

# **De invloed van songteksten en semantische betekenis op de maskeringseffectiviteit van achtergrondmuziek bij gesproken zinnen**

Noortje Akkermans - s1014317

15 juli 2020

BA Taalwetenschap

Radboud Universiteit Nijmegen

Begeleidster: Dr. Susanne Brouwer

Tweede beoordelaar: Dr. Louis ten Bosch

**Radboud Universiteit**



## Voorwoord

Deze scriptie is geschreven ter afsluiting van mijn bacheloropleiding Taalwetenschap aan de Radboud Universiteit Nijmegen. Ik begon in maart 2020 met de eerste besprekingen en het opzetten van het onderzoek, en ik ben tot halverwege juli 2020 bezig geweest met het uitvoeren van het onderzoek en het schrijven van het werkstuk. Het onderzoek is zeer interessant en relevant voor iedereen die in het dagelijks leven gesprekken voert in rumoerige omgevingen of in aanwezigheid van achtergrondgeluid, of geïnteresseerd is in de cognitieve verwerking van gesproken taal.

Tijdens dit project heb ik veel geleerd en heb ik veel inzichten opgedaan over wetenschappelijk onderzoek doen en waar mijn interesses liggen binnen de Taalwetenschap. De psycholinguïstische kant van de Taalwetenschap blijkt nu wel degelijk wat mij het meest trekt. Gedurende de hele opleiding, en gedurende het schrijfproces van deze scriptie, heb ik ook veel ervaringen opgedaan. In dit afsluitende werkstuk komen deze verschillende dingen samen.

Mijn begeleidster Susanne Brouwer heeft mij enorm geïnspireerd met haar ervaringen, kennis en enthousiasme voor haar werk en vakgebied. Zelfs tijdens de bijzondere situatie waardoor de universiteit tijdelijk volledig gesloten was en fysiek overleg niet kon plaatsvinden. Allerlei alternatieve werkwijzen werden geopperd en benut waardoor ik zonder problemen door kon werken en niet al te veel vertraging hoefde op te lopen. Ik wil haar daarvoor, en voor de fijne en zelfs gezellige samenwerking en begeleiding, ontzettend bedanken. Ook ben ik mijn familie en vrienden dankbaar voor hun enthousiasme en interesse, en Joep ook voor zijn hulp tijdens dit traject. Met de steun en inzichten van anderen en door mijn enthousiasme te kunnen delen wogen deze laatste loodjes toch een stuk minder zwaar. Ik kijk terug op een leuke en leerzame ervaring, en een goede afsluiting van een mooie studie.

Ik wens u veel leer- en leesplezier toe.

Noortje Akkermans

Nijmegen, 15 juli 2020

## Inhoudsopgave

Voorwoord .....	I
Samenvatting.....	1
Inleiding .....	1
<i>Maskering door songteksten</i> .....	3
<i>Semantische betekenis en maskeringseffectiviteit</i> .....	4
<i>Het huidige onderzoek</i> .....	7
<i>Hypotheses</i> .....	8
Methode .....	9
<i>Participanten</i> .....	9
<i>Materiaal</i> .....	9
<i>Procedure</i> .....	12
<i>Scoren</i> .....	12
<i>Data-analyse</i> .....	13
Resultaten.....	13
Discussie .....	14
<i>Aanwezigheid van songteksten</i> .....	14
<i>Semantische betekenis van songteksten</i> .....	15
<i>Invloed van SNR niveau</i> .....	16
<i>Discussiepunten</i> .....	17
<i>Suggesties voor toekomstig onderzoek</i> .....	18
<i>Conclusies</i> .....	19
Referenties .....	19
Bijlage A .....	23

## Samenvatting

Achtergrondgeluid is onderdeel van onze dagelijkse bezigheden. De aanwezigheid van achtergrondgeluid tijdens conversatie resulteert in maskering van een spraaksignaal waardoor de verwerking hiervan bemoeilijkt wordt. Voorgaand onderzoek heeft aangetoond dat verschillende factoren van achtergrondgeluid invloed hebben op de effectiviteit van de maskering. Volgens de *target-masker linguistic similarity hypothesis* (Brouwer et al., 2012) is gelijkenis tussen het doelsignaal en het achtergrondgeluid een belangrijke voorspeller van maskeringseffectiviteit. Het doel van het huidige onderzoek is tweeledig. Er is onderzocht wat de invloed is van (1) de aanwezigheid van songteksten in achtergrondmuziek, en (2) de semantische betekenis van die songteksten op de maskeringseffectiviteit van achtergrondmuziek op betekenisvolle gesproken zinnen. 36 moedertaalsprekers van het Nederlands hebben deelgenomen aan een zinsherkenningstaak, waarbij zij luisterden naar gesproken Nederlandse zinnen met drie varianten van hetzelfde nummer op de achtergrond op twee verschillende SNR niveaus. De varianten bevatten geen songteksten, semantisch betekenisvolle songteksten, of semantisch betekenisloze songteksten. Resultaten lieten een sterk effect zien van zowel aanwezigheid van songteksten als de semantische betekenis ervan. De maskeringseffectiviteit van achtergrondmuziek op betekenisvolle gesproken zinnen is het hoogst als de muziek semantisch betekenisvolle songteksten bevat. Bij semantisch betekenisloze songteksten is de maskeringseffectiviteit significant lager, maar nog steeds hoger dan bij afwezigheid van songteksten. Mogelijk speelt hierbij bekendheid met de songteksten een rol. Doordat de semantisch betekenisvolle songteksten bekender zijn bij de participanten, zijn deze mogelijk moeilijker te onderdrukken dan de semantisch betekenisloze songteksten. De bevindingen van het huidige onderzoek tonen aan dat zowel de aanwezigheid van songteksten als de semantische betekenis invloed hebben op de maskeringseffectiviteit van achtergrondmuziek bij gesproken zinnen. De resultaten ondersteunen de *target-masker linguistic similarity hypothesis*, en tonen aan dat deze hypothese ook opgaat voor songteksten in achtergrondmuziek.

## Inleiding

Mensen zijn al vanaf de geboorte in staat om gesproken taal te onderscheiden van andersoortige geluiden (Vouloumanos, Hauser, Werker, & Martin, 2010; Vouloumanos & Werker, 2007). Veel conversaties kunnen dan ook plaatsvinden in rumoerige omgevingen zonder dat daarbij al te veel communicatieproblemen ontstaan, zoals in restaurants of cafés waar vaak muziek wordt gedraaid. Toch zal elke luisteraar wel een situatie hebben meegemaakt waarin het bijzonder lastig was om een gesprekspartner te verstaan of om zichzelf verstaanbaar te maken. In het huidige onderzoek wordt ingegaan op specifieke factoren van achtergrondmuziek en hoe deze de verstaanbaarheid van gesproken taal beïnvloeden.

Achtergrondgeluid interfereert met de perceptie van gesproken taal. Dit achtergrondgeluid kan muziek, lawaai, ruis of geroezemoes van andere pratende mensen zijn. Het maskeert in meerdere of mindere mate het spraaksignaal waar de luisteraar zich op wil richten – dit zal verder het doelsignaal genoemd worden. De sterkte, of effectiviteit, van de maskering is van verschillende factoren afhankelijk, zoals bijvoorbeeld de luidheid van het achtergrondgeluid ten opzichte van die van het doelsignaal (bijv. Brouwer, 2017; Brouwer, Van Engen, Calandruccio, & Bradlow, 2012; Brungart, 2001; Brungart, Simpson, Ericson, & Scott, 2001; Cooke, Garcia Lecumberri, & Barker, 2008). Deze verhouding tussen de luidheid van deze twee verschillende signalen wordt ook wel de *signal-to-noise ratio* (SNR) genoemd. Ook een toename in het aantal sprekers op de achtergrond (Calandruccio, Buss, & Bowdrie, 2017; Simpson & Cooke, 2005; Van Engen & Bradlow, 2007), of een oplopende complexiteit van het achtergrondgeluid (Başkent, Van Engelshoven, & Galvin, 2014; Scharenborg & Larson, 2018a, 2018b) kan voor een effectievere maskering van het doelsignaal zorgen.

Er wordt onderscheid gemaakt tussen twee soorten spraakmaskering (Pollack, 1975). Enerzijds kan er energetische maskering optreden, wat ook wel maskering op een lager niveau wordt genoemd. Energetische maskering ontstaat als een achtergrondgeluid akoestische en temporele eigenschappen van het doelsignaal maskeert waardoor niet alle informatie van het doelsignaal beschikbaar is voor de luisteraar (Durlach et al., 2003a; Mattys, Brooks, & Cooke, 2009). Dit maakt dat de luisteraar het doelsignaal niet volledig kan waarnemen en verwerken. Alle typen achtergrondgeluid, of geluiden naast het doelsignaal, leveren per definitie in meerdere of mindere mate energetische maskering van het doelsignaal op.

Naast energetische maskering kan er ook informatiele maskering, of maskering op een hoger niveau, optreden. Informatiele maskering wordt ook wel cognitieve of linguïstieke maskering genoemd, en verwijst naar de resterende maskering die plaatsvindt nadat er rekening is gehouden met de energetische maskering. Deze vorm van maskering kan teweeggebracht worden door een grotere cognitieve belasting in het scheiden van twee geluidsstromen tijdens de verwerking (Brungart, 2001; Mattys et al., 2009). Dit kan ontstaan als zowel het doelsignaal als het achtergrondgeluid beide waarneembaar zijn voor de luisteraar, maar de luisteraar moeite heeft met het scheiden van de twee signalen of het isoleren van de informatie in het doelsignaal.

Er is al veel onderzoek gedaan naar de verschillende factoren die de effectiviteit van informatiele maskering beïnvloeden, en welke mechanismen luisteraars kunnen gebruiken om in de aanwezigheid van meerdere geluidssignalen het doelsignaal te isoleren. Zo blijkt bijvoorbeeld dat ruimtelijke informatie goed gebruikt kan worden om het doelsignaal te onderscheiden van interfererende geluidsstromen (bijv. Cherry, 1953; Freyman, Balakrishnan, & Helfer, 2001; Jones & Litovsky, 2008; Litovsky, 2012; Viswanathan, Kokkinakis, & Williams, 2018). Als de luisteraar weet uit welke richting het doelsignaal komt, kunnen signalen uit andere richtingen relatief goed onderdrukt worden. Verder is er al veel bewijs gevonden voor de zogenaamde *target-masker linguistic similarity hypothesis* (Brouwer et al., 2012), die stelt dat verwerking van het doelsignaal (*target*) moeilijker wordt door een toename aan informatiele maskering naarmate de interfererende geluidsstroom (*masker*) er meer op lijkt. Naarmate twee geluidsstromen meer op elkaar lijken, wordt het immers lastiger om deze uit elkaar te houden. Zo is er meer informatiele maskering als bijvoorbeeld twee concurrerende sprekers vergelijkbare stemeigenschappen of hetzelfde geslacht hebben (Brungart, 2001; Brungart et al., 2001; Leibold, Buss, & Calandruccio, 2018). Daarnaast is ook de taal waarin wordt gesproken een factor die informatiele maskering beïnvloedt, met effectievere maskering naarmate de taal (Brouwer & Bradlow, 2014; Brouwer et al., 2012; Calandruccio et al., 2017; Garcia Lecumberri & Cooke, 2006; Van Engen, 2010; Van Engen, 2012; Van Engen & Bradlow, 2007) of zelfs het accent in het doelsignaal en de interfererende spraakstroom dichter bij elkaar liggen (Brouwer, 2017, 2019; Calandruccio, Dhar, & Bradlow, 2010).

Een veelvoorkomende vorm van spraakmaskering is achtergrondmuziek, wat in veel openbare ruimtes zoals restaurants, cafés, en winkels aanwezig is. Bij de aanwezigheid van achtergrondmuziek vindt er per definitie energetische maskering plaats. De muziek overlapt met het doelsignaal waardoor dit minder verstaanbaar wordt. Achtergrondmuziek kan daarnaast ook informatieel maskeren. Deze informatiele maskering kan ontstaan door overeenkomsten in spectro-temporele of structurele eigenschappen tussen spraak en muziek (Başkent et al., 2014; Brungart, 2001; Russo & Pichora-Fuller, 2008; Shi & Law, 2010). Bovendien is in muziek ook vaak zang aanwezig, wat voor nog een extra laag aan zowel energetische als informatiele maskering kan zorgen (Başkent et al., 2014; Scharenborg & Larson, 2018a, 2018b). Met de aanwezigheid van zang is ook de gelijkenis tussen het doelsignaal en de achtergrondmuziek groter. Doordat songteksten een talige eigenschap aan de achtergrondmuziek toevoegen kan het scheiden van de twee signalen verder bemoeilijkt worden.

Ook bekendheid met het doelsignaal en/of het achtergrondgeluid is een belangrijke factor die de effectiviteit van informationele maskering kan beïnvloeden (bijv. Tun, O’Kane, & Wingfield, 2002; Van Engen, 2010). Als de luisteraar bekend is met de taal in het achtergrondgeluid, bijvoorbeeld wanneer dit zijn/haar moedertaal is, is het moeilijker om dit te onderdrukken en derhalve moeilijker om te richten op het doelsignaal (zie ook Bialystok, Craik, & Ryan, 2006; Emmorey, Luk, Pyers, & Bialystok, 2008; Jiang, 2015). Moedertaalsprekers van het Engels ervaren minder maskering van interfererende spraakstromen in een tweede of onbekende taal ten opzichte van hun moedertaal (Calandruccio et al., 2010; Garcia Lecumberri & Cooke, 2006; Van Engen, 2010; Van Engen & Bradlow, 2007). Dit zou gedeeltelijk verklaard kunnen worden door akoestische of fonetische verschillen tussen talen, waardoor het scheiden van de twee signalen makkelijker wordt. Wat echter opvallend is, is dat dit effect ook wordt gevonden voor twee talen die op dit gebied redelijk dicht bij elkaar liggen (Nederlands en Engels; Brouwer et al., 2012). Dit suggereert dat alleen bekendheid met fonetische eigenschappen van een taal niet voldoende is om bekendheidseffecten te verklaren. Mogelijk speelt ook de taalvaardigheid een rol. Voor luisteraars die volledig onbekend zijn met de taal zal de semantische inhoud van de achtergrondspraak betekenisloos zijn, in tegenstelling tot die van het doelsignaal. Hierdoor zal de semantische betekenis van de achtergrondspraak niet interfereren met het doelsignaal en zal er minder informationele maskering plaatsvinden.

Resultaten uit voorgaand onderzoek lijken te suggereren dat de *target-masker linguistic similarity hypothesis* ook opgaat voor semantische betekenis van achtergrondspraak (Brouwer et al., 2012, maar zie Calandruccio, Buss, Bencheck, & Jett, 2018; Tun et al., 2002). Als zowel het doelsignaal als het achtergrondgeluid semantisch betekenisvol is, en deze twee signalen op dit gebied dus een grote gelijkheid laten zien, is de maskeringseffectiviteit van het achtergrondgeluid hoger. Het is de vraag in hoeverre dit gegeneraliseerd kan worden naar songteksten in achtergrondmuziek. Door de talige eigenschap en semantische inhoud van songteksten zou semantische betekenis in achtergrondmuziek, net als bij achtergrondspraak, invloed kunnen hebben op de maskeringseffectiviteit. De huidige studie zal hier verder op ingaan om meer inzicht te verkrijgen in (1) het effect van songteksten in achtergrondmuziek, en (2) de invloed van semantische betekenis van deze songteksten, op de verwerking van het doelsignaal.

### *Maskering door songteksten*

Onderzoek naar de invloed van achtergrondmuziek op bijvoorbeeld aandacht wijst uit dat de aanwezigheid van songteksten een nadelig effect heeft op het concentratievermogen (bijv. Abril & Flowers, 2007; Shih, Huang, & Chiang, 2012). In het onderzoek van Shih en collega’s voerden participanten een visuele aandachtstaak uit met muziek op de achtergrond. Participanten waarbij de muziek geen songteksten bevatte (instrumentele muziek) presteerden beter dan participanten die achtergrondmuziek met songteksten hoorden. Daarnaast verschilden de prestaties bij instrumentele achtergrondmuziek ook minder van de prestaties in een baselineconditie (geen achtergrondmuziek). Het verschil tussen de baseline en de songtekst conditie was groter. Uit deze studie blijkt dat achtergrondmuziek een nadelig effect heeft op concentratievermogen, maar dat de aanwezigheid van songteksten daarbij een grote rol speelt.

Er zijn slechts enkele studies uitgevoerd die hebben gekeken naar de maskeringseffectiviteit op spraak van muziek met en zonder songteksten (Başkent et al., 2014; Russo & Pichora-Fuller, 2008; Scharenborg & Larson, 2018a, 2018b). Uit het onderzoek van Russo en Pichora-Fuller (2008) blijkt dat instrumentele achtergrondmuziek net als een interfererende spraakstroom informationele maskering kan veroorzaken, maar dat de laatste een hogere maskeringseffectiviteit heeft. Dit is waarschijnlijk te verklaren door een grotere gelijkheid tussen het doelsignaal en het achtergrondgeluid, waardoor de twee signalen

moeilijker van elkaar te scheiden zijn voor de luisteraar. Daarnaast kan ook meespelen dat de interfererende spraakstroom meer informatie bevat, waardoor het meer kan afleiden van het doelsignaal dan de instrumentele achtergrondmuziek. Dit impliceert dat talige informatie een belangrijke rol speelt in de maskeringseffectiviteit van achtergrondgeluid. Veel muziek bevat songteksten, wat deze talige informatie met zich meebrengt. Dit zou kunnen betekenen dat achtergrondmuziek met songteksten meer maskering teweegbrengt dan achtergrondmuziek zonder songteksten.

Het effect van de aanwezigheid van songteksten is onderzocht door Scharenborg en Larson (2018a, 2018b). In hun beide studies werd onderzocht hoe goed participanten gesproken woorden konden verstaan met verschillende soorten achtergrondmuziek. Zij vonden dat muziek met songteksten een effectievere maskering teweegbracht dan instrumentele muziek. De resultaten van Scharenborg en Larson (2018b) sluiten echter niet geheel aan bij die uit een onderzoek van Başkent et al. (2014), waarin geen duidelijk effect van aanwezigheid van songteksten gevonden werd. In dit onderzoek bleek complexiteit (tempo en aantal instrumenten) van het muziekstuk een sterkere voorspeller te zijn van de maskeringseffectiviteit dan de aanwezigheid van talige informatie uit songteksten. Er werd in Başkent et al. bijvoorbeeld geen verschil in prestaties gevonden tussen een laag-complex instrumenteel muziekstuk en een laag-complex muziekstuk met songteksten. Ook werd het doelsignaal gemiddeld genomen veel minder accuraat verstaan met hoog-complexe muziekstukken op de achtergrond dan wanneer de achtergrondmuziek een laag-complexe muziekstuk was. Scharenborg en Larson vonden ook een effect van complexiteit, maar in hun onderzoek had de aanwezigheid van songteksten juist een groter effect op de maskeringseffectiviteit van de achtergrondmuziek. In Başkent et al. bleek echter wel dat van de twee complexe muziekstukken degene met songteksten een veel hogere maskeringseffectiviteit had dan het instrumentele stuk, hoewel dit mogelijk ook te maken heeft met de bekendheid van de muziekstukken. Het complexe stuk met songteksten gebruikt in Başkent et al. (Dromen Zijn Bedrog van Marco Borsato) was zeer bekend bij de participanten, in tegenstelling tot de andere muziekstukken in die studie. Dit maakt dat de muziek in deze conditie mogelijk meer afleidt en de verwerking van het doelsignaal verder bemoeilijkt (vgl. het effect van bekendheid met achtergrondspraak; bijv. Brouwer et al., 2012; Calandruccio et al., 2010; Garcia Lecumberri & Cooke, 2006; Van Engen, 2010; Van Engen & Bradlow, 2007). De muziek in beide studies van Scharenborg en Larson was veel minder tot niet bekend bij de luisteraars. De effecten van aanwezigheid van songteksten die Scharenborg en Larson vonden zijn niet beïnvloed door bekendheidseffecten.

Samengevat, in zowel Başkent et al. (2014) als de onderzoeken van Scharenborg en Larson (2018a, 2018b) bleek het effect van songteksten op maskeringseffectiviteit van achtergrondmuziek groter of zelfs uniek voor moeilijkere luisteromstandigheden. In het onderzoek van Başkent et al. bleek alleen een effect van songteksten in de complexe (en bekendere) achtergrondmuziek, en in de onderzoeken van Scharenborg en Larson bleek het songteksteffect sterker bij moeilijkere SNRs, waarin over het algemeen ook minder goed gepresteerd werd. De aanwezigheid van songteksten in achtergrondmuziek lijkt dus wel degelijk een hogere maskeringseffectiviteit teweeg te brengen, vooral in moeilijkere luisteromstandigheden.

### *Semantische betekenis en maskeringseffectiviteit*

Naar de invloed van semantische betekenis in songteksten op de maskeringseffectiviteit van achtergrondmuziek is voor zover bekend nog geen onderzoek gedaan. Een studie van De Groot en Smedinga (2014) vond wel een korte-termijn effect van bekendheid met de taal in songteksten op het concentratie- en leervermogen van participanten. Dit zou kunnen betekenen dat bekendheid en semantische betekenis van songteksten ook een rol spelen in de

maskeringseffectiviteit van achtergrondmuziek, maar dit is nog niet zo specifiek onderzocht. Er is echter wel een aantal studies waarin is gekeken naar semantische betekenis in achtergrondspraak en de invloed op de verwerking van het doelsignaal (Brouwer et al., 2012; Calandruccio et al., 2018; Tun et al., 2002). In het onderzoek van Tun et al. is bijvoorbeeld gekeken naar hoe goed luisteraars het doelsignaal konden verstaan in verschillende condities. Zij kregen in het doelsignaal zowel betekenisvolle als betekenisloze zinnen te horen<sup>1</sup>, en ook de interfererende spraakstroom kon zowel betekenisvol als betekenisloos (bekende taal met zinnen in willekeurige woordvolgorde, of een voor de luisteraar volledig onbekende taal) zijn. Uit deze studie bleek dat voor jonge luisteraars de semantische betekenis van achtergrondspraak geen invloed had op hun accuratesse in het herkennen van zinnen in het doelsignaal. Prestaties bij betekenisloze en betekenisvolle achtergrondspraak waren gelijk.

Het is echter belangrijk om te noemen dat in dit onderzoek voor elk item de stem van de spreker in het doelsignaal eerder aanving dan de stem van het interfererende spreker. Dit heeft er mogelijk voor gezorgd dat luisteraars zich volledig konden richten op het doelsignaal voordat de interfererende spraakstroom hier invloed op ging uitoefenen (zie ook Brouwer & Bradlow, 2015). Ook kregen deze participanten in een van de condities het doelsignaal zonder enig achtergrondgeluid te horen, waardoor zij deze ook in de andere condities mogelijk beter konden herkennen. Vorig onderzoek toont aan dat bekendheid met stemeigenschappen het herkennen van gesproken taal faciliteert doordat de luisteraar zich aan deze eigenschappen kan aanpassen (bijv. Bradlow, Nygaard, & Pisoni, 1999; Brungart et al., 2001; Kleinschmidt & Jaeger, 2015; Nygaard, Sommers, & Pisoni, 1994; Palmeri, Goldinger, & Pisoni, 1993). Door de hoge bekendheid met het doelsignaal hebben de verschillende typen achtergrondgeluid mogelijk minder effect kunnen uitoefenen op de verwerking ervan.

Daarnaast is in deze studie ook niet gekeken naar de rol van mate van gelijkheid tussen de twee spraakstromen. Volgens de *target-masker linguistic similarity hypothesis* is de gelijkheid tussen het doelsignaal en het achtergrondgeluid een belangrijke voorspeller van hoe succesvol het doelsignaal verstaan wordt (Brouwer et al., 2012). Deze hypothese stelt dat zowel taal-gerelateerde, stimulus-gerelateerde en luisteraar-gerelateerde factoren invloed hebben op hoe groot die gelijkheid is. Talen kunnen in meerdere of mindere mate op elkaar lijken. Een grotere gelijkheid tussen de taal in het doelsignaal en de taal in de interfererende spraakstroom zal leiden tot grotere moeilijkheden in het scheiden van de twee spraakstromen en meer informationele maskering. De luisteraar-gerelateerde factoren hebben betrekking op bijvoorbeeld de kennis en ervaring van de luisteraar met de taal die gesproken wordt in ofwel het doelsignaal, ofwel de interfererende spraakstroom. Dit is nauw verbonden met voorspelbaarheid van de spraakstroom, waarvoor over het algemeen is gevonden dat een voorspelbaar of herkenbaar signaal makkelijker te volgen (doelsignaal) en moeilijker te onderdrukken (achtergrondgeluid) is dan een onbekend signaal (bijv. Calandruccio et al., 2017; Russo & Pichora-Fuller, 2008; Van Engen & Bradlow, 2007). Met stimulus-gerelateerde factoren worden bijvoorbeeld fonetische eigenschappen of semantische inhoud van de spraakstromen bedoeld. Volgens deze hypothese is er meer informationele maskering naar mate de verwerking voor de twee spraakstromen dichter bij elkaar ligt of van hetzelfde niveau is. Dit zou bijvoorbeeld ook het geval kunnen zijn als beide spraakstromen voor de luisteraar semantisch betekenisvol zijn.

In het onderzoek van Brouwer et al. (2012) is gekeken naar de verwerking van het doelsignaal met verschillende soorten interfererende spraakstromen, en de invloed van talige gelijkheid tussen de twee spraakstromen. Deze interfererende spraakstromen waren semantisch betekenisvolle zinnen (bijv. *Rice is often served in round bowls*) of semantisch betekenisloze

---

<sup>1</sup> Een voorbeeld van een betekenisvolle doelzin uit Tun et al. (2002) is “*More attention has been paid to diet but mostly in relation to disease and to the growth of young children.*” en een voorbeeld van een betekenisloze doelzin is “*Window to down hold to boy appearance think with again room the beat go in there beside some is was*”.



zinnen (bijv. *The great car met the milk*), in het Engels of Nederlands. In lijn met de *target-masker linguistic similarity hypothesis* werd er een negatieve invloed gevonden van semantische gelijkenis tussen het doelsignaal en de interfererende spraakstroom op de herkenning van zinnen in het doelsignaal.

Tegenstrijdige resultaten zijn gevonden in een recent onderzoek van Calandruccio et al. (2018). Hierin is opnieuw gekeken naar de invloed van semantische betekenis op verwerking van het doelsignaal. Daarbij is extra gecontroleerd op syntactische structuur en aantal syllaben per woord in de twee spraakstromen. Er zijn in dit onderzoek geen effecten gevonden van semantische betekenis die aansluiten bij de resultaten in Brouwer et al. (2012). Verschillen tussen condities lijken eerder te verklaren door de mate van variatie in syntactische structuur, waarbij veel variatie voor hogere stimulus onzekerheid zorgt en zo meer informationele maskering teweeg kan brengen (zie ook Brouwer & Bradlow, 2014; Durlach et al., 2003a, 2003b; Russo & Pichora-Fuller, 2008).

Wat echter belangrijk is om te noemen, is dat in het onderzoek van Calandruccio et al. (2018) iets makkelijkere SNR niveaus zijn gebruikt. Betekenisvolle doelzinnen werden gepresenteerd op SNR niveaus van -1 dB en -3 dB, waar dit in het onderzoek van Brouwer et al. (2012) voor een vergelijkbare populatie (volwassen moedertaalsprekers van het Engels) de SNR niveaus van -3 dB en -5 dB waren. Uit voorgaand onderzoek blijkt dat SNR niveau een belangrijke rol speelt. Voor bepaalde typen maskering wordt alleen een effect gevonden bij moeilijkere SNRs (Scharenborg & Larson, 2018a, 2018b; Van Engen & Bradlow, 2007) of is het effect sterker bij moeilijkere SNRs dan bij makkelijkere SNRs (Gittleman & Van Engen, 2018; Russo & Pichora-Fuller, 2008; Scharenborg & Larson, 2018b; Van Engen, 2010). Dit zou kunnen verklaren waarom het onderzoek van Calandruccio et al., in tegenstelling tot dat van Brouwer et al., geen effect van semantische betekenis vond.

Naast gelijkenis tussen het doelsignaal en de achtergrondspraak bleek in het onderzoek van Brouwer et al. (2012) ook bekendheid met het achtergrondgeluid een belangrijke rol te spelen. Voor Engelse luisteraars bleek semantische betekenis van achtergrondspraak in een onbekende taal (Nederlands) geen invloed te hebben. Betekenisvolle Nederlandse spraak en betekenisloze Nederlandse spraak zijn voor Engelse luisteraars zonder enige kennis van het Nederlands beide semantisch betekenisloos. Nederlandse luisteraars, die ook bekend waren met het Engels, lieten wel een effect zien van semantische betekenis voor beide achtergrondtalen. Alleen een gelijkenis tussen het doelsignaal en de achtergrondspraak in semantische betekenis is niet genoeg om een effect teweeg te brengen. De luisteraar moet bekend genoeg zijn met de taal om een verschil tussen betekenisvolle en betekenisloze achtergrondspraak te horen.

Een bekendheidseffect is ook in ander voorgaand onderzoek naar voren gekomen. Een onbekende taal in het achtergrondgeluid (ook als het alleen onbekend klinkt doordat het achterstevoren wordt afgespeeld; Calandruccio et al., 2017) brengt een minder effectieve maskering van het doelsignaal teweeg (Calandruccio et al., 2010; Garcia Lecumberri & Cooke, 2006; Van Engen, 2010; Van Engen & Bradlow, 2007). Dit kan, naast een lagere gelijkenis tussen het doelsignaal en het achtergrondgeluid, mogelijk verklaard worden door het feit dat een onbekende taal voor de luisteraar semantisch betekenisloos is – in tegenstelling tot de bekende taal in het doelsignaal. Als beide signalen op semantisch niveau verwerkt moeten worden, resulteert dit in een zwaardere cognitieve belasting en meer moeite om de twee signalen te scheiden. Hierdoor zal de taal in het doelsignaal minder goed herkend worden.

Er zijn tot nu toe tegenstrijdige bevindingen voor de invloed van semantische betekenis van achtergrondspraak, maar voor songteksten tonen De Groot en Smedinga (2014) wel een effect van bekendheid aan. Dit impliceert dat luisteraars gevoelig zijn voor semantische betekenis en bekendheid met de taal in songteksten. Deze resultaten en het feit dat songteksten net als gesproken taal semantische inhoud heeft, suggereren dat de semantische betekenis en

bekendheid van songteksten net als gesproken taal invloed kunnen hebben op herkenning van het doelsignaal.

### *Het huidige onderzoek*

In dit onderzoek wordt de invloed van (1) songteksten in achtergrondmuziek en (2) de semantische betekenis daarvan op het herkennen van het doelsignaal onderzocht. Hiervoor zijn de volgende twee onderzoeksvragen opgesteld:

1. Is er een verschil in maskering op gesproken zinnen door achtergrondmuziek met songteksten ten opzichte van achtergrondmuziek zonder songteksten?
2. Heeft semantische betekenis van songteksten invloed op de maskeringseffectiviteit van achtergrondmuziek bij gesproken zinnen?

Allereerst wordt er specifiek gekeken naar het effect van aanwezigheid van songteksten. Het onderzoek door Scharenborg en Larson (2018a, 2018b) en Başkent et al. (2014) heeft hier al inzichten in gegeven, maar in het huidige onderzoek wordt gecontroleerd voor eventuele additionele invloeden. In Başkent et al. zijn voor de verschillende muziekcondities verschillende muziekstukken gebruikt. Verschillen tussen de muziekcondities kunnen derhalve naast de complexiteit en de songtekst-eigenschap ook teweeggebracht zijn door algehele verschillen in bijvoorbeeld genre, melodie, instrumenten, zangstem, timbre, toonhoogte of temporele eigenschappen. Door Scharenborg en Larson (2018a) is dit al gedeeltelijk ondervangen door de achtergrondmuziek voor de condities met songteksten en de condities zonder songteksten uit dezelfde muziekstukken te halen. In hun volgende onderzoek (Scharenborg & Larson, 2018b) is hier nog beter voor gecontroleerd door de twee condities te maken met precies hetzelfde nummer; de originele versie voor de conditie met songtekst en de instrumentale versie voor de conditie zonder songtekst. Deze laatste methode wordt ook in het huidige onderzoek toegepast, aangezien hiermee volledig gecontroleerd kan worden voor andere factoren in achtergrondmuziek. Bovendien wordt in het huidige onderzoek gekeken naar herkenning van gesproken zinnen, in plaats van herkenning van gesproken losse woorden zoals in de vorige onderzoeken. Dit is een natuurlijkere luistersituatie, aangezien luisteraars doorgaans gesprekken voeren waarin zij volledige zinnen moeten verstaan in plaats van losse woorden.

Daarnaast wordt ook gekeken naar de semantische betekenis van de songteksten. Voorgaand onderzoek heeft al wat licht geworpen op de invloed van semantische betekenis in van achtergrondspraak mogelijk speelt op de herkenning van een doelsignaal. Zoals eerder genoemd is het echter nog niet duidelijk of deze bevindingen gegeneraliseerd kunnen worden naar songteksten in achtergrondmuziek. Het doel van het huidige onderzoek is om hier meer inzicht in te krijgen. Daarbij wordt in dit onderzoek ook rekening gehouden met bekendheidseffecten. Een bekende taal in achtergrondspraak brengt een effectievere maskering van het doelsignaal teweeg (Brouwer et al., 2012; Calandruccio et al., 2010, 2017; Garcia Lecumberri & Cooke, 2006; Van Engen, 2010; Van Engen & Bradlow, 2007). Resultaten van De Groot en Smedinga (2014) en Scharenborg en Larson (2018a) lijken ook voor songteksten in deze richting te wijzen. Scharenborg en Larson vonden een iets minder effectieve maskering van het doelsignaal bij achtergrondmuziek met songteksten in een bij de participanten minder bekende variant van het Engels, namelijk een West-Afrikaanse variëteit. Ook kan een taal alleen maar semantisch betekenisvol zijn voor de luisteraar als deze bekend is met de taal. Door in de achtergrondmuziek een onbekende taal te gebruiken die semantisch betekenisloos is voor de luisteraar kan in het huidige onderzoek gekeken worden naar de invloed van semantische betekenis van songteksten op herkenning van het doelsignaal.

In het huidige onderzoek hebben moedertaalsprekers van het Nederlands deelgenomen aan een zinsherkenningstaak. Er is gekeken naar hoe accuraat de luisteraars gesproken zinnen herkennen in drie verschillende typen achtergrondmuziek. Dit is op twee SNR niveaus onderzocht, aangezien vorig onderzoek heeft aangetoond dat het SNR niveau kan interacteren met het effect van een type maskering (bijv. Russo & Pichora-Fuller, 2008; Scharenborg & Larson, 2018a, 2018b; Van Engen, 2010; Van Engen & Bradlow, 2007). Er zijn drie varianten van het nummer Hot N Cold van Katy Perry gebruikt. De drie varianten verschillen enkel op het gebied van songteksten en zijn door dezelfde zangeres ingezongen. Deze varianten zijn (1) de instrumentele versie zonder songteksten, (2) de originele versie met Engelstalige songteksten, en (3) een versie met semantisch betekenisloze songteksten. Deze laatste variant is een versie waarin de songteksten zijn vertaald naar en gezongen in Simlish. Simlish is een fictieve taal uit het levenssimulatie spel De Sims. Voor dit computerspel werd gestreefd naar een unieke taal die universeel gebruikt kon worden en spelers de ruimte gaf om hun eigen fantasie en improvisatie te gebruiken (Barnes, 2020; Boland, 2010, pp. 17-22). Op een paar geregistreerde woorden na (Great Big Story, 2019) kent Simlish geen vaste semantische betekenissen.<sup>2</sup> Verschillende populaire nummers zijn vertaald naar en opgenomen in Simlish. In deze Simlish versies lijkt de zang op het origineel (The Sims, 2009) maar hebben de songteksten geen semantische betekenis. Het enige dat een herleidbare vorm van betekenis kan overbrengen in Simlish is de fonologie of expressie in de productie (Great Big Story, 2019; Portnow, 2011, p.147-148). Voorgaand onderzoek gebruikte bestaande, betekenisvolle woorden in (voor de luisteraar) onbekende talen of betekenisloze woordvolgordes voor de betekenisloze conditie. Door in het huidige onderzoek Simlish te gebruiken kan semantisch betekenisloze taal vergeleken worden met semantisch betekenisvolle taal, zonder de kans op herkenning van losse, betekenisvolle woorden in de betekenisloze conditie of grote fonetische verschillen tussen de condities.

### *Hypotheses*

Op basis van resultaten van Scharenborg en Larson (2018a, 2018b) en Başkent et al. (2014) wordt verwacht dat achtergrondmuziek met songteksten een effectievere maskering van het doelsignaal is dan achtergrondmuziek zonder songteksten. De songteksten zullen een extra laag aan maskering toevoegen door de extra aanwezigheid van een zangstem en talige informatie in de teksten. Dit effect zal naar verwachting ook groter zijn in de moeilijkere SNR conditie.

Over de semantisch betekenis van de songteksten lijkt voorgaand onderzoek te suggereren dat een hoge gelijkheid tussen het doelsignaal en het achtergrondgeluid resulteert in effectievere maskering van het doelsignaal. Songteksten die semantisch betekenisloos zijn zullen naar verwachting een minder effectieve maskering teweegbrengen, omdat deze niet, net als het doelsignaal, op semantisch niveau verwerkt hoeven worden. Daarbij zal de semantisch betekenisloze variant van de songteksten naar verwachting minder bekend zijn bij de luisteraars dan de originele, betekenisvolle songteksten, waardoor deze een minder effectieve maskering van het doelsignaal teweeg zal brengen.

---

<sup>2</sup> Dit gegeven en het feit dat Simlish ook andere eigenschappen (bepaalde *constraints*) van natuurlijke taal mist (Portnow, 2011, p.147-148) maakt dat het niet zo zeer als een echte taal bestempeld kan worden (zie ook Johansson (2005, pp. 5-11) voor een overzicht van verschillende definities en vereisten van taal). Ten behoeve van het leesgemak wordt hier wel gebruik gemaakt van deze term.

## Methode

### *Participanten*

Aan dit onderzoek hebben 36 moedertaalsprekers van het Nederlands (22 vrouwen, 14 mannen) van gemiddeld 23;2 jaar oud vrijwillig deelgenomen. Geen van de participanten had naar eigen zeggen last van gehoorproblemen. Responses van nog 24 extra deelnemers zijn niet meegenomen in de analyses vanwege niet voltooiën van het volledige experiment (18), gehoorproblemen (1), geen beschikking over een koptelefoon en rustige luisteromgeving (3), of een leeftijd van boven de 35 jaar (2). Er is bewust gekozen om geen participanten ouder dan 35 jaar mee te nemen in de analyses, om zo eventuele leeftijdseffecten te voorkomen (bijv. Başkent et al., 2014; Russo & Pichora-Fuller, 2008; Tun et al., 2002).

Participanten werd gevraagd hun taalvaardigheidsniveau in het Engels op vier verschillende deelvaardigheden te waarderen op een schaal van 1 (beginner) tot 5 (moedertaalniveau). Gemiddeld werd een score van 4.3 ( $SD = .676$ ) gegeven voor luistervaardigheid in het Engels, en geen van de participanten rapporteerde hiervoor een score lager dan een 3 (gemiddeld). Ook voor leesvaardigheid ( $M = 4.3$ ;  $SD = .692$ ), spreekvaardigheid ( $M = 3.8$ ;  $SD = .710$ ) en schrijfvaardigheid ( $M = 4.0$ ;  $SD = .654$ ) in het Engels werden hoge scores gerapporteerd. Op basis hiervan kan worden aangenomen dat voor deze participanten de Engelse songteksten ook daadwerkelijk als betekenisvol ervaren werden.

Uit een afsluitende vragenlijst bleek verder dat een overgrote meerderheid (30 van de 36 participanten) Simlish niet had herkend in de achtergrondmuziek. Dit suggereert dat Simlish en de Simlish variant van de songteksten naast semantisch betekenisloos ook niet bekend zijn bij de participanten. Participanten waren wel bekend met de muziek zelf. Slechts 2 participanten gaven aan het nummer niet te kennen of herkend te hebben. Van de overige participanten, die aangaven het nummer wel herkend te hebben, kon de overgrote meerderheid (30 van de 34) zowel de naam van het nummer als de artiest correct benoemen.

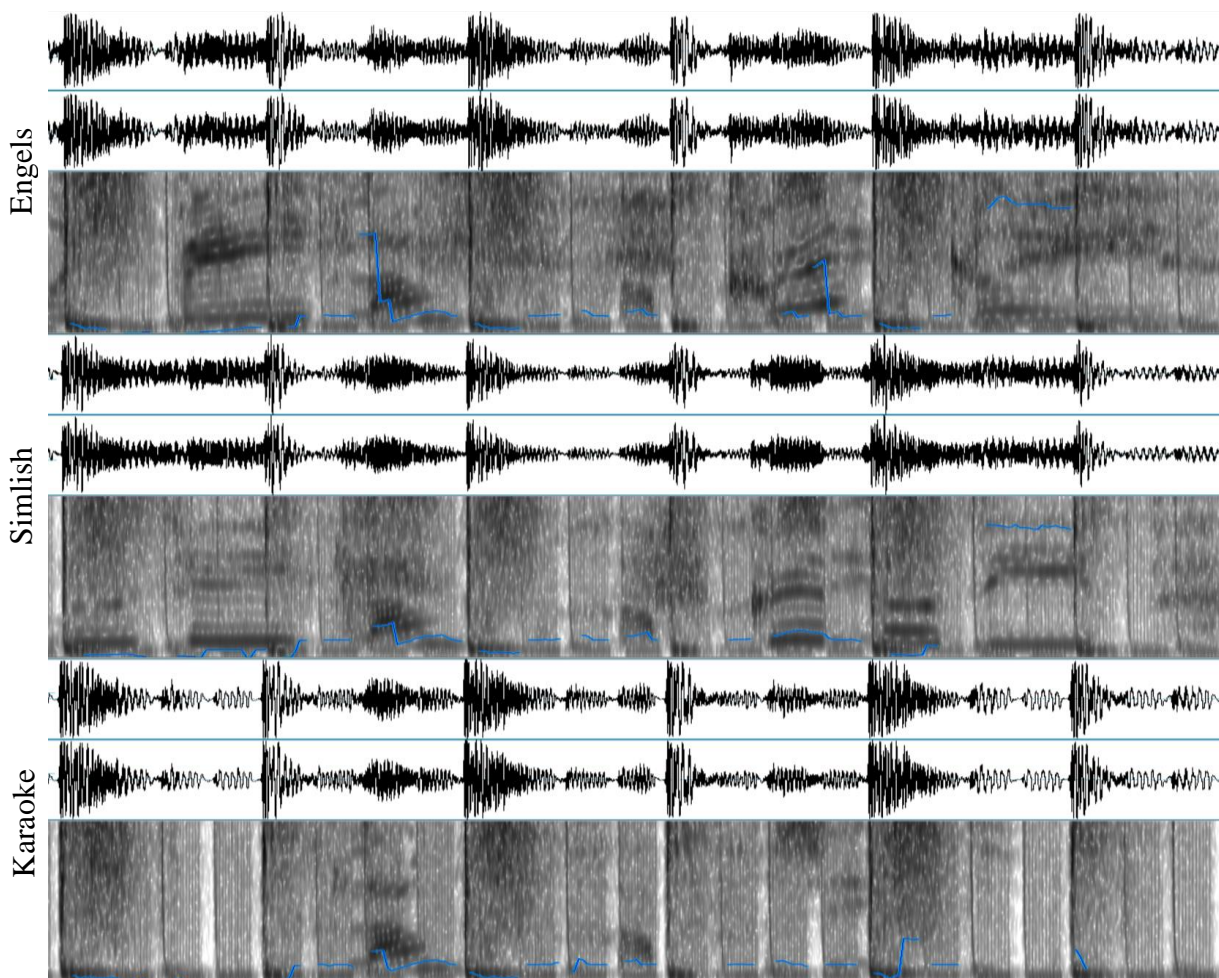
### *Materiaal*

Voor de doelzinnen is een selectie van zinnen uit de lijsten 1, 7, 8 en 9 van de herziene Bamford-Kowal-Bench test (BKB; Bamford & Wilson, 1979; Bench, Kowal, & Bamford, 1979) gebruikt. De originele zinnen zijn vertaald naar het Nederlands door een moedertaalspreker en gecontroleerd door een andere moedertaalspreker, beiden met C2 vaardigheidsniveau in het Engels. Elke lijst bestaat uit 16 semantisch betekenisvolle zinnen met drie of vier sleutelwoorden per zin. Voor dit experiment zijn per lijst 15 zinnen geselecteerd, door uit elke lijst één van de zinnen met vier sleutelwoorden weg te laten. Een voorbeeld van een van de originele zinnen is “The CLOWN had a FUNNY FACE” (de sleutelwoorden zijn aangegeven met hoofdletters). De bijbehorende Nederlandse vertaling is “De CLOWN had een GRAPPIG GEZICHT”. De vertaalde Nederlandse zinnen zijn ingesproken door een vrouwelijke moedertaalspreker van het Nederlands (identiek aan de spreekster uit Brouwer et al., 2012) in een geluidsdichte cabine (22050 Hz, 24 bit). De spreekster las de zinnen op van papier en werd gevraagd om in gesprekstijl te spreken.

Om zowel naar de invloed van songteksten als naar de invloed van semantische betekenis te kunnen kijken zijn drie muziekcondities gemaakt. Voor de achtergrondmuziek in deze drie condities zijn drie varianten van het nummer Hot N Cold van de Amerikaanse zangeres Katy Perry gebruikt. Door voor alle condities hetzelfde nummer te gebruiken wordt er gecontroleerd voor andere muzikale factoren die invloed kunnen uitoefenen op maskeringseffectiviteit, zoals bekendheid met de muziek (Russo & Pichora-Fuller, 2008; Scharenborg & Larson, 2018b). Allereerst is de originele versie gebruikt met Engelse,

semantisch betekenisvolle songteksten gezongen door Katy Perry. Voor de semantisch betekenisloze songteksten is een variant gebruikt waarin Katy Perry de songteksten in Simlish zingt. Zoals genoemd in de inleiding, is Simlish een semantisch betekenisloze fantasietaal uit het levenssimulatiespel De Sims. De Simlish variant is afgezien van de semantiek van de songteksten identiek aan het origineel. Een voorbeeld van een Simlish variant van een zin uit de originele songteksten is “*Vous chikanip laka gurl chika claps*”,<sup>3</sup> de vertaling van “*You change your mind like a girl changes clothes*”. Zie Bijlage A voor de volledige songteksten van de twee varianten naast elkaar. Tot slot is de karaoke- of instrumentale versie van het nummer gebruikt, waarin de zangstem volledig is weggehaald. Deze versie is afgezien van de zang en songteksten ook identiek aan het origineel. Deze drie varianten van de achtergrondmuziek resulteerden in de drie condities Engels, Simlish en Karaoke respectievelijk. De nummers zijn ieder verkregen van YouTube (respectievelijk Dslite804, 2008; The Sims, 2008; Affandi, 2009) als .wav files.

De *long-term average spectra* van de muziekbestanden zijn genormaliseerd met behulp van Praat (Boersma & Weenink, 2020) om ongelijkheid in energetische maskering tussen de muziekcondities te minimaliseren. In Figuur 1 zijn de golfvormen en spectrogrammen van de doelzin “De CLOWN had een GRAPPIG GEZICHT” in de drie muziekcondities te zien. De drie condities verschillen amper van elkaar, maar een verschil tussen de conditie zonder en de condities met songteksten is wel zichtbaar.



**Figuur 1** – Golfvormen en spectrogrammen van de doelzin “De clown had een grappig gezicht” in alle drie de muziekcondities. Condities van boven naar beneden: Engels, Simlish, Karaoke.

<sup>3</sup> Deze uitgeschreven songtekst is verkregen van [https://sims.fandom.com/wiki/Songs\\_in\\_Simlish](https://sims.fandom.com/wiki/Songs_in_Simlish) op 26 juni 2020.

De doelzinnen werden aangeboden op 60 dB. De achtergrondmuziek is op twee verschillende geluidsniveaus aangeboden, namelijk op 72 dB en 75 dB, resulterend in respectievelijk een makkelijkere SNR van -12 dB en een moeilijker SNR van -15 dB. Pilots wezen uit dat moeilijker SNRs van -15 dB en -17 dB resulteerden in vloereffecten voor alle condities. Deze effecten werden niet meer teruggevonden bij SNRs van -12 dB en -15 dB. De verschillende SNRs zijn in twee aparte blokken aangeboden. Elke participant kreeg eerst het blok met 30 items in de makkelijkere SNR en daarna het blok met 30 items in de moeilijker SNR. In elk blok werden alle drie de muziekcondities aangeboden, met telkens 10 doelzinnen per muziekconditie binnen een blok. Voor elke participant was de volgorde van de BKB-zinnen gelijk, met eerst lijst 8 en 9 in blok 1 en daarna lijst 1 en 7 in blok 2. De condities zijn pseudowillekeurig aangeboden.

Voor elke doelzin werd een willekeurig nieuw stukje uit de achtergrondmuziek gebruikt, waardoor bij elke doelzin een uniek stukje uit de achtergrondmuziek hoorde. Daarbij werd er wel op gelet dat er altijd een stukje werd gekozen waarin (in de songtekst condities) de zangstem te horen was om ervoor te zorgen dat er daadwerkelijk gezongen tekst aanwezig was voor die condities. Doordat de volgorde van de doelzinnen gelijk bleef voor alle participanten, bleef ook de volgorde van de stukjes achtergrondmuziek gelijk. Deze konden wel verschillen in muziekconditie (Karaoke, Simlish of Engels). Doelzinnen werden met behulp van het programma Audacity© met de stukjes achtergrondmuziek gecombineerd. Per item werd de doelzin voorafgegaan en gevolgd door 500 milliseconden achtergrondmuziek. De gemiddelde duur van een item was 2.5 seconden.

Door het willekeurig kiezen van de stukjes achtergrondmuziek werd het nummer niet chronologisch doorlopend afgespeeld maar juist onvoorspelbaar en door elkaar gehusseld. Voorgaand onderzoek wijst uit dat bekendheid met achtergrondmuziek voor jongere luisteraars een faciliterend effect kan hebben op het concentratievermogen (Mori, Naghsh, & Tezuka, 2014) en herkenning van het doelsignaal (Russo & Pichora-Fuller, 2008). Russo en Pichora-Fuller speculeren dat luisteraars mogelijk gebruik kunnen maken van hun verwachtingen en voorspellingen over de bekende muziek. Met informatie over het verdere verloop van de muziek zouden zij op strategische momenten selectief kunnen filteren voor het doelsignaal. Door de muziek niet chronologisch maar gehusseld af te spelen is er minder invloed van bekendheid met de muziek, en gaat daarnaast de stimulus onzekerheid omhoog. Gehusselde muziek resulteert in moeilijker luisteromstandigheden (Shi & Law, 2010) waardoor effecten van songteksten en semantische betekenis naar verwachting duidelijker zullen zijn.

Om prestaties per doelzin over verschillende muziekcondities te kunnen vergelijken is voor elke doelzin een vast stukje uit het nummer gekozen als achtergrondmuziek. Dat wil zeggen dat één doelzin in de drie verschillende muziekcondities met precies hetzelfde stukje uit het nummer gecombineerd werd. Dit maakt ook dat er makkelijk gecontroleerd kon worden of een stukje achtergrondmuziek niet meerdere keren gebruikt werd – om leereffecten te voorkomen – en dat de verhouding van refrein tot couplet per blok en per conditie gelijk gehouden kon worden. Naar verwachting zijn de prestaties bij zinnen met een stukje refrein op de achtergrond slechter dan bij zinnen met een stukje couplet op de achtergrond. Het refrein is namelijk van hogere complexiteit dan een couplet vanwege een complexer ritme en meer instrumenten, wat een hogere maskeringseffectiviteit teweegbrengt (Başkent et al., 2014; Scharenborg & Larson, 2018a, 2018b). Daarnaast is mogelijk de songtekst in het refrein bekender dan dat in de coupletten, wat eveneens voor meer maskering van het doelsignaal kan zorgen (Brouwer et al., 2012; Calandruccio et al., 2010; Garcia Lecumberri & Cooke, 2006; Van Engen, 2010; Van Engen & Bradlow, 2007). Om voor deze mogelijke effecten van complexiteit en bekendheid met het refrein te controleren is voor elke conditie (elke combinatie van type achtergrondmuziek en SNR niveau) een gelijk aantal doelzinnen gecombineerd met een refreinstukje. Pilot tests hebben uitgewezen dat 5 refreinstukjes per conditie (10 items) te

veel interfereerde met de algehele prestaties doordat de taak te moeilijk werd. Het aantal refreinstukjes per conditie is derhalve vastgesteld op 3 van die 10 items, waarbij prestaties geen vloereffecten meer lieten zien.

Voorafgaand aan de experimentele taak kregen de participanten een oefensessie van zes items om bekend te worden met de taak en de doelstem. De oefenzinnen zijn geselecteerd uit lijst 20 van BKB. Deze zinnen zijn ingesproken door dezelfde spreker als de teststimuli en zijn op hetzelfde geluidsniveau van 60 dB aangeboden. Tijdens de oefensessie kregen participanten voor alle drie de muziekcondities twee doelzinnen aangeboden. De SNR van de oefenitems was -5 dB voor de eerste twee zinnen, -8 dB voor de middelste twee zinnen en -10 dB voor de laatste twee zinnen. Zo werd langzaam opgebouwd naar de moeilijkere SNRs in de experimentele items.

### *Procedure*

Het experiment werd online afgenomen met behulp van het programma Qualtrics (Provo, UT). Participanten kregen een link naar het experiment en konden deze zelfstandig invullen op hun eigen laptop, desktop of smartphone. Hen werd gevraagd om in een rustige, stille omgeving te gaan zitten en een goede koptelefoon te gebruiken. Voorafgaand aan het experiment ondertekenden participanten digitaal een toestemmingsformulier. Daarin werd toestemming gevraagd om hun (geanonimiseerde) gegevens en responses te verzamelen. Ook werden zij gewezen op het recht om op ieder gewenst moment hun deelname te beëindigen zonder daar een verklaring voor te hoeven geven, en om binnen twee weken een verzoek in te kunnen dienen om hun gegevens te laten verwijderen.

De instructie die de participanten kregen was om te luisteren naar de gesproken zinnen met achtergrondmuziek en de zin die ze hadden gehoord in te typen met behulp van het toetsenbord. De participanten werd gevraagd om alles wat ze hadden gehoord op te schrijven, ook al waren ze er niet helemaal zeker over of hadden ze alleen delen van de zin verstaan. Na het invullen van een antwoord konden participanten telkens zelf naar het volgende item gaan door op een pijltje te klikken. Gedurende het experiment kregen de participanten geen feedback over hun responses. Elk item kon maar één keer afgespeeld worden. De totale duur van het experiment was ongeveer 15 tot 20 minuten.

### *Scoren*

De responses van de participanten werden door één beoordelaar gescoord op basis van het aantal sleutelwoorden dat correct waargenomen was. Fouten die duidelijk te herleiden waren als typfouten of spelfouten werden goed gerekend, maar mogelijke spelfouten die resulteerden in een ander bestaand woord (bijvoorbeeld 'kam' in plaats van 'kan', 'voel' in plaats van 'viel') werden fout gerekend. Daarnaast werden homofonen (bijvoorbeeld 'licht' in plaats van 'ligt') of kleine variaties, zoals toevoeging of weglating van een meervoudssuffix, niet fout gerekend, maar fonetische variaties (bijvoorbeeld 'drinkt' in plaats van 'dronk', 'koopt' in plaats van 'kocht') wel.

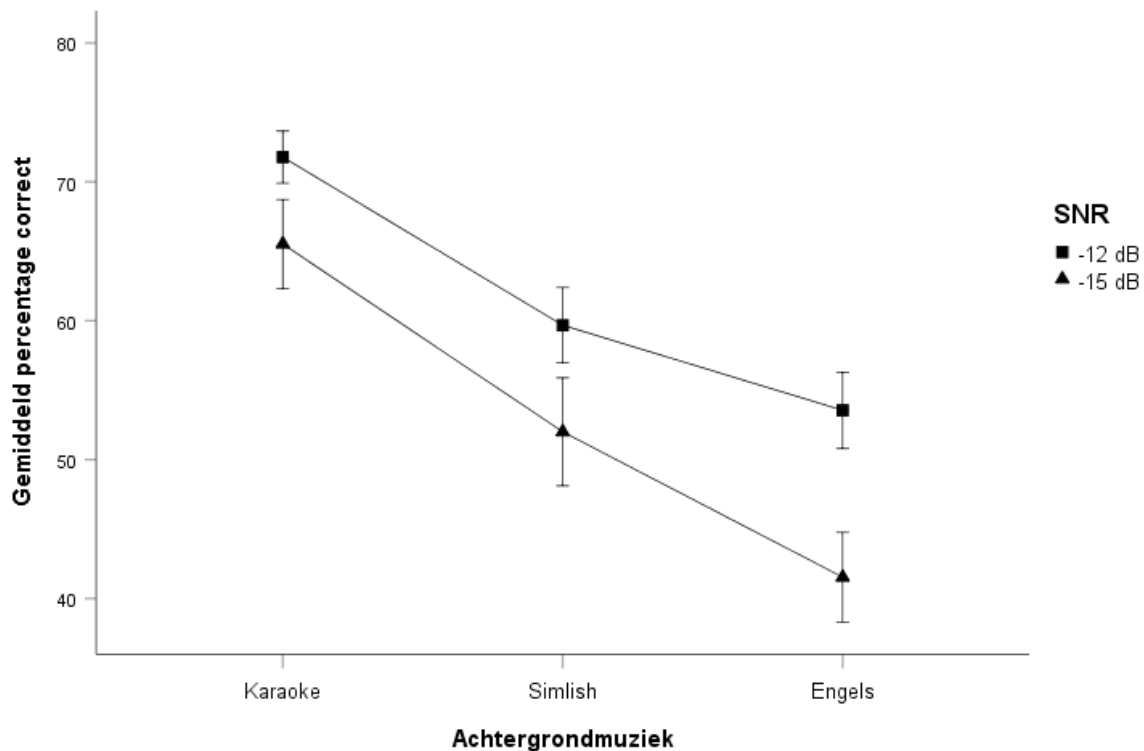
Voor elke participant werd per conditie (drie varianten achtergrondmuziek en twee SNRs) het aantal correct herkende sleutelwoorden opgeteld en hiermee werd een percentage correct herkende sleutelwoorden berekend. Zo waren er voor elke participant zes scores beschikbaar, een voor elke combinatie van een type achtergrondmuziek en een SNR. Deze percentages werden afgerond op één decimaal.

## Data-analyse

Een variantieanalyse met de binnen-subject factoren Achtergrondmuziek (Karaoke, Simlish, Engels) en SNR (-12 dB, -15 dB) is uitgevoerd om te kijken naar mogelijke hoofd- en interactie-effecten op het gemiddelde percentage correct herkende sleutelwoorden. Door het toevoegen van Helmert contrasten kon ten eerste de score bij achtergrondmuziek zonder songteksten (Karaoke) vergeleken worden met de score in de condities met songteksten (Simlish en Engels samen). In het tweede contrast werd de score bij semantisch betekenisloze songteksten (Simlish) vergeleken met de score bij semantisch betekenisvolle songteksten (Engels). Uitkomsten van deze contrasten geven inzicht in respectievelijk een effect van aanwezigheid van songteksten in achtergrondmuziek, en de semantische betekenis van de songteksten.

## Resultaten

Figuur 2 toont de prestaties van de participanten per type achtergrondmuziek en SNR niveau. In de Karaoke conditie – achtergrondmuziek zonder songteksten – werd bij een SNR van -12 dB gemiddeld 71.8% ( $SD = 5.62$ ) van de sleutelwoorden in de gesproken zinnen correct herkend. Bij een SNR van -15 dB was dit 65.5% ( $SD = 9.61$ ). In de Simlish conditie – achtergrondmuziek met semantisch betekenisloze songteksten – werd gemiddeld 59.7% ( $SD = 8.15$ ) van de sleutelwoorden correct herkend bij een SNR van -12 dB, en 52.0% ( $SD = 11.66$ ) bij een SNR van -15 dB. In de Engels conditie – achtergrondmuziek met semantisch betekenisvolle songteksten – werd bij een SNR van -12 dB gemiddeld 53.6% ( $SD = 8.26$ ) van de sleutelwoorden correct herkend, en bij een SNR van -15 dB was dit 41.5% ( $SD = 9.68$ ).



**Figuur 2** – Gemiddelde scores in percentage correct herkende sleutelwoorden in de condities Karaoke, Simlish en Engels bij een SNR van -12 dB (vierkanten) en -15 dB (driehoeken). Foutbalken representeren 2 maal de standaardfout, oftewel een 95% betrouwbaarheidsinterval.



Een 3x2 Repeated Measures ANOVA met de binnen-subject factoren Achtergrondmuziek (Karaoke, Simlish, Engels) en SNR (-12 dB, -15 dB) en de afhankelijke variabele percentage correct herkende sleutelwoorden is uitgevoerd. Mauchly's test toonde aan dat de assumptie van sfericiteit voor Achtergrondmuziek niet is geschonden,  $\chi^2(2) = 1.996, p = .369$ . Er werd een significant hoofdeffect van Achtergrondmuziek gevonden,  $F(2, 70) = 116.233, p < .001, \eta^2_p = .769$ . Helmert contrasten lieten zien dat het percentage correct bij achtergrondmuziek zonder songteksten (Karaoke) significant hoger was,  $F(1, 35) = 256.211, p < .001, \eta^2_p = .880$ , dan bij achtergrondmuziek met songteksten (Simlish en Engels samen), en dat daarnaast het percentage correct bij achtergrondmuziek met semantisch betekenisloze songteksten (Simlish) significant hoger was,  $F(1, 35) = 28.733, p < .001, \eta^2_p = .451$ , dan bij achtergrondmuziek met semantisch betekenisvolle songteksten (Engels). Er werd ook een significant hoofdeffect van SNR gevonden,  $F(1, 35) = 76.530, p < .001, \eta^2_p = .686$ , waarbij het percentage correct bij een SNR van -12 dB significant hoger was dan bij een SNR van -15 dB. Tot slot bleek het interactie-effect tussen Achtergrondmuziek en SNR marginaal significant,  $F(2, 70) = 2.812, p = .067, \eta^2_p = .074$ . Dit leek na een visuele inspectie van Figuur 2 vooral gedreven door de Engels conditie, waarin een sterkere afname in percentage correct te zien is bij een SNR van -15 dB ten opzichte van een SNR van -12 dB.

## Discussie

In het huidige onderzoek is gekeken naar de invloeden van (1) de aanwezigheid van songteksten in achtergrondmuziek en (2) de semantische betekenis van die songteksten, op de maskeringseffectiviteit van achtergrondmuziek op gesproken zinnen. Nederlandstalige luisteraars hebben een zinsherkenningstaak met drie typen achtergrondmuziek uitgevoerd. De typen achtergrondmuziek verschilden enkel op het gebied van songteksten. Songteksten waren semantisch betekenisvol, semantisch betekenisloos of volledig afwezig. Verwacht werd dat (1) achtergrondmuziek met songteksten een effectievere maskering teweegbrengt dan achtergrondmuziek zonder songteksten, en dat (2) de semantisch betekenisloze en minder bekende songteksten het doelsignaal minder effectief maskeren dan de bekende, semantisch betekenisvolle songteksten. De resultaten laten een sterk effect zien van zowel de aanwezigheid van songteksten, als de semantische betekenis ervan. Luisteraars zijn beter in het herkennen van semantisch betekenisvolle zinnen in het doelsignaal als de songteksten semantisch betekenisloos zijn dan wanneer deze semantisch betekenisvol zijn. De prestaties van de luisteraars zijn nog beter als de achtergrondmuziek helemaal geen songteksten bevat.

### *Aanwezigheid van songteksten*

De bevinding dat de aanwezigheid van songteksten in de achtergrondmuziek een nadelig effect heeft op de herkenning van gesproken zinnen, is in lijn met de verwachtingen op basis van eerdere bevindingen van Scharenborg en Larson (2018a, 2018b). De aanwezigheid van songteksten in achtergrondmuziek brengt een extra laag aan maskering op het doelsignaal met zich mee. Songteksten kunnen zowel energetische als informationele maskering teweegbrengen, maar in het huidige experiment is gecontroleerd op mogelijke verschillen in energetische maskering tussen drie typen achtergrondmuziek. Het gevonden verschil in zinsherkenning tussen de Karaoke conditie en de songtekst condities (Simlish en Engels) is voornamelijk te verklaren door een verschil in informationele maskering. Hoewel instrumentele muziek ook voor informationele maskering kan zorgen (bijv. Brungart, 2001; Russo & Pichora-Fuller, 2008) voegen songteksten een extra element toe, namelijk talige informatie. Deze extra informatie kan zorgen voor meer afleiding van het doelsignaal (vgl. Abril & Flowers, 2007;

Shih et al., 2012). Door de aanwezigheid van songteksten is er in de achtergrondmuziek meer informatie te verwerken, en heeft deze informatie meer overeenkomsten met de gesproken taal in het doelsignaal, wat resulteert in een effectievere maskering van het doelsignaal (bijv. Brouwer et al., 2012). Hiermee is de eerste hypothese bevestigd.

### *Semantische betekenis van songteksten*

In het huidige onderzoek zijn interessante nieuwe bevindingen gedaan over de invloed van semantische betekenis van songteksten op de maskeringseffectiviteit van achtergrondmuziek. De resultaten tonen aan dat maskering van achtergrondmuziek met semantisch betekenisloze songteksten significant minder effectief is dan dat van achtergrondmuziek met semantisch betekenisvolle songteksten. Dit sluit aan bij eerdere bevindingen van De Groot en Smedinga (2014), waaruit bleek dat semantische betekenis van songteksten een negatieve invloed heeft op concentratie- en leervermogen van luisteraars. Mogelijk kunnen luisteraars zich minder goed concentreren op de zinsherkenningstaak als zij betekenisvolle songteksten horen in de achtergrondmuziek dan wanneer zij betekenisloze songteksten horen. Ook sluiten de resultaten van het huidige onderzoek aan bij de bevindingen van Brouwer et al. (2012) over de maskeringseffectiviteit van tweespreker achtergrondspraak. De gevonden effecten van semantische betekenis op de maskeringseffectiviteit blijken in het huidige onderzoek gegeneraliseerd te kunnen worden naar songteksten in achtergrondmuziek.

In de inleiding zijn al verschillende mogelijke verklaringen aangedragen voor het feit dat resultaten van Tun et al. (2002) en Calandruccio et al. (2018) niet bij deze bevindingen van Brouwer et al. (2012) aansluiten. Het is echter wel opvallend dat na controle voor verschil in stimulus onzekerheid tussen de condities er geen invloed van semantische betekenis meer werd gevonden in Calandruccio et al., maar wel in het huidige onderzoek. In Calandruccio et al. wordt voor stimulus onzekerheid gecontroleerd door de syntactische structuur (variabel of constant) van de verschillende achtergrondcondities mee te nemen in de analyse. Veel variëteit in de syntactische structuur zou namelijk voor een hogere stimulus onzekerheid kunnen zorgen. In het huidige onderzoek is ook gecontroleerd voor stimulus onzekerheid, door deze voor de verschillende achtergrondcondities gelijk te houden. In elke conditie wordt de achtergrondmuziek in een pseudowillekeurige, onverwachte volgorde afgespeeld en is het ritme en de melodie van de achtergrondmuziek gelijk. Toch vindt Calandruccio et al. geen effect van semantische betekenis, terwijl dit effect wel optrad in het huidige onderzoek.

Een mogelijke verklaring is de rol van bekendheid. In het huidige onderzoek zijn verschillende talen gebruikt. Brouwer et al. (2012) vond dat bekendheid met de taal in het doelsignaal en/of het achtergrondgeluid interacteert met het effect van semantische betekenis. Engelse luisteraars horen geen verschil tussen semantische betekenisvolle en semantische betekenisloze Nederlandse achtergrondspraak. Omdat zij niet bekend zijn met het Nederlands zijn beide condities voor hen betekenisloos. Engelse luisteraars presteerden dan ook slechter bij Engelse (bekende) achtergrondspraak dan bij Nederlandse (onbekende) achtergrondspraak. Door Calandruccio et al. is slechts één taal gebruikt voor de achtergrondspraak, namelijk de moedertaal van de luisteraars. In het huidige onderzoek is voor de semantisch betekenisloze conditie een voor de luisteraars (relatief) onbekende taal gebruikt. De betekenisloze inhoud en relatief lage bekendheid met de taal zouden samen mogelijk voor een groter effect gezorgd kunnen hebben.

Dit zou ook kunnen verklaren waarom het verschil tussen de semantisch betekenisvolle en semantisch betekenisloze conditie in het huidige onderzoek groter is dan dat in Brouwer et al. (2012). Voor de Nederlandse luisteraars die naar Nederlandse doelzinnen met Engelse achtergrondspraak luisterden werd in Brouwer et al. maar een klein verschil tussen de

betekenisvolle en betekenisloze conditie gevonden. Deze twee condities verschilden niet in bekendheid met de taal, zoals in het huidige onderzoek.

Een ander verschil met het onderzoek van Brouwer et al. (2012), en van Calandruccio et al. (2018), is dat de semantisch betekenisloze conditie anders is vormgegeven. De zinnen in de semantisch betekenisloze condities in Brouwer et al. en Calandruccio et al. bestonden uit bestaande woorden in dezelfde taal als de zinnen in de betekenisvolle conditie. Wat deze zinnen semantisch betekenisloos maakte was de combinatie van die woorden. Dit vormde vreemde zinnen waar geen zinvolle betekenis aan te koppelen is. Voorbeelden zijn “*The great car met the milk*” (uit Brouwer et al.) en “*The girl lives purple teens*” (uit Calandruccio et al.). Ondanks dat deze zinnen als geheel niet betekenisvol zijn in het Engels zijn de individuele woorden dat wel. De songteksten in het huidige onderzoek zijn, per definitie zelfs, volledig betekenisloos. Aan Simlish is geen systeem of vaste, herkenbare semantische betekenis te koppelen. Dit maakt ook gelijk dat de taal minder bekend is bij de luisteraars. Woorden die niks betekenen zullen minder bekend klinken dan bestaande woorden in je moedertaal, ook al is de combinatie van die woorden vreemd.

De keuze voor Simlish als betekenisloze conditie betekent echter wel dat in het huidige onderzoek de talen in de twee condities verschillend zijn. Verschillende talen zullen in veel gevallen ook verschillen in de maskering die ze teweegbrengen (bijv. Brouwer & Bradlow, 2014; Brouwer et al., 2012; Calandruccio et al., 2017; Van Engen & Bradlow, 2007). Wel zijn de verschillende talen in het huidige onderzoek geproduceerd door dezelfde zangeres, waardoor effecten door verschillen in stemeigenschappen zijn uitgesloten (in tegenstelling tot de genoemde studies, met verschillende sprekers voor de verschillende talen). Daarnaast lijken de Simlish songteksten fonetisch gezien op het origineel. Hoewel Simlish niet hetzelfde is als het Engels, is het verschil tussen de twee varianten van songtekst in het huidige onderzoek mogelijk kleiner dan het verschil tussen twee verschillende achtergrondtalen in voorgaand onderzoek.

De verschillen in prestaties tussen de semantisch betekenisvolle en semantisch betekenisloze condities in het huidige onderzoek wijzen op een effect van semantische betekenis van songteksten in achtergrondmuziek op herkenning van gesproken zinnen. Dit sluit aan bij de opgestelde hypothese. Bekendheid met de taal en de songteksten in de achtergrondmuziek zou hier mogelijk ook een invloed op kunnen hebben. De bekende songteksten in de bekende taal vormden een effectievere maskering op het doelsignaal. Een lagere bekendheid is echter vaak ook een indicatie van semantische betekenis, aangezien een onbekende taal voor luisteraars per definitie betekenisloos is.

### *Invloed van SNR niveau*

Uit de resultaten bleek een effect van SNR op de maskeringseffectiviteit van achtergrondmuziek op gesproken zinnen. Dit effect sluit aan bij voorgaand onderzoek (bijv. Brouwer et al., 2012; Brungart, 2001; Brungart et al., 2001; Cooke et al., 2008). Een moeilijkere SNR, waarbij het achtergrondgeluid harder staat, maakt het doelsignaal moeilijker te herkennen.

Er werd ook verwacht dat SNR invloed zou kunnen hebben op de effecten van songteksten en semantische betekenis. Voorgaand onderzoek toont aan dat sommige effecten uniek zijn of sterker zijn in moeilijkere luistercondities, zoals moeilijkere SNRs (bijv. Gittleman & Van Engen, 2018; Russo & Pichora-Fuller, 2008; Van Engen, 2010; Van Engen & Bradlow, 2007). In het onderzoek van Scharenborg en Larson (2018a, 2018b) werd alleen voor de moeilijkere SNRs een sterk effect van songtekst gevonden. Dit wordt in het huidige onderzoek niet gerepliceerd. In de beide SNR condities was er een vergelijkbaar effect van songteksten. Een mogelijke verklaring hiervoor is dat Scharenborg en Larson in hun onderzoek veel makkelijkere SNR niveaus hebben gebruikt. De moeilijkste SNR in hun onderzoek was 0 dB,

met daarnaast nog twee makkelijkere SNRs van +5 dB en +15 dB. Mogelijk is voor deze makkelijkere SNR niveaus gekozen omdat in het onderzoek van Scharenborg en Larson ook complexe muziekstukken getest zijn wat de taak zelfs in makkelijkere SNRs aanzienlijk moeilijker maakt (zie ook Başkent et al., 2014). Desalniettemin is dit een groot verschil met de moeilijker SNRs van -12 dB en -15 dB in het huidige onderzoek. Deze moeilijker luisteromstandigheden hebben mogelijk een algeheel sterker effect van songteksten teweeggebracht. Ook zijn de verschillen tussen de SNR niveaus in het huidige onderzoek kleiner dan die in het onderzoek van Scharenborg en Larson, waardoor er mogelijk minder verschil tussen de makkelijkere en moeilijker conditie zat.

Hoewel er voor beide SNR niveaus een effect van songteksten gevonden is, bleek het interactie-effect van Achtergrondmuziek en SNR marginaal significant. Een visuele inspectie van de gemiddelde scores (Figuur 2) wijst uit dat dit waarschijnlijk vooral te verklaren is door de Engels conditie. In deze conditie neemt de gemiddelde score sterker af voor de moeilijker SNR dan voor de makkelijkere SNR. Dit suggereert dat het effect van semantische betekenis ook sterker is in de moeilijker SNR. Mogelijk zou een groter verschil tussen de SNR niveaus een volledig significant interactie-effect van Achtergrondmuziek en SNR teweeg hebben gebracht. Toekomstig onderzoek zal dit moeten uitwijzen.

### *Discussiepunten*

In het huidige onderzoek is geen rekening gehouden met ervaringen van participanten met muziek. Het zou verwacht kunnen worden dat participanten die veel ervaring hebben met muziek maken of muziek luisteren (tijdens het concentreren op werk) minder maskering van de achtergrondmuziek ervaren dan participanten die minder vaak met muziek bezig zijn. Scharenborg en Larson (2018b) vonden echter geen invloed van muzikale ervaringen op prestaties van hun participanten, waardoor deze factoren niet zijn meegenomen in het huidige onderzoek.

Ook is het belangrijk om op te merken op welke manier de scores in het huidige onderzoek zijn berekend. Er is alleen gekeken naar een aantal sleutelwoorden per gesproken zin in het doelsignaal. Door responses van participanten te vergelijken met de originele sleutelwoorden kon worden vastgesteld hoeveel sleutelwoorden correct zijn herkend. Er is echter niet specifiek gekeken naar de mate van accuratesse. Zowel volledig correcte vormen als licht gereduceerde vormen of homofonen bijvoorbeeld werden volledig goed gerekend. Het is niet bijgehouden of gedefinieerd wat voor soort fouten er zijn gemaakt, of wat de mate van accuratesse was. Dit zou een vollediger beeld kunnen geven van wat voor soort invloeden achtergrondgeluiden kunnen hebben op herkenning van het doelsignaal, en welke fonemen bijvoorbeeld zwaarder lijden onder de maskering door het achtergrondgeluid. Ook is niet meegenomen hoe de overige woorden in de doelzinnen – naast de sleutelwoorden – zijn verstaan, of hoeveel van deze woorden zijn weggelaten of juist toegevoegd. Dergelijke analyses kunnen inzicht bieden in welke posities in de zin of welke soort woorden vooral gevoelig zijn voor maskeringseffecten bijvoorbeeld. Het kan verder ook inzichtelijk maken welke informatie luisteraars zelf toevoegen of nodig hebben om een boodschap volledig te begrijpen. Dit waren echter niet de focusgebieden van de huidige studie. Toekomstig onderzoek zou bovenstaande redenties in overweging kunnen nemen voor het opstellen van een scoringsprotocol.

Er zijn in het huidige onderzoek specifieke keuzes gemaakt voor het scoren van bepaalde varianten van de sleutelwoorden. Homofonen zijn wel goed gerekend maar duidelijke fonetische variaties van een sleutelwoord niet. Dat betekent bijvoorbeeld dat voor het sleutelwoord ‘ligt’ de respons ‘licht’ wel goed is gerekend, maar voor het sleutelwoord ‘dronk’ de respons ‘drinkt’ niet goed is gerekend. Het zou beargumenteerd kunnen worden dat dit andersom moet zijn, aangezien de semantische inhoud en betekenis van ‘dronk’ en ‘drinkt’ zo

goed als gelijk zijn en die van ‘ligt’ en ‘licht’ niet. Een persoon die ‘licht’ verstaat in plaats van ‘ligt’ heeft dus mogelijk niet de boodschap in het doelsignaal correct begrepen. In het huidige onderzoek ligt echter de focus op het bepalen of participanten gesproken zinnen (en woorden in die zinnen) correct kunnen verstaan. Als een participant ‘drinkt’ invult in de doelzin “Het KIND DRONK wat MELK” kan het zijn dat hij/zij dit werkwoord zelf heeft ingevuld om een logisch lopende zin te maken maar eigenlijk het woord niet heeft verstaan. Om die reden zijn fonetische variaties van een sleutelwoord in het huidige onderzoek niet goedgekeurd.

Een laatste en zeer belangrijk punt om te noemen voor het interpreteren van het huidige onderzoek is dat de gebruikte wijze van afname een aantal negatieve eigenschappen kan hebben. Het experiment is online afgenomen, wat voordelen heeft zoals een snellere werving van participanten en een kortere afnameperiode aangezien er niet bij elke afname een experimenteleider aanwezig hoeft te zijn. Het brengt echter ook nadelen met zich mee. Zo is het, ondanks duidelijke schriftelijke instructies, niet met alle zekerheid aan te nemen dat participanten volledig begrepen hebben wat de taak is en dat ze de taak naar genoegzaam uitvoeren. Participanten kunnen voor en tijdens het experiment geen vragen stellen aan een experimenteleider, en er kan ook niet gecontroleerd worden of participanten niet afgeleid worden of overleggen met anderen. Daarnaast zal het vaker gebeuren dat participanten de taak niet afmaken waardoor een groter aandeel van de verkregen data – groter dan in offline afnames – niet bruikbaar is voor analyses, zoals ook het geval was in het huidige onderzoek. Desalniettemin is in dit onderzoek extra gelet op responses van de participanten om te herleiden of zij de taak goed hadden begrepen en het experiment volledig hadden voltooid, alvorens hun responses mee te nemen in de analyses. Hiermee wordt zo goed mogelijk voor de mogelijke negatieve eigenschappen van deze wijze van afname gecontroleerd.

### *Suggesties voor toekomstig onderzoek*

Het onderzoek naar maskeringseffectiviteit van achtergrondmuziek met songteksten is schaars. Meer onderzoek kan interessante inzichten bieden in de verschillende factoren van achtergrondmuziek die een rol spelen in de maskeringseffectiviteit van achtergrondgeluid op gesproken zinnen. Het huidige onderzoek biedt – voor zover bekend – de eerste inzichten in de invloed van semantische betekenis van songteksten op maskeringseffectiviteit van achtergrondmuziek bij gesproken zinnen.

Voor gesproken taal als achtergrondgeluid is al veel bekend over de verschillende factoren en hoe deze de maskeringseffectiviteit beïnvloeden. Zo is er een belangrijke rol weggelegd voor de taal die in de achtergrond wordt gesproken, en de mate van gelijkenis met de taal in het doelsignaal. Het huidige onderzoek en dat van Scharenborg en Larson (2018a) lijken al een invloed van de taal van songteksten te suggereren. Toekomstig onderzoek zou kunnen kijken naar mogelijke verschillen tussen typen achtergrondmuziek met songteksten in verschillende talen. Hoewel Engelse songteksten in Nederland erg gebruikelijk zijn in achtergrondmuziek – Nederlandstalige muziek bijvoorbeeld zal in veel openbare ruimtes veel minder gedraaid worden dan Engelstalige muziek – is het interessant om ook te kijken naar herkenning van het doelsignaal bij Nederlandse songteksten. Beide talen zijn erg bekend bij veel Nederlandse luisteraars, maar mogelijk op verschillende manieren. Waar voor gesproken taal Nederlands waarschijnlijk het meest bekend is, zou voor songteksten mogelijk het Engels bekender zijn. Door deze twee (of andere) talen met elkaar te vergelijken kan gekeken worden naar mogelijke verschillende invloeden van bekendheid met de achtergrondmuziek, en in hoeverre de *target-masker linguistic similarity hypothesis* ook geldt voor de taal van songteksten in achtergrondmuziek.

Afgezien van de taal is in het huidige onderzoek ook slechts één liedje en één zangstem getest. De huidige resultaten zouden nog gegeneraliseerd moeten worden naar andere liedjes

en/of andere stemmen. Mogelijk speelt bijvoorbeeld genre of het soort instrument dat gebruikt is ook een rol in hoe sterk het effect van songtekst of semantische betekenis is (vgl. een effect van complexiteit; Baškent et al., 2014; Scharenborg & Larson, 2018a, 2018b). Door meerdere liedjes en zangstemmen te testen kan worden onderzocht hoe sterk de gevonden effecten zijn en wat mogelijke andere invloeden hierop zijn.

Tot slot is in het huidige onderzoek – en in dat van Baškent et al. (2014) en Scharenborg en Larson (2018a, 2018b) – de achtergrondmuziek niet chronologisch over de experimentele items aangeboden. Dit maakt de muziek minder bekend en minder voorspelbaar. Zoals Russo en Pichora-Fuller (2008) aantonen kunnen jongvolwassen luisteraars hun voorspellingen over het verloop van de muziek gebruiken om het herkennen van het doelsignaal te faciliteren. Mogelijk zijn de effecten in het huidige en voorgaand onderzoek naar maskeringseffectiviteit van achtergrondmuziek uniek of sterker in deze moeilijkere luisteromstandigheid. Door een hogere stimulus onzekerheid hebben de aanwezigheid van songteksten en de semantische betekenis mogelijk een sterker effect gehad op de herkenning van het doelsignaal. In natuurlijke luistersituaties zal achtergrondmuziek niet op deze manier maar op chronologische wijze worden afgespeeld. Toekomstig onderzoek zou kunnen kijken of de gevonden effecten van songteksten en semantische betekenis ook voor chronologisch afgespeelde achtergrondmuziek opgaan.

### *Conclusies*

Uit het huidige onderzoek blijkt een sterk effect van de aanwezigheid van songteksten op de maskeringseffectiviteit van achtergrondmuziek op gesproken zinnen. Herkenning van gesproken zinnen in het doelsignaal is minder moeilijk bij instrumentele achtergrondmuziek dan bij achtergrondmuziek met songteksten. Ook blijkt de semantische betekenis van de songteksten invloed te hebben op de maskeringseffectiviteit van de achtergrondmuziek. Bij een semantisch betekenisvol doelsignaal is achtergrondmuziek met semantisch betekenisvolle songteksten een effectievere maskering dan achtergrondmuziek met semantisch betekenisloze songteksten. Hierbij kan bekendheid met de taal en de songteksten ook een rol gespeeld hebben. Aangezien de semantisch betekenisvolle songteksten ook bekender waren bij de participanten dan de semantisch betekenisloze songteksten, kan bekendheid mogelijk hebben bijgedragen aan een hogere maskeringseffectiviteit. De resultaten van het huidige onderzoek ondersteunen de *target-masker linguistic similarity hypothesis*, en laten zien dat deze hypothese ook opgaat voor songteksten. Ook voor songteksten in achtergrondmuziek geldt dat een hogere gelijkenis tussen het doelsignaal en het achtergrondgeluid en bekendheid met het achtergrondgeluid resulteert in een hogere maskeringseffectiviteit.

### **Referenties**

- Abril, C.R., & Flowers, P.J. (2007). Attention, Preference, and Identity in Music Listening by Middle School Students of Different Linguistic Backgrounds. *Journal of Research in Music Education*, 55(2), 204-219. doi: 10.1177/002242940705500303
- Affandi, R. (2009, 21 juni). *Katy Perry – Hot N Cold ( Instrumental / karaoke / lyrics )* [Videobestand]. Geraadpleegd van <https://www.youtube.com/watch?v=rRVdEQxqNKs>
- Audacity Team. (2020). *Audacity®: Free Audio Editor and Recorder (2.3.3)* [Computerprogramma]. Geraadpleegd van <https://www.audacityteam.org/>
- Bamford, J., & Wilson, I. (1979). Methodological considerations and practical aspects of the BKB sentence lists. In J. Bench & J. Bamford (Eds.), *Speech-hearing tests and the spoken*

- language of hearing-impaired children* (pp.148-187). Londen, Verenigd Koninkrijk: Academic Press.
- Barnes, A. (2020, 4 februari). *The Sims turns 20: Creator Will Wright reflects on the battle he waged to get one of the best games of all time made*. Geraadpleegd op 20 juni, van <https://www.gamesradar.com/the-making-of-the-sims/>
- Başkent, D., Van Engelshoven, S., & Galvin, J.J. (2014). Susceptibility to interference by music and speech maskers in middle-aged adults. *The Journal of the Acoustical Society of America*, 135(3), 147-153. doi: 10.1121/1.4865261
- Bench, J., Kowal, A., and Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, 13, 108–112.
- Bialystok, E., Craik, F.I.M., & Ryan, J. (2006). Executive control in a modified anti-saccade task: Effects on aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6), 1341-1354. doi: 10.1037/0278-7393.32.6.1341
- Boersma, P. & Weenink, D. (2020). *Praat: doing phonetics by computer* (6.1.09) [Computerprogramma]. Geraadpleegd van <https://www.praat.org/>
- Boland, E. (2010). *The Sims: The Complete Guide*. Geraadpleegd op 20 juni 2020, van <https://books.google.nl/books?id=5fkLAgAAQBAJ>
- Bradlow, A.R., Nygaard, L.C., & Pisoni, D.B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics*, 61(2), 206-219. doi: 10.3758/BF03206883
- Brouwer, S. (2017). Masking release effects of a standard and a regional linguistic variety. *The Journal of the Acoustical Society of America*, 142(2), 237-243. doi: 10.1121/1.4998607
- Brouwer, S. (2019). The role of foreign accent and short-term exposure in speech-in-speech recognition. *Attention, Perception, & Psychophysics*, 81(6), 2053-2062. doi: 10.3758/s13414-019-01767-8
- Brouwer, S., & Bradlow, A.R. (2014). Contextual variability during speech-in-speech recognition. *The Journal of the Acoustical Society of America*, 136(1), EL26-EL32. doi: 10.1121/1.4881322
- Brouwer, S., & Bradlow, A.R. (2015). The effect of target-background synchronicity on speech-in-speech recognition. *Proceedings of the 18<sup>th</sup> International Congress of Phonetic Sciences 2015 (ICPhS XVIII)*, Glasgow.
- Brouwer, S., Van Engen, K.J., Calandruccio, L., & Bradlow, A.R. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. *The Journal of the Acoustical Society of America*, 131(2), 1449-1464. doi: 10.1121/1.3675943
- Brungart, D.S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, 109(3), 1101-1109. doi: 10.1121/1.1345696
- Brungart, D.S., Simpson, B.D., Ericson, M.A., & Scott, K.R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, 110(5), 2527-2538. doi: 10.1121/1.1408946
- Calandruccio, L., Dhar, S., & Bradlow, A.R. (2010). Speech-on-speech masking with variable access to the linguistic content of the masker speech. *The Journal of the Acoustical Society of America*, 128(2), 860-869. doi: 10.1121/1.3458857
- Calandruccio, L., Buss, E., & Bowdrie, K. (2017). Effectiveness of Two-Talker Maskers That Differ in Talker Congruity and Perceptual Similarity to the Target Speech. *Trends in Hearing*, 21, 1-14. doi: 10.1177/2331216517709385
- Calandruccio, L., Buss, E., Bencheck, P., & Jett, B. (2018). Does the semantic content or syntactic regularity of masker speech affect speech-on-speech recognition? *The Journal of the Acoustical Society or America*, 144(6), 3289-3302. doi: 10.1121/1.5081679

- Cherry, E.C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, 25(5), 975-979. doi: 10.1121/1.1907229
- Cooke, M., Garcia Lecumberri, M.L., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America*, 123(1), 414-427. doi: 10.1121/1.2804952
- De Groot, A.M.B., & Smedinga, H.E. (2014). Let The Music Play! A Short-Term but No Long-Term Detrimental Effect of Vocal Background Music with Familiar Language Lyrics on Foreign Language Vocabulary Learning. *Studies in Second Language Acquisition*, 36(4), 681-707. doi: 10.1017/S0272263114000059
- Dslite804. (2008, 10 oktober). *Katy Perry Hot N Cold With Lyrics* [Videobestand]. Geraadpleegd van <https://www.youtube.com/watch?v=lz0JazR4-sI>
- Durlach, N.I., Mason, C.R., Kidd, Jr., G., Arbogast, T.L., Colburn, H.S., & Shinn-Cunningham, B. G. (2003a). Note on informational masking. *The Journal of the Acoustical Society of America*, 113(6), 2984-2987. doi: 10.1121/1.1570435
- Durlach, N.I., Mason, C.R., Shinn-Cunningham, B.G., Arbogast, T.L., Colburn, H.S., & Kidd, Jr., G. (2003b). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *The Journal of the Acoustical Society of America*, 114(1), 368-379. doi: 10.1121/1.1577562
- Emmorey, K., Luk, G., Pyers, J.E., & Bialystok, E. (2008). The Source of Enhanced Cognitive Control in Bilinguals: Evidence from Bimodal Bilinguals. *Psychological Science*, 19(12), 1201-1206. doi: 10.1111/j.1467-9280.2008.02224.x
- Freyman, R.L., Balakrishnan, U., & Helfer, K.S. (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America*, 109(5), 2112-2122. doi: 10.1121/1.1354984
- Garcia Lecumberri, M.L. & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America*, 119(4), 2445-2454. doi: 10.1121/1.2180210
- Gittleman, S., & Van Engen, K.J. (2018). Effects of noise and talker intelligibility on judgements of accentedness. *The Journal of the Acoustical Society of America*, 143(5), 3138-3145. doi: 10.1121/1.5038653
- Great Big Story. (2019, 12 november). *How the Language From the Sims Was Created* [Videobestand]. Geraadpleegd in april 2020, van <https://www.youtube.com/watch?v=FGsbeTV76YI>
- Jiang, N. (2015). Six decades of research on lexical representation and processing in bilinguals. In J. W. Schwieter (Red.), *The Cambridge Handbook of Bilingual Processing* (pp. 29-84). Cambridge, Verenigd Koninkrijk: Cambridge University Press. doi:10.1017/CBO9781107447257.002
- Johansson, S. (2005). *Origins of Language: Constraints on hypotheses*. Geraadpleegd op 23 juni 2020, van <https://books.google.nl/books?id=G49GjIXj8DMC>
- Jones, G.L., & Litovsky, R.Y. (2008). Role of masker predictability in the cocktail party problem. *The Journal of the Acoustical Society of America*, 124(6), 3818-3830. doi: 10.1121/1.2996336
- Kleinschmidt, D.F., & Jaeger, T.F. (2015). Robust Speech Perception: Recognize the Familiar, Generalize to the Similar, and Adapt to the Novel. *Psychological Review*, 122(2), 148-203. doi: 10.1037/a0038695
- Leibold, L.J., Buss, E., & Calandruccio, L. (2018). Developmental Effects in Masking Release for Speech-in-Speech Perception Due to a Target/Masker Sex Mismatch. *Ear & Hearing*, 39(5), 935-945. doi: 10.1097/AUD.0000000000000554
- Litovsky, R.Y. (2012). Spatial release from masking. *Acoustics Today*, 8(2), 18-25.



- Mattys, S.L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, 59(3), 203-243. doi: 10.1016/j.cogpsych.2009.04.001
- Mori, F., Naghsh, F.A., & Tezuka, T. (2014). The Effect of Music on the Level of Mental Concentration and its Temporal Change. *Proceedings of the 6<sup>th</sup> International Conference on Computer Supported Education*, 34-42. doi: 10.5220/0004791100340042
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5(1), 42-46. doi: 10.1111/j.1467-9280.1994.tb00612.x
- Palmeri, T.J., Goldinger, S.D., & Pisoni, D.B. (1993). Episodic Encoding of Voice Attributes and Recognition Memory for Spoken Words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2), 309-328. doi: 10.1037/0278-7393.19.2.309
- Pollack, I. (1975). Auditory informational masking. *Journal of the Acoustical Society of America*, 57, S5. doi: 10.1121/1.1995329
- Portnow, J. (2011). Gaming Languages and Language Games. In M. Adams (Red.), *From Elvish to Klingon: Exploring Invented Languages* (pp. 135-160). Geraadpleegd in april 2020, van <https://books.google.nl/books?id=fiGODwAAQBAJ>
- Qualtrics, versie 2020. Copyright © 2020 Qualtrics®, Provo, Utah, Verenigde Staten. Geraadpleegd van <https://www.qualtrics.com/>
- Russo, F.A., & Pichora-Fuller, M.K. (2008). Tune In or Tune Out: Age-Related Differences in Listening to Speech in Music. *Ear & Hearing*, 29(5), 746-760. doi: 10.1097/AUD.0b013e31817bdd1f
- Scharenborg, O., & Larson, M. (2018a). Investigating the Effect of Music and Lyrics on Spoken-Word Recognition. arXiv:1803.05058 [cs.SD]
- Scharenborg, O., & Larson, M. (2018b). The Conversation Continues: The Effect of Lyrics and Music Complexity of Background Music on Spoken-Word Recognition. In B. Yegnanarayana (Red.) *Interspeech 2018: 2-6 September 2018, Hyderabad* (pp. 2280-2284). doi: 10.21437/Interspeech.2018-1088
- Shi, L.-F., & Law, Y. (2010). Masking effects of speech and music: Does the masker's hierarchical structure matter? *International Journal of Audiology*, 49(4), 296-308. doi: 10.3109/14992020903350188
- Shih, Y.-N., Huang, R.-H., & Chiang, H.-Y. (2012). Background music: Effects on attention performance. *Work*, 42(4), 573-578. doi: 10.3233/WOR-2012-1410
- Simpson, S.A., & Cooke, M. (2005). Consonant identification in N-talker babble is a nonmonotonic function of N. *The Journal of the Acoustical Society of America*, 118(5), 2775-2778. doi: 10.1121/1.2062650
- The Sims. (2008, 18 september). *Katy Perry Hot n Cold in Simlish using Sims 2 Apartment Life* [Videobestand]. Geraadpleegd van <https://www.youtube.com/watch?v=ucYv1zX13zU&t>
- The Sims. (2009, 16 november). *Simlish Behind the Scenes* [Videobestand]. Geraadpleegd in april 2020, van <https://www.youtube.com/watch?v=yW4nfveKW5s>
- Tun, P.A., O'Kane, G., & Wingfield, A. (2002). Distraction by Competing Speech in Young and Older Adult Listeners. *Psychology and Aging*, 17(3), 453-467. doi: 10.1037/0882-7974.17.3.453
- Van Engen, K.J. (2010). Similarity and familiarity: Second language sentence recognition in first- and second-language multi-talker babble. *Speech Communication*, 52(11/12), 943-953. doi: 10.1016/j.specom.2010.05.002
- Van Engen, K.J. (2012). Speech-in-speech recognition: A training study. *Language and Cognitive Processes*, 27(7/8), 1089-1107. doi: 10.1080/01690965.2012.654644

- Van Engen, K.J., & Bradlow, A.R. (2007). Sentence recognition in native- and foreign-language multi-talker background noise. *The Journal of the Acoustical Society of America*, *121*(1), 519-526. doi: 10.1121/1.2400666
- Viswanathan, N., Kokkinakis, K., & Williams, B.T. (2018). Listeners Experience Linguistic Masking Release in Noise-Vocoded Speech-in-Speech Recognition. *Journal of Speech, Language, and Hearing Research*, *61*(2), 428-435. doi: 10.1044/2017\_JSLHR-H-17-0215
- Vouloumanos, A., Hauser, M.D., Werker, J.F., & Martin, A. (2010). The Tuning of Human Neonates' Preference for Speech. *Child Development*, *81*(2), 517-527. doi: 10.1111/j.1467-8624.2009.01412.x
- Vouloumanos, A., & Werker, J.F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Developmental Science*, *10*(2), 159-171. doi: 10.1111/j.1467-7687.2007.00549.x

## Bijlage A

In Tabel 1 zijn de volledige songteksten van het nummer Hot N Cold van Katy Perry weergegeven. Aan de linkerkant staat de originele songtekst in het Engels, en aan de rechterkant staat de vertaling in Simlish.<sup>4</sup> Voor zover bekend is de officiële Simlish vertaling, die gebruikt is tijdens de opname van het nummer, niet beschikbaar. In verschillende blogs en YouTube filmpjes is een eigen poging gedaan om de Simlish songteksten uit te schrijven, maar deze lijken allemaal in meerdere of mindere maten van elkaar te verschillen. De songteksten die hier zijn weergegeven dienen ter illustratie van Simlish, en in hoeverre deze songteksten overeenkomen of juist afwijken van de originele Engelse teksten.

**Tabel 1** – De songteksten van het nummer Hot N Cold van Katy Perry in de originele taal (Engels) en in Simlish.

Engels	Simlish
You change your mind like a girl changes clothes	Vous chikanip laka gurl chika claps
Yeah you PMS like a bitch, I would know	Va vous zee e rex laka brich, aba nup
And you always think, always speak cryptically	Om vous olweh zonk, ibu seek cryptozeek
I should know that you're no good for me	Ashanup za vous ku da vina
'Cause you're hot then you're cold	Kavahap ninap cou
You're yes then you're no	Fahit tinap blo
You're in then you're out	Vaip tinip aw
You're up then you're down	Vanip tinip taw
You're wrong when it's right	Va longe rishwai
It's black and it's white	Zi hurkinish why
We fight, we break up	Wi fout, we wi dap
We kiss, we make up	Wi kip, we ni daw
(You) you don't really wanna stay, no	(Vous) vous nu really wanna sti, oh
(You) but you don't really wanna go, oh	(Vous) vous nu really wanna cri, oh
'Cause you're hot then you're cold	Kavahap ninap cou
You're yes then you're no	Fahit tinap blo
You're in then you're out	Vaip tinip aw
You're up then you're down	Vanip tunip taw

<sup>4</sup> De teksten die hier zijn weergegeven zijn gebaseerd op de songteksten zoals verkregen van [https://sims.fandom.com/wiki/Songs\\_in\\_Simlish](https://sims.fandom.com/wiki/Songs_in_Simlish) op 1 juli 2020.

We used to be just like twins, so in sync  
The same energy now's a dead battery  
Used to laugh 'bout nothing, now you're plain boring  
I should know that you're not gonna change

'Cause you're hot then you're cold  
You're yes then you're no  
You're in then you're out  
You're up then you're down  
You're wrong when it's right  
It's black and it's white  
We fight, we break up  
We kiss, we make up  
(You) you don't really wanna stay, no  
(You) but you don't really wanna go, oh  
'Cause you're hot then you're cold  
You're yes then you're no  
You're in then you're out  
You're up then you're down

Someone, call the doctor  
Got a case of love bipolar  
Stuck on a roller coaster  
Can't get off this ride

You change your mind  
Like a girl changes clothes

'Cause you're hot then you're cold  
You're yes then you're no  
You're in then you're out  
You're up then you're down  
You're wrong when it's right  
It's black and it's white  
We fight, we break up  
We kiss, we make up

You're hot then you're cold  
You're yes then you're no  
You're in then you're out  
You're up then you're down  
You're wrong when it's right  
It's black and it's white  
We fight, we break up  
We kiss, we make up  
(You) you don't really wanna stay, no  
(You) but you don't really wanna go, oh  
'Cause you're hot then you're cold  
You're yes then you're no  
You're in then you're out  
You're up then you're down

Wa yip tu bey juk ku twip, su ba sip  
Ta si enerjonk nawza dwip batteronk  
Yiptulorp ba nothing, no yoh playy amorengk  
Ashanup za u kudi ki ye

Kawahap ninap cou  
Fahit tinap blo  
Vaip tinip aw  
Vanip tinip taw  
Va longe rishwai  
Zi hurkinish why  
Wi fout, we wi dap  
Wi kip, we ni daw  
(Vous) vous nu really wanna sti, oh  
(Vous) vous nu really wanna cri, oh  
Kawahap ninap cou  
Fahit tinap blo  
Vaip tinip aw  
Vanip tunip taw

Sugnorg, cul di docturg  
Gutta case o luv bicoler  
Stucowa roli nowster  
Can't rinup za wey

Vous chikanip  
Laka gurl chika claps

Kawahap ninap cou  
Fahit tinap blo  
Vaip tinip aw  
Vanip tinip taw  
Va longe rishwai  
Zi hurkinish why  
Wi fout, we wi dap  
Wi kip, we ni daw

Vahap ninap cou  
Fahit tinap blo  
Vaip tinip aw  
Vanip tinip taw  
Va longe rishwai  
Zi hurkinish why  
Wi fout, we wi dap  
Wi kip, we ni daw  
(Vous) vous nu really wanna sti, oh  
(Vous) vous nu really wanna cri, oh  
Kawahap ninap cou  
Fahit tinap blo  
Vaip tinip aw  
Vanip tunip taw