

# The Voice of Jena: BVMT vs. JVMT

## Die Erfassung der Diskriminationsleistung bei der Wahrnehmung von Stimmen

Laura Ekkawi, Chiara Spielmann, Elisa Lehmann, Tim Jesgarzewsky, Oliver Kresin  
Leitung: Helene Kreysa



FRIEDRICH-SCHILLER-  
UNIVERSITÄT  
JENA

### Einleitung

Tagtäglich kommen wir mit den vielfältigsten Stimmen unserer Mitmenschen in Kontakt. Doch wie einfach fällt es uns tatsächlich, sie zu unterscheiden? Auch Mühl et al. (2017) beschäftigten sich mit dem Themenfeld zur Sprachwahrnehmung und entwickelten den Bangor Voice Matching Test (BVMT), dieser zielt in diesem Zusammenhang darauf ab, die Stimmwahrnehmungsfähigkeit in einem standardisierten Testverfahren zu ermitteln. Neben der Replizierung des BVMT entwickelten wir den Jena Voice Matching Test (JVMT) und wendeten ihn auf unsere deutsche Stichprobe an. Mit dem JVMT verfolgten wir die Fragen, ob es möglich ist, einen ähnlich validen Test zur Stimmenunterscheidung mit ganzen Sätzen zu schaffen und ob ein subjektives Kriterium, wie die geschätzte Distinktheit von Stimmen, eine gute Itemauswahl ermöglicht.

### Hypothesen

- I. BVMT und JVMT messen eine vergleichbare Fähigkeit
- II. Diese Fähigkeit ist weitgehend unabhängig von der sprachlichen Herkunft der Probanden
- III. Da der JVMT längere Stimmbeispiele enthält, sollten die Probanden durchschnittlich bessere Ergebnisse erzielen als im BVMT
- IV. Die Stimmkombinationen mit weniger distinkt beurteilten Stimmen sollten im Mittel weniger oft korrekt beurteilt werden als distinkt beurteilte Stimmen
- V. Frauenstimmen sind schwerer zu unterscheiden als Männerstimmen

### Methode

#### Probanden:

Die Daten von  $N = 83$  deutschsprachige ProbandInnen konnten in die Auswertung einfließen (71% weiblich, mittleres Alter = 21.55 Jahre,  $SD = 3.43$ ).

#### Material und Design:

##### Bangor Voice Matching Test (BVMT)

- Präsentation von zwei kurzen Silben, die entweder von dem selben oder von verschiedenen Sprechern stammen
- Nach der Präsentation soll eingeschätzt werden, ob die Sprecher identisch oder verschieden sind
- Dabei ist das Verhältnis an zu unterscheidenden Stimmenpaaren (insgesamt 80) über Geschlechter und Aufgaben (gleiche Stimme / verschiedene Stimmen) ausgeglichen (Pro Geschlecht 40 Items, davon jeweils 20 gleich / verschieden)

##### Jena Voice Matching Test (JVMT)

- Äquivalentes Prinzip, jedoch bestand das Stimulusmaterial aus kompletten Sätzen von 61 Sprechern aus dem Jenaer Speaker Set (Zäske et al., eingereicht)
- Zur Bildung geeigneter Stimmpaare ließen wir im Wintersemester von Probanden bewerten, welche Stimmen besonders gut aus einer Menschenmenge herauszuhören wären (In-The-Crowd-Distinktheit) und welche Stimmen leicht unterscheidbar erscheinen (Deviation-Based-Distinktheit)
- Es folgte eine Mittelung der beiden Distinktheitsbewertungen für jede Stimme und die Erstellung einer Rangfolge der Stimmen bezüglich dieser subjektiv eingeschätzten Distinktheit
- Verwendung der 15 am wenigsten distinkt bewerteten Stimmen als Stimulusmaterial
- Für Frauen stammten diese aus dem 1. Quartil, für Männer aus dem 2. und 3. Quartil der gemeinsamen Rangfolgeliste
- Auch in diesem Test war das Verhältnis von Männer- und Frauenstimmen ausgeglichen, genauso wie es dieselbe Anzahl von insgesamt 80 Trials mit denselben (40) und verschiedenen (40) Sprechern gab
- Reihenfolge der Blöcke und Trails wurde randomisiert
- Ein am Ende des Experiments präsentierter Fragebogen erfasste mögliche Strategien der Probanden, sowie ob weibl. / männl. Stimmen oder lange / kurze Äußerungen leichter zu unterscheiden waren

#### Prozedur:

- Beide verwendeten Tests starteten mit einem zentralen Fixationskreuz (1000 ms), daraufhin folgten nacheinander die beiden Stimmproben. Auf dem Antwortbildschirm mussten die Probanden auf ein Antwort-Feld (selber Sprecher / verschiedene Sprecher) klicken, wonach der nächste Trial nach 500 ms begann. Abschließend füllten sie den Fragebogen aus

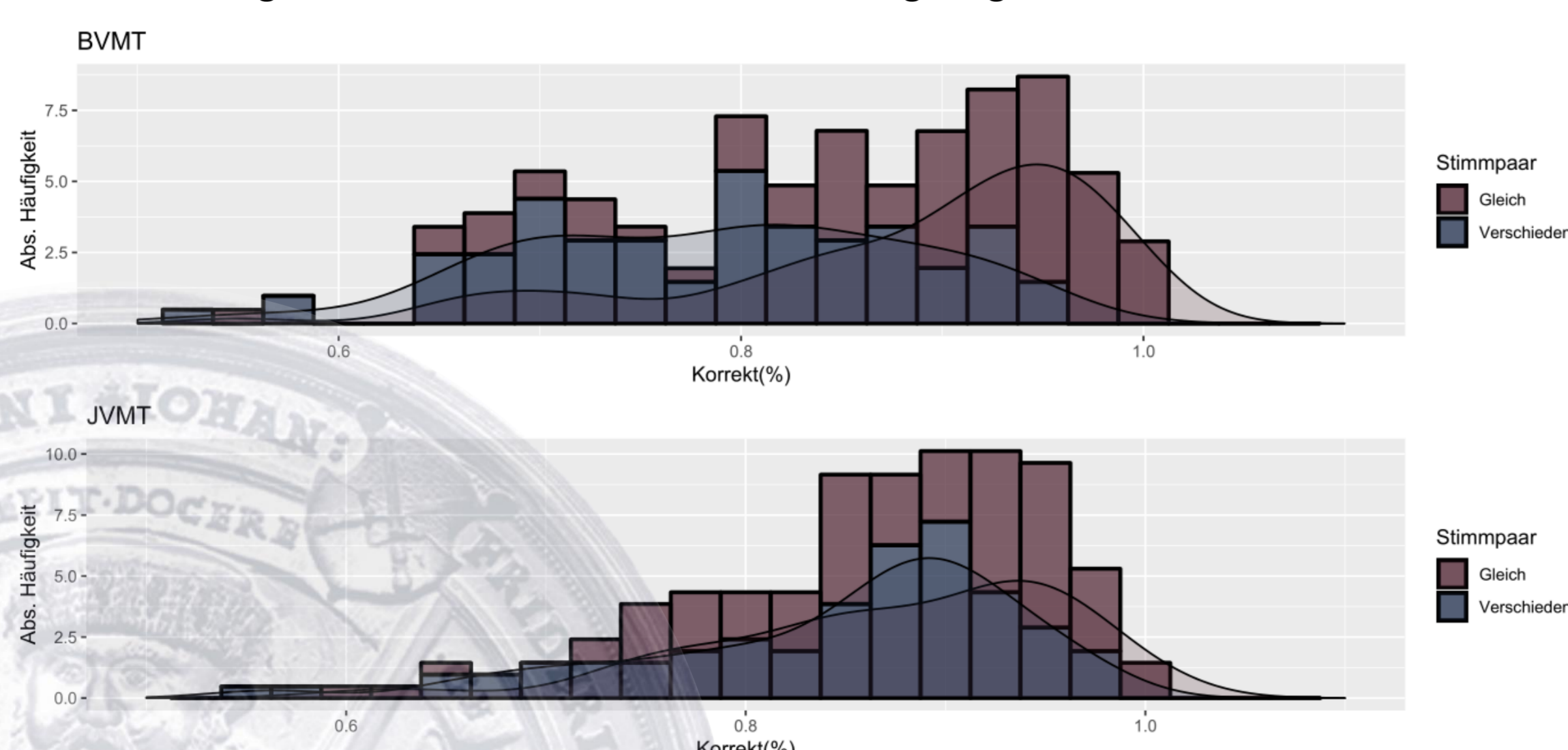


Abb. 1: Lediglich die Items mit gleichem Sprecher werden teilweise von allen Probanden richtig beantwortet

### Ergebnisse

- I. Für den Vergleich zwischen den Tests wurden die jeweiligen Testergebnisse der Probanden in beiden Tests einzeln gemittelt und einander gegenübergestellt. Es zeigte sich ein mittelstarker Zusammenhang zwischen den Testergebnissen in beiden Tests:  $r(81) = .62, p < .001$ . (Pearson-Produkt-Moment-Korrelation)

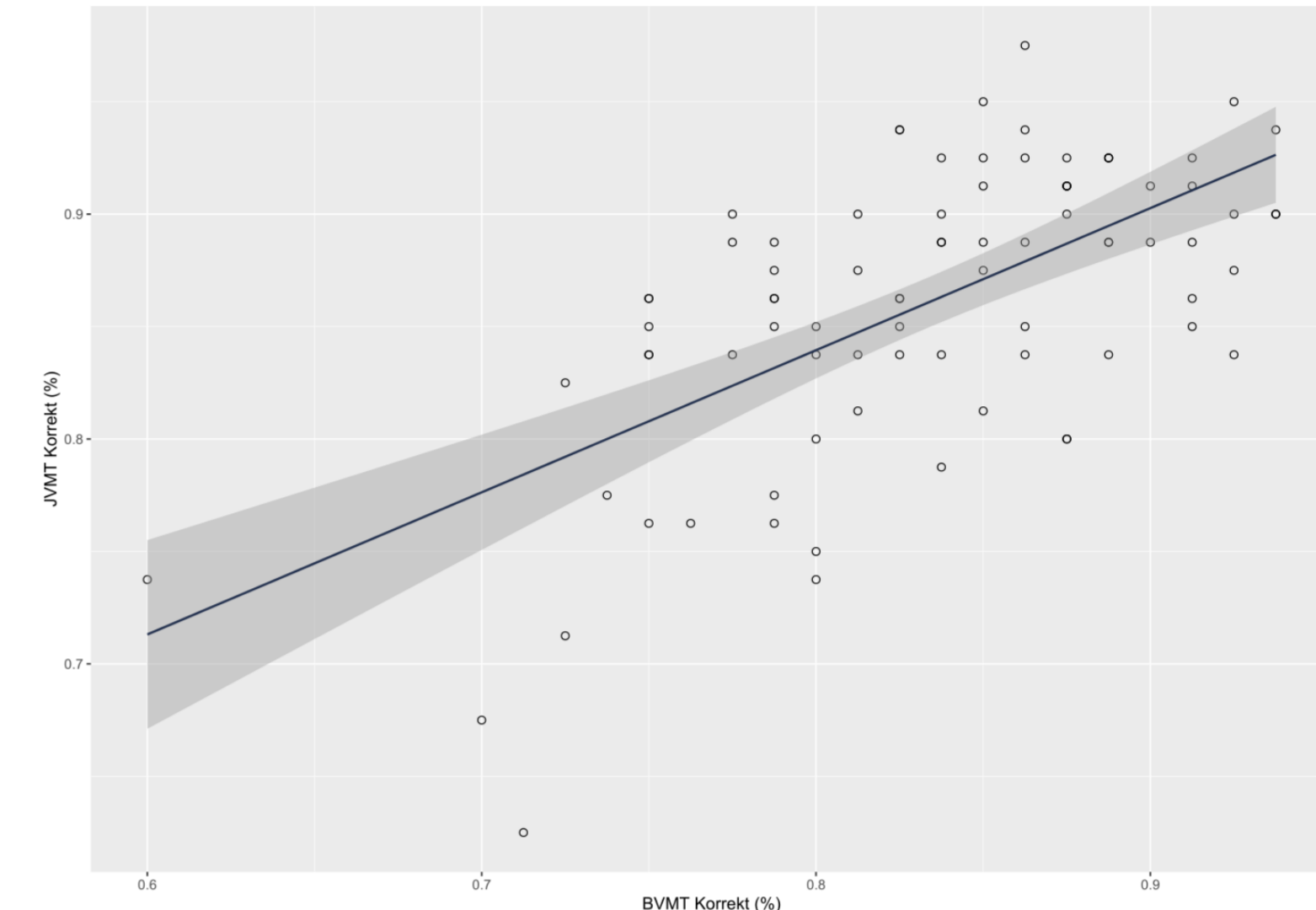


Abb. 2: Der mittelstarke Zusammenhang von JVMT und BVMT spricht für eine zugrundeliegende konvergente Validität

- II. Nach einem nicht signifikanten Levene-Test ( $F(1,230) = .19, p = .65$ ) konnte von gleicher Varianz bei dem in Jena und in Bangor ausgeführten BVMT ausgegangen werden. Die mittlere Testleistung im Jenaer BVMT ( $M = 83.28\%, SD = 6.46$ ) unterschied sich nicht signifikant von der Normierungsstichprobe ( $M = 84.57\%, SD = 7.20; t(78) = -0.27, p = .78$ ). (unabhängiger 2-Gruppen-T-Test)
- III. Die Testergebnisse beider Tests (JVMT, BVMT) wurden verglichen. Im Durchschnitt erzielten Probanden im JVMT eine etwas bessere Leistung als im BVMT ( $M = 86.02\%, SD = 6.50; t(82) = -4.46, p < .001$ ). Zudem erlebten die Probanden die Zuordnung von Sätzen als leichter als die Zuordnung von Silben:  $t(82) = -9.91, p < .001$ . (T-Test für abhängige Stichproben)

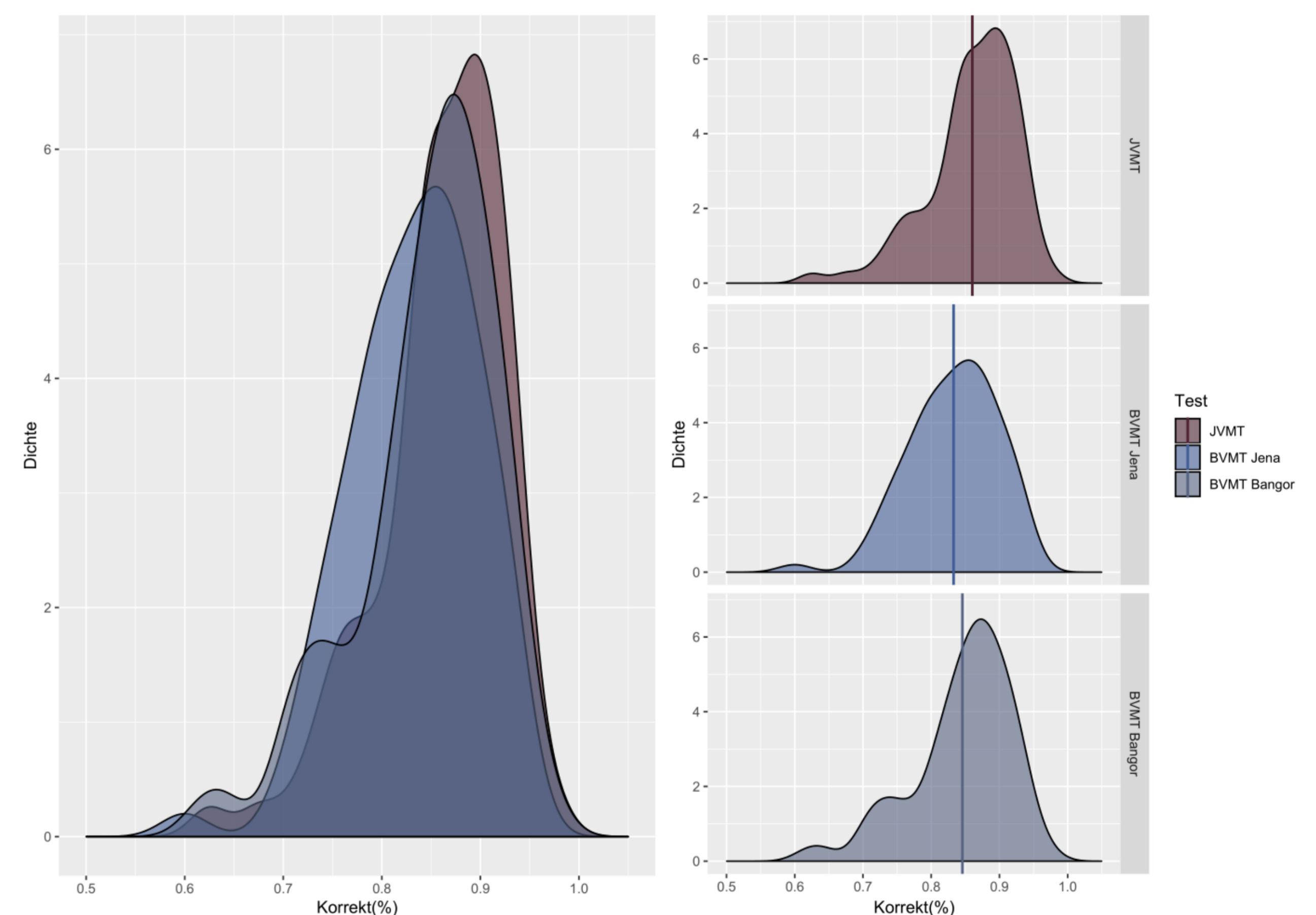


Abb. 3: Die Ähnlichkeit der Dichteverteilungen deutet auf ähnliche Leistungsverteilungen der Probanden hin. Weiterhin sind die Mittelwerte markiert

- IV. Zwischen der mittleren Distinktheit der präsentierten Stimmen (Distinktheitswerte im Wintersemester erhoben) und der korrekten Zuordnung der gleichen Stimmenpaare gab es keinen Zusammenhang:  $r_s(4427.9) = .01, p = .93$ . (Spearman-Rangkorrelationskoeffizient)
- V. Nach einem nicht signifikanten Levene-Test ( $F(1,78) = .19, p = .65$ ), wurde auch ein Vergleich der Genauigkeit bei der Stimmenzuordnung für Frauen- und Männerstimmen nicht signifikant ( $t(78) = 0.27, p = .78$ ). (unabhängiger 2-Gruppen-T-Test) In ihrer subjektiven Einschätzung empfanden Probanden die Zuordnung der Frauenstimmen aber tatsächlich als schwerer als die der Männerstimmen ( $t(82) = 5.35, p < .001$ ). (T-Test für abhängige Stichproben)

- Weiterhin führten wir die Berechnung der internen Konsistenz für den JVMT durch. Über den gesamten Test ergab sich ein Cronbachs  $\alpha$  von .73, was für eine akzeptable interne Reliabilität spricht und vergleichbar ist, mit der des BVMT (Mühl et al., 2007)

### Diskussion

Unsere Hauptziele waren es, den BVMT in einer deutschsprachigen Stichprobe durchzuführen und die Fähigkeit der Stimmerkennung auch mittels ganzer Sätze zu erfassen. Die Ergebnisse des BVMT wurden repliziert. Somit kann der BVMT auch im deutschsprachigen Raum als Test zur Stimmerkennung angewendet werden. Allerdings lässt sich aufgrund des niedrigen Alters beider Stichproben keine Aussage darüber treffen, ob die von JVMT und BVMT gemessene Fähigkeit mit dem Alter variiert. Dass eine längere Stimmexposition zur erleichterten Zuordnung der Stimmen führt, wird durch die erhobenen Daten unterstützt. Die mittlere Distinktheit jeder Stimme hatte keinen Einfluss auf ihre korrekte Erkennung, wenn sie mit sich selbst präsentiert wurde.

Unsere Erkenntnisse tragen zum Forschungsfeld der Stimmwahrnehmung bei, wobei sich weitergehend die Fragen aufdrängen, welche Merkmale einer Stimme besonders zu ihrer Unterscheidbarkeit beitragen und welche Strategien Hörer zur Differenzierung nutzen.

#### Literatur

- Mühl, C., Sheil, O., Jarutytė, L., & Bestelmeyer, P. E. G. (2017). *The Bangor Voice Matching Test: A standardized test for the assessment of voice perception ability*. Behav Res <https://doi.org/10.3758/s13428-017-0985-4>
- Zäske, R., Skuk, V. G., Golle, J., & Schweinberger, S. R. (2018). *The Jena Speaker Set (JESS) – A database of voice stimuli from unfamiliar young and old adult speakers*. Manuskript eingereicht zur Publikation.