

Der Einfluss videobasierter Kommunikation (Zoom) auf die Vokalqualität

Malte Belz, Melissa Ebert, Miriam Müller,
Jianqi Sun, Megumi Terada, Qiang Xia

Humboldt-Universität zu Berlin

Während der COVID19-Pandemie hat die Nutzung von Diensten zur audio-visuellen Kommunikation in Echtzeit stark zugenommen. Damit einher geht die Frage, ob die Verwendung audiovisueller Kommunikationsdienste einen Einfluss auf die gesprochene Sprache hat. Während für die technische Übertragung verschiedene Einflüsse auf das Signal festgestellt werden [1, 2], ist bisher unklar, ob der Wechsel von einer kopräsenten in eine videokonferenzbasierte Situation Auswirkungen auf die phonetische Realisierung hat. Im Rahmen eines Seminars untersuchen wir explorativ die Qualität von Vokalen des Deutschen in gelesenen Wortlisten in einer kopräsenten (*face-to-face*) und einer videobasierten Dialogsituation (*per Zoom*). Die Wortlisten sind Teil des Korpus BeDiaCo [3] und wurden am Anfang und Ende des Experiments in Anwesenheit der jeweils anderen Versuchsperson erhoben (entweder kopräsent oder über Zoom zugeschaltet). Zwischen den zwei Wortlisten lagen spontane gesprochene Dialoge von ca. 45 Minuten, die hier nicht untersucht werden. Die 20 Sprecher:innen sind miteinander vertraut (Mitbewohner:innen, Partner:innen, Geschwister) und wurden mit separaten Mikrofonen sowohl in der kopräsenten als auch in der videobasierten Dialogsituation aufgenommen, die einige Tage später stattfand.

Für beide Dialogsituationen wurde für 15 Monophthonge und zwei Reduktionsvokale ein gemeinsamer Zentroid im Vokalraum berechnet und anschließend die euklidische Distanz zu den neun äußeren Vokalen gemessen. In der Zoom-Situation werden sechs dieser neun Vokale signifikant reduziert artikuliert. Am stärksten zeigte sich der Effekt bei [a:]. Mittelgroße Effekte ließen sich bei [u:], [o:] und [y:] beobachten, kleine Effekte zeigten sich bei [i:] und [ɛ:]. Keine signifikanten Veränderungen waren bei [e:], [ɔ] und [ø:] festzustellen. Systematisch sind somit die hohen Vokale und der tiefe Vokal von einer Reduktion betroffen, während der Effekt für die ober- und untermittelhohen Vokale inkonsistent ist.

Drei Gründe könnten diese Ergebnisse erklären. (1) Nach der Theorie zur Hyper- und Hypoartikulation [4] variiert die Artikulation auf einem Kontinuum von klarer zu weniger klarer Sprache – Sprecher:innen bemühen sich dabei so weit, wie es für Hörer:innen zum Verständnis nötig ist. Die Reduktion der Vokale in der Zoom-Situation könnte auf eine geringere Involviertheit im Dialog zurückgehen („an internal, even emotional connection individuals feel which binds them to other people“ [5, S.27]), da sich die Personen nur über Bildschirme und Lautsprecher sehen und hören, was zu einer geringeren kommunikativen Anstrengung und somit zu Hypoartikulation führen kann. (2) Auch niedrigere Lautstärke könnte eine Rolle spielen: Das Signal gelangt in der Zoom-Situation direkt vom Mikrophon zu den Kopfhörern der anderen Versuchsperson, während es in Kopräsenz einen Raum zur ihr durchqueren muss. Das Wissen um diese direkte Perzeption könnte wiederum zu einer leiseren Produktion führen. Geringere Lautstärke kann nach [6] zu einem reduzierten [a:] führen, aber nicht unbedingt zu systematisch reduzierten hohen Vokalen (in unseren Daten sind jedoch [a:] und hohe Vokale systematisch reduziert); ein reduzierter Vokalraum ist aber nicht notwendigerweise auf geringere Lautstärke zurückzuführen. (3) Der Effekt könnte auch aus der internen technischen Aufbereitung des Signals durch Zoom resultieren [1, 2]. Die Versuchspersonen passen sich möglicherweise dem über Zoom gehörten Signal an.

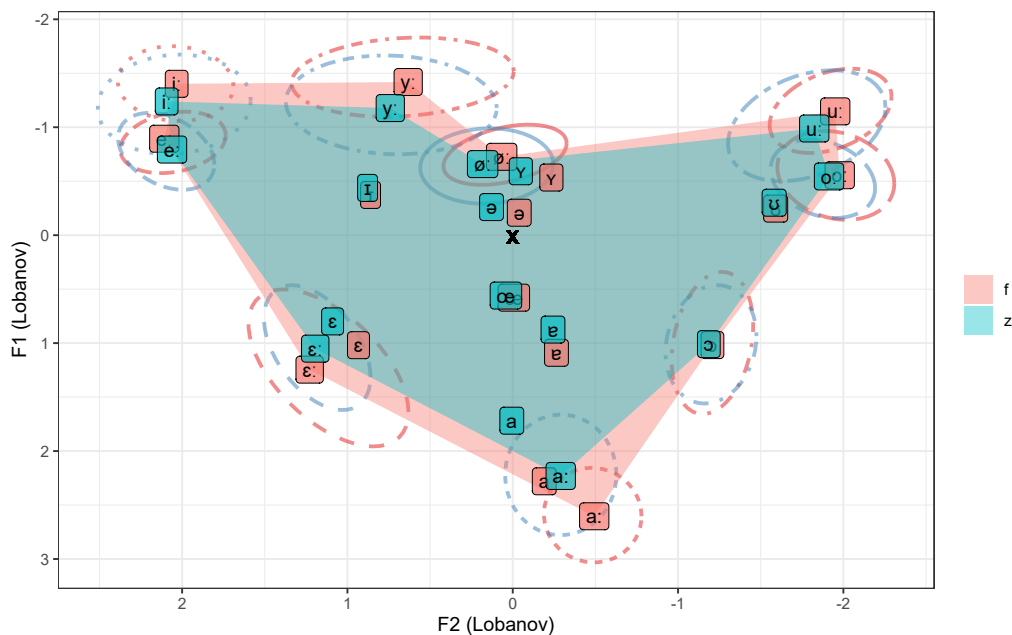


Abbildung 1: Formantraum für gelesene Vokale in der kopräsenten (f) und der videobasierten Situation (z). Die Ellipsen geben 68% der Vokaldistribution der äußeren Vokale wieder. Die Kästchen stellen die Mediane der Distribution dar.

Literatur

- [1] Jeremy Calder, Rebecca Wheeler, Sarah Adams, Daniel Amarelo, Katherine Arnold-Murray, Justin Bai, Meredith Church, Josh Daniels, Sarah Gomez, Jacob Henry, Yunan Jia, Brienna Johnson-Morris, Kyo Lee, Kit Miller, Derrek Powell, Caitlin Ramsey-Smith, Sydney Rayl, Sara Rosenau, and Nadine Salvador. Is zoom viable for sociophonetic research? A comparison of in-person and online recordings for vocalic analysis. *Linguistics Vanguard*, 2022.
- [2] Valerie Freeman and Paul De Decker. Remote sociophonetic data collection: Vowels and nasalization over video conferencing apps. *The Journal of the Acoustical Society of America*, 149(2):1211–1223, February 2021.
- [3] Malte Belz, Christine Mooshammer, Alina Zöllner, and Lea-Sophie Adam. Berlin Dialogue Corpus (BeDiaCo): Version 2, 2021.
- [4] Björn Lindblom. Explaining phonetic variation: A sketch of the h&h theory. In William J. Hardcastle and Alain Marchal, editors, *Speech Production and Speech Modelling*, NATO ASI Series, pages 403–439. Springer, Dordrecht, 1990.
- [5] Deborah Tannen. *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse*. Studies in Interactional Sociolinguistics. Cambridge University Press, 2007.
- [6] Laura L. Koenig and Susanne Fuchs. Vowel formants in normal and loud speech. *Journal of speech, language, and hearing research : JSLHR*, 62(5):1278–1295, 2019.