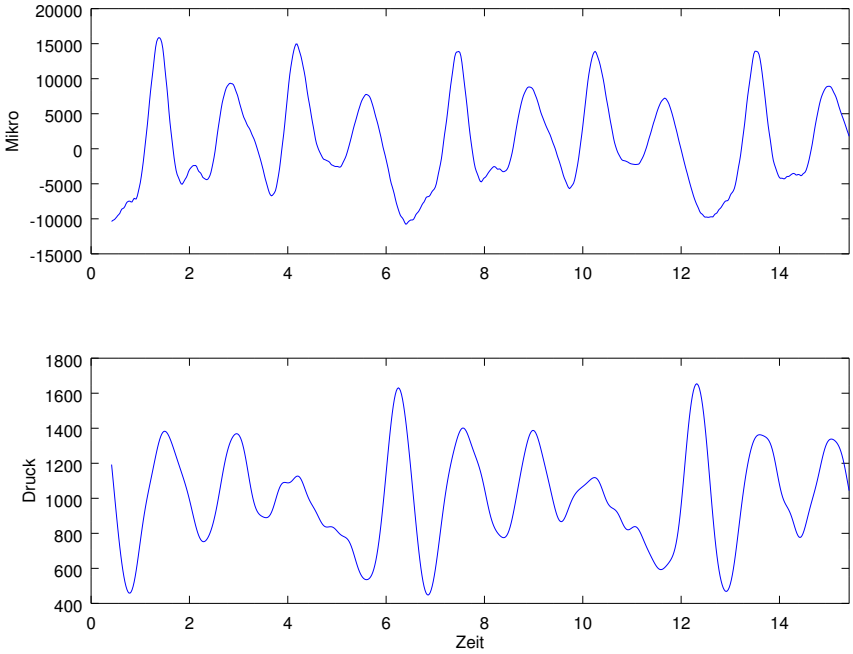

Abbildung 6 – Äußerung [psa] (Signalausschnitt aus Abbildung 3), Zeit in Millisekunden