Der Einfluss emotionaler Sprachprosodie auf die assoziative Leistung im Compound Remote Associates Test

Cleopatra Christina Moshona[1], André Fiebig[1]

[1] *Fachgebiet Technische Akustik, Technische Universität Berlin, 10587 Berlin, E-Mail: c.moshona@tu-berlin.de*

Einleitung und Fragestellung

Der in Studien vielfach belegte Irrelevant Speech Effect (ISE) beschreibt das Phänomen, dass die Arbeitsgedächtnisleistung, speziell die serielle Wiedergabe, unter der Exposition von Sprachschall erheblich beeinträchtigt wird [3, 4], auch wenn der betreffende Hintergrundschall für die eigentliche Tätigkeit irrelevant ist. Als Erklärung für die besondere Störwirkung von Sprache wird oft das Changing-State-Kriterium [5], d. h. die kontinuierliche Zustandsänderung des Sprachsignals, herangezogen. Dieses erscheint jedoch unzureichend. So konnte gezeigt werden, dass sinnvolle [7] und verständliche [15] Sprache, synthetisierte Sprache mit zeitlich invertierten Formantverläufen [16] sowie Äußerungen mit negativer Valenzsemantik [11] additive Störfaktoren für die Gedächtnisleistung zu sein scheinen. Auch emotionale Sprachprosodie, die im Vergleich zu neutralen Pendants höhere Schwankungsstärke-Werte aufweist, beeinträchtigt die serielle Wiedergabe [19]. Es bleibt jedoch unklar, welche konkreten (spektral-temporale) Merkmale von Sprache die beschriebene Beeinträchtigung verursachen bzw. verstärken.

Der Einfluss interferierenden Sprachschalls auf kognitive Bereiche jenseits des Kurzzeitgedächtnisses, speziell auf verbal-assoziative Leistungen, ist bisher wenig erforscht. Der (Compound) Remote Associates Test (RAT bzw. cRAT) [1, 9] operationalisiert die assoziative Leistung, verstanden als problemlösende Fähigkeit, anhand der Lösungsrate von Worträtseln. Neuere Untersuchungen mittels des cRAT zeigen, dass auch assoziative Prozesse für den ISE anfällig sein können, weil durch den Hintergrundschall semantische Such- und Auffindprozesse sowie die Subvokalisierung gestört werden. So sinkt die Lösungsrate unter Sprachschallexposition deutlich im Vergleich zur Ruhe [20]. Im Rahmen der nachfolgenden experimentellen Untersuchung wurde überprüft, inwiefern die Störwirkung des ISE anhand des deutschsprachigen cRAT [17] für neutrale und emotional behaftete, ärgerliche Sprachschalle repliziert werden kann. Ärgerliche Sprechweise ist von besonderen prosodischen und artikulatorischen Merkmalen geprägt, die u. U. erhöhte kognitive Kapazitätsressourcen binden könnten. Darunter zählen eine extreme Form der Betonung, ein erhöhtes Sprechtempo, eine erhöhte Artikulationsgenauigkeit mit Overshooting-Effekten, eine höhere Lautstärke, eine deutliche Anhebung der mittleren Grundfrequenz sowie der ersten beiden Formanten, eine erhöhte F_0-Range, steile F_0-Konturen und abrupte Richtungswechsel, ein sehr steiles Einschwingverhalten von Vokalen in den Frequenzbändern über 600 Hz [10] sowie eine generelle Verlagerung des energetischen Schwerpunktes zum oberen Teil des Spektrums [2]. Wir nahmen daher an, dass die Lösungsrate in Schallbedingungen mit ärgerlicher Prosodie niedriger sein wird als in Schallbedingungen mit neutraler Prosodie und in Ruhe. Wir fragten uns außerdem, inwiefern, falls ein Effekt emotionaler Prosodie zu beobachten ist, dieser auch dann auftritt, wenn die höheren energetischen Bereiche des Spektrums gedämpft sind. Sollte sich ein Effekt trotz der Dämpfung einstellen, würde dies für die Bedeutung zeitlicher Strukturen beim ISE sprechen.

Experimentelle Untersuchung

Die Untersuchung wurde mit einer Stichprobe von 30 Versuchspersonen (10 weiblich, $MW_{Alter} = 30.9$ Jahre, $SD_{Alter} = 7.6$, $Range_{Alter} = 21 - 57$ Jahre) unter kontrollierten Laborbedingungen an der Technischen Universität Berlin durchgeführt. Die Stichprobe setzte sich überwiegend aus Studierenden und Mitarbeitenden des Fachgebiets Technische Akustik der Technischen Universität Berlin zusammen. Alle Versuchspersonen (VP) gaben ihre schriftliche Einverständniserklärung zur Studienteilnahme.

Das Testmaterial bestand aus 36 ausgewählten Worträtseln aus der deutschsprachigen Version des cRAT, die in [17] von 29-60% (MW: 40%) der VP in unter 60 Sekunden gelöst wurden. Es handelte sich somit um eine anspruchsvolle Aufgabe. Die Worträtsel umfassen Wortketten, bestehend aus drei Substantiven, z. B. „Wurm - Wetter - Tonne“. Die Aufgabe besteht darin, ein viertes Substantiv zu finden, welches sich mit jedem der vorgegebenen Wörter zu einem sinnvollen Kompositum zusammensetzen lässt, hier: „*Regen*wurm - *Regen*wetter - *Regen*tonne“. Dabei kann das Lösungswort entweder vor oder hinter den vorgegebenen Wörtern stehen. Es wurden nur Worträtsel ohne Dopplungen in den Triaden-Konstituenten inkludiert, die anhand der berichteten Lösungsraten manuell in drei Schwierigkeitsgruppen (leicht, mittel, schwer) à zwölf Items unterteilt wurden. Das Experiment wurde programmatisch in Matlab implementiert und die Worträtsel mittels einer graphischen Benutzeroberfläche an einem Computer zur Bearbeitung präsentiert. Die Darbietung der Audiodateien erfolgte über ein kalibriertes Wiedergabesystem per Kopfhörer (Sennheiser HD 650).

Die 36 Worträtsel wurden in einem 4x3 *within-subjects* Design in vier Bedingungen (3x Sprachschall, 1x Ruhesituation zur Kontrolle), bestehend aus jeweils drei Items aus jeder Schwierigkeitsgruppe, präsentiert. Es wurden folgende Schallbedingungen dargeboten: neutrale und

ärgerliche Sprechweise, angeglichen auf eine Lautstärke von 70dB(A) sowie tiefpassgefilterte, ärgerliche Sprechweise mit einer Lautstärke von 55dB(A). Durch die Filterung resultiert eine Dämpfung des Spektrums, wodurch der auditive Eindruck eines „durch die Wand“ kommenden Schallereignisses entsteht. Die Reihenfolge der Expositions-Bedingungen erfolgte randomisiert. Die Häufigkeit des Vorkommens der gewählten c-RAT Items pro Bedingung wurde über die Anzahl der VP ausbalanciert, sodass jedes Worträtsel etwa gleich oft in jeder Bedingung vorkam. Die Abfolge der Worträtsel innerhalb der einzelnen Bedingungen war zufällig. Das Zeitlimit für die Bearbeitung jedes Worträtsels betrug 45 Sekunden. Jede Bedingung dauerte somit maximal 6.75 Minuten, sodass die Bearbeitung aller Worträtsel insgesamt bis zu 27 Minuten in Anspruch nahm. Das Sprachmaterial wurde der Berliner Emo-DB [12], einer Datenbank zehn kurzer Sätze, die von zehn Schauspielern in sechs Emotionen eingesprochen worden sind, entliehen. Es wurden nur ärgerliche und neutrale Äußerungen inkludiert, für die Zwecke des Experiments zu längeren Äußerungen mit Pausen von 650 ms zwischen den einzelnen Sätzen konkateniert und als Schleife rückwärts abgespielt, um diese unverständlich zu machen und den Einfluss der Semantik auszuschließen. Bei der Verkettung wurde das Schema $n_{1,1}, n_{1,2} \ldots n_{2,1}, n_{2,2} \ldots$ verwendet, wobei der erste Index für die einzelnen Sprecher und der zweite Index für die einzelnen Sätze steht.

Vor dem Start des eigentlichen Experiments konnten sich die VP anhand einiger Beispielrätsel mit der Aufgabe vertraut machen. Sie wurden instruiert den Hintergrundschall möglichst zu ignorieren, da dieser für die nachfolgende zu bearbeitende Aufgabe irrelevant sei. Zwischen den einzelnen Bedingungen wurden die VP angehalten auf einer neunstufigen Skala (1 = sehr schlecht, 9 = sehr gut) anzugeben, wie gut sie sich in der jeweiligen Bedingung auf die zu bearbeitende Aufgabe konzentrieren konnten. Zur Erfassung der individuellen Stimmung wurden Self-Assessment-Manikin (SAM) [6] Piktogramme eingesetzt. Die VP sollten auf zwei neunstufigen Skalen (1 = unangenehm, 9 = angenehm) bzw. (1 = unaufgeregt, 9 = aufgeregt) angeben, wie sie sich zum jeweiligen Zeitpunkt fühlten (Valenz, Erregung). Am Ende des Experiments wurde außerdem die individuelle Lärmempfindlichkeit anhand des LEF-K Fragebogens [8] erhoben. Der Wertebereich des LEF-K liegt zwischen 0 und 27 Punkten, wobei ein hoher Wert eine hohe Lärmempfindlichkeit bezeichnet.

Ergebnisse

Lösungsrate: Die durchschnittlichen Lösungsraten für alle cRAT-Items in Abhängigkeit der Expositionsbedingung sind Abbildung 1 zu entnehmen. Die insgesamt niedrige Lösungsrate bestätigt, dass es sich um eine anspruchsvolle Aufgabe handelte. Wie aus der Grafik zu erkennen, wurden in allen Schallbedingungen zwar weniger Worträtsel gelöst als in Ruhe und in der ärgerlichen 70dB(A)-Bedingung insgesamt die wenigsten, die Unterschiede zwischen den Bedingungen sind jedoch marginal und statistisch nicht signifikant, vgl. Tabelle 1. Betrachtet man jedoch die durchschnittlichen Lösungsraten differenziert nach Schwierigkeitsgrad, zeigt der gepaarte, einseitige t-Test bei den mittleren cRAT-Items etwas ausgeprägtere Unterschiede, derart, dass in den beiden Schallbedingungen $\text{Ärger}_{70dB(A)}$ ($p = 0.03$) sowie $\text{Ärger}_{55dB(A)}$ ($p = 0.04$) signifikant weniger Rätsel gelöst worden sind als in der neutralen Bedingung, vgl. Abbildung 2. Für leichte und schwere cRAT-Items ergaben sich zwischen den Bedingungen keine signifikanten Unterschiede.

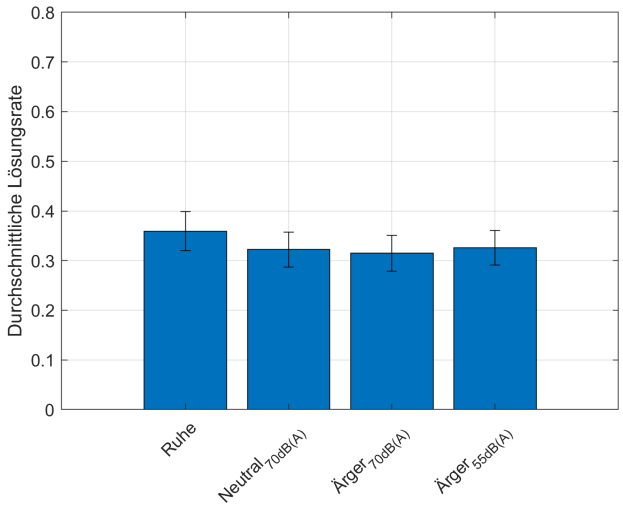

Abbildung 1: Mittlere relative Lösungsraten für alle cRAT-Items in Abhängigkeit der Expositionsbedingung. Die Fehlerbalken repräsentieren den Standardfehler der Mittelwerte innerhalb der VP.

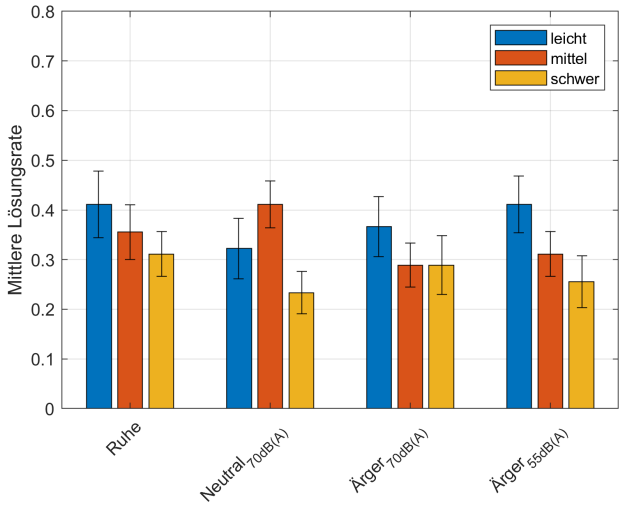

Abbildung 2: Mittlere relative Lösungsraten für leichte, mittlere und schwere cRAT-Items in Abhängigkeit der Expositionsbedingung. Die Fehlerbalken repräsentieren den Standardfehler der Mittelwerte innerhalb der VP.

LEF-K: Basierend auf den ermittelten LEF-K-Scores wurden die VP anhand eines Median-Splits in eine sensible (LEF-K > 13.5, $n = 13$) und eine robuste (LEF-K < 13.5, $n = 17$) Subgruppe unterteilt, die sich, wie in Abbildung 3 dargestellt, in ihren entsprechenden LEF-K-

Mittelwerten (LEF-K$_{sen}$ = MW: 17.53, SD: 3.04 und LEF-K$_{rob}$ = MW: 9.18, SD: 2.90) signifikant ($p < 0.00$) voneinander unterschieden. Eine genauere Betrachtung der Lösungsrate innerhalb der einzelnen Subgruppen bestätigte den allgemeinen Trend, dass in der Ruhesituation mehr Rätsel gelöst worden sind als in den Schallbedingungen, vgl. Tabelle 1. Dieser Trend war für die sensible Gruppe am deutlichsten zwischen der Ruhe- und der ärgerlichen, 55dB(A)-Bedingung ausgeprägt, lag jedoch knapp unter der Signifikanzgrenze ($p = 0.06$).

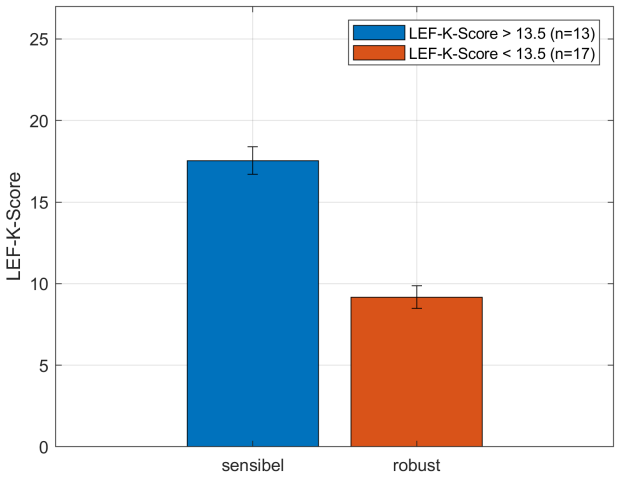

Abbildung 3: Mittlere LEF-K-Scores, unterteilt nach sensibel (blau) und robust (rot). Die Fehlerbalken repräsentieren den Standardfehler der Mittelwerte innerhalb der VP.

Selbstberichtete Konzentration: Die teilnehmenden VP gaben an sich in der Ruhesituation am besten (MW: 6.97, SD: 1.81) und in der ärgerlichen, 70dB(A)-Bedingung (MW: 3.23, SD: 1.76) am schlechtesten auf die zu bearbeitende Aufgabe konzentrieren zu können, vgl. Tabelle 1. Für sensible VP verlagerte sich der Wertebereich nach unten, d. h. stärkere Beeinträchtigung der Konzentration, für robuste VP nach oben. Die Korrelation nach Pearson ergab einen schwachen, signifikanten positiven Zusammenhang zwischen Lösungsrate und Konzentration-Score ($r = 0.21, p = 0.02$).

SAM: Wie aus Tabelle 1 zu entnehmen, fühlten sich die VP in der Ruhesitation am angenehmsten (MW: 6.37, SD: 2.06) und in der ärgerlichen, 70dB(A)-Bedingung am unangenehmsten (MW: 4.57, SD: 2.14). Auch hier verschob sich der Wertbereich für sensible VP nach unten und für robuste VP nach oben. Zwischen Lösungsrate und SAM-Valenz konnte kein statistischer Zusammenhang festgestellt werden, jedoch zeigte sich eine moderate, signifikante positive Korrelation ($r = 0.57, p < 0.00$) zwischen Konzentrations- und Valenz-Score. Analog gaben die VP an in der Ruhesituation am wenigsten aufgeregt (MW: 2.87, SD: 1.78) und in der ärgerlichen, 70dB(A)-Bedingung am meisten aufgeregt zu sein (MW: 4.20, SD: 1.95). Unerwarteterweise ergaben sich jedoch für die robusten VP höhere Erregungs-Scores als für die sensiblen VP. Auch hier zeigte sich kein signifikanter Zusammenhang zwischen Lösungsrate und SAM-Erregung. Es ergab sich jedoch eine schwache, signifikante negative Korrelation ($r = -0.27, p < 0.00$) zwischen Konzentrations- und Erregungs-Score sowie zwischen der Valenz- und Erregungsbewertung ($r = -0.31, p < 0.00$). Wie zu erwarten, zeigte sich bei der Bewertung der Valenz und Erregung für Ärger$_{55dB(A)}$ ein Neutralisierungseffekt im Vergleich zu Ärger$_{70dB(A)}$, vgl. Tabelle 1, was wiederum bestätigt, dass die Dämpfung des Spektrums einen Einfluss auf die Ausgeprägtheit der wahrgenommenen Emotion hat, da viel akustische Information, die für Ärger charakteristisch ist, verloren geht.

Diskussion

Die Ergebnisse vorheriger Studien [20] konnten in der vorliegenden experimentellen Untersuchung nicht eindeutig repliziert werden. Es zeigte sich jedoch ein genereller Trend in Richtung der beiden Forschungshypothesen, d. h. a) dass die Störwirkung des ISE für assoziative Leistungen unter Sprachschallexposition höher ist als in Ruhe und b) dass emotionale Prosodie für Worträtsel mittlerer Schwierigkeit einen additiven Störeffekt auf die Lösungsrate hat. Mögliche Erklärungen für die fehlende statistische Signifikanz könnten darin liegen, dass der Störeffekt des ISE aufgrund der Unverständlichkeit des Sprachsignals abgeschwächt wurde. Es könnte jedoch auch sein, dass die Wahl der cRAT-Items einen sehr starken Einfluss auf die Testleistung hat, wie auch die differenzierte Betrachtung des Schwierigkeitsgrads vermuten lässt. Die Tatsache, dass sich bei [20] nach der Wahl eines modifizierten Item-Sets angelehnt an [14] klarere Ergebnisse einstellten, spricht dafür, dass der cRAT für die Erfassung der Schallexpositionswirkung im Rahmen assoziativer Prozesse nicht zuverlässig funktioniert. Es erscheint plausibel, dass die cRAT-Leistung insgesamt eher primär von subjektiven Faktoren wie die individuelle Arbeitsgedächtnisspanne und die Aufmerksamkeitskontrolle [18] abhängt. Des Weiteren könnten sekundäre Faktoren wie z. B, die Lärmempfindlichkeit die Testleistung beeinflussen. Es kann jedoch auch nicht ausgeschlossen werden, dass assoziative Prozesse grundsätzlich gegen die Störwirkung des ISE weniger anfällig sind als serielle Wiedergabeleistungen.

Die vorliegende Untersuchung zeigte ferner, dass die Expositionsbedingung einen deutlichen Einfluss auf die wahrgenommene, selbstberichtete Konzentrationsfähigkeit und auf die Stimmung hat. Dies könnte bedeutsame Implikationen für die individuelle Disposition und somit auch für die Bereitschaft von Personen sich mit bestimmten Aufgaben auseinanderzusetzen haben. Obwohl die subjektive Einschätzung in Bezug auf die Konzentration und die eigentlichen Test-Leistungen korrelativ zusammenhingen, kann jedoch nicht ausgeschlossen werden, dass diese Relation eher auf ein konsistentes Verhalten seitens der VP zurückzuführen ist als auf einen kausalen Zusammenhang zwischen der selbstberichteten Konzentration und der Testleistung. So könnte es sein, dass die VP in den Bedingungen, in denen sie nach ihrem Empfinden schlecht abgeschnitten haben, auch eine entsprechend niedrigere Konzentrationsfähigkeit an-

gaben. Dafür spricht auch, dass die Lösungsrate weder mit dem Valenz-, noch mit dem Erregungs-Score korrelierte. Aufgrund der ermittelten Ergebnisse liefert die vorliegende Untersuchung einen Beitrag für ein besseres Verständnis der Wirkung von Schallexposition auf assoziative Leistungen. Um den Zusammenhang zwischen spezifischen Störschalleigenschaften und assoziativen Leistungen im Detail nachvollziehen zu können, bedarf es jedoch weiterer Untersuchungen.

Tabelle 1: Mittlere relative Lösungsraten und selbstberichtete Konzentrations- sowie SAM Valenz- und Erregungs-Scores pro Expositionsbedingung für alle VP und differenziert nach sensiblen und robusten VP.

		Expositionsbedingung			
		R	**N**	**Ä70**	**Ä55**
Lösungsrate	alle	0.36	0.32	0.31	0.33
	sen.	0.39	0.35	0.35	0.29
	rob.	0.33	0.30	0.29	0.35
Konzentration	alle	6.97	3.50	3.23	4.97
	sen.	6.46	2.92	2.69	4.38
	rob.	7.35	3.94	3.65	5.41
SAM-V	alle	6.37	4.87	4.57	5.30
	sen.	5.69	4.08	3.92	4.38
	rob.	6.88	5.47	5.06	6.00
SAM-E	alle	2.87	3.77	4.20	3.50
	sen.	2.31	3.15	4.08	3.31
	rob.	3.29	4.24	4.29	3.65

Literatur

[1] Mednick, S. A.: The associative basis of the creative process. Psychological Review (1962), 69, 220–232

[2] Williams, C. E.: Emotions and speech: some acoustical correlates, The Journal of the Acoustical Society of America (1972), 52, 1238-50

[3] Colle, H. A., Welsh, A.: Acoustic masking in primary memory. Journal of Verbal Learning and Verbal Behaviour (1976), 15, 17-31

[4] Salamé, P., Baddeley, A. D.: Disruption of short-term memory by unattended speech: Implications for the structure of working memory. Journal of Verbal Learning & Verbal Behavior (1982), 21, 150–164

[5] Jones, D. M., Macken, W. J.: Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. Journal of Experimental Psychology: Learning, Memory, & Cognition (1993), 19, 369-381

[6] Bradley, M., Lang, P.: Measuring emotion: The self-assessment manikin and the semantic differential. In: Journal of Behavior Therapy and Experimental Psychiatry (1994), 25(1), 49–59

[7] LeCompte, D. C., Neely, C. B., Wilson, J. R.: Irrelevant speech and irrelevant tones: the relative importance of speech to the irrelevant speech effect. Journal of Experimental Psychology: Learning, Memory and Cognition (1997), 23(2), 472-483

[8] Ellermeier, W., Zimmer, K.: Ein Kurzfragebogen zur Erfassung der Lärmempfindlichkeit. Umweltpsychologie (1998), 2(2), 54-63

[9] Bowden, E. M., Beeman-Jung, M.: Normative data for 144 compound remote associate problems. Behavior Research Methods, Instruments & Computers (2003), 35(4), 634-639

[10] Paeschke, A.: Prosodische Analyse emotionaler Sprechweise, Mündliche Kommunikation, Band 1, Berlin Logos Verlag (2003), 226-228

[11] Buchner A. et al.: Valence of distractor words increases the effects of irrelevant speech on serial recall. Memory and Cognition (2004), 32(5), 722-731

[12] Burkhardt et al.: A Database of German Emotional Speech, Proceedings Interspeech (2005), Lissabon, Portugal

[13] Hughes, R. W., Vachon, F., Jones, D. M.: Disruption of short-term memory by changing and deviant sounds: Support for a duplex-mechanism account of auditory distraction. Journal of Experimental Psychology: Learning, Memory, and Cognition (2007), 33(6), 1050–1061

[14] Sandkühler, S., Bhattacharya, J.: Deconstructing insight: EEG correlates of insightful problem solving, PLoS One (2008), 3(1), e1459, 1–12. https://doi.org/10.1371/journal.pone.0001459

[15] Little, J. S., Martin, F. H., Thomson, R. H. S.: Speech versus non-speech as irrelevant sound: controlling acoustic variation. Biological Psychology (2010), 85(1), 62-70

[16] Viswanathan, N., Dorsi J., George, S.: The role of speech-specific properties of the background in the Irrelevant Sound Effect. The Quarterly Journal of Experimental Psychology (2014), 67(3), 581–589

[17] Landmann, N. et al.: Entwicklung von 130 deutschsprachigen Compound Remote Associate (CRA)-Worträtseln zur Untersuchung kreativer Prozesse im deutschen Sprachraum. Psychologische Rundschau (2014), 65(4), 200-211, Hogrefe Verlag Göttingen

[18] Chein, J. M., Weisberg, R. W.: Working memory and insight in verbal problems: Analysis of compound remote associates. Memory and Cognition (2014), 42(1), 67–83

[19] Kattner, F., Ellermeier, W.: Emotional prosody of task-irrelevant speech interferes with the retention of serial order. Journal of experimental psychology: Human perception and performance (2018), 44(8), 1303-1312

[20] Marsh, J. E. et al.: The susceptibility of compound remote associate problems to disruption by irrelevant sound: a Window onto the component processes underpinning creative cognition? Journal of Cognitive Psychology (2021), https://doi.org/10.1080/20445911.2021.1900201