

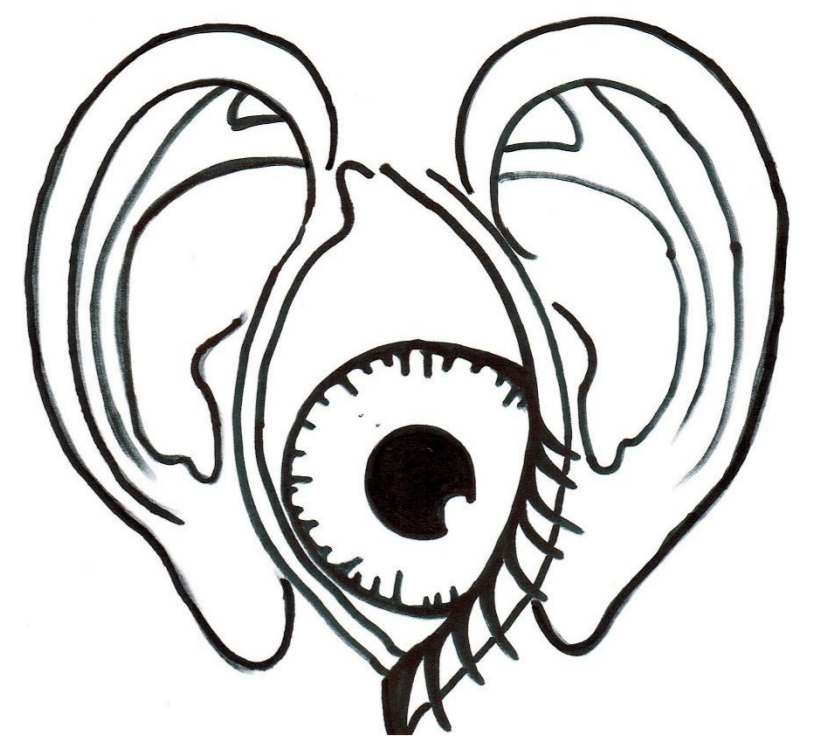


seit 1558

# Das Auge hört mit

## Eine Untersuchung zum Zusammenhang zwischen Kongruenzeffekt und McGurk-Effekt

Isabel Baderschneider, Eugen Depperschmidt, Madeleine Frister, Henrike Rieß  
Leitung: Jürgen M. Kaufmann, Denise Müller



### Einleitung

Gesichter und Stimmen machen uns so einzigartig wie unser Fingerabdruck. Wie wir die vielfältigen Informationen beider Modalitäten verarbeiten, ist im Modell der Gesichter-Stimmen-Integration (Campanella, & Belin, 2007) festgehalten, beginnend auf einem niedrigen Level der Verarbeitung bis hin zur Identifikation der Person auf einem höheren Level. So entspricht beispielsweise der McGurk-Effekt (McGurk & MacDonald, 1976) einem Low-Level-Effekt, die audiovisuelle Integration von Stimme und Gesicht einem High-Level-Effekt. Für die audiovisuelle Integration konnten Schweinberger et al. (2007) einen sogenannten Kongruenzeffekt nachweisen.

**Doch gibt es einen Zusammenhang zwischen den Verarbeitungsprozessen auf dem niedrigen Level mit denen auf dem höheren Level, also zwischen dem McGurk- und dem Kongruenzeffekt?**

**Unsere Hypothese, dass sich bei Personen, die besonders stark die audiovisuellen Hinweise integrieren (und damit deutlich den Kongruenzeffekt aufweisen), auch eher der McGurk-Effekt bestätigen lässt, geht davon aus.** Zur Überprüfung dieser Annahme wurde ein Computereperiment generiert, in welchem den Probanden Gesichter und Stimmen von insgesamt 8 Prominenten in kongruenten und inkongruenten Kombinationen gleichzeitig präsentiert wurden. Die Aufgabe bestand darin, den Prominenten anhand der Stimme so schnell wie möglich zu identifizieren. Abschließend wurde den Probanden ein Video zur Veranschaulichung des McGurk-Effekts gezeigt, mit der Aufforderung, die verstandene Silbe anzugeben. Wir erwarteten, dass Personen, die deutlich langsamere Reaktionszeiten oder weniger korrekte Antworten in den inkongruenten Durchgängen aufweisen, im Video eher eine neue Silbe wahrnehmen, die einer Art Mischung aus den Silben der Mundbewegung und der darüber gelegten Tonspur entspricht.

### Methode

#### Stichprobe

- $N = 25$ ; Alter:  $M = 24.7$ ,  $SD = 9.7$ ; 13 weiblich

#### Stimulusmaterial

- jeweils 4 Bilder von 8 prominenten Persönlichkeiten (Heike Makatsch, Angela Merkel, Heidi Klum, Nena, Franz Beckenbauer, Thomas Gottschalk, Joachim Löw, Otto Waalkes; s. Abb. 1)
- ein Bild von Bart Simpson (s. Abb. 1) um sicherzustellen, dass die Probanden die visuellen Stimuli beachten
- jeweils 4 Tonspuren zu jedem Prominenten (mittlere Dauer: 2450 ms)
- Video zum McGurk-Effekt (Dauer: 45000 ms; s. Abb. 2)



Abbildung 1: Auswahl aus dem verwendeten Bildmaterial



Abbildung 2: Video zum McGurk-Effekt

#### Prozedur

- Lernphase zur Tastenzuordnung:
  - pro Durchgang wurde der Name eines Prominenten auf dem Bildschirm präsentiert
  - die Probanden sollten jeweils die entsprechende Taste drücken
- Übungsphase:
  - 9 Durchgänge
  - Ablauf wie Testphase
- Testphase (s. Abb. 3):
  - 4 Blöcke mit jeweils 128 kongruenten, 128 inkongruenten Durchgängen und 24 Testdurchgängen
  - kongruente Durchgänge: zeitgleiche Präsentation eines Gesichts mit der zu dieser Person gehörigen Stimme
  - inkongruente Durchgänge: zeitgleiche Präsentation eines Gesichts mit einer nicht passenden Stimme
  - Testdurchgänge: Präsentation des Bildes von Bart Simpson ohne Tonspur
  - die Identifizierung der Sprecher erfolgte per Tastendruck
  - in den Testdurchgängen sollte ebenfalls mit Tastendruck reagiert werden
- Präsentation des Videos zum McGurk-Effekt:
  - die Probanden sollten nach Anschauen des Videos die verstandene Silbe in ein Textfeld eingeben

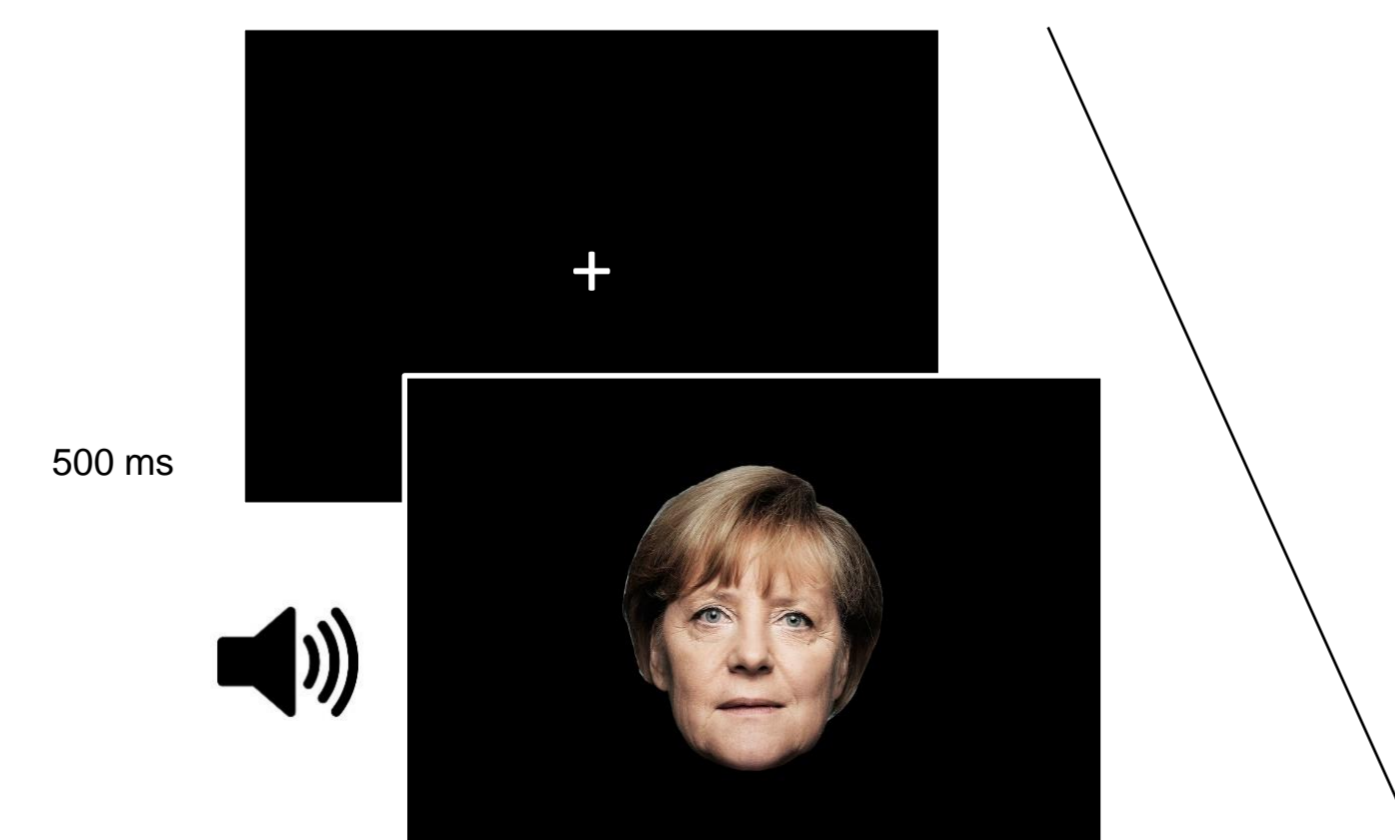


Abbildung 3: schematische Darstellung des Trialablaufs in der Testphase

### Literatur

- Campanella, S. & Belin, P. (2007). Integrating Face and Voice in Person Perception. *Trends in Cognitive Sciences*, Vol. 11, No. 12, 535 - 542
- McGurk, H. & MacDonald, J. (1976). Hearing Lips and Seeing Voices. *Nature*, Vol. 246, 746 - 748
- Schweinberger, S. R., Robertson, D. & Kaufmann, M. J. (2007). Hearing facial identities. *The Quarterly Journal of Experimental Psychology*, Vol. 60, No. 10, 1446 - 1456

### Ergebnisse

#### Kongruenzeffekt

- t-Test für verbundene Stichproben mit den Variablen: mittlere RT der kongruenten  $M = 1664$  ms ( $SD = 185$  ms) und der inkongruenten Bedingung  $M = 1844$  ms ( $SD = 155$  ms),  $t(24) = -11.292$ ,  $p < .001$  (s. Abb. 4a)
- t-Test für verbundene Stichproben mit den Variablen: mittlere ACC der kongruenten  $M = 0.940$  ( $SD = 0.058$ ) und der inkongruenten Bedingung  $M = 0.873$  ( $SD = 0.084$ ),  $t(24) = 5.919$ ,  $p < .001$  (s. Abb. 4b)

#### Untersuchung des Zusammenhangs zwischen Kongruenzeffekt und McGurk-Effekt

- Antwort = /ba/ → Nichtfusionierer ( $N = 19$ )
- Antwort = nicht /ba/ → Fusionierer ( $N = 6$ )
- Mittlere Differenz Fusionierer:
  - mittlere RT kongruent – mittlere RT inkongruent:  $M = -194$  ms ( $SD = 50$  ms) (s. Abb. 5a)
  - mittlere ACC kongruent – mittlere ACC inkongruent:  $M = 0.080$  ( $SD = 0.048$ ) (s. Abb. 5b)
- Mittlere Differenz Nichtfusionierer:
  - mittlere RT kongruent – mittlere RT inkongruent:  $M = -175$  ms ( $SD = 70$  ms) (s. Abb. 5a)
  - mittlere ACC kongruent – mittlere ACC inkongruent:  $M = 0.063$  ( $SD = 0.061$ ) (s. Abb. 5b)
- t-Test für unabhängige Stichproben mit den Variablen mittlere Kosten Fusionierer und Nichtfusionierer
  - RT  $t(23) = .505$ ,  $p = .618$
  - ACC  $t(23) = -.609$ ,  $p = .548$

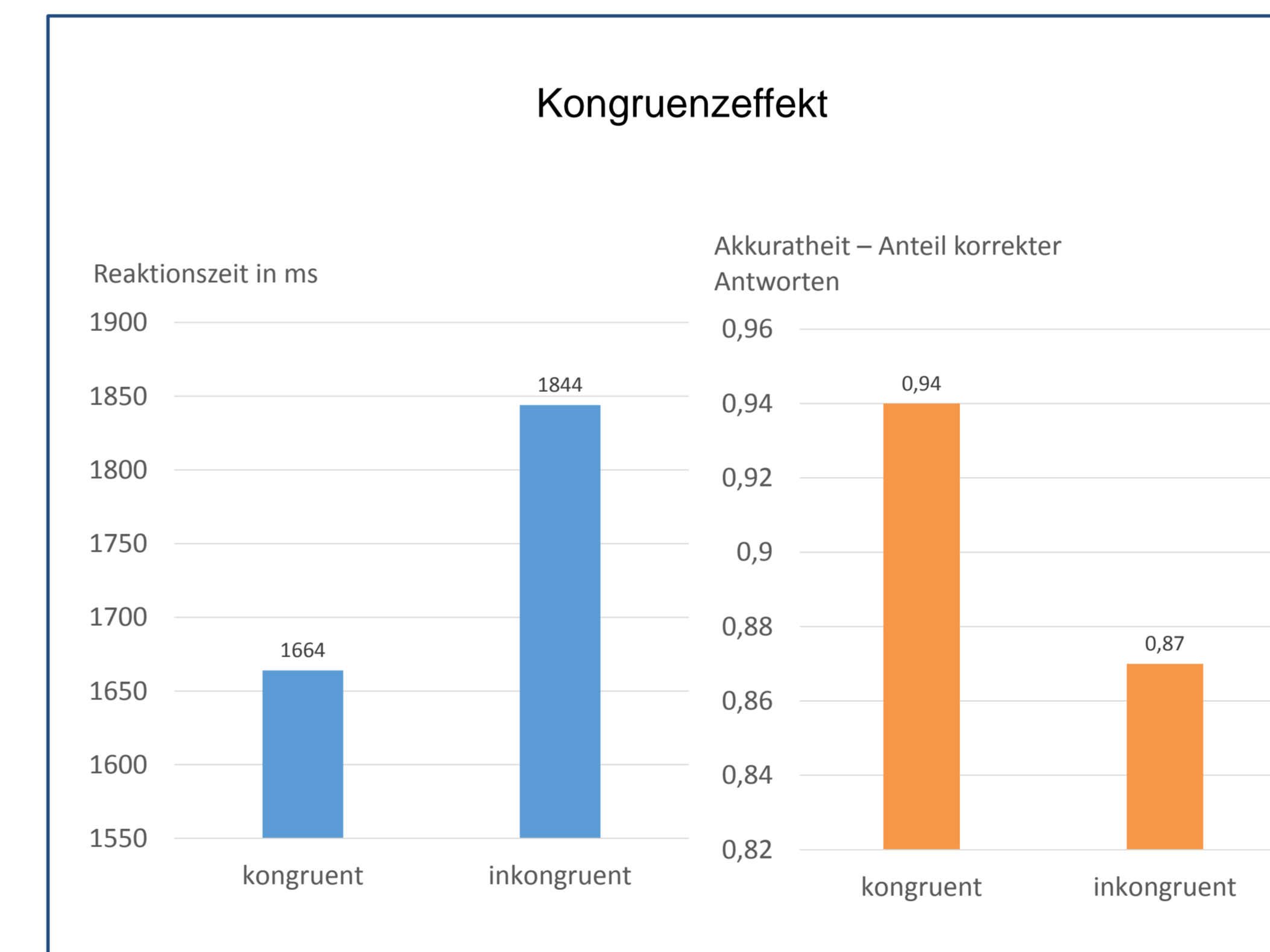


Abbildung 4a: Reaktionszeiten in der kongruenten und inkongruenten Bedingung

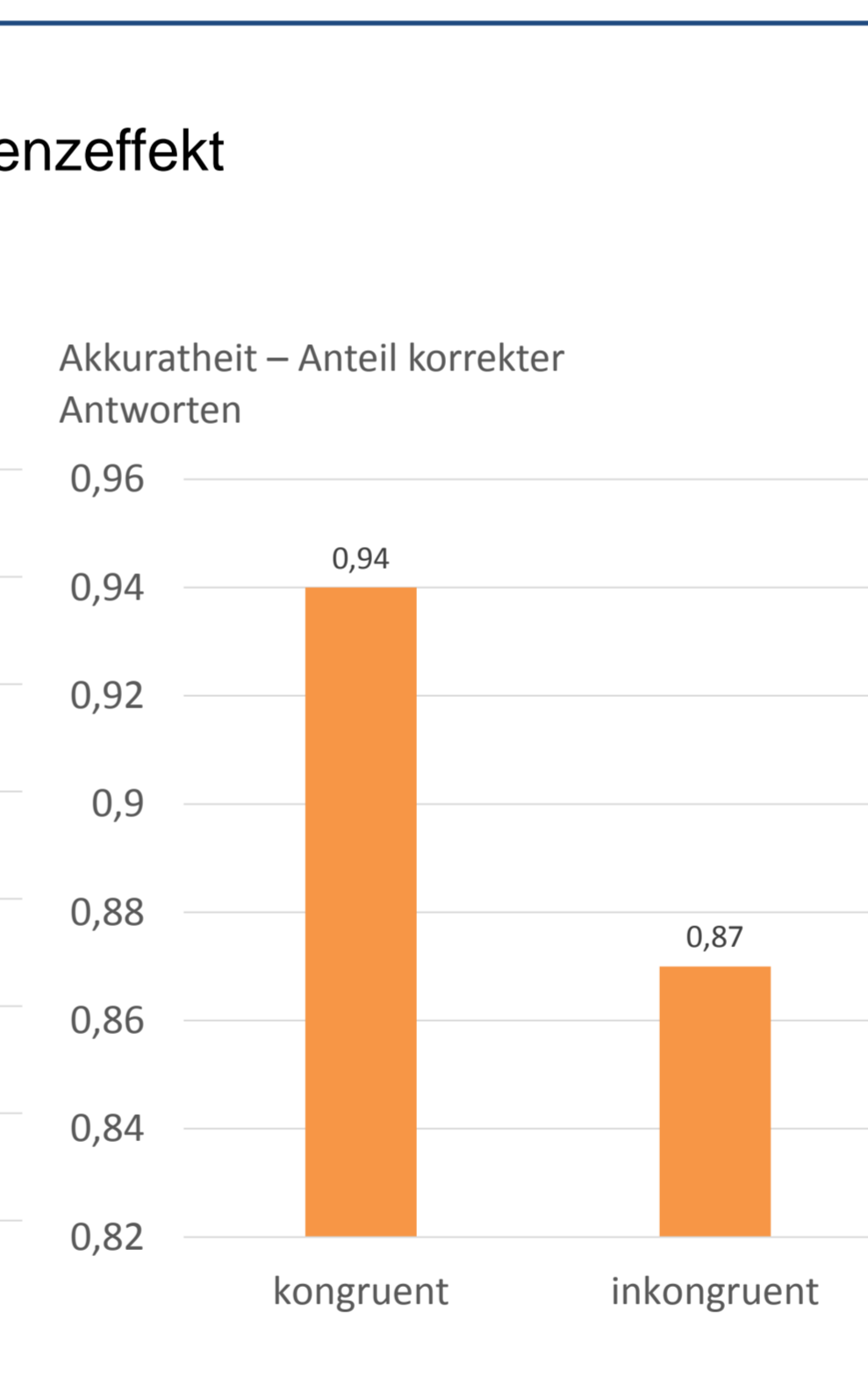


Abbildung 4b: Akkurtheit in der kongruenten und inkongruenten Bedingung

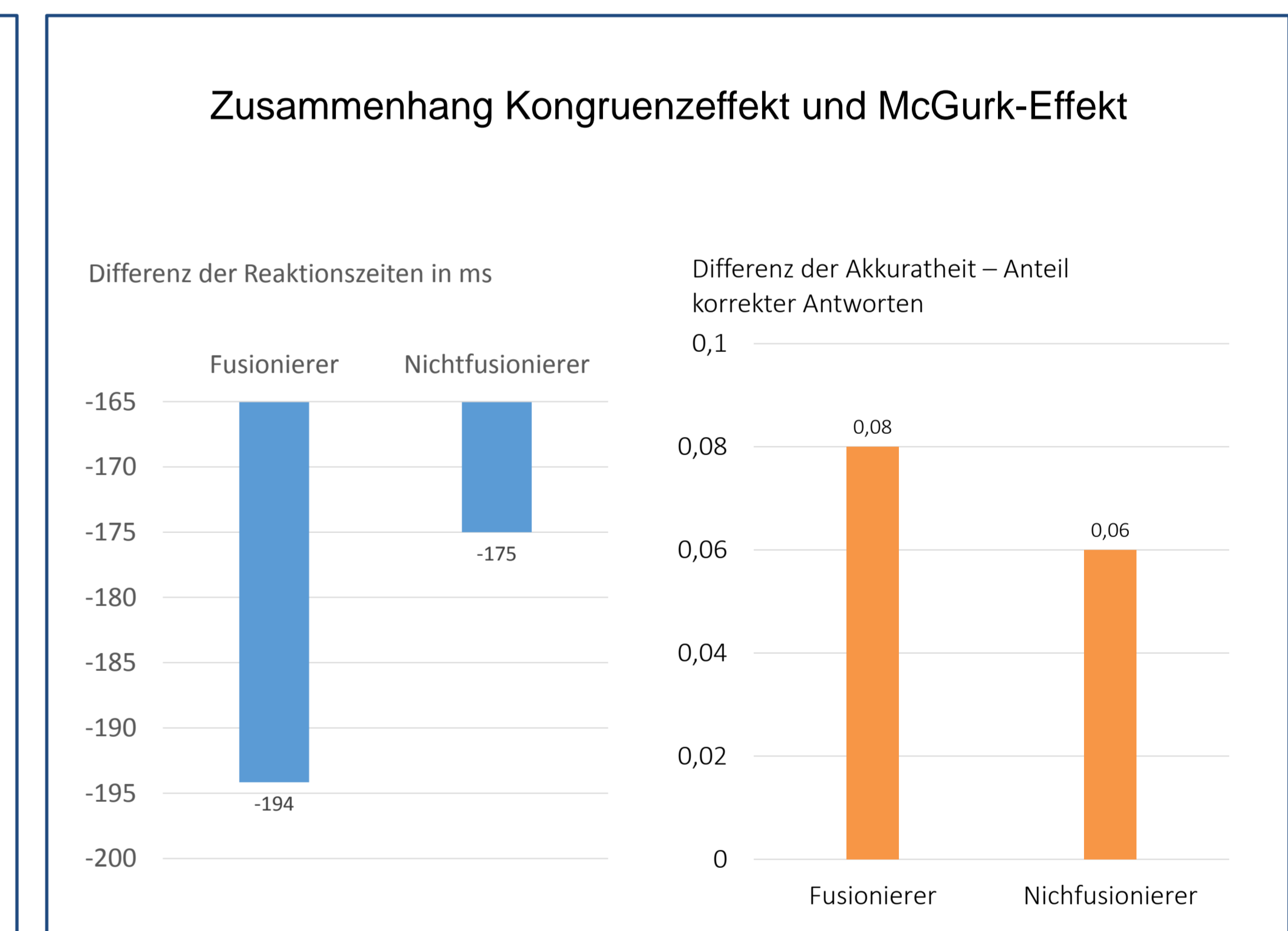


Abbildung 5a: Differenz der Reaktionszeiten in kongruenter und inkongruenter Bedingung bei Fusionierern und Nichtfusionierern

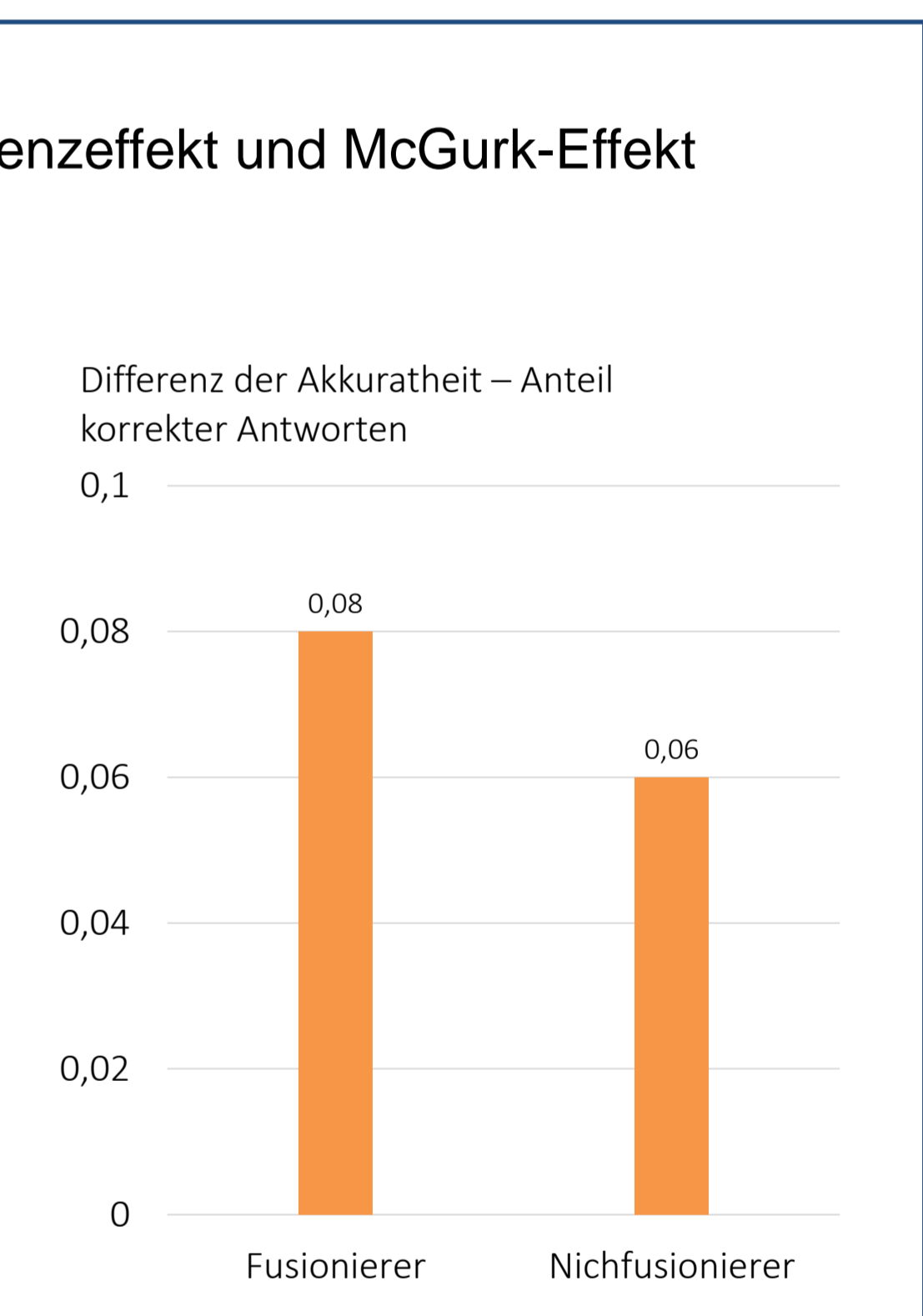


Abbildung 5b: Differenz der Akkurtheit in kongruenter und inkongruenter Bedingung bei Fusionierern und Nichtfusionierern

### Diskussion

#### Hypothesen

- ✓ Der Kongruenzeffekt (High-Level-Effekt) bei audiovisueller Integration von Gesicht und Stimme konnte repliziert werden (Schweinberger et al., 2007).
- × Ein Zusammenhang zwischen diesem High-Level-Effekt und dem McGurk-Effekt (Low-Level-Effekt) konnte in unserem Experiment nicht nachgewiesen werden.

Da sich bei der Testung ergab, dass nur wenige Probanden die gehörte Sprachsequenz und die Lippenbewegung während der Integration fusionierten, konnten wir den McGurk-Effekt, der auf einem niedrigen Verarbeitungslevel stattfinden soll (Campanella & Belin, 2007), nur teilweise nachweisen. Dieses Ergebnis könnte in der Qualität und der sehr lang gewählten Darstellungsdauer des Videos begründet liegen. Ein Motivationsverlust während der Videopräsentation am Ende des Durchlaufs hat womöglich auch dazu geführt, dass einige Versuchspersonen die Augen schlossen statt sie geöffnet zu halten. Bei Wiederholung des Experiments wäre ein Positionieren des Tests für den McGurk-Effekt an den Anfang der Untersuchung empfehlenswert, um eine hohe Konzentrationskapazität der Probanden zu gewährleisten. Zudem würde der Einsatz mehrerer kürzer präsentierter Videosequenzen zur Testung des McGurk-Effekts die Reliabilität der Studie erhöhen und somit auch einen größeren Spielraum bei der Interpretation der Testergebnisse erlauben. Der Kongruenzeffekt zeigt, dass zur Identifizierung von Personen sowohl auditive als auch visuelle Informationen hinzugezogen werden, welche interferieren können. Dies spricht für das Modell der Gesichter-Stimmen-Integration. Inwiefern sich die Stärke audiovisueller Integration bereits auf einem niedrigen Level unimodaler Verarbeitungsmechanismen zeigt und welche Unterschiede Menschen dabei aufweisen, bleibt weiterhin ein Thema für die Forschung.