

Kurzer Erfahrungsbericht im Umgang mit dem Spracherfassungstool von recapp

(Fast) alles neu macht recapp... Hier möchte ich kurz aus meiner Optik die ersten Erfahrungen mitteilen, wie sich der Einsatz des Spracherfassungstools auf die Erstellung des Wort-Protokolls auswirkte. Es ist selbstredend, dass dies nicht nur meiner persönlichen Wahrnehmung entspricht, sondern nachvollziehbar eine grosse Auswirkung auf meine bisherige Vorgehensweise ausübt und ich somit auch derjenige bin, der direkt davon betroffen ist. Ebenso klar ist, nach einem einmaligen Test noch nicht alles rund laufen kann. Noch hat technisch wie auch bezüglich der Aufnahmequalität nicht alles zu 100% geklappt. Dabei kam es einerseits u.a. zu Abgrenzungsproblemen zwischen Votum x oder Votum y sowie zu einem etwas seltsamen Vertauschen der Reihenfolge von Traktandum Nr. 1 und Nr. 2, was ich mir so nicht wirklich erklären kann (Aufnahme von Traktandum Nr. 1 war unter Traktandum Nr. 2 erfasst und umgekehrt, im Wort-Protokoll entsprechend korrigiert).

In der Umsetzung der Transkription in das bislang gewohnte Format gibt es mehrere Formatierungsprobleme; entweder sind diese jeweils im Nachhinein anzupassen, oder sie könnten allenfalls bereits im Erfassungstool von recapp besser vorbereitet werden, was es entsprechend abzuklären gilt. Dies betrifft einerseits die Nummerierung der Traktanden und dabei insbesondere die Paginierung innerhalb eines Traktandums sowie andererseits den offiziellen Wortlaut der Beschlüsse (individuelle Bezugnahme auf Traktandum und die jeweilige Seite innerhalb des Traktandums) inklusive der offiziellen Signatur der Stadtratssekretärin (dieser Nachtrag ist allerdings einfach zu erledigen oder entsprechend anzupassen und somit ebenso erledigt).

Im Weiteren fällt auf, dass eine undeutliche, eventuell auch genuschelte Aussprache, unabhängig des Dialekts, zu mehr Missinterpretationen führt, was bei einer herkömmlichen Transkription etwas besser abgefedert werden kann. Dementsprechend wäre es meine Empfehlung, die Stadträt:innen bei dem nächsten Test höflich darauf hinzuweisen, etwas bewusster als gewohnt auf eine deutliche und klare Aussprache zu achten. Da gab es zum Teil lustige Erfassungsfehler (aus Ressortvorsteher wurde «Ressortforscher»... oder dann wurde daraus alternativ auch ein «Rösselvorsteher» aka Pferdeflüsterer 😊; der Gemeinderat wurde zum «Mainrot», eine Mischung aus Englisch und Dialekt, oder der Tiefbau wurde als «Teufel» verstanden). Dies sind Dinge, mit denen ich gut umgehen kann, das heisst, die mit der entsprechenden Aufmerksamkeit auch korrigiert werden können.

Gesprochenes Wort zu verschriftlichen stellt im Ansatz immer auch ein Stückweit einen Eingriff in die Sprache der redenden Person dar, nicht zuletzt auch wegen den deutlichen Unterschieden zwischen Dialekt und Hochsprache in Grammatik und Wortstellung. Dies führt auch zur Tatsache, ohne dies als Wertung zu verstehen, dass sich die Gilde der Transkriptor:innen im eigentlichen Sinn auch als Wortveredler:innen verstehen. Der Umstand, dass ich anhand eines bereits bestehenden, nicht von mir persönlich transkribierten Textes (mit einer abweichenden Wortstellung, da das Verb nachgestellt wird sowie der Vergangenheitsform im Perfekt anstatt im Präteritum) Ergänzungen, Anpassungen und entsprechenden Umstellungen vorzunehmen habe, erschwert mir das «freie» Transkribieren, dh. zuerst das Gesagte abzuhören, den weiteren Verlauf zu erkennen und dann zu verschriftlichen.

Mit der bereits vorliegenden «Übersetzung» kommt es zu einer «Einengung» meiner Arbeit; gerade auch dann, wenn die Aussage mit Falschinterpretationen verzerrt wiedergegeben wurde. Es ist in diesem Fall einfacher, ein leeres Blatt vor sich zu haben und dann die Aussage schriftlich in eine Form zu bringen, als Korrekturen am bestehenden Text vorzunehmen. Das hat insbesondere auch damit zu tun, dass viele Redner:innen während der Rede noch auf einen neuen Gedanken kommen, Einschübe vornehmen und dann nicht mehr genau wissen, mit was sie den Satz begannen. Ein völlig verständlicher Vorgang, den ich dann jeweils zu «glätten» habe. Da gibt es dann plötzlich Satzfragmente, die alleinstehend einfach keinen Sinn ergeben. Auch Schachtelsätze sind in diesem Zusammenhang immer eine Herausforderung, um dennoch in Anpassung der Wortstellung der beabsichtigten Aussage gerecht zu werden. Entsprechend ist es angesagt, die geäusserten Gedanken etwas in aufgeschlüsselter Form auseinanderzunehmen und so dem Sinn der Aussage anzupassen. Selbstverständlich ist das Teil meiner Arbeit und soll an dieser Stelle in keiner Weise als Lamento verstanden werden, sondern allein als Erklärung dienen. Das hat auch nichts mit dem neuen Tool zu tun und ist bei einer Transkription immer so. Das ist für mich ja gerade das Interessante und Spannende, dies – ohne die getätigte Aussage zu manipulieren – zu glätten und so verständlich aufzuschreiben, damit jemand, der nicht an der Sitzung teilnahm, versteht, was gesagt werden wollte.

Diese «Einengung» mit bereits in Schriftform vorliegenden Text erschwert mir wie gesagt die Glättung des Textes, sodass ich zum Teil dazu gezwungen war, auf das übliche Audio-File zurückzugreifen, um eine sinnvolle Transkription zu realisieren. Wenn es um den Anspruch geht, nebst dem Wort-Protokoll auch über eine verbindliche Tonaufnahme zwecks späteren Abhörens einer spezifischen Sequenz zu verfügen, so wird es dabei aber immer zu einer gewissen Abweichung von Audio-Aufnahme und Wort-Protokoll kommen, was aber – mit dem Anspruch, dass die Aussage verständlich sein sollte – leicht nachvollziehbar ist. Das sind wie gesagt grundsätzliche Feststellungen, die nichts mit dem Tool zu tun haben. Ebenso etwas mühsamer zu handhaben ist der Umstand, dass sich das erst nachträgliche Einfügen von Chart- und Folienabbildungen (also erst, wenn der Text im recap-Tool erfasst und heruntergeladen wurde) etwas aufwendiger gestaltet, weil das Einfügen nicht aus dem Redefluss erfolgt, sondern im Text diejenige Passage gesucht werden muss, wo das Einfügen Sinn ergibt und der Bezug hergestellt werden kann. Zuletzt geht es dann auch darum mit Argusaugen Grammatik und Orthografie des Tools, was nicht selber geschrieben wurde, auf deren Richtigkeit zu kontrollieren, was ein entsprechend noch aufmerksames Lektorat nötig macht.

Wie weit der Einsatz dieses Tools den Transkriptionsaufwand wie auch deren Qualität nachhaltig beeinflusst, kann zum jetzigen Zeitpunkt noch nicht abschliessend beurteilt werden, gilt es dafür doch die weiteren Tests abzuwarten wie auch zu analysieren, wie es insgesamt den Erwartungen entspricht, die mit einer Anschaffung seitens des Stadtrats damit verbunden sind. Aber genau aus diesem Grund wird es ja auch getestet.

Daneben gibt es dann auch noch konkrete Fragen an David Imseng; sei es, dass uns oder mir nochmals aufgezeigt werden sollte, wie man Sprecher:innen einfügt oder zumindest umbenennt oder auch Segmente und Traktanden verschiebt. Daneben ist mir zudem aufgefallen, dass in der Webversion aufgrund eines ausführlichen Votums mit längerem Text gegen Ende des Textes jedes Mal zuerst nach oben gescrollt werden muss, um das Audio-File zu stoppen oder zu starten. Das ist insofern mühsam, als dass die Tonspur vorab weiterläuft, bis nach oben gescrollt werden konnte um es zu stoppen. Eventuell lässt sich dies komfortabler gestalten.

Die hier skizzierten Beobachtungen entsprechen meine ersten, persönlichen Erfahrungen, dies unabhängig vom Entscheid zum späteren, definitiven Einsatz dieses Tools, und insbesondere auch hoffentlich ohne negativen Einfluss auf die Qualität der Transkription. Ich versuche wie immer weiterhin mein Bestes zu geben und dabei mein Verständnis für die politische Debatte zu nutzen, damit meine Arbeit weiterhin zur vollen Zufriedenheit des Auftraggebers ausfällt. Gleichwohl wäre es auch interessant zu erfahren, wie die ersten Erfahrungen seitens der Stadträt:innen ausfällt, wenn man mit der neu zur Verfügung stehenden Tonspur ein bestimmtes Votum nochmals anhören möchte.

Reinach, 25. Juli 2025



Thomas Thurnherr, tete-services