

Individuelle Unterschiede beim Sprachverstehen im Störschall: Welche Rolle spielt der Irrelevant Sound Effect?

Katharina Rost¹, Wolfgang Ellermeier², Florian Kattner² & Daniel Oberfeld-Twistel¹

¹ Johannes Gutenberg-Universität Mainz

² Technische Universität Darmstadt

Einleitung

Gespräche, die im Hintergrund geführt werden, sind auf den ersten Blick irrelevant, beeinflussen aber sowohl die Wahrnehmung als auch die Gedächtnisfähigkeit unbeteiligter Dritter.

So zeigt die Forschung zum sogenannten *Cocktail Party Listening*, dass es schwieriger ist, einen Gesprächspartner zu verstehen, wenn im Hintergrund eine weitere oder mehrere Unterhaltungen stattfinden. Diese Störung der Sprachwahrnehmung wird unter anderem auf eine energetische und informationale Maskierung des Zielschalls durch den Störschall sowie auf Aufmerksamkeitsprozesse zurückgeführt [1]. Wie sehr das Sprachverständnis durch die Hintergrundgeräusche beeinträchtigt wird, ist individuell sehr unterschiedlich: So unterscheiden sich auch normalhörende Probanden deutlich darin, zu welchem Prozentsatz sie die Sätze eines imaginären Gesprächspartners korrekt wiedergeben können, während sie gleichzeitig irrelevanten Hintergrundgeräuschen und -gesprächen ausgesetzt sind, die räumlich versetzt präsentiert werden [2][3]. Diese Unterschiede wurden in früheren Studien bereits u.a. auf Unterschiede in der basalen Hörleistung [4], der binauralen Sensitivität für die zeitliche Feinstruktur [2] und kognitive Faktoren wie Gedächtniskapazität und Fähigkeit zur Zuweisung selektiver Aufmerksamkeit [2][5][6] zurückgeführt. Die Ergebnisse sind dabei nicht immer eindeutig und können die inter-individuelle Varianz nur zu einem gewissen Teil aufklären.

Unsere Studie ging der Frage nach, ob der Einfluss des Hintergrundschalls auf die Gedächtnisleistung eine Rolle für die individuellen Unterschiede im Sprachverstehen unter Störschall spielt. Langjährige Forschung zum Irrelevant Sound Effect zeigt, dass nicht-stationärer, aufgabenirrelevanter Hintergrundschall die Leistung in einer Gedächtnisaufgabe (serial recall) beeinträchtigt [7] und dass auch hier deutliche individuelle Unterschiede bestehen [8]. Überraschenderweise wurde bislang aber nicht untersucht, ob individuelle Unterschiede im Irrelevant Sound Effect eine Rolle für das Sprachverstehen im Störschall spielen. Im hier vorgestellten Experiment wurde deshalb die individuelle Leistung im Sprachverstehen unter Störschall (Cocktail Party Listening) als auch in der Gedächtnisleistung unter Störschall (Irrelevant Sound Effect) bestimmt. Die Messung des Irrelevant Sound Effect wurde sowohl mit visuell als auch mit auditiv präsentierten Items durchgeführt, um einen Einfluss der Präsentationsart untersuchen zu können. Darüber hinaus wurden mit der individuellen Hörschwelle des besseren Ohres, der mittleren Hörschwellenasymmetrie, der binauralen Sensitivität für die zeitliche Feinstruktur, der Kurzzeitgedächtnisleistung in einer Bedingung ohne

Hintergrundschall und der Lärmempfindlichkeit weitere potenzielle Einflussfaktoren gemessen.

Methode

Versuchspersonen

An der Studie nahmen 67 normalhörende Versuchspersonen (10 Männer und 57 Frauen, 17-37 Jahre, $M = 22.85$ Jahre, $SD = 4.37$ Jahre) mit Deutschkenntnissen auf muttersprachlichem Niveau teil. Die Durchführung des Experiments erfolgte gemäß der Richtlinien der Declaration of Helsinki.

Gemessene Parameter

1. Mittlere Hörschwelle und Hörschwellenasymmetrie. Alle auditiven Stimuli wurden mittels Sennheiser HDA300 Kopfhörern in einer schallgeschützten Kabine dargeboten. Die Hörschwelle der Versuchspersonen wurde mittels Békésy-Tracking [9] für beide Ohren im Frequenzbereich zwischen 100 Hz und 16 kHz gemessen.

Die mittleren Hörschwellen (pure tone average) wurden separat für rechtes und linkes Ohr in an den Oktavfrequenzen 250 Hz, 500 Hz, 1 kHz, 2 kHz und 4 kHz zentrierten Terzbändern errechnet. Der Asymmetriewert wurde für dieselben Frequenzen als Betrag der Differenz der Hörschwellen zwischen rechtem und linkem Ohr berechnet und die mittlere HL-Asymmetrie genauso wie die mittlere Hörschwelle des besseren Ohres als Kennzahl für die Auswertung verwendet.

2. Messung der binauralen Sensitivität für die zeitliche Feinstruktur. Zur Messung der binauralen Sensitivität für die zeitliche Feinstruktur wurde der TFS-LF-Test [10] verwendet. In einem adaptiven Zwei-Intervall-Verfahren mit einer 3-down, 1-up-Regel wurden der Versuchsperson in jedem Trial nacheinander zwei Intervalle mit jeweils vier Sinustönen (500 Hz, 400 ms) präsentiert. In einem der beiden Intervalle waren alle Töne diotisch mit einer interauralen Phasendifferenz von 0 Grad; in dem anderen Intervall wurden der zweite und vierte Ton mit einer interauralen Phasendifferenz größer als 0 Grad präsentiert. Die Phasenverschiebung trat mit gleicher Wahrscheinlichkeit im ersten oder zweiten Intervall auf. Die Versuchsperson hatte die Aufgabe, das Intervall zu identifizieren, in der der zweite und vierte Ton aus einer anderen Richtung kamen.

Für jede Versuchsperson wurde die IPD-Detektionsschwelle in drei Versuchsblöcken gemessen. Als Maß für die binaurale Sensitivität für die zeitliche Feinstruktur wurde der arithmetische Mittelwert der IPD-Schwellen der drei Blöcke berechnet.

3. Messung des Irrelevant Sound Effect und der Kurzzeitgedächtnisleistung. Der Irrelevant Sound Effect wurde sowohl für visuell als auch für auditiv präsentierte zu memorierende Ziffernfolgen gemessen. In jedem Trial

wurden acht der Ziffern 0-9 ohne die 7 in zufälliger Reihenfolge auf dem Bildschirm (in der visuellen Irrelevant Sound Effect-Aufgabe) oder auditiv (in der auditiven Irrelevant Sound Effect-Aufgabe) präsentiert. Im Anschluss an die letzte Ziffer folgte ein 4,8-sekündiges Rehearsal-Intervall; danach musste die Versuchsperson die Ziffern in einer Matrix auf dem Bildschirm in der richtigen Reihenfolge anklicken (serial recall).

In der Hälfte der Trials gab es keinen Störschall; in der anderen Hälfte wurden zwei Störsprecher präsentiert, die zeitgleich mit der ersten Ziffer begannen und bis zum Ende des Rehearsal-Intervalls zu hören waren. Als Grundlage für die Störsprecher diente standardisiertes Sprachmaterial aus dem Oldenburger Satztest (OLSA; Hörtech gGmbH, Oldenburg) [11]. Für die Irrelevant Sound Effect-Aufgaben wurden 50 verschiedene Störschalle verwendet, für die für jeden der zwei Störsprecher jeweils sieben Sätze aus dem OLSA zufällig ausgewählt, aneinandergesetzt und die entstandenen Sprachaufnahmen einheitlich auf 12 Sekunden Dauer mit einer Anstiegs- und Abfallszeit von jeweils 200 ms gekürzt wurden. Die Sprachschalle hatten einen energieäquivalenten Dauerschallpegel von $L_{Aeq} = 55$ dB. Per virtueller Akustik wurden ihnen raumbezogene binaurale Impulsantworten (BRIRs) eines aechonischen Raums mit einem Azimutwinkel von 25 Grad nach rechts (für den einen Sprecher) bzw. 25 Grad nach links (für den anderen Sprecher) aufgeprägt und somit der Eindruck erzeugt, die Sprecher würden sich vorn rechts und links der Versuchsperson befinden.

Bei der Messung des Irrelevant Sound Effect für visuell präsentierte Ziffernfolgen wurden die zu lernenden Ziffern zentral in Schwarz auf einem hellgrauen Bildschirmhintergrund angezeigt. Sie wurden mit einer zeitlichen Dauer von jeweils 900 ms unmittelbar nacheinander präsentiert.

Für die Messung des Irrelevant Sound Effect auditiv präsentierte Ziffernfolgen wurden Sprachaufnahmen von Ziffern verwendet, die von einer weiblichen Sprecherin vorgelesen wurden [12]. Die Ziffern wurden ebenfalls mit einem Inter-Onset-Intervall von 900 ms und auf einem Azimutwinkel von 0 Grad präsentiert; dadurch wurde die Sprecherin direkt gegenüber der Versuchsperson positioniert. Der Pegel der Zielsprecherin war mit $L_{Aeq} = 62$ dB um etwa 7 dB höher als der Pegel der Störschalle, um eine energetische Maskierung weitgehend zu vermeiden.

Pro Irrelevant Sound Effect-Variante (visuell, auditiv) gab es zwei Blöcke mit jeweils 25 Trials in Stille und 25 Trials mit Störsprechern. Der individuelle Irrelevant Sound Effect wurde als Differenz aus der mittleren Anzahl an richtiger Position erinnerter Ziffern in Stille und der mittleren Anzahl an richtiger Position erinnerter Ziffern bei Sprache berechnet, getrennt nach den beiden Aufgabentypen.

Die mittlere Anzahl korrekt erinnerter Ziffern in Stille wurde als Maß für die basale Kurzzeitgedächtnisleistung verwendet.

4. Messung der Satzidentifikationsleistung unter Störschall (Cocktail Party Listening). Die Satzidentifikationsleistung wurde mittels eines Matrixtests unter Verwendung von Sprachmaterial aus dem Oldenburger Satztest [11] gemessen. Es wurde mittels virtueller Akustik

erneut eine unterschiedliche räumliche Position des Zielsprechers und von zwei Störsprechern simuliert, die der räumlichen Situation in der auditiven Irrelevant Sound-Aufgabe entsprach (Zielsprecher auf 0° , Störsprecher auf $\pm 25^\circ$). Für die beiden Störsprecher wurde dieselbe Stimme aus dem OLSA-Sprachmaterial wie für den Zielsprecher verwendet. Alle Sprecher starteten gleichzeitig und produzierten je einen Satz aus dem OLSA-Sprachmaterial, mit einem Pegel von jeweils 60 dB SPL (Störsprecher) und 58 dB SPL (Zielsprecher).

Es wurden drei Versuchsblöcke mit jeweils 50 Trials präsentiert. Der erste Block wurde dabei als Übungsblock genutzt. Als Maß für die individuelle Satzidentifikationsleistung wurde das arithmetische Mittel aus der prozentual richtig wiedergegebene Anzahl an Wörtern aus dem zweiten und dritten Block verwendet.

5. Messung der Lärmempfindlichkeit. Zur Messung der individuellen Lärmempfindlichkeit wurde der Lärmempfindlichkeitsfragebogen von Zimmer und Ellermeier [13] verwendet.

Versuchsablauf

Jede Versuchsperson nahm an zwei Versuchssitzungen mit jeweils etwa eineinhalb bis zwei Stunden Dauer teil. In der ersten Sitzung fanden 1.) die Messung der Hörschwelle, 2.) die Messung der binauralen Sensitivität für die zeitliche Feinstruktur und 3.) die Messung des visuellen Irrelevant Sound-Effect statt, während in der zweiten Sitzung 1.) die Messung der Satzidentifikationsleistung unter Störschall, 2.) die Messung der Lärmempfindlichkeit und 3.) die Messung des auditiven Irrelevant Sound Effect stattfanden.

Ergebnisse und Diskussion

Um einen Einfluss der gemessenen Faktoren auf das Cocktail Party Listening zu untersuchen, wurde eine multiple lineare Regressionsanalyse mit der Satzidentifikationsleistung im Störschall als Kriterium und allen gemessenen Parametern sowie dem Alter der Versuchsperson als Prädiktoren durchgeführt. Alle Variablen waren z-standardisiert. Im Rahmen einer Regressionsdiagnostik [14] wurden sechs Datensätze als Ausreißer identifiziert und von der Analyse ausgeschlossen.

Die Ergebnisse sind in Tabelle 1 dargestellt. Es fand sich kein signifikanter Zusammenhang zwischen der Satzidentifikationsleistung im Störschall und dem Irrelevant Sound Effect, weder für die Variante mit visuell noch für die Variante mit auditiv präsentierten Ziffernfolgen.

Für zwei mit auditiver Lokalisation verknüpfte Parameter (binaurale Asymmetrie der Hörschwellen, IPD-Detektionsschwellen) gab es hingegen einen signifikant negativen Zusammenhang zur Satzidentifikationsleistung im Störschall: Je größer die interaurale Hörschwellenasymmetrie bzw. der zur Detektion notwendige interaurale Phasenunterschied war, desto geringer war die Satzidentifikationsleistung. Dies ist nachvollziehbar, da die Sprecher in der Satzidentifikationsaufgabe auf verschiedenen räumlichen Positionen präsentiert wurden, und deckt sich mit früheren Ergebnissen [z.B. 2]. Die Analyse zeigte außerdem einen signifikant positiven Zusammenhang zwischen der

basalen Kurzzeitgedächtnisleistung (in den Bedingungen ohne Störschall) und der Satzidentifikationsleistung im Störschall. Zusammenhänge zwischen Cocktail-Party Listening und der Arbeitsgedächtnisleistung wurden auch in verschiedenen früheren Studien gefunden, jedoch nicht konsistent [6].

Die Regressionskoeffizienten für die mittlere Hörschwelle des besseren Ohres, den Lärmempfindlichkeitscore und das Alter der Versuchsperson unterschieden sich nicht signifikant von 0.

Tabelle 1: Ergebnis der Regressionsanalyse mit der Satzidentifikationsleistung als Kriterium. $N = 61$

Prädiktor	β	p
Intercept	0.060	0.489
Alter	0.189	0.074
Irrelevant Sound Effect für visuelle Items	0.142	0.176
Irrelevant Sound Effect für auditive Items	-0.145	0.152
mittlere Hörschwelle des besseren Ohres	-0.041	0.652
HL-Asymmetrie	-0.303	0.002
IPD-Detektionsschwelle	-0.266	0.011
Kurzzeitgedächtnisleistung in Stille	0.507	<.001
Lärmempfindlichkeits-Score	-0.106	0.283

Zusammenfassend bestätigt unsere Studie den Einfluss von überschwelligen Hörleistungen und kognitiven Fähigkeiten auf das Sprachverstehen im Störschall. Die Störbarkeit des Kurzzeitgedächtnisses durch irrelevante Hintergrundsprache (Irrelevant Sound Effect) scheint hingegen keine wesentliche Rolle zu spielen.

Literatur

- [1] Bronkhorst, A. W.: The cocktail-party problem revisited: early processing and selection of multi-talker Sound. *Attention, Perception, & Psychophysics*, 77(5) (2015), 1465-1487
- [2] Oberfeld, D., & Klöckner-Nowotny, F.: Individual differences in selective attention predict Sound identification at a cocktail party. *eLife*, 5 (2016), e16747
- [3] Ruggles, D., & Shinn-Cunningham, B.: Spatial selective auditory attention in the presence of reverberant energy: individual differences in normal-hearing listeners. *Journal of the Association for Research in Otolaryngology*, 12(3) (2011), 395-405
- [4] Bronkhorst, A. W.: The cocktail party phenomenon: A review of research on Sound intelligibility in multiple-talker conditions. *Acta Acustica united with Acustica*, 86(1) (2000), 117-128
- [5] Akeroyd, M. A.: Are individual differences in sound reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, 47(2) (2008), 53-571
- [6] Füllgrabe, C., & Rosen, S. (2016). On the (un)importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers in Psychology*, 7.
- [7] Ellermeier, W., & Zimmer, K.: The psychoacoustics of the irrelevant sound effect. *Acoustical Science and Technology*, 35(1) (2014), 10-16
- [8] Ellermeier, W., & Zimmer, K.: Individual differences in susceptibility to the "irrelevant Sound effect". *The Journal of the Acoustical Society of America*, 102(4) (1997), 2191-2199
- [9] Békésy, G. V.: A new audiometer. *Acta Oto-Laryngologica*, 35(5-6) (1947), 411-422
- [10] Hopkins, K., & Moore, B. C.: Development of a fast method for measuring sensitivity to temporal fine structure information at low frequencies. *International journal of audiology*, 49(12) (2010), 940-946
- [11] Wagener, K., Kühnel, V., & Kollmeier, B.: Entwicklung und Evaluation eines Satztests für die deutsche Sprache, I: Design des Oldenburger Satztests (Development and evaluation of a German sentence test I: Design of the Oldenburg sentence test). *Zeitschrift für Audiologie*, 38(1) (1999), 4-15
- [12] Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., & Maess, B.: Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience*, 32(36) (2012), 12376-12383
- [13] Zimmer, K. & Ellermeier, W.: Ein Kurzfragebogen zur Erfassung der individuellen Lärmempfindlichkeit. *Umweltpsychologie*, 2 (1998), 54-63
- [14] Belsley, D.A., Kuh, E., Welsch, R.E.: *Regression diagnostics: Identifying influential data and sources of collinearity*. Wiley Series in Probability and Statistics. Hoboken, N.J: Wiley, 1980