KURZBERICHT ZUM VERWENDUNGSNACHWEIS

BMFTR-Verbundprojekt „Anti-Stotter“ – Teilvorhaben PPB

Zuwendungsempfänger	Otto-Friedrich-Universität Bamberg (PPB)
Förderkennzeichen / Laufzeit	16SV9298 | 01.11.2024–28.02.2026
Teilvorhaben	Empirische Wirk-, Usability- und Anwendungsvalidierung des Anti-Stotter-Systems

1. Ursprüngliche Aufgabenstellung sowie wissenschaftlicher und technischer Stand

Das Teilvorhaben PPB hatte die Aufgabe, die von der PsyAIHance GmbH (PPA) entwickelte neuartige Technologie zur unmittelbaren Verbesserung der Sprechflüssigkeit bei chronischem Stottern empirisch, nutzerbezogen und anwendungsnah zu validieren. Die Technologie konvertiert die Eigenstimme in Echtzeit mit sehr geringer Latenz und spielt sie als ego-dystones, räumlich externalisiertes auditives Feedback zurück. PPB übernahm damit die Brückenfunktion zwischen KI-basierter Stimmkonvertierung, experimentalpsychologischer Operationalisierung, Usability-Prüfung und Prüfung des angenommenen Wirkmodells im Hinblick auf eine spätere klinische Validierung. Im Zentrum standen objektive und subjektive Effekte auf Sprechflüssigkeit, Sprachverständlichkeit, Sprechanstrengung, Akzeptanz und Nutzungsbereitschaft sowie individuelle Wirkbedingungen, Parametrisierung und Wirkungserhalt. Im Projektverlauf wurde die Validierung zudem um die soziale Wahrnehmungsdimension des veränderten Sprechens erweitert.

Wissenschaftlich knüpfte das Vorhaben an Verfahren des Altered Auditory Feedback (AAF) an. Klassische Ansätze wie Delayed Auditory Feedback (DAF) oder Frequency-Altered Feedback können die Sprechflüssigkeit kurzfristig verbessern, weisen jedoch wesentliche Grenzen auf: stark unterschiedliche Responsivität zwischen Personen, mögliche Gewöhnung und Wirkungsabnahme, Sprechverlangsamung sowie Einschränkungen der Natürlichkeit und Alltagstauglichkeit. Der Anti-Stotter-Ansatz adressiert diese Limitationen durch eine personalisierte „Anti-Voice“, geringe Latenz, räumliche 3D-Audio-Externalisierung, kontinuierliche Veränderung der Zielstimme und eine differenzierte Parametrisierung. Das BAMLiD der Universität Bamberg stellte hierfür eine multimodale Laborumgebung mit synchroner Audio- und Multi-Camera-Aufzeichnung, kontinuierlicher Erfassung der Sessionparameter und standardisierten Selbstauskünften bereit.

2. Ablauf des Vorhabens

Der Demonstrator V1 wurde im Februar 2025 an PPB übergeben, in das BAMLiD integriert und für kontrollierte ON/OFF-Vergleiche vorbereitet. Studie 1 erreichte den Proof-of-Concept-Meilenstein M4 im April 2025 planmäßig. Zusätzlich wurden die erhobenen deutschsprachigen Audiodaten durch mehrere Beurteiler hinsichtlich konkreter Stottereregnisse annotiert und dem Konsortialpartner Universität Magdeburg (PPC) für das projektspezifische Fine-Tuning der automatischen Stottererkennung zur Verfügung gestellt.

Der Demonstrator V2 wurde nach konsortial abgestimmter Verzögerung am 27.10.2025 übernommen. Studie 2 untersuchte erstmals die personalisierte Anti-Voice-Bestimmung und die Continuous-Change-Funktion, die perzeptuelle Eigen-/Zielstimmen-Diskrimination, Akzeptabilität sowie individuelle Responder-Muster. Die Befunde wurden unmittelbar in technische Anforderungen für V3 übersetzt, insbesondere hinsichtlich Sprachverständlichkeit, Stimmnatürlichkeit, Artefaktfreiheit, Zielstimmenvielfalt, niedriger Veränderungsraten und robuster Signaltrennung. V3 wurde am 20.12.2025 übergeben.

Studie 3 wurde im Februar 2026 als kontrolliertes Einzelfalldesign mit zwei kontrastierenden Symptomprofilen durchgeführt. Die ursprünglich vorgesehene 14-tägige In-vivo-Erprobung wurde – nach Anzeige und Genehmigung durch den Projektträger – in ein Labor-Setting überführt, da im Projektzeitraum keine über Wochen alltagstauglich einsetzbare mobile Version zur Verfügung stand. Für jede Versuchsperson wurde ein voller Testtag angesetzt; die Erhebung dauerte rund zweieinhalb Stunden. Nach Baselines im freien Sprechen ohne Intervention und mit konventionellem DAF (60 ms) wurden individuelle Best-Fit-Konfigurationen bestimmt und unter drei Lautheitsstufen validiert. In der abschließenden Stabilitätsphase wurden Best-Fit und DAF in ausbalancierter Reihenfolge jeweils über ca. 20 Minuten durchgehenden Sprechens geprüft; jeder der beiden Blöcke war zur Verlaufsanalyse in drei Teilblöcke gegliedert. Die Prüfung lag nach rund zwei Stunden Systemnutzung am Ende der Sitzung und adressierte damit den Wirkungserhalt unter hoher Beanspruchung. Die Sprechperformanz wurde videobasiert durch unabhängige Beurteiler ausgewertet.

Studie 4 wurde im Konsortialtreffen am 17.10.2025 mit dem Projektträger als methodisch begründetes Ersatzdesign für die geplante V3-In-vivo-Usability-Studie abgestimmt und im Dezember 2025 abgeschlossen. Laienhörerinnen und -hörer bewerteten Videoaufnahmen aus Studie 1 im direkten Vergleich zwischen VC-ON und VC-OFF. Dadurch konnte neben der klinisch beziehungsweise performanzbezogen orientierten Perspektive erstmals die soziale Wahrnehmung der Sprechenden untersucht werden. Alle Anpassungen betrafen die zeitliche und methodische Umsetzung, nicht die wissenschaftliche Zielsetzung; sie wurden in enger Abstimmung mit PPA, PPC und VDI/VDE-IT vorgenommen.

3. Wesentliche Ergebnisse und Zusammenarbeit mit anderen Stellen

Studie 1 lieferte den ersten belastbaren innerprojektlichen Wirksamkeitsnachweis. Unter VC-ON verbesserten sich sämtliche sechs sprechbezogenen Zielgrößen konsistent gegenüber VC-OFF; für das zusammenfassende Veränderungsmaß ergab sich eine posteriore Wahrscheinlichkeit von 98 % für eine Verbesserung. In Einzelfällen war die Wirkung so ausgeprägt, dass klinisch relevantes Stotterverhalten unter ego-dystonem VC-Feedback als vollständig eliminiert eingestuft werden konnte. Gleichzeitig wurde eine ausgeprägte interindividuelle Variabilität sichtbar, sodass Responder und Non-Responder unterschieden werden konnten. Rund 70 % der Versuchspersonen äußerten trotz erhöhter kognitiver Anstrengung eine grundsätzliche Kaufbereitschaft für eine mobile Ausführung. Als zentrale Entwicklungsanforderungen wurden höhere Sprachverständlichkeit, natürlichere Stimmqualität und geringere Artefaktbelastung identifiziert.

Studie 2 bestätigte die wahrnehmungsbezogene Grundannahme der personalisierten Anti-Voice: Die Zielstimme wurde signifikant von der eigenen Stimme diskriminiert. Die größte perzeptuelle Distanz entstand bei geschlechtsübergreifenden Konvertierungen; innerhalb desselben Geschlechts war der verfügbare Zielstimmenraum noch weniger ausgeprägt. Für Continuous Change zeigte sich, dass subtile, niedrige Veränderungsraten als angenehmer erlebt wurden, der Variationsraum für eine dauerhafte Stimulus-Neuheitswahrnehmung aber erweitert werden sollte. Die vertiefte Einzelfallauswertung zeigte zudem eine Responder-Logik: Bereits bei moderater Stotterausprägung war ein deutliches Ansprechen erkennbar, während bei milder Symptomatik weniger Verbesserungsspielraum bestand. Situative Selbstauskünfte erwiesen sich als wenig sensitiv; für Studie 3 wurden daher objektive Sprechmaße wieder als primäre Zielgrößen herangezogen.

Studie 3 zeigte, dass eine individuell optimierte Konfiguration entscheidend ist. Als tragende Basiskombination erwiesen sich Stimmkonvertierung, räumliche Externalisierung und dynamische Mikrozeitvariation; beide Versuchspersonen bevorzugten zusätzlich eine Beimischung der eigenen Stimme. Diese Konstellation weist strukturelle Parallelen zum etablierten Choral-Speech-Effekt auf und begründet eine zentrale Hypothese für die Anschlussforschung. Unter der jeweiligen Best-Fit-Konfiguration lag der Anteil gestotterter Silben bei 1,1 % beziehungsweise 0,7 %, gegenüber 12,0 % beziehungsweise 4,0 % im freien Sprechen ohne Intervention. Der Anteil blieb über den gesamten ca. 20-minütigen Best-Fit-Testblock auf diesem Niveau. Unter DAF lagen die Werte bei 2,8 % beziehungsweise 1,7 % und stiegen über die drei Teilblöcke leicht an, während sie unter Best-Fit stabil blieben. Damit fiel die Wirkung in beiden Fällen günstiger aus als beim Stand der Technik; ein Nachweis der Langzeitwirkung über Wochen oder Monate ist damit ausdrücklich noch nicht erbracht.

Die ergänzende Kontrastanalyse zeigte die erreichbare Wirkung auch jenseits eines reinen Frequenzmaßes: Bei der Versuchsperson mit schwerer Symptomatik bewerteten fünf unabhängige Beurteiler die ON-Sequenz in allen sieben Subskalen des von PPB entwickelten multimodalen Beobachtungsfragebogens günstiger als OFF; besonders deutlich war der Unterschied bei Verständlichkeit, Sprachflüssigkeit sowie wahrgenommener Anstrengung und Spannung. Eine automatische Spracherkennung konnte den Inhalt unter OFF nicht angemessen transkribieren, unter ON dagegen vollständig und fehlerfrei. Der Fallkontrast zwischen offener und verdeckter Symptomatik zeigte zudem, dass Wirkung nicht allein über gestotterte Silben erfasst werden sollte: Bei verdecktem Stottern können der Wegfall von Vermeidung, längere und elaboriertere Äußerungen und geringere Sprechanstrengung sensitivere Endpunkte sein. Der Beobachtungsfragebogen umfasst 34 Items auf sieben Subskalen.

Studie 4 ergänzte die soziale Außenperspektive. Unter VC-ON wurden die Sprechenden tendenziell als weniger angespannt beziehungsweise unsicher wahrgenommen. Zugleich zeigte sich bei V1, dass eine leichte Sprechverlangsamung die positiven Effekte der Stotterreduktion sozial überlagern und in einzelnen Fällen ungünstigere Zuschreibungen von Intelligenz oder Vertrauenswürdigkeit auslösen konnte. Diese Befunde sind V1- und laborbezogen zu rahmen; auch das sichtbar getragene Laborequipment stellt einen möglichen Einflussfaktor dar. In der V3-Kontrastanalyse fiel die Subskala soziale Wahrnehmung unter aktivierter Konvertierung dagegen durchgehend günstiger aus. Damit bestehen Hinweise, dass die Demonstratorlinie das bei delay-basierten Verfahren relevante Spannungsfeld zwischen sprechförderlicher Wirkung und sozialer Stigmatisierung reduzieren kann.

Die Zusammenarbeit im Verbund war eng, kooperativ und iterativ: PPB überführte empirische Befunde fortlaufend in Anforderungen an PPA, während PPC den annotierten deutschsprachigen Datensatz für die automatische Stottererkennung nutzte. Kontakte zu Betroffenenverbänden und Behandelnden verbesserten die Rekrutierung. Insgesamt liefern die vier Studien eine kohärente Evidenzbasis für das Potenzial ego-dystoner Echtzeit-Stimmkonvertierung und präzisieren zugleich die notwendigen Randbedingungen: Personalisierung, hohe Verständlichkeit und Natürlichkeit, objektive Endpunkte und eine multidimensionale Erfassung von Flüssigkeit, Vermeidung, Sprechanstrengung, kommunikativer Handlungsfähigkeit und sozialer Wirkung. Die Befunde beruhen auf Laborerhebungen mit kleinen Stichproben; für V3 auf zwei gezielt kontrastierenden Einzelfällen. Größere kontrollierte Stichproben und längere alltagsnahe Nutzungszeiträume bleiben daher Aufgabe der klinischen Anschlussvalidierung. Die wirtschaftliche Verwertung liegt primär bei PPA; PPB schafft hierfür die wissenschaftliche Evidenzgrundlage für klinische, regulatorische und versorgungsbezogene Anschlusswege.

Sachbericht zum Verwendungsnachweis
Anti-Stotter | Teilvorhaben PPB

Zuwendungsempfänger / ausführende Stelle	Otto-Friedrich-Universität Bamberg (PPB)
Förderkennzeichen	16SV9298
Vorhabenbezeichnung	Anti-Stotter
Teilvorhaben PPB	Empirische Wirk-, Usability- und Anwendungsvalidierung des Anti-Stotter-Systems
Laufzeit des Vorhabens	01.11.2024 bis 28.02.2026

1. Aufzählung der wichtigsten wissenschaftlich-technischen Ergebnisse und anderer wesentlicher Ereignisse

1.1 Einordnung des PPB-Teilvorhabens

Das Teilvorhaben der Otto-Friedrich-Universität Bamberg (PPB) fokussierte im Verbundvorhaben Anti-Stotter auf die empirische, nutzerbezogene und anwendungsnahe Validierung der von PsyAIHance GmbH (PPA) entwickelten neuartigen Technologie zur unmittelbaren Verbesserung der Sprechflüssigkeit von Menschen, die von chronischem Stottern betroffen sind. Die Technologie beruht auf einer KI-basierten Stimmkonvertierung in Echtzeit (engl.: Voice Conversion, VC): Die eigene Stimme der nutzenden Person wird während des Sprechens in eine Zielstimme mit anderer Stimmidentität umgewandelt und unmittelbar über Kopfhörer zurückgespielt. Die Person hört ihre eigene Äußerung damit so, als spräche sie ein anderer Mensch. Der Eindruck einer anderen Stimme wird durch weitere akustische Gestaltungsmerkmale des Feedbacksignals verstärkt, unter anderem durch räumliches Audio: Das Signal wird außerhalb des Kopfes verortet, während die eigene Stimme sonst als von innen kommend wahrgenommen wird. Das Zusammenwirken dieser Merkmale, die die eigene Stimme als fremd erscheinen lassen, wird im Folgenden als **ego-dystones VC-Feedback** bezeichnet. Dem Vorhaben liegt die Annahme zugrunde, dass diese Fremdheitswahrnehmung der eigenen Stimme die Sprechflüssigkeit unmittelbar verbessert.

Während PPA die technologischen Demonstratoren in sukzessiven Stufen entwickelte, übernahm PPB deren empirische Erschließung: die Überführung in laborpsychologisch prüfbare Fragestellungen, die Untersuchung von Nutzererleben und Handhabbarkeit sowie die Prüfung der zugrunde liegenden Wirkweise im Hinblick auf eine spätere klinische Validierung. Aus den Befunden wurden jeweils gezielte Gestaltungsempfehlungen an PPA abgeleitet, die in die Entwicklung der folgenden Demonstratorstufe eingingen.

Die Studienarbeiten zu Studie 1, Studie 2 und Studie 4 wurden durch den Projektangestellten Herrn Dr. Niklas Döbler (WiMi) sowie ergänzend durch Herrn Thomas Kollmann (Stud.HK), jeweils unter Supervision von Prof. Dr. Claus-Christian Carbon, durchgeführt. Studie 3 wurde maßgeblich von Prof. Dr. Claus-Christian Carbon durchgeführt.
Die Studienzuordnung wurde gegenüber dem Zwischennachweis 2025 präzisiert. Maßgeblich ist die in diesem Bericht verwendete Nummerierung: Studie 1 bis Studie 4.

1.2 Übernahme und Laborintegration des Demonstrators V1

V1 wurde im Februar 2025 an PPB übergeben und im BAMLiD (Bamberger Living Lab) in Betrieb genommen.

Aus PPB-Sicht war V1 die erste bereitgestellte Demonstratorstufe, um die im Antrag beschriebene Wirkung des ego-dystonen Stimmfeedbacks auf den Sprechfluss bei Stottern experimentell zu prüfen. Hierfür wurde V1 in das BAMLiD-Versuchsumfeld integriert, das eine multimodale Erfassung der Sprechperformanz und der psychophysiologischen Begleitprozesse erlaubt: synchron geführte Audio-Aufnahmen, Multi-Camera-Aufzeichnung der Versuchspersonen sowie kontinuierliche Aufzeichnung der V1-Session-Parameter (ON/OFF-Bedingung, Konvertierungs-Konfiguration). Ergänzend kamen Versuchspersonen-Selbstauskünfte über LimeSurvey (kognitive Anstrengung, Akzeptanz, Bedingungs-Nachbefragung) sowie eine im Anschluss an die Testung ausgehändigte Schwierigkeits-Skala zur subjektiven mentalen Anstrengung und Stotterschwere zum Einsatz.
In strukturierten Interviews kommunizierten Versuchsleitung und Versuchsperson über Mikrofone und kabelgebundene Kopfhörer; beide hörten die Stimme des Gegenübers, sodass die dialogische Interaktion vollständig erhalten blieb. Die experimentelle Manipulation betraf ausschließlich, wie die Versuchsperson sich selbst hörte: Sie nahm ihre eigene, in Echtzeit stimmkonvertierte und im 3D-Audio-Feld externalisierte Stimme wahr, während die Versuchsleitung ihre eigene Stimme und die der Versuchsperson unverändert hörte. Über die V1-Lab-GUI wurden die Bedingungswechsel zwischen ON- und OFF-Phasen gesteuert. Damit waren die Voraussetzungen für die erste Proof-of-Concept-Studie (Studie 1) geschaffen.

1.3 Studie 1 — Proof of Concept (PoC)-Effizienz von V1 (AP 2.1, M4)
PPB hat den Meilenstein M4 im April 2025 planmäßig erreicht. Die Studie diente der ersten empirischen Prüfung, ob das ego-dystone VC-Feedback unter Laborbedingungen einen sprechförderlichen Nutzen entfaltet.

Studie 1 zeigte, dass das ego-dystone VC-Feedback bereits in der frühen Entwicklungsphase (V1) die Sprechperformanz signifikant gegenüber der OFF-Bedingung verbessert. Der Befund wird durch eine ergänzende bayesianische Auswertung gestützt und erweist sich dabei als über alle Zielgrößen konsistent: Sämtliche sechs sprechbezogenen Zielgrößen weisen in dieselbe Richtung: der Anteil gestotterter Silben (PCTSS), der SSI-4 als standardisiertes Maß der Stotterschwere, die Dauer der Unterbrechungen, die expertenbewertete Sprechnatürlichkeit sowie die subjektive Stotterschwere im Interview und beim Vorlesen (Abbildung A1). Für das zusammenfassende Veränderungsmaß über alle Zielgrößen hinweg ergibt sich eine posteriore Wahrscheinlichkeit von 98 Prozent für eine Verbesserung unter dem VC-Feedback (Abbildung A2, siehe Auch Abbildung A3).

In Einzelfällen war die Wirkung so ausgeprägt, dass das klinisch relevante Stotterverhalten unter ego-dystonem VC-Feedback als vollständig eliminiert eingestuft werden konnte.
Zugleich wurde — konsistent mit dem Forschungsstand zum Altered Auditory Feedback (AAF) — eine ausgeprägte interindividuelle Varianz in der Responsivität beobachtet. Auf Basis einer ergänzenden versuchspersonenindividuellen Auswertung ließen sich Responder (Wirkung bereits auf der aktuellen Entwicklungsstufe) und Non-Responder (Wirkung auf der aktuellen Entwicklungsstufe noch unzureichend) trennscharf identifizieren: Bei Respondern stellte sich die Wirkung innerhalb weniger Sekunden nach Zuschalten des VC-Feedbacks ein, während Non-Responder auch bei fortgesetzter Nutzung keine merkliche Verbesserung zeigten. Das Responder-/Non-Responder-Muster wird in den nachfolgenden Auswertungen systematisch berücksichtigt (zunächst in der V2-Erprobung).
Die Versuchspersonen äußerten — bei gleichzeitig berichteter erhöhter kognitiver Anstrengung beim Sprechen unter dem frühen Demonstrator-Stand — eine grundsätzliche Bereitschaft, die Technologie in einer mobilen Ausführung käuflich zu erwerben (70 % Kaufbereitschaft). Dass diese Bereitschaft bereits gegenüber einem laborbetriebenen Demonstrator besteht, verweist auf einen hohen ungedeckten Bedarf an einer Lösung, die Betroffene im Sprechmoment unterstützt.

PPB hat im Berichtszeitraum eine über die im Antrag dokumentierten Arbeitspakete hinausgehende

Konsortialleistung erbracht: Die in Studie 1 erhobenen Audiodaten wurden am Lehrstuhl der Universität Bamberg von mehreren Beurteilern hinsichtlich des Vorliegens von Stottereignissen annotiert. Der daraus resultierende deutschsprachige Datensatz wurde dem Konsortialpartner Universität Magdeburg (PPC) zum Zweck des projektspezifischen Fine-Tunings des KI-Modells zur automatischen Stottererkennung übergeben. Der Datensatz schließt eine im Forschungsfeld dokumentierte Lücke: Öffentlich verfügbare Disfluency-Korpora sind überwiegend englischsprachig; die im PPB erhobenen deutschsprachigen Daten erweitern die Trainingsbasis für die ML-basierte Stottererkennung im Deutschen. Damit ist die für die Fertigstellung der einschlägigen PPC-Arbeitspakete benötigte Fine-Tuning-Datenbasis bereitgestellt.

Das zentrale Ergebnis aus Studie 1 umfasst zwei Befunde: erste belastbare Hinweise auf die sprechfördernde Wirkung des ego-dystonen VC-Feedbacks sowie die Erkenntnis, dass das Nutzererleben über die Stotterreduktion hinaus entscheidend durch Sprachverständlichkeit und perzeptuelle Qualität des Feedbacksignals geprägt ist — erkennbar auch an der erhöhten kognitiven Anstrengung unter dem frühen Demonstratorstand. Hieraus folgte die Empfehlung an PPA, für V2 und V3 Sprachverständlichkeit, Stimmnatürlichkeit und Artefaktfreiheit als primäre endnutzerseitige Akzeptanzkriterien zu priorisieren.

1.4 Studie 2 — V2-Übernahme und Lab-Erprobung des Nutzererlebens (AP 3.1, M6)

Der V2-Demonstrator wurde am 27.10.2025 mit zeitlichem Verzug — konsortial abgestimmt und vom Projektträger abgenommen — an PPB übergeben und im BAMLiD einer explorativen qualitativen wie quantitativen Erprobung mit Versuchspersonen unterzogen.

Für die PPB-Studienreihe war V2 ein wichtiger Zwischenschritt: Mit der personalisierten Anti-Voice-Bestimmung und der Continuous-Change-Funktion wurden die beiden Kernmechanismen des Ansatzes erstmals experimentell zugänglich. Damit konnte PPB von den binären ON/OFF-Vergleichen unter V1 zu parametrisch graduierten Versuchsdesigns übergehen.

Die in V2 implementierte Anti-Voice-Bestimmung wählt die Zielstimme anhand ihres perzeptuellen Abstands zur Eigenstimme aus. Dieser Abstand ist im Projektansatz eine zentrale Größe und wird damit erstmals systematisch einstellbar und in seiner Wirkung untersuchbar. Ergänzend erlaubt V2 die Parametrisierung der Geschwindigkeit und des Verlaufs des Zielstimmenwechsels. Damit wird der Wechsel selbst zur experimentellen Variablen und die Frage prüfbar, welche Änderungsrate erforderlich ist, damit das Feedback über die Anwendungsdauer als fremd wahrgenommen wird.

Perzeptuelle Diskriminationstests (AB-Vergleich Eigen- gegen Zielstimme). Die V2-Erprobung lieferte für die personalisierte Anti-Voice-Bestimmung den zentralen Proof-of-Concept (Machbarkeitsnachweis): Die Versuchspersonen diskriminierten die systemseitig bereitgestellte Zielstimme statistisch signifikant von der eigenen Stimme. Damit ist die wahrnehmungsbezogene Voraussetzung für die im Projektansatz prognostizierte sprechfördernde Wirkung erfüllt.

Eine Auflösung nach Konvertierungstyp zeigte, dass die perzeptuelle Unähnlichkeit zwischen Eigen- und Zielstimme — also die von den Versuchspersonen beurteilte Verschiedenheit beider Stimmen — primär durch geschlechtsübergreifende Konvertierungen (m → w, w → m) getragen wird, während sie bei Konvertierungen innerhalb desselben Geschlechts (m → m, w → w) geringer ausfiel. Diese Asymmetrie verweist auf eine in V2 noch ausbaufähige Variationsbreite des Zielstimmenpools: Eine hinreichende Unähnlichkeit, unabhängig von der Geschlechterkonstellation, setzt einen breiteren Pool verfügbarer Zielstimmen voraus, der in V2 noch nicht realisiert war.

Continuous Change. Für die Continuous-Change-Funktion (kontinuierliche Veränderungsfunktion) ergab die V2-Erprobung eine Präzisierung der methodischen Anforderungen: Der verfügbare perzeptuelle Variationsraum interpolierter Zielstimmen erwies sich als noch nicht hinreichend differenziert, um eine konstante Stimulus-Neuheitswahrnehmung sicherzustellen. Ergänzend zeigte die subjektive Akzeptabilitätsbewertung, dass subtile graduelle Veränderungen — also eine niedrige Change-Rate — als angenehmer erlebt werden. Daraus ergibt sich für die Folgeentwicklung die doppelte, perzeptuell

differenziert formulierte Anforderung.

Als explorative Zusatzvariable wurde die subjektive Selbsteinschätzung der Stotterschwere unter ON- und OFF-Bedingung erhoben. Auf Mittelwertbasis blieb die Datenlage uneindeutig. Ausschlaggebend war die Zusammensetzung der Stichprobe, in der Versuchspersonen mit milder Stotterausprägung überwogen. Die Rekrutierung war bewusst offen über das gesamte Spektrum angelegt, da bei einem neuartigen Wirkansatz nicht vorauszusetzen war, dass die aus der AAF-Forschung bekannte Abhängigkeit des Effekts vom Ausgangsniveau in gleicher Weise gilt. Der Befund bestätigt diese Übertragbarkeit: Die zugrundeliegenden Wirkmodelle setzen eine instabile auditorisch-motorische Selbststeuerung voraus, die durch externe Rückkopplung funktional entlastet wird — bei milder Symptomatik ist diese Instabilität gering ausgeprägt, sodass der Spielraum für eine akute Verbesserung entsprechend kleiner ausfällt. Hinzu kommt eine Einschränkung, die sich erst in dieser Erprobung zeigte: Während Studie 1 belegt hatte, dass Betroffene ihre allgemeine Stotterschwere valide einschätzen, erwies sich die Selbstauskunft für situative Veränderungen einzelner Sprechepisoden als deutlich weniger sensitiv. Die Einschätzung im unmittelbaren Sprechvollzug bildet den tatsächlichen Bedingungsunterschied damit nur eingeschränkt ab. Fachgespräche, die PPB im Berichtszeitraum mit Vertreterinnen und Vertretern der klinisch-logopädischen Versorgung geführt hat, bestätigten diese Beobachtung.

Dies gab Anlass, die Auswertung um vertiefte einzelfallbezogene Analysen zu ergänzen (Abbildungen A3-A5). Diese zeigten eine ausgeprägte Responder-Logik: Bereits bei moderater Stotterausprägung war ein deutliches Ansprechverhalten auf die VC-Manipulation erkennbar, während die Sprechperformanz bei milder und sehr milder Ausprägung weitgehend unverändert blieb. Da die Stichprobe überwiegend aus letzteren bestand, konnte der Effekt auf Aggregatebene mit der gegebenen Stichprobengröße nicht gezeigt werden.

Aus der V2-Erprobung folgten zwei methodische Konsequenzen für die weitere Untersuchung. Erstens wurde die Rekrutierung für Studie 3 auf Versuchspersonen mit mindestens moderater Ausprägung nach SSI-4 beschränkt. Zweitens wurden objektive Maße der Sprechperformanz wieder als primäre Zielgrößen herangezogen, da die situative Selbsteinschätzung als alleiniges Wirksamkeitsmaß nicht hinreichend trennscharf ist. Für die Ausführung der Technologie leitete PPB zudem die Gestaltungsempfehlung ab, der nutzenden Person eine objektive Rückmeldung zur aktuellen Sprechflüssigkeit zu geben und damit das Selbstmonitoring zu unterstützen. Eine solche Rückmeldung macht den Nutzen des Systems für die anwendende Person unmittelbar sichtbar.

Aus den Erfahrungen der Studien 1 und 2 leitete PPB eine Anforderungsanalyse für das laborfähige System ab, die in die V3-Entwicklung einfloss. Im Vordergrund standen zwei Anforderungen: eine weiter verbesserte Sprachverständlichkeit und Stimmnatürlichkeit, da Konvertierungsartefakte die untersuchten Zielvariablen überlagern können; sowie eine zuverlässige Trennung von Versuchspersonen- und Versuchsleitungsstimme (Übersprechungseffekte). Letztere ist über den Laborbetrieb hinaus relevant, da sie in jeder realen Gesprächssituation gegeben sein muss. Die Befunde wurden in enger Abstimmung mit PPA in die technische Weiterentwicklung zu V3 aufgenommen.

1.5 Studie 3 — Individuelle Parameterwahl des ego-dystonen Voice-Conversion-Feedbacks und Stabilität des Sprechflüssigkeitseffekts unter V3 (AP 2.2, M9 — durchgeführt Februar 2026; Ablaufanpassung angezeigt und genehmigt)

Fragestellung. Studie 3 untersucht drei Fragen: erstens, welche Parameterkonfiguration je Person die günstigste Wirkung erzeugt; zweitens, welche Wirkung unter dieser Konfiguration erreicht wird; drittens, ob die Wirkung über eine längere zusammenhängende Sprechdauer stabil bleibt oder Hinweise auf Adaptationserscheinungen zeigt, wie sie für konventionelles Altered Auditory Feedback (AAF) dokumentiert sind. Die dritte Frage ist für die Alltagstauglichkeit der Technologie zentral: Die im Antrag als zentrale Limitierung bestehender AAF-Verfahren benannte rasch einsetzende Gewöhnung an das

veränderte Feedback und die dadurch nachlassende Wirkung bestimmen, ob eine Versorgungslösung über eine längere Anwendungsdauer hinweg trägt.

Untersuchte Wirkmechanismen. Der bereitgestellte Demonstrator V3 verfügt nach Angabe des Entwicklungspartners PPA über zwei Funktionen, die der Gewöhnung entgegenwirken sollen: eine sich kontinuierlich wandelnde Stimmidentität im Feedbacksignal — die im Antrag beschriebene Kernlösung — sowie eine dynamische Mikrozeitvariation des Feedbacksignals (Time Warping), die PPA über den vereinbarten Leistungsumfang hinaus explorativ ergänzt hat aufgrund der Rückmeldungen von PPB bzgl. Demonstratorstufen V1 und V2. Beiden liegt dasselbe Grundprinzip zugrunde: Sie verhindern ein gleichbleibendes Reizmuster, an das sich die maßgeblichen Hirnmechanismen, die Wahrnehmung ermöglichen, anpassen könnten. Ob dieses Prinzip trägt, ist eine empirische Frage; Studie 3 prüft die kombinierte Wirkung beider Funktionen erstmals unter kontrollierten Bedingungen.

Bestimmung der Prüfkonfiguration. Vorgeschaltet war die Bestimmung derjenigen Konfiguration, unter der die Überprüfung des Wirkungserhalts erfolgen sollte: Der Wirkungserhalt ist sinnvoll nur für diejenige Einstellung zu prüfen, die für die betreffende Person am besten funktioniert — also für die Konfiguration, unter der die Sprechflüssigkeit situativ merklich verbessert ist und die zugleich als angenehm genug für eine längere Anwendung beurteilt wird. Dazu wurden die im Demonstrator V3 verfügbaren Systemkombinationen im direkten Vergleich erprobt und nach jedem Block sowohl von der Versuchsperson als auch von der Versuchsleitung beurteilt.

Erkenntniswert dieses Schritts. Der Parameterraum von V3 unterscheidet sich grundlegend von dem bisheriger Verfahren: Herkömmliche Anti-Stotter-Hilfsmittel bieten wenige Justierungsmöglichkeiten — im Wesentlichen den zeitlichen Versatz des Feedbacks (Delay) und dessen Tonhöhenverschiebung (Pitch Shift) —, während der modular aufgebaute Demonstrator V3 die Feinjustierung von rund zwanzig akustischen Parametern erlaubt, die die Eindrucksbildung einer „anderen Stimme“ beeinflussen. Da der VC-basierte Wirkansatz neuartig ist und der Forschungsstand keine Herleitung einer geeigneten Parametrisierung zulässt, ist dieser Raum ohne empirische Vorarbeit weder für die Anwendung einstellbar noch für eine automatische Kalibrierung zugänglich.

Anpassung des Studienablaufs. Der Studienablauf wurde auf Basis der Ergebnisse der vorgehenden Studienergebnissen gegenüber der Antragsplanung angepasst: Die ursprünglich vorgesehene 14-tägige In-vivo-Erprobung wurde in ein Labor-Setting überführt. Konsortial und mit dem Projektträger abgestimmt wurde, dass die Ausführung des Demonstrators als über Wochen alltagstauglich einsetzbare mobile Lösung aus fachlich zwingenden Gründen nicht Gegenstand des Vorhabens ist; eine mehrwöchige In-vivo-Erprobung war damit im Projektzeitraum nicht durchführbar. Das Labor-Setting erlaubt demgegenüber die fein abgestufte und reproduzierbare Variation der Experimentalparameter bei deutlich höherer interner Validität sowie einer effizienteren Nutzung der Testzeit je Versuchsperson. Die Änderung wurde dem Projektträger am 01.12.2025 angezeigt und genehmigt.

Studiendesign. Die Untersuchung wurde gezielt als Einzelfalldesign mit zwei Versuchspersonen angelegt. Maßgeblich war die Fragestellung selbst: Gesucht ist die je Person günstigste Parameterkonfiguration — eine personenspezifische Ausrichtung, die sich nicht über mehrere Personen hinweg mitteln lässt. Ein aggregierendes Gruppendesign hätte über eine große Spannweite individueller Reaktionen gemittelt und damit genau die parameterbezogenen Effekte überdeckt, die hier den Untersuchungsgegenstand bilden. Im Einzelfalldesign durchläuft jede Versuchsperson demgegenüber sämtliche Bedingungen und dient als ihre eigene Kontrolle, sodass der Vergleich innerhalb der Person erfolgt.
Das gewählte Design ermöglicht so für jede Versuchsperson innerhalb einer Sitzung eine Reihe systematisch kontrastierter Bedingungen und macht Verlaufseffekte messbar, wie sie für die Prüfung des Wirkungserhalts erforderlich sind. Die Adaptationsfrage wurde bewusst innerhalb einer Sitzung erhoben, da sich Unterschiede in Tagesform und aktueller Symptomatik zwischen zwei Terminen bei zwei

Versuchspersonen nicht ausmitteln lassen. Die Auswahl der Versuchspersonen folgte einem strengen Kontrastkriterium: Die beiden Fälle bilden die klinisch bedeutsamen Erscheinungsformen ab — eine stark ausgeprägte, offene Symptomatik sowie eine moderate Oberflächensymptomatik bei ausgeprägter aktiver Vermeidung, in der Literatur als offenes und verdecktes Stottern (overt und covert stuttering) beschrieben. Damit lässt sich prüfen, ob die Wirkung bei offener und bei verdeckter Symptomatik gleichermaßen auftritt und auf welchen Ebenen sie jeweils sichtbar wird.

Ablauf und Erhebung. Für jede Versuchsperson wurde ein voller Testtag angesetzt; die Erhebung selbst dauerte rund zweieinhalb Stunden. Vorab wurden eine ausführliche Anamnese zur Symptomatik — Schweregrad, Erscheinungsform und Vermeidungsverhalten —, die Tagesform (Schlaf, Stressbelastung, subjektive Stotterschwere am Testtag) sowie eine Pegel- und Komfortprüfung erhoben. Es folgte eine Eingewöhnungsphase mit freiem Sprechen unter dem System ohne Wertung, damit die späteren Bedingungsunterschiede nicht durch die Erstkonfrontation mit dem veränderten Feedback überlagert werden. Die Erhebung gliederte sich in vier aufeinander aufbauende Phasen.

Erste Phase: Vergleichsgrundlagen. Erhoben wurden zwei Baselines — freies Sprechen ohne Feedback-Intervention zur Erfassung der unbehandelten Sprechflüssigkeit sowie zeitverzögertes auditives Feedback (Delayed Auditory Feedback, DAF) mit einer Verzögerung von 60 ms als Stand der Technik im Bereich der Stotterintervention. Die zweite Baseline erlaubt den Vergleich nicht nur mit der Nichtbehandlung, sondern auch mit dem in kommerziellen Anti-Stotter-Hilfsmitteln am weitesten verbreiteten Verfahren.

Zweite Phase: Bestimmung der individuellen Best-Fit-Konfiguration. In vier aufeinanderfolgenden Blöcken wurde geprüft, welche Konfiguration je Person am günstigsten ist. Ausgangspunkt war die im Antrag beschriebene Kernlösung: ego-dystones Stimmkonvertierungs-Feedback (Anti-Stimme) mit räumlicher Externalisierung — einer 3D-Audio-Wiedergabe über kopfbezogene Übertragungsfunktionen (Head-Related Transfer Function), die das Feedbacksignal außerhalb des Kopfes verortet und damit zur Wahrnehmung als fremde Stimme beiträgt — sowie kontinuierlicher Wandlung der Zielstimme. Innerhalb der Kernlösung wurden die Parameter graduell justiert, insbesondere die Überblendungsgeschwindigkeit zwischen den Zielstimmen.
Auf diese Grundlage wurden drei optionale Funktionen zugeschaltet: eine dynamische Mikrozeitvariation des Feedbacksignals (Time Warping), die Beibehaltung der natürlichen Stimmlage (sog. „F0“-Erhaltung) sowie die Einblendung der Eigenstimme. Geprüft wurde damit sowohl in welcher Einstellung die Kernlösung am günstigsten wirkt als auch ob eine der optionalen Funktionen diese Wirkung verbessert. Nach jedem Block beurteilte die Versuchsperson die Bedingung anhand fünf Dimensionen — Sprechflüssigkeit, Anstrengung, Natürlichkeit und Angenehmheit, Irritation und Ablenkung sowie Nutzungsbereitschaft — und benannte im Freikommentar den wahrgenommenen Hauptunterschied gegenüber dem Vorblock. Aus diesen Beurteilungen und der begleitenden Beobachtung der Versuchsleitung wurde je Versuchsperson diejenige Konfiguration ausgewählt, unter der die Sprechflüssigkeit merklich verbessert und die zugleich als angenehm genug für eine längere Anwendung beurteilt wurde. Damit ist die Personalisierung, die im Antrag als Kernlösung beschrieben ist, erstmals prozedural prüfbar: Die Zielkonfiguration wird nicht vorgegeben, sondern von der nutzenden Person im direkten Vergleich ausgewählt und graduell bestimmt.

Dritte Phase: Validierung. Die ausgewählte Best-Fit-Konfiguration wurde unter drei Lautheitsstufen geprüft, um die Breite des Bereichs zu bestimmen, in dem sie trägt.

Vierte Phase: Wirksamkeit bei fortgesetzter Anwendung und im Vergleich zum Stand der Technik. Die abschließende Phase adressiert die Kernfrage von Studie 3, ob die Wirkung der eingesetzten Technologie nach rund 2 ½ Stunden zusammenhängender Systemnutzung, in deren Verlauf potentiell eine Gewöhnung an das veränderte Feedback auftreten könnte, bestehen bleibt. Über jeweils ca. 20 Minuten durchgehenden Sprechens wurde das Sprechverhalten unter der individuell bestimmten Best-Fit-

Konfiguration mit dem unter konventionellem DAF verglichen. Beide Bedingungen wurden in drei Teilblöcke gegliedert, um den Verlauf innerhalb der jeweiligen Bedingung zu erfassen. Die Reihenfolge beider Bedingungen wurde zwischen den Versuchspersonen ausbalanciert. Die Phase lag am Ende der Sitzung; die Prüfung erfolgte damit unter der höchsten im Studienverlauf erreichten Beanspruchung. Nach jedem der beiden Blöcke wurde erfragt, ob die Versuchsperson das Sprechen im Verlauf leichter oder schwerer empfunden hatte, ob der Effekt nachließ, welcher Abschnitt am günstigsten war und ob die Versuchsperson unter dieser Bedingung länger hätte sprechen wollen. Abschließend wurden beide Bedingungen direkt gegenübergestellt — hinsichtlich Gesamtbewertung, Natürlichkeit, Anstrengung auf Dauer und Alltagspräferenz.

Beurteilung der Sprechperformanz. Die Sitzungen wurden videodokumentiert; die Auswertung der Sprechperformanz erfolgte nachträglich anhand dieses Materials, da eine verlässliche Erfassung während des laufenden Gesprächs nicht möglich ist. Für die Verlaufsprüfung der Stabilitätsphase wurde der Anteil gestotterter Silben (Percentage of Stuttered Syllables, PCTSS) je Teilblock durch zwei Beurteiler unabhängig voneinander ausgezählt (Abbildung A6). Als Frequenzmaß ist er weitgehend unabhängig von der Einschätzung der Beurteilenden.

Für die Kontrastanalyse ausgewählter Sequenzen wurde der im Projekt von PPB entwickelte multimodale Beobachtungsfragebogen mit 34 Items auf sieben Subskalen eingesetzt; die Bewertung erfolgte durch fünf Beurteiler unabhängig voneinander (Abbildung A7). Die Bewertungskriterien wurden vor der Datenerhebung festgelegt.

Befund 1 — Individuelle Parameterwahl und Eigenstimmanteil

Als tragende Basiskombination erwies sich die Verbindung aus Stimmkonvertierung, räumlicher Externalisierung und dynamischer Mikrozeitvariation. Beide Versuchspersonen wählten unabhängig voneinander eine Konfiguration, die diese Basis um eine Beimischung der eigenen Stimme in das Feedbacksignal ergänzt, und präferierten diese Bedingung deutlich gegenüber den Varianten ohne Eigenstimmanteil.

Von den beiden Erweiterungen der Basiskombination wurde die Eigenstimmen-Beimischung bei beiden Versuchspersonen günstiger beurteilt als die Beibehaltung der natürlichen Stimmlage. Beide Varianten nähern das Feedback der Eigenstimme an − die eine über deren Präsenz im Signal, die andere über die erhaltene Tonhöhe. Dass die Beimischung bevorzugt wurde, spricht dafür, dass die gleichzeitige Präsenz einer zweiten Stimme den Ausschlag gibt.

Dieser Befund ist über die Akzeptanzfrage hinaus bedeutsam. Ein Feedbacksignal, das die eigene Stimme gemeinsam mit einer perzeptuell distanten Zielstimme enthält, entspricht strukturell der Konstellation des sog. *Choral Speech* — des in der Stotterforschung am besten belegten flüssigkeitsinduzierenden Effekts, der bislang die tatsächliche Anwesenheit eines zweiten, physikalischen Sprechers voraussetzt. Die konvergente Präferenz beider Versuchspersonen liefert damit einen Hinweis darauf, dass die ego-dystone Stimmkonvertierung in ihrer präferierten Konfiguration den Choral-Speech-Effekt technisch nachbildet. Der Befund ist in zwei Fällen hypothesengenerierend und begründet die Eigenstimmen-Beimischung als vorrangigen Parameter der Folgeforschung.

Befund 2 —Erreichbarer Zielzustand unter V3 (Kontrastanalyse)

Zur Bestimmung der Wirkung im Einzelfall wurden zwei Sequenzen aus dem Videomaterial der Versuchsperson mit schwerer Symptomatik gegenübergestellt: eine Episode mit typischer Stottersymptomatik unter OFF gegenüber einer Episode flüssigen Sprechens unter ON. Die Auswahl bildet damit nicht den durchschnittlichen Bedingungsunterschied ab, sondern zeigt, welcher Zustand unter der Intervention erreichbar ist — und damit den Nutzen für Betroffene.
Die Bewertung erfolgte durch fünf Beurteiler unabhängig voneinander mit dem im Projekt entwickelten multimodalen Beobachtungsfragebogen. Sämtliche Beurteiler bewerteten die ON-Sequenz in allen sieben Subskalen günstiger; die individuellen Differenzen im Gesamtwert lagen zwischen 1,74 und 5,18

Skalenpunkten (Abbildung A6). Der deutlichste Unterschied ergab sich in der Verständlichkeit, wo vier der fünf Beurteiler unter ON den Skalenmaximalwert vergaben. Ebenfalls ausgeprägt fielen die Unterschiede in Sprachflüssigkeit und Artikulation sowie in der wahrgenommenen Anstrengung und Spannung aus; soziale Wahrnehmung, Körpersprache und Mimik wurden unter ON durchgehend günstiger beurteilt. Ergänzend wurden beide Sequenzen einer automatischen Spracherkennung zugeführt. Unter OFF ergab sich keine dem gesprochenen Inhalt entsprechende Transkription; unter ON wurde der Inhalt vollständig und fehlerfrei transkribiert. Der Unterschied in der Verständlichkeit zeigt sich damit auch unabhängig von menschlicher Beurteilung — nicht als graduelle Verbesserung, sondern als Übergang von maschinell nicht zugänglicher zu vollständig erfasster Sprache. Der Zielzustand der Intervention — flüssiges Sprechen auch bei schwerer Symptomatik — wurde unter V3 damit erreicht.

Befund 3 — Wirksamkeit unter fortgesetzter Anwendung

Unter der Best-Fit-Konfiguration lag der Anteil gestotterter Silben bei 1,1 Prozent (VP1) und 0,7 Prozent (VP2), gegenüber 12,0 beziehungsweise 4,0 Prozent im freien Sprechen ohne Intervention (Abbildung A6). Der Anteil blieb über den gesamten Testblock auf diesem Niveau. Da die Prüfung am Ende der ca. 2 ½-stündigen Sitzung lag, erfolgte sie unter der höchsten im Studienverlauf erreichten Beanspruchung.
Die im Antrag beschriebene Kernlösung erweist sich damit in ihrer individuell bestimmten Konfiguration auch bei fortgesetzter, beanspruchender Anwendung als wirksam.

Befund 4 — Vergleich mit dem Stand der Technik

Unter DAF lag der Anteil gestotterter Silben bei 2,8 Prozent (VP1) und 1,7 Prozent (VP2) und damit ebenfalls deutlich unter den Werten des freien Sprechens ohne Intervention. Unter der Best-Fit-Konfiguration fiel der Anteil gestotterter Silben bei beiden Versuchspersonen jedoch noch niedriger aus als unter DAF, was auf eine stärkere sprechfördernde Wirkung des ego-dystonen VC-Feedbacks hindeutet. Im Verlauf der drei Teilblöcke stieg der Anteil unter DAF leicht an, während er unter der Best-Fit-Konfiguration unverändert blieb. Damit ist die Frage des Wirkungserhalts erstmals im direkten Vergleich mit dem verbreitetsten AAF-Verfahren adressiert. Das Befundmuster entspricht der im Antrag beschriebenen anti-adaptiven Wirkannahme: Während das statische Feedback des Vergleichsverfahrens im Verlauf an Wirkung verliert, bleibt die Wirkung des kontinuierlich veränderten VC-Feedbacks erhalten. Der durch ein experimentelles Setting inhärent limitierte Testzeitraum erlaubt allerdings keine abschließende Zuordnung des Unterschieds zu einem bestimmten Mechanismus; er bildet die Dauer des Wirkungserhalts ab, die unter kontrollierten Laborbedingungen (nach intensiver Anwendung des Systems) ohne Ermüdungseffekte erreichbar ist. Ein Nachweis der Langzeitwirkung über Wochen oder Monate ist damit ausdrücklich nicht erbracht; er bleibt In-vivo-Anschlussuntersuchungen vorbehalten.

Befund 5 — Fallkontrast bei offener und verdeckter Symptomatik

Bei der Versuchsperson mit stark ausgeprägter offener Symptomatik ist die Wirkung unmittelbar an der Sprechflüssigkeit sichtbar. Offenes Stottern ist die weitaus häufiger dokumentierte Form; für sie sind etablierte Flüssigkeitsmaße geeignet. Bei der Versuchsperson mit moderater Symptomatik und ausgeprägter Vermeidung fällt der Zuwachs an Flüssigkeit zwar geringer aus, die Wirkung zeigt sich jedoch im Wegfall der Vermeidungshandlungen: Statt ausweichender und kurzer Antworten treten längere, vollständigere und inhaltlich elaboriertere Äußerungen. Die Aussprache fällt klarer aus, und der Gesamteindruck ist deutlich gelöster und selbstsicherer bei zugleich verringerter Anstrengung.
Der Fallkontrast ist damit selbst ein wesentlicher Befund: Das Verfahren wirkt bei beiden Versuchspersonen, jedoch auf unterschiedlichen Ebenen. Für die Endpunktwahl künftiger Wirksamkeitsprüfungen folgt daraus, dass neben Flüssigkeitsmaßen auch Maße für das Vermeidungsverhalten und die Sprechanstrengung zu berücksichtigen sind.

Die Befunde wurden PPA als Grundlage für die Ausgestaltung des abschließenden Demonstrators V4 übergeben.
Das im Antrag ebenfalls definierte Kriterium zur Nachhaltigkeit — eine Wirkungsminderung von weniger als 20 Prozent bei täglicher Anwendung über zwei Wochen — konnte nicht geprüft werden, da eine mehrwöchige In-vivo-Erprobung im Projektzeitraum (s.o.) nicht durchführbar war.

Für die Anschlussforschung ergeben sich vier vorrangige Fragestellungen:
— die systematische Untersuchung des Eigenstimmanteils und der dynamischen Mikrozeitvariation des Feedbacksignals als Wirkparameter des Choral-Speech-Effekts;
— die Erhärtung des Wirkungserhalts in einer In-vivo-Untersuchung über Wochen und Monate;
— die Erfassung von Vermeidungsverhalten als eigenständige Zielgröße bei Fällen mit verdeckter Symptomatik.

1.6 Studie 4 — Soziale Wahrnehmung der Studie-1-Versuchspersonen durch unbeteiligte Dritte; Ersatzdesign zur V3-In-Vivo-Usability-Studie (AP 3.2, M10)
Im Verlauf der Studienkonzeption wurde das Design der Studie 4 von einer antragsgemäßen In-Vivo-Usability-Studie zu V3 (AP 3.2) auf eine Studie zur sozialen Wahrnehmung der Versuchspersonen der Studie 1 durch unbeteiligte Dritte (Laienhörer*innen) umgewidmet; die Anpassung wurde im Konsortiumstreffen am 17.10.2025 abgestimmt und vom Projektträger abgenommen.

Die Umwidmung war methodisch indiziert: Die antragsgemäße V3-In-Vivo-Usability-Studie hätte eine Frühphasen-Mobile-Version des Demonstrators vorausgesetzt, die im Projektverbund zugunsten einer technisch ausgereifteren Lab-Version umgestellt wurde (Begründung siehe Bericht von PPA). Die Laienhörer-Bewertung an V1-Material erlaubte stattdessen die zeitkritische Erhebung von Befunden zur sozialen Wahrnehmung, die unmittelbar in die V3-Ausgestaltung einflossen.
Auf Grundlage der in Studie 1 erhobenen Videoaufzeichnungen führte PPB Studie 4 zur sozialen Wahrnehmung der Versuchspersonen von Studie 1 mittels unbeteiligter Dritter (Laienhörer) durch. Beurteilt wurden die Versuchspersonenaufnahmen im direkten Bedingungsvergleich zwischen VC-ON (mit Sprechfeedback) und VC-OFF (ohne Sprechfeedback).

Befund 1:
Konsistent mit den in Studie 1 beobachteten Verbesserungen des Sprechflusses wurden die Versuchspersonen unter VC-ON von Laienhörern tendenziell als weniger angespannt oder unsicher beim Sprechen wahrgenommen — ein indirekter Hinweis darauf, dass die VC-Intervention nicht nur den Sprechfluss objektiv verbessert, sondern auch die nach außen sichtbare Sprechanstrengung mindert.

Befund 2:
Die soziale Wahrnehmung wurde unter V1 weniger durch die — auch von Laien erkannte — Stotterreduktion als durch eine unter VC-Sprechfeedback eingetretene leichte Sprechverlangsamung geprägt; in einzelnen Fällen wurden Versuchspersonen als weniger vertrauenswürdig oder intelligent eingestuft. Auf der Grundlage sozialpsychologischer Theorien ist es plausibel, dass ein verlangsamter Sprechfluss als verminderte Kompetenz interpretiert wird. Die Sprechverlangsamung ist vermutlich nicht allein latenzinduziert: Die V1-Latenz von ca. 60 ms lag unterhalb der in der AAF-Literatur etablierten kritischen Schwelle, ab der das Sprechtempo durch zeitverzögerte Feedbackwiedergabe beeinflusst wird. Als Mechanismus liegt nahe, dass die in der V1-Frühphase noch nicht hinreichende Sprachverständlichkeit den kognitiven Aufwand der Versuchspersonen erhöhte und zu einer kompensatorischen Sprechverlangsamung führte. Hinzu kommt, dass sich die Bedingungen im Videomaterial nicht allein durch das Sprechverhalten unterschieden: Das in der ON-Bedingung getragene Laborequipment war für die Beurteilenden sichtbar. Ein Einfluss auf die Personenwahrnehmung ist damit nicht auszuschließen**.**

Methodische Einordnung: Die Befunde sind als V1-spezifisch und laborbezogen zu rahmen; für V3 wurden sie nicht mehr beobachtet: In der Kontrastanalyse der Studie 3 fiel die Subskala soziale Wahrnehmung unter aktivierter Konvertierung durchgehend günstiger aus als in der Vergleichsbedingung (siehe Anhang). Die Befunde sind allerdings für die gesamte Klasse delay-basierter Anti-Stotter-Verfahren strukturell bedeutsam, wie in der Versorgungs-Einordnung weiter ausgeführt wird.
Die erhobenen Hörerurteile zeigen zudem, dass der klinisch bestimmte Stotterschweregrad und die

Wahrnehmung durch Laienhörer nicht zwingend übereinstimmen. Damit wird empirisch sichtbar, dass Stottern nicht allein über klinische Symptommaße beschrieben werden kann, sondern auch eine eigenständige soziale Wahrnehmungsdimension besitzt.
Zudem wurden Personen mit ausgeprägtem Stottern bereits ohne technische Intervention durch Laienhörer tendenziell als weniger intelligent oder weniger vertrauenswürdig eingeschätzt (Abbildungen A8 bis A10). Dieser Befund steht im Einklang mit dem Forschungsstand zur Stigmatisierung stotternder Personen und bestätigt ihn im vorliegenden experimentellen Design. Für PPB ist dies von erheblicher Relevanz, da die Bewertung möglicher Interventionseffekte nicht auf Sprechflüssigkeitsmaße allein beschränkt werden sollte, sondern auch die soziale Wirkung berücksichtigen muss.

Versorgungs-Einordnung. Klassische Anti-Stotter-Verfahren beruhen auf Delayed Auditory Feedback (DAF) und wirken durch eine zeitverzögerte Rückkopplung der Eigenstimme, die eine Verlangsamung des Sprechens nach sich zieht. Genau diese Verlangsamung erzeugt, wie Studie 4 zeigt, soziale Beurteilungsnachteile und schreibt damit das Stotter-Stigma fort. In V3 sind die von PPB als ursächlich identifizierten Bedingungen — Sprachverständlichkeit und Stimmnatürlichkeit — substanziell verbessert (vgl. Abschnitt 1.5); die Kontrastanalyse zeigt für V3 keine entsprechenden Beurteilungsnachteile mehr. Damit überwindet die Demonstrator-Linie das DAF-typische Trade-off zwischen sprechförderlicher und sozialer Wirkung und adressiert ein zentrales Versorgungsziel des Projekts, die Stigmafreiheit des Systems.

1.7 Konsortiale Abstimmung und Schnittstellenarbeit

Die PPB-Studienreihe war im Berichtsjahr durch eine eng getaktete, iterativ angelegte Abstimmung mit PPA gekennzeichnet. Aus experimentalpsychologischer Sicht war diese Verzahnung methodisch erforderlich: Die in Studie 1, Studie 2 und Studie 4 erhobenen Befunde sollten nicht nachgelagert in einen abgeschlossenen Entwicklungsstand einfließen, sondern bereits während der Demonstratorentwicklung versuchsplanerisch abgebildet werden — insbesondere in den Lab-GUI-Anforderungen, in der Parameterzugänglichkeit sowie in der gemeinsamen Operationalisierung der zentralen Wirkhypothesen. Die regelmäßigen Präsenzabstimmungen mit PPA fungierten in diesem Sinn weniger als projektkoordinierende Routine, sondern als methodische Transferschnittstelle zwischen empirischen PPB-Befunden und den daraus abzuleitenden Anforderungen an Demonstratorparameter und Studiendesigns der jeweils nachfolgenden Studienstufe.

Das Konsortialtreffen am 17.10.2025 in Berlin diente für PPB darüber hinaus der konsensualen Bestätigung der Programmänderungen mit dem Projektträger: Die Zeitplananpassungen für V2 und V3, die V4-Scope-Änderung sowie die daraus resultierenden Implikationen für die PPB-Studienlage wurden dort dem VDI/VDE-IT (Dr. Muth) vorgestellt und konsensual bestätigt. Auf dieser Grundlage konnte M9 methodisch vertretbar in die Folgeperiode verschoben (vgl. Abschnitt 1.5) und M10 in einer angepassten Studienform abgeschlossen werden (vgl. Abschnitt 1.6)

1.8 Wissenschaftliche Gesamtwürdigung

Erkenntnisstand nach vier Studien. Das PPB-Teilvorhaben hat die im Antrag beschriebene Wirkannahme über vier aufeinander aufbauende Studien und vier Demonstratorstufen hinweg empirisch erschlossen. Am Anfang stand die Frage, ob ein ego-dystones VC-Feedback überhaupt sprechfördernd wirkt; am Ende steht ein Erkenntnisstand, der die Wirkung nicht nur belegt, sondern ihre Bedingungen benennt: welche Parameter sie tragen, bei welchen Betroffenen sie sich auf welcher Ebene zeigt, wie sie sich gegenüber dem Stand der Technik verhält und mit welchen Maßen sie weiterführend erfasst werden sollte.

Wirksamkeit. Studie 1 hat den Proof-of-Concept auf dem frühen Demonstratorstand V1 erbracht: Sämtliche sechs sprechbezogenen Zielgrößen wiesen in dieselbe Richtung, die posteriore Wahrscheinlichkeit für eine Verbesserung lag bei 98 Prozent. Studie 3 hat diesen Befund auf V3-Basis mit

einem objektiven Frequenzmaß bestätigt und quantifiziert: Der Anteil gestotterter Silben sank bei beiden Versuchspersonen um mehr als 80 Prozent. Bemerkenswert ist dabei die Konvergenz unterschiedlicher Messzugänge: Sowohl menschliche Beurteilungen als auch die automatische Spracherkennung zeigten deutliche Vorteile der ON-Bedingung. Das im Antrag definierte Kriterium für den Proof-of-Concept der sprechfördernden Wirkung ist damit in beiden untersuchten Fällen erfüllt.

Wirkmechanismus. Studie 2 hat die wahrnehmungsbezogene Voraussetzung des Ansatzes bestätigt: Die personalisiert bestimmte Zielstimme wird von der Eigenstimme signifikant unterschieden. Studie 3 hat darüber hinaus einen Hinweis auf den zugrundeliegenden Mechanismus geliefert, der über den Projektansatz hinausweist. Beide Versuchspersonen wählten unabhängig voneinander eine Konfiguration mit Beimischung der eigenen Stimme — eine Konstellation, die strukturell dem Choral Speech entspricht, dem in der Stotterforschung am besten belegten flüssigkeitsinduzierenden Effekt. Dass dieser Effekt bislang die Anwesenheit eines zweiten Sprechers voraussetzt, macht seine technische Nachbildung zu einer eigenständigen Forschungsfrage.

Vergleich mit dem Stand der Technik. Studie 3 hat das Verfahren erstmals direkt gegen Delayed Auditory Feedback gestellt. Unter beiden Bedingungen sank der Anteil gestotterter Silben deutlich; unter der individuell bestimmten Konfiguration fiel er niedriger aus. Im Verlauf zeigte sich zudem ein unterschiedliches Stabilitätsverhalten, das der im Antrag beschriebenen anti-adaptiven Wirkannahme entspricht. Der Beobachtungszeitraum erlaubt keine abschließende Zuordnung des Mechanismus; der Nachweis der Langzeitwirkung bleibt der Anschlussforschung vorbehalten.

Zielgruppe und Wirkebenen. Zwei Befunde grenzen ein, für wen und in welcher Form die Technologie wirkt. Studie 2 zeigte, dass das Ansprechverhalten mit dem Ausgangsniveau der Symptomatik zusammenhängt und bereits bei moderater Ausprägung deutlich hervortritt. Studie 3 ergänzte, dass sich die Wirkung bei offener und bei verdeckter Symptomatik auf unterschiedlichen Ebenen zeigt: dort in der Sprechflüssigkeit, hier im Wegfall des Vermeidungsverhaltens. Beide Befunde haben unmittelbare Konsequenzen für die Endpunktwahl bei künftigen Wirksamkeitsprüfungen.

Soziale Wirkungsdimension. Studie 4 zeigt zunächst, dass die unter VC-ON beobachtete Verbesserung des Sprechflusses auch nach außen hin wahrnehmbar war: Die Versuchspersonen wurden von Laienhörerinnen und -hörern tendenziell als weniger angespannt und weniger unsicher wahrgenommen. Die Studie hat zudem sichtbar gemacht, dass klinisch bestimmter Schweregrad und Wahrnehmung durch Unbeteiligte auseinanderfallen können und dass stotternde Personen bereits ohne Intervention ungünstiger beurteilt werden. Damit ist die soziale Wahrnehmung als eigenständige Zielgröße einer Anti-Stotter-Intervention ausgewiesen.

Dieser Befund gewinnt an Versorgungsrelevanz im Vergleich zu den etablierten delay-basierten Verfahren: Die Wirkung dieser Verfahren ist konstitutiv mit einer Verlangsamung des Sprechens verbunden, und genau diese Verlangsamung erzeugt soziale Beurteilungsnachteile. Es entsteht ein funktionaler Zielkonflikt — die Sprechflüssigkeit verbessert sich, während das veränderte Sprechmuster den Gewinn sozial teilweise wieder aufzehrt. Für die Demonstratorlinie besteht dieser Zielkonflikt nicht: Die unter V1 beobachtete Sprechverlangsamung ging nicht auf die Feedbacklatenz zurück, sondern auf die damals noch unzureichende Sprachverständlichkeit des konvertierten Signals. Mit deren Verbesserung entfiel sie; unter V3 wurde sie nicht mehr beobachtet. Die sprechfördernde Wirkung wird damit ohne systematisch erforderliche Sprechverlangsamung erreicht.

Methodischer Ertrag. Über die Wirkbefunde hinaus hat das Teilvorhaben drei übertragbare Beiträge geliefert: den multimodalen Beobachtungsfragebogen als Instrument zur Erfassung von Sprechperformanz und sozialer Wirkung, den annotierten deutschsprachigen Disfluency-Datensatz, der eine dokumentierte Lücke in der Trainingsbasis für die maschinelle Stottererkennung schließt, sowie das Single-Case-Verfahren der nutzergesteuerten Parameterauswahl, das die im Antrag zentrale Personalisierung erstmals prozedural prüfbar macht.

Reichweite und Grenzen. Die Befunde beruhen auf Laborerhebungen mit kleinen Stichproben, die

Ergebnisse zu V3 auf einem Einzelfalldesign mit zwei gezielt kontrastierend ausgewählten Versuchspersonen. Sie sind hypothesengenerierend und belegen die Erreichbarkeit der im Antrag gesetzten Schwellen, nicht aber deren Häufigkeit in der Zielpopulation. Der Nachweis unter Verwendung größerer Stichproben, unter Alltagsbedingungen und über längere Zeiträume ist Gegenstand der klinischen Validierung. Für diese liegen mit dem Erkenntnisstand des Teilvorhabens nunmehr die Grundlagen vor: eine belegte Wirkung, ein eingegrenzter Parameterraum, eine begründete Zielgruppenhypothese und ein Instrumentarium, das die relevanten Wirkebenen abbildet. Aus PPB-Sicht ist damit eine belastbare wissenschaftliche Grundlage geschaffen, um den Ansatz in nachfolgenden kontrollierten Studien mit größeren Stichproben, längeren Anwendungszeiträumen und einem multidimensionalen Endpunktmodell zu prüfen, das über die im SSI-4 erfassten Ereignishäufigkeiten hinaus auch Sprachverständlichkeit, Sprechanstrengung, Vermeidungsverhalten, kommunikative Handlungsfähigkeit und soziale Wirkung berücksichtigt.

Aus PPB-Sicht ergibt sich für den Berichtszeitraum ein konsistentes empirisches Bild: Die Wirksamkeit des VC-basierten Anti-Stotter-Ansatzes ist gestützt, seine Wirkbedingungen sind präzisiert, und die Anforderungen an die weitere technologische, wissenschaftliche und versorgungsbezogene Entwicklung sind identifiziert.

Über die Studien 1 bis 4 hinweg zeichnet sich dabei eine durchgängige Erkenntnislinie ab: vom ersten Wirksamkeitsnachweis über die Identifikation relevanter Wirk- und Störfaktoren und die daraus abgeleitete technologische Optimierung bis zur Untersuchung der sozialen Konsequenzen veränderter Sprechperformanz.

2. Vergleich des Vorhabenstands mit der ursprünglichen bzw. mit Zustimmung des ZG geänderten Arbeits-, Zeit- und Kostenplanung

2.1 Zusammenfassung der PPB-relevanten Abweichungen

Die PPB-relevanten Abweichungen im Berichtszeitraum ergeben sich vor allem aus den im Konsortium abgestimmten Verschiebungen der PPA-Demonstratoren V2 und V3 sowie aus der Anpassung des V4-Scope. Für PPB betrafen diese Änderungen nicht die wissenschaftliche Zielsetzung des Teilvorhabens, sondern die zeitliche und methodische Umsetzung einzelner Studien.

M4 wurde im April 2025 planmäßig erreicht. Die V1-basierte Proof-of-Concept-Studie lieferte erste belastbare Hinweise auf eine sprechfördernde Wirkung des ego-dystonen VC-Feedbacks sowie auf zentrale Akzeptanzanforderungen, insbesondere auf Sprachverständlichkeit und perzeptuelle Feedbackqualität.

M6 wurde aufgrund der verspäteten V2-Übergabe nicht zum ursprünglichen Zeitpunkt im August 2025, sondern erst nach Übergabe des V2-Demonstrators am 27.10.2025 im Rahmen der V2-Erprobung bearbeitet. Inhaltlich wurde damit die in Abschnitt 1.4 beschriebene laborbasierte Untersuchung des Nutzererlebens, der personalisierten Anti-Voice-Bestimmung und der Continuous-Change-Funktion durchgeführt. Die Verzögerung ist somit auf die spätere Verfügbarkeit des V2-Demonstrators zurückzuführen, nicht auf eine Änderung der PPB-Zielsetzung.

M9 wurde aufgrund der erst am 20.12.2025 erfolgten V3-Übergabe auf Februar 2026 verschoben und dort durchgeführt. Aus PPB-Sicht ist diese Verschiebung methodisch begründet, da die geplante Untersuchung der V3-Parametrisierung und der längerfristigen Anwendbarkeit nur auf einem Demonstratorstand sinnvoll durchführbar ist, der die in Studie 1 und Studie 2 identifizierten Konfundierungen hinreichend reduziert. Die Durchführung auf V3 erhöht daher die Aussagekraft der Studie gegenüber einer Durchführung auf einem früheren Entwicklungsstand.

M10 wurde im Dezember 2025 im Rahmen eines mit dem Projektträger abgestimmten Ersatzdesigns auf V1-Basis erreicht. Anstelle der ursprünglich vorgesehenen V3-In-vivo-Usability-Studie führte PPB Studie 4 zur sozialen Wahrnehmung der Studie-1-Versuchspersonen durch unbeteiligte Dritte durch. Diese Anpassung erlaubte es, trotz der V3-Verschiebung wissenschaftlich relevante Befunde zur

Außenwahrnehmung von Stottern und von VC-unterstütztem Sprechen noch im Berichtszeitraum zu erheben.
Die Zeitplan- und Designanpassungen wurden im Konsortialtreffen am 17.10.2025 mit VDI/VDE-IT abgestimmt und anschließend formal bestätigt. Aus PPB-Sicht handelt es sich damit *nicht* um eine Änderung der Zielsetzung, sondern um eine methodisch begründete Anpassung des Arbeitsplans an den jeweils verfügbaren Demonstratorstand.

2.2 PPB-relevante Meilensteine

M	Bezeichnung	Verantwortung	Plan	Beschreibung	Status
M3	Erprobung Demonstrator V1	PPA / PPB-Empfang	Feb. 2025	V1 wurde im Februar 2025 an PPB übergeben, im BAMLiD in Betrieb genommen und in die laborpsychologische Versuchsumgebung integriert.	Planmäßig
M4	PoC-Effizienz von V1	PPB, Studie 1	Apr. 2025	Erreicht. Die V1-basierte PoC-Studie wurde abgeschlossen und lieferte erste belastbare Hinweise auf die sprechfördernde Wirkung des ego-dystonen VC-Feedbacks sowie auf zentrale Akzeptanzanforderungen.	Planmäßig
M5	Erprobung Demonstrator V2	PPA / PPB-Empfang	Juni 2025	V2 wurde am 27.10.2025 an PPB übergeben; Verzögerung um 4 Monate.	Zeitplananpassung abgestimmt
M6	Usability von V2 unter Laborbedingungen	PPB, Studie 2	Aug. 2025	Nach V2-Übergabe im Rahmen der V2-Erprobung bearbeitet. Untersucht wurden unter anderem personalisierte Anti-Voice-Bestimmung, Continuous Change, perzeptuelle Diskrimination und nutzerbezogene Akzeptanzbedingungen.	Inhaltlich im Berichtszeitraum bearbeitet; terminlich verzögert
M7	Erprobung Demonstrator V3	PPA / PPB-Empfang	Okt. 2025	V3 wurde am 20.12.2025 an PPB übergeben; Verzögerung um zwei Monate. Interne PPB-Abnahme- und Erprobungssessions an nicht betroffenen Personen wurden vorbereitet beziehungsweise begonnen.	Zeitplan- und Scope-Anpassung abgestimmt
M9	V3-Parametrisierung und längerfristige Anwendbarkeit	PPB, Studie 3	Dez. 2025	Erreicht im Februar 2026. Die Verschiebung folgte aus der V3-Übergabe am 20.12.2025; V3 diente als methodisch geeignete Studiengrundlage.	Erreicht; Verschiebung abgestimmt
M10	Usability von V3 in vivo, angepasstes Ersatzdesign	PPB, Studie 4	Dez. 2025	Erreicht im Dezember 2025 als abgestimmtes Ersatzdesign auf V1-Basis. Durchgeführt wurde	Erreicht; Ersatzdesign abgestimmt

M	Bezeichnung	Verantwortung	Plan	Beschreibung	Status
				eine Studie zur sozialen Wahrnehmung der Studie-1-Versuchspersonen durch Laienhörerinnen und Laienhörer.	

Meilensteine (M). M3, M5 und M7 sind Meilensteine der Gesamtvorhabensplanung und bezeichnen die Übergabe der Demonstratorstufen an PPB. Die Meilensteine des Teilvorhabens PPB sind M4, M6, M9 und M10.

2.3 Plan-Ist-Begründung der Studienarbeiten

Studie 1 / M4

Plan: Proof-of-Concept-Effizienz von V1 im April 2025.
Ist: Durchführung und Meilensteinerreichung im April 2025.
Die Studie lieferte die empirische Grundlage für die Bewertung des unmittelbaren sprechfördernden Nutzens des ego-dystonen VC-Feedbacks unter V1-Bedingungen. Zugleich wurden zentrale Akzeptanzanforderungen identifiziert, insbesondere Sprachverständlichkeit, Stimmnatürlichkeit und perzeptuelle Feedback-Qualität. Diese Befunde wurden in die Anforderungen an V2 und V3 überführt.
Keine Planabweichung.

Studie 2 / M6

Plan: Usability von V2 unter Laborbedingungen im August 2025.
Ist: Bearbeitung nach verspäteter V2-Übergabe am 27.10.2025.
Die Verzögerung ergibt sich unmittelbar aus der verzögerten Bereitstellung des V2-Demonstrators. Inhaltlich wurde die Studie entsprechend der in Abschnitt 1.4 dargestellten V2-spezifischen Fragestellungen durchgeführt: personalisierte Anti-Voice-Bestimmung, Continuous-Change-Funktion, perzeptuelle Diskrimination, subjektive Akzeptabilität sowie explorative Auswertung der Responsivität.
Die Abweichung betrifft daher den Zeitpunkt, nicht die Zielsetzung der Studie. Methodisch war die Durchführung auf V2-Basis erforderlich, da die zentralen Funktionen von V2 — insbesondere Anti-Voice und Continuous Change — unter V1 noch nicht zugänglich waren.

Studie 3 / M9

Plan: Untersuchung der V3-Parametrisierung und der längerfristigen Anwendbarkeit bis Dezember 2025.
Ist: Verschiebung auf Februar 2026 aufgrund der V3-Übergabe am 20.12.2025; Durchführung im Februar 2026.
Die Verschiebung ist aus PPB-Sicht fachlich begründet. Die in Studie 1 und Studie 2 identifizierten Konfundierungen — insbesondere eingeschränkte Sprachverständlichkeit, kompensatorische Sprechverlangsamung und akzeptanzbezogene Sekundäreffekte — hätten eine valide Untersuchung der Parametrisierung auf einem früheren Demonstratorstand erschwert. Die Durchführung auf dem V3-Demonstrator ist daher nicht nur eine terminliche Anpassung, sondern auch eine Maßnahme zur Sicherung der internen Validität der Untersuchung.
Die Verschiebung wurde mit VDI/VDE-IT abgestimmt; die Studie wurde im Februar 2026 durchgeführt (vgl. Abschnitt 1.5).

Studie 4 / M10

Plan: In-vivo-Usability von V3 bis Dezember 2025.
Ist: Abschluss im Dezember 2025 als abgestimmtes Ersatzdesign auf V1-Basis.
Eine In-vivo-Usability-Studie hätte eine über Wochen alltagstauglich einsetzbare mobile Ausführung des Demonstrators vorausgesetzt, die nicht Gegenstand des Vorhabens ist; hinzu kam die Verzögerung der V3-Übergabe. Studie 4 wurde daher auf die soziale Wahrnehmung der Versuchspersonen aus Studie 1 durch unbeteiligte Dritte umgestellt. Dieses Ersatzdesign erlaubte es, eine für das PPB-Teilvorhaben zentrale Fragestellung noch im Berichtszeitraum zu bearbeiten: wie stotterndes und VC-unterstütztes Sprechen von Laienhörerinnen und Laienhörern sozial bewertet wird.

Die Anpassung ist konsistent mit Abschnitt 1.6 als methodisch begründete Umwidmung von M10 zu verstehen. Sie ersetzt eine spätere V3-basierte Nutzungsprüfung nicht vollständig, liefert jedoch eigenständige, wissenschaftlich relevante Befunde zur sozialen Wahrnehmungsdimension von Stottern und zum VC-Feedback.

Die Verzögerung ergibt sich unmittelbar aus der verzögerten Bereitstellung des V2-Demonstrators. Inhaltlich wurde die Studie entsprechend der in Abschnitt 1.4 dargestellten V2-spezifischen Fragestellungen durchgeführt: personalisierte Anti-Voice-Bestimmung, Continuous-Change-Funktion, perzeptuelle Diskrimination, subjektive Akzeptabilität sowie explorative Auswertung der Responsivität. Die Abweichung betrifft daher den Zeitpunkt, nicht die Zielsetzung der Studie. Methodisch war die Durchführung auf V2-Basis erforderlich, da die zentralen Funktionen von V2 — insbesondere Anti-Voice und Continuous Change — unter V1 noch nicht zugänglich waren.

2.4 Personelle, strukturelle und kostenbezogene Aspekte

Aus PPB-fachlicher Sicht ergeben sich aus den im Berichtszeitraum eingetretenen Demonstratorverschiebungen keine Änderungen der wissenschaftlichen Zielsetzung oder des bewilligten Arbeitsumfangs. Die PPB-seitigen Anpassungen betreffen die zeitliche und methodische Durchführung einzelner Studien und wurden durch Umwidmung, Ersatzdesign und zusätzliche Auswertungsschritte kompensiert.

Die Studienarbeiten wurden im Berichtszeitraum durch die in Abschnitt 1 genannten Projektmitarbeiter durchgeführt und ausgewertet. Die zusätzlichen Arbeiten im Zusammenhang mit Studie 4 stellen eine über den ursprünglichen Studienplan hinausgehende methodische Vertiefung dar, die jedoch innerhalb der PPB-seitigen wissenschaftlichen Zielsetzung liegt.
Für die verbindliche kostenbezogene Bewertung ist der zahlenmäßige Nachweis der Otto-Friedrich-Universität Bamberg maßgeblich. Nach fachlicher Bewertung ergibt sich aus den technischen Demonstratorverschiebungen *keine* Änderung der PPB-Zielerreichbarkeit.

2.5 Gesamtbewertung aus PPB-Sicht

PPB hat das Teilvorhaben im Rahmen der mit VDI/VDE-IT abgestimmten Planänderungen erfüllt.
M4 wurde planmäßig erreicht. M6 wurde aufgrund der verspäteten V2-Übergabe terminlich verzögert, im Berichtszeitraum jedoch inhaltlich im Rahmen der V2-Erprobung bearbeitet. M10 wurde im abgestimmten Ersatzdesign abgeschlossen. M9 wurde aufgrund der späten V3-Übergabe methodisch begründet verschoben und im Februar 2026 erreicht.

Die Verschiebung von M9 ist aus PPB-Sicht nicht als Defizit der Studienlogik zu bewerten, sondern als Maßnahme zur Sicherung der Aussagekraft. Parametrisierung und längerfristige Anwendbarkeit des VC-basierten Feedbacks ließen sich auf Grundlage des technisch verbesserten V3-Demonstrators valider untersuchen als auf früheren Demonstratorständen, bei denen Sprachverständlichkeit, Stimmnatürlichkeit und akzeptanzbezogene Sekundäreffekte noch stärker als Störvariablen bzw. Konfundierungen wirksam waren.

Mit Studie 3 hat PPB zudem eine methodische Erweiterung vorgenommen: Für die Bewertung der Sprechqualität und des Gesamteindrucks der sprechenden Person wurde ein eigener multimodaler Beobachtungsfragebogen entwickelt und erfolgreich eingesetzt. Das Instrument erfasst Veränderungen, die über die reine Stotterhäufigkeit hinausgehen, und ist für weitere Demonstratorstufen sowie für verwandte Forschungskontexte anschlussfähig.
Insgesamt haben die durchgehend abgestimmten Anpassungen die Zielerreichung des PPB-Teilvorhabens ermöglicht. Sie betreffen die zeitliche und methodische Umsetzung, nicht die wissenschaftliche Zielsetzung. Aus PPB-Sicht ist der Vorhabenstand damit konsistent mit dem abgestimmten geänderten Arbeitsplan.

3. Haben sich die Aussichten für die Erreichung der Ziele des Vorhabens innerhalb des angegebenen Ausgabenzeitraums gegenüber dem ursprünglichen Antrag geändert?

Die Aussichten für die Erreichung der Ziele des PPB-Teilvorhabens hatten sich gegenüber dem ursprünglichen Antrag nicht grundlegend verändert. Aus experimentalpsychologischer Sicht bestanden weiterhin beste Voraussetzungen, um die zentralen Fragestellungen des Teilvorhabens empirisch zu prüfen.

Die Summe der erreichten Arbeitsergebnisse bildet eine ideale Grundlage, um den angenommenen Wirkmechanismus, die subjektive und objektive Anwendbarkeit sowie relevante Parameter des veränderten auditiven Feedbacks systematisch zu beschreiben. Methodisch ist dabei bedeutsam, dass gegenüber früheren Demonstratorständen mehrere potenzielle Konfundierungen reduziert wurden, insbesondere hinsichtlich Sprachverständlichkeit, Stimmnatürlichkeit, prosodischer Kontrollierbarkeit und mikrofonabhängiger Einflüsse. Dies erhöhte zudem die interne Validität der durchgeführten Untersuchungen.

Die Verschiebung von M9 wirkte sich aus PPB-Sicht nicht wesentlich auf die Zielerreichbarkeit aus, da sie mit der Verzögerung des V3-Demonstrators zusammenhing und dieser zum Ende des Berichtszeitraums als Studiengrundlage verfügbar war. Die weiteren Arbeiten konnten somit in der Folgeperiode auf Basis eines abgestimmten Demonstratorstands durchgeführt werden.

4. Sind inzwischen von dritter Seite Ergebnisse bekannt geworden, die für die Durchführung des Vorhabens relevant sind?

Ja. Im Berichtszeitraum sind von dritter Seite mehrere Ergebnisse bekannt geworden, die für die Durchführung des Vorhabens relevant sind. Diese stellen die Zielsetzung des Vorhabens nicht in Frage, bestätigen jedoch die Bedeutung der im Projekt vorgesehenen Personalisierung, Latenzoptimierung und kontrollierten Evaluation.

Aktuelle Übersichtsarbeiten und empirische Studien zu Altered Auditory Feedback bestätigen weiterhin kurzfristige positive Effekte von AAF auf die Sprechflüssigkeit bei stotternden Personen, zeigen zugleich aber eine ausgeprägte interindividuelle Variabilität. Eine aktuelle systematische Übersichtsarbeit mit Metaanalyse berichtet über positive Kurzzeiteffekte von AAF, weist jedoch darauf hin, dass die Effektgröße zwischen Personen stark variiert; die Kombination aus Delayed Auditory Feedback und Frequency-Altered Feedback scheint besonders wirksam zu sein, während jüngere Personen weniger zuverlässig ansprechen als Erwachsene. Für das Vorhaben ergibt sich daraus die Relevanz individueller Parametrierung und nutzeradaptiver Feedbackstrategien; die PPB-Studien wurden entsprechend ausschließlich mit erwachsenen Versuchspersonen durchgeführt.
Darüber hinaus stützen neuere Arbeiten zur sensomotorischen Integration die Annahme, dass auditorisches Feedback bei stotternden Personen anders verarbeitet wird als bei nicht stotternden Kontrollpersonen. Eine 2025 veröffentlichte Studie zeigte, dass Delayed Auditory Feedback bei stotternden Erwachsenen die artikulatorische Variabilität verringerte, während es diese bei nicht stotternden Erwachsenen erhöhte. Dies ist für das Vorhaben relevant, weil AAF damit nicht nur als unspezifische Störung oder Maskierung, sondern als gezielte Modulation der auditorisch-motorischen Rückkopplung verstanden werden kann.

Für den Voice-Conversion-Teil des Vorhabens ist insbesondere die Arbeit zu sog. Digital Speech Makeup relevant. Diese zeigt, dass Voice-Conversion-basiertes auditorisches Feedback technisch und psychologisch als veränderte Selbstwahrnehmung eingesetzt werden kann. Gleichzeitig wird dort hervorgehoben, dass Latenz ein kritischer Faktor ist; das Proof-of-Concept-System wies dort etwa 200 ms Latenz auf und war damit für eine Anwendung im Sprechvollzug nicht geeignet. Streaming-fähige, praktisch fast verzögerungsfreie,Voice-Conversion-Verfahren liegen bislang ausschließlich als Forschungsarbeiten vor; ein auf die ego-dystone Wirkung ausgerichtetes Echtzeitsystem ist außerhalb

dieses Vorhabens nicht verfügbar. Die von PPA bereitgestellten Demonstratoren erreichen eine Latenz von rund 60 ms und erfüllen damit die Anforderung an eine Rückkopplung im Sprechvollzug.
Zudem sind im Bereich der automatischen Stottererkennung und der stotterbezogenen Sprachdatensätze neue technische Ergebnisse erzielt worden, die für die objektive Evaluation und mögliche spätere Modellanpassungen relevant sind. Beispielsweise wurde mit AS-70 ein Datensatz für stotternde Sprache in Mandarin auf dem für dieses Thema wichtige Leitkongress Interspeech im Jahre 2024 vorgestellt ; die Autorinnen und Autoren berichten, dass Fine-Tuning aktueller Automatic Speech Recognition (ASR)-Modelle wie Whisper und HuBERT die Verarbeitung stotternder Sprache verbessert. Dass ein solcher Datensatz eigens erstellt werden musste, unterstreicht den Bedarf, dem auch die im PPB-Teilvorhaben erstellte deutschsprachige Annotation begegnet (vgl. Abschnitt 1.3). Für die Auswertung ist außerdem relevant, dass eventbasierte Annotationen gegenüber starren Zeitintervallen bessere Modelle für die automatische Stottererkennung ermöglichen; zugleich weisen diese Arbeiten darauf hin, dass solche Verfahren für Echtzeitanwendungen wegen Segmentierungs- und Latenzanforderungen sorgfältig angepasst werden müssen.
Zusammenfassend erfordern die bekannt gewordenen Dritt-Ergebnisse keine grundlegende Änderung der Projektplanung. Sie bestätigen vielmehr zentrale technische und wissenschaftliche Annahmen des Vorhabens: die Relevanz individueller Parametrierung, die das Vorhaben mit der personalisierten Anti-Voice-Bestimmung adressiert; die Bedeutung des Wirkungserhalts über die Anwendungsdauer, auf die der Ansatz mit der kontinuierlichen Veränderung der Zielstimme ausgerichtet ist; sowie die Anforderungen an Latenz, Natürlichkeit und objektive Evaluation. Dies steht im Einklang mit den Erkenntnissen, die PPB im Projektzeitraum aus eigenen empirischen Befunden gewonnen hat.

5. Sind oder werden Änderungen in der Zielsetzung notwendig?

Nein. Für die von PPB verantworteten Arbeiten sind im Berichtszeitraum keine Änderungen der Zielsetzung notwendig geworden; nach derzeitigem Stand sind solche Änderungen auch nicht vorgesehen. Die wissenschaftliche Zielsetzung des PPB-Anteils bleibt unverändert: Es soll empirisch untersucht werden, unter welchen Bedingungen verändertes auditives Feedback und Voice Conversion sprechflüssigkeitsbezogene, wahrnehmungsbezogene und anwendungsbezogene Effekte bei stotternden Personen zeigen.

Aus PPB-Sicht steht weiterhin die experimentell-psychologische Prüfung der zugrunde liegenden Wirkannahmen im Vordergrund. Dazu gehören insbesondere die kontrollierte Untersuchung relevanter Einflussgrößen, die Erfassung objektiver und subjektiver Zielgrößen sowie die methodisch belastbare Interpretation möglicher Effekte des veränderten auditiven Feedbacks. Ebenso bleibt die nutzer- und anwendungsbezogene Perspektive zentral, insbesondere mit Blick auf Durchführbarkeit, Akzeptanz, Belastung und Anschlussfähigkeit der späteren Untersuchungen an alltagsnähere Anwendungskontexte. Diese Zielsetzungen bestehen unverändert fort.
Angepasst wurde nicht die Zielsetzung, sondern ausschließlich die zeitliche und methodische Umsetzung einzelner Studien in Abhängigkeit vom jeweils verfügbaren Demonstratorstand. So wurde M6 mit Studie 2 zusammengeführt, M9 auf Februar 2026 verschoben und M10 als Ersatzdesign auf V1-Basis abgeschlossen. Diese Anpassungen dienten dazu, die Studien so anzulegen, dass die erhobenen Daten für die Fragestellungen des PPB-Anteils interpretierbar bleiben. Sie stellen daher keine inhaltliche Neuausrichtung dar, sondern eine methodische Anpassung an die technischen Rahmenbedingungen des Berichtszeitraums.

Auch die konsortial abgestimmte Umstellung des Demonstrators V4 von einer mobilen auf eine laborbasierte Ausführung erfordert aus PPB-Sicht keine Änderung der Zielsetzung. Für die PPB-Fragestellungen ist entscheidend, dass die experimentellen Bedingungen hinreichend kontrollierbar sind und relevante Parameter systematisch variiert werden, um Rückschlüsse auf den Wirkmechanismus, die Nutzbarkeit und spätere alltagsnähere Anwendungen ziehen zu können. Diese Voraussetzungen blieben durch die angepasste Demonstratorplanung erhalten. Berührt ist allein die Untersuchung unter

Alltagsbedingungen, die entsprechend in die Anschlussforschung verlagert wurde.
Zusammenfassend ist festzuhalten, dass keine Änderung der Zielsetzung des PPB-Anteils notwendig war. Die vorgenommenen Anpassungen betreffen ausschließlich die Umsetzung der Studienplanung, nicht aber die wissenschaftlichen Ziele von PPB.

6. Fortschreibung des Verwertungsplans

6a. Erfindungen, Schutzrechtsanmeldungen und erteilte Schutzrechte
Im PPB-Teilvorhaben sind im Berichtszeitraum keine eigenständigen Erfindungen, Schutzrechtsanmeldungen oder erteilten Schutzrechte entstanden. Die im PPA-Sachbericht dargestellte Patentfamilie zur ego-dystonen Stimmkonvertierung liegt auf PPA-Seite. Die Regelungen zu Schutzrechten und Verwertung ergeben sich aus der Kooperationsvereinbarung vom 14.02.2025.

6b. Wirtschaftliche Erfolgsaussichten nach Projektende
Die wirtschaftliche Verwertung der F&E-Ergebnisse liegt gemäß der Kooperationsvereinbarung beim gewerblichen Verwertungsträger PPA. Der PPB-Beitrag ist für diese Verwertung dennoch zentral, weil belastbare empirische Evidenz zu Wirkung, Akzeptanz, Usability und Parametrisierung eine Voraussetzung für regulatorische Anschlussfähigkeit, klinische Validierung und spätere Versorgungspfade ist.
Aus PPB-Sicht erhöhen die im Berichtszeitraum gewonnenen Studienergebnisse die wirtschaftlichen Erfolgsaussichten mittelbar: Die früh identifizierte Verständlichkeits- und Akzeptanzanforderung führte zur Entwicklung einer besseren technischen Demonstratorqualität; die V3-basierte Untersuchung hat die Wirksamkeits- und Nutzungsaussagen auf ein belastbareres Niveau gehoben; die V4-Parametrisierung kann auf empirischen Befunden statt auf rein technischen Annahmen beruhen.

Besonders relevant für die spätere Umsetzung sind die von PPA dargestellten Versorgungspfade für eine app-basierte, DiGA-fähige Lösung sowie eine mittelfristige Embedded-TWS-Variante. Die im Vorhaben aufgebaute methodische und apparative Infrastruktur sowie die entwickelten Erhebungsinstrumente stehen für die anschließende klinische Validierung zur Verfügung.

6c. Wissenschaftliche und technische Erfolgsaussichten nach Projektende
Die wissenschaftlichen Erfolgsaussichten sind aus PPB-Sicht sehr günstig. Die im Berichtszeitraum abgeschlossenen Studien schaffen eine Grundlage für wissenschaftliche Publikationen zur ego-dystonen Stimmkonvertierung, zur Rolle von Sprachverständlichkeit und Stimmnatürlichkeit für die Akzeptanz, zur technischen Nachbildung des Choral-Speech-Effekts sowie zur Vermeidung von Habituationsprozessen durch kontinuierliche Veränderung der Zielstimme.

6d. Wissenschaftliche und wirtschaftliche Anschlussfähigkeit für eine mögliche nächste Phase
Für die Anschlussforschung bieten sich drei Linien an: die Aufklärung des Wirkmechanismus hinter der beobachteten Nachbildung des Choral-Speech-Effekts, insbesondere des Beitrags des Eigenstimmanteils und der dynamischen Mikrozeitvariation; die Prüfung des Wirkungserhalts sowie der Habituationsresistenz über längere Anwendungszeiträume in einer In-vivo-Untersuchung; sowie die Entwicklung des multidimensionalen Endpunktmodells für die Versorgungsforschung, das neben Flüssigkeitsmaßen auch Vermeidungsverhalten und Sprechanstrengung erfasst.
PPB steht für diese Anschlusslinien als zentraler wissenschaftlicher Partner zur Verfügung, da die Universität Bamberg die laborbasierte experimentalpsychologische, wahrnehmungspsychologische und nutzerbezogene Validierung des Systems tragen kann. Die im PPA-Bericht skizzierte mögliche Anschlussstruktur mit PsyAIHance, PPB, der University of Edinburgh/CSTR und weiteren neuropsychologisch ausgerichteten Forschungspartnern ist aus PPB-Sicht fachlich plausibel und zielführend.

Darüber hinaus hat der Lehrstuhl durch die Zusammenarbeit mit Betroffenenverbänden und Behandelnden ein Netzwerk aufgebaut, das die Rekrutierung von Versuchspersonen erheblich erleichtert. Während im Vorlauf von Studie 1 noch aufwendige Rekrutierungsarbeit erforderlich war, konnte für die späteren Studien auf diese Kontakte zurückgegriffen werden. Für künftige Forschungsvorhaben steht damit ein belastbarer Zugang zur Zielgruppe zur Verfügung — bei einer Prävalenz von rund einem Prozent ist dies eine wesentliche Voraussetzung für eine effiziente Studiendurchführung.

Abbildungsanhang

Zuwendungsempfänger / ausführende Stelle	Otto-Friedrich-Universität Bamberg (PPB)
Förderkennzeichen	**16SV9298**
Vorhabenbezeichnung	Anti-Stotter
Teilvorhaben PPB	Empirische Wirk-, Usability- und Anwendungsvalidierung des Anti-Stotter-Systems
Laufzeit des Vorhabens	01.11.2024 bis 28.02.2026
Berichtsstand	31.08.2026