

**6.3. Annotation** Für jede der 30 Variablen wurde entsprechend der definierten Subvariablen ein „regulärer Ausdruck“, d.h. ein Algorithmus entwickelt, der eine automatisierte Zusammenstellung aller für die Untersuchung relevanten Belege mittels des Analysetools Exakt ermöglichte. Diese mit Exakt erstellten Konkordanzlisten bildeten die Grundlage für die Annotation, die in zwei Arbeitsschritten erfolgte. Im ersten Schritt wurde für jeden Beleg die Sigle eingetragen, der er nach der Variablendefinition zuzuordnen war (Vorannotation). Während dieser Schritt für die Vorlesetexte lediglich ein einziges Mal durchgeführt werden musste, da hier alle Gewährspersonen dieselben Tokens realisiert haben, erforderte jedes Tischgespräch und jedes Interview eine gesonderte Vorannotation. Im zweiten Schritt erfolgte die eigentliche Annotation durch Abhören der entsprechenden Gesprächspassage in der mit Exakt verknüpften Audiodatei und Notation der aktuell realisierten Lautvariante.

Die Zuordnung der Belege zu den Varianten erfolgte aus arbeitsökonomischen Gründen rein ohrenphonetisch (phonetische Messungen wurden nur bei der Variable V5 durchgeführt). Bei den Variablen, in denen sich zwei auditiv gut differenzierbare Varianten gegenüberstehen, wie im Falle von *das/dat* (Frikativ vs. Plosiv), *Tag/Tach* (Plosiv vs. Frikativ), *nicht/nich* (Plosiv vs. Nullrealisierung) oder *Baum/Boom* (Diphthong vs. Monophthong), war eine Zuordnung meist problemlos möglich. Eine größere Herausforderung stellte die Variantenzuordnung bei Variablen mit kontinuierlichen Übergängen dar. Zur Vermeidung von unsicheren Annotationen wurde hierbei grundsätzlich dem Prinzip gefolgt, nur eindeutig wahrnehmbare Abweichungen vom Standard zu annotieren und Zweifelsfälle der Standardrealisierung zuzuordnen. Zur Qualitätssicherung wurde innerhalb der Arbeitsteams über Probeanalysen und gemeinsames „Einhören“ eine Verständigung über Variantengrenzen herbeigeführt.

Zur intersubjektiven Kontrolle wurden doppelte Annotationen oder nachträgliche Korrekturdurchläufe durchgeführt. Trotz dieser Maßnahmen ließ sich das aus der dialektologischen Empirie hinreichend bekannte Problem der divergierenden Wahrnehmung unterschiedlicher Bearbeiter nicht vollständig beseitigen. Die daraus resultierenden Schwankungen bewegen sich allerdings in einem tolerierbaren Bereich von einigen Prozentpunkten und beeinträchtigen somit nicht die Zuverlässigkeit der Ergebnisse zur arealen und situativen Variation. Auf spezifische Schwierigkeiten im Zusammenhang mit der Annotation einzelner Variablen wird ggf. in den Kommentaren eingegangen. Die Annotation erfolgte auf diese Weise für alle Stichproben aus den Erhebungssituationen Tischgespräch, Interview und Vorleseausprache. Im Rahmen von gemeinsamen Projektsitzungen wurden die vorgenommenen Annotationen einer kritischen Revision unterzogen und die Daten von möglichen Annotationsfehlern und ungeeigneten Belegen bereinigt. So wurden Belege in Kontexten, in denen die betreffende Variante auditiv schlecht wahrnehmbar ist, nicht in die Analyse mit einbezogen. Des Weiteren führen die automatisierten Suchroutinen in Einzelfällen zu Belegen, die auf den ersten Blick den Suchkriterien entsprechen, aber aus bestimmten Gründen dennoch nicht ausgewertet werden können, etwa Vokale in nebetonigen Silben (z.B. ist das nebetonige *o* in *probieren* im Unterschied zu dem haupttonigen *o* in *Probe* z.B. nicht diphthongierungsanfällig) oder Belege, bei denen kein Kontrast zum Standard vorliegt (z.B. das Suffix *-ig*, bei dem die Frikativaussprache des *g* keine Abweichung vom Standard darstellt). Auf diese Weise wurden die Daten durch die manuelle Beseitigung von Annotationsfehlern und Herausnahme ungeeigneter Belege bereinigt.