

**HUR FÖRMEDLAR ROBOTKARAKTÄRER I
STAR WARS EMOTION, VALENS OCH
AROUSAL GENOM ICKE-VERBALA LJUD?**

Vilka skillnader och likheter finns i
kommunikationen mellan olika
robotkaraktärer i Star Wars?

**HOW DO ROBOT CHARACTERS IN STAR
WARS CONVEY EMOTION, VALENCE, AND
AROUSAL THROUGH NONVERBAL
SOUNDS?**

What differences and similarities exist in the
communication between different robot
characters in Star Wars?

Examensarbete i medier, estetik och berättande
Grundnivå 15 högskolepoäng
Vårtermin 2026

Linus Josefsson
Axel Lindbom

Handledare: Jamie Fawcus
Examinator: Anders Sjölin

Sammanfattning

Den här studien undersöker hur robotar i *Star Wars* kommunicerar emotion genom icke-verbala ljud. Med stöd i tidigare forskning om bland annat vokala mönster, valens och arousal analyseras hur akustiska parametrar som frekvensomfång (Hz), grundton (-fo) och ljudintensitet (dB) bidrar till att förmedla emotion utan tal. Materialet består av tjugosex scener från *Star Wars: The Force Awakens*, *Star Wars: The Clone Wars* samt spelen *Star Wars Jedi: Fallen Order* och *Star Wars Jedi: Survivor*. Ljuden analyserades med spektrogram och bedömdes utifrån Russells affektmodell.

Resultatet visar att ljudets intensitet (dB) är starkt kopplad till arousal, där högre ljudnivåer indikerar en högre grad av aktivering. Tonförändringar och frekvensomfång (Hz) påverkar också hur ljud tolkas, ofta på sätt som liknar mänsklig intonation. Resultaten indikerar vidare att ljuddesign i film är mer unik, medan spel och serier ofta återanvänder ljud. Framtida studier kan använda ett bredare urval och involvera fler forskare för att fördjupa analysen.

Nyckelord: Robotar, Star Wars, Ljud, Valens, Arousal

Innehållsförteckning

1	Introduktion	1
2	Bakgrund	2
2.1	Icke-verbal kommunikation	2
2.2	Vokala Mönster	2
2.3	Valens och arousal	4
2.4	Hertz och toner	5
2.5	Star Wars och Ben Burt	5
3	Problemformulering	7
3.1	Metodbeskrivning	7
3.2	Metodproblem och begränsningar	7
3.3	Urval	8
3.4	Analysmetod	8
3.5	Avgränsning	9
4	Genomförande/Analys	9
4.1	Introduktion av materialet	10
4.2	Datainsamling och bearbetning	10
4.3	Analys	11
4.3.1	Övergripande mönster	11
4.3.2	Närvarosignalering	13
4.3.3	Uppmärksamhetssignaler	14
4.3.4	Valens	16
4.3.5	Arousal	16
4.3.6	H/L Mönster	17
4.4	Slutsats	18
5	Sammanfattning och diskussion	21
5.1	Sammanfattning	21
5.2	Diskussion	21
5.3	Samhälleliga och etiska aspekter	22
5.4	Framtida arbete	23
	Referenser	25

1 Introduktion

Den här studien undersöker hur robotar i *Star Wars* kommunicerar med icke-verbala ljud och hur de kan förmedla emotion utan ord. *Star Wars* är relevant för studien eftersom det innehåller flera robotkaraktärer som kommunicerar utan tal, där ljuddesignen spelar en central roll i hur emotion, valens och arousal uppfattas av publiken. Studien kan även bidra till kunskap och inspiration till samtida ljuddesigners som arbetar med icke-verbala robotkaraktärer.

Studien bygger på tidigare forskning om icke-verbal kommunikation, vokala mönster, valens och arousal samt akustiska egenskaper hos ljud. Studien fokuserar på akustiska parametrar såsom frekvens, grundton (fo) och ljudintensitet (dB) i syfte att analysera hur emotion kan kommuniceras utan verbalt språk.

Till studien samlades ljudklipp från tjugosex scener i *Star Wars: The Force Awakens*, *Star Wars: The Clone Wars* (S5E11) samt spelen *Star Wars Jedi Fallen Order* och *Star Wars Jedi Survivor*. Ljuden analyserades med spektrogram för att mäta grundton (-fo), frekvensomfång (Hz), ljudintensitet (dB) och H/L-mönster. Forskarna observerade även den visuella kontexten i scenerna samtidigt som de lyssnade på ljuden. Varje ljud bedömdes utifrån Russells (1980) affektmodell och dokumenterades i tabeller. Resultaten jämfördes sedan för att förstå hur robotarna kommunicerar emotion utan ord, samt likheter och skillnader mellan de olika robotarna.

2 Bakgrund

2.1 Icke-verbal kommunikation

Icke-verbal kommunikation kan definieras som all kommunikation som sker på något annat sätt än med ord, exempelvis ögonkontakt, kroppsspråk eller tonläge (Wrench, Punyanunt-Carter & Thweatt, 2025). Vidare menar Wrench et al. (2025) att majoriteten av den information som överförs mellan två eller fler personer vid en interaktion kommer från icke-verbal kommunikation. Det kan vara hastigheten på talet, tonen på rösten, ögonkontakt eller avsaknad av ögonkontakt. Hastigheten på talet och tonen på rösten är särskilt relevant för studien då de utvalda icke-verbala robotarna saknar kroppsspråk och verbalt språk.

Enligt Wrench, Punyanunt-Carter och Thweatt (2025) är icke-verbal kommunikation starkt kulturellt baserad. De menar att icke-verbal kommunikation inte är ett universellt språk utan skiljer sig mellan kulturer. Hur vi förstår icke-verbala ljud grundas i vilken kultur vi är uppväxta i. När man träder in i en ny kulturell kontext blir det därför nödvändigt att förstå och anpassa sig till lokala traditioner och normer för att korrekt kunna tolka och använda icke-verbala uttryck. Kultur är något som formas och överförs genom socialt samspel inom den grupp man tillhör, snarare än något individuellt. (Wrench, Punyanunt-Carter och Thweatt, 2025). Det är en viktig att ha i åtanke vid analys av ljudmaterial.

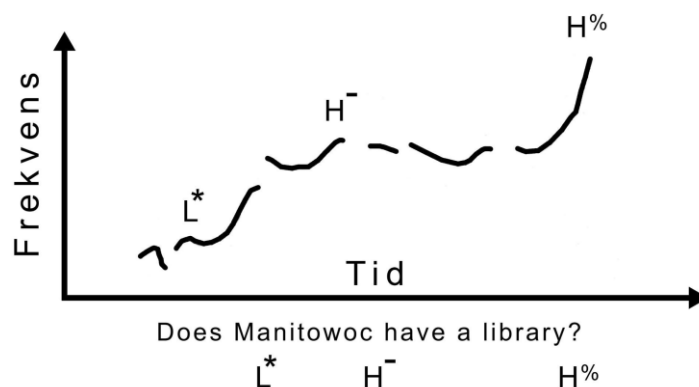
2.2 Vokala Mönster

I psykologisk forskning kopplas icke-verbala robotljud ofta till djurkommunikation, eftersom djur använder ljud för att förmedla information. Exempelvis kan stora kattdjur indikera stress med ett skrikande ljud och aggression kan identifieras med ett morrande ljud. På samma sätt som människor använder ljud för att uttrycka emotion (Bethel & Murphy, 2008). Bethel och Murphy menar att människor kan förmedla emotion genom tonläge och vokala mönster. Även om den som lyssnar inte förstår orden så kan den med hjälp av tonläge och det vokala mönstret förstå vad motpartens känslomässiga tillstånd är (Bethel & Murphy, 2008).

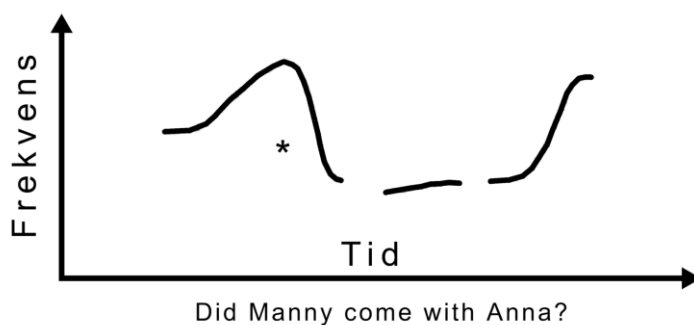
I en senare studie undersökte Crumpton och Bethel (2016) hur människor reagerar på icke-verbala robotar där ljudet manipuleras med hjälp av akustiska parametrar såsom tonförändring, varaktighet och ljudintensitet (dB). Resultatet visade att deltagarna uppfattade emotionella uttryck fast de saknade ansikte eller kroppsrörelser. Detta resultat stärker

Murphys och Bethels (2008) resonemang som argumenterar att tonläge och vokala mönster kan fungera som ett sätt för icke-verbala robotar att förmedla emotion.

För att bättre förstå tonhöjd och intonation hos människor presenterar Pierrehumbert (1980) en modell för det engelska språket, som beskriver hur rösten går upp och ner i tonhöjd när vi pratar. För att lättare förstå sig på begreppen har hon delat upp det i två kategorier: H = hög ton, L = låg ton. När vi talar sätts toner ihop till olika mönster. Vissa ord i en mening är mer betydelsefulla än andra, vilket gör att dessa ord får särskild betoning när vi talar. I slutet av en mening finns också speciella toner som visar på om vi ställer en fråga eller är klara med att prata. Modellen beskriver hur toner är organiserade i språket och hur det faktiskt låter när vi pratar. Enligt Pierrehumbert (1980) räcker det endast med två toner, hög och låg, för att förklara hur det engelska språket låter.



Figur 1. (Pierrehumbert, 1980, s. 265) illustrerar hur en fråga avslutas med en hög ton.



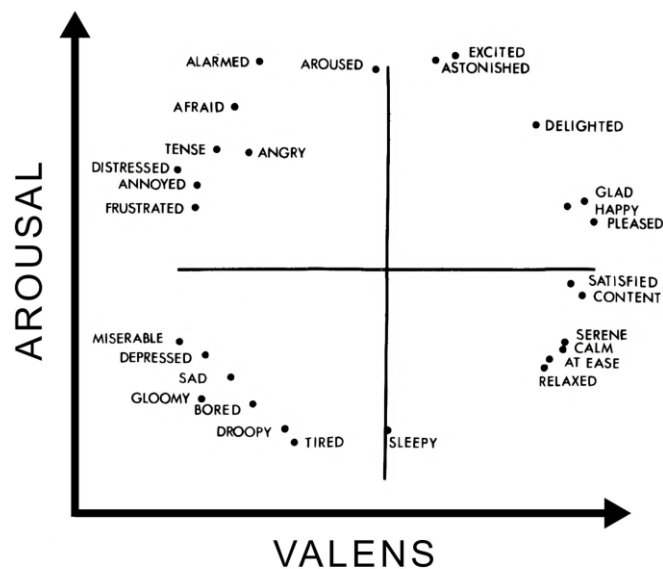
Fall rise accent, "Did Manny come with Anna?" A peak in the F0 contour occurs in the stressed syllable of "Manny" (). After coming down from peak, the contour rises slowly and finishes with a sharp rise at the end of the phrase. This type of accent is often used for simple yes/no questions.*

Figur 2. (Taylor, 1992, s. 16) markerar en viktig betoning i ordet (*).

2.3 Valens och arousal

I Russells (1980) artikel *A circumplex model of affect* introduceras en modell för att beskriva emotion utifrån två dimensioner: valens och arousal. Han presenterar ett cirkulärt affektdiagram där arousal representeras vertikalt och valens horisontellt. Valens anger om en känsla är positiv eller negativ och utgör en av de grundläggande dimensionerna i emotionsforskning. Arousal beskriver känslans energinivå, om den är låg (lugn, trött) eller hög (upphetsad, stressad). Två ljud kan alltså båda ha en positiv valens men samtidigt skilja i energinivå, något som arousal dimensionen fångar.

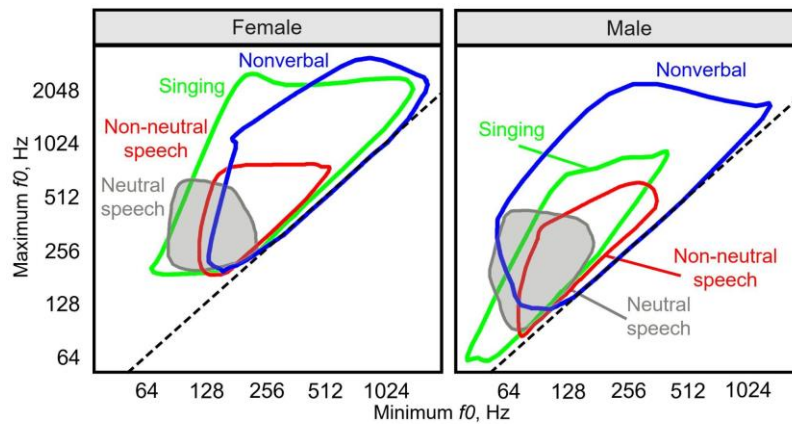
Vidare i Russells (1980) artikel presenteras 28 olika affektord som är utplacerade i ett cirkeldiagram. Positionen de har i cirkeldiagrammet påverkas av ordets valens och arousal. Genom att använda sig av Russells (1980) modell kan ljudet kodas i en tabell utifrån deras affektiva egenskaper. Först beskrivs ljudet valens, om den är positiv eller negativ, därefter om arousal är låg eller hög. Därefter kan passande affektord identifieras utifrån ljudets kombination av valens och arousal. Denna modell utgör därför en central grund för analysen av emotion i icke-verbala robotljud och möjliggör koppling till specifika affektiva tillstånd.



Figur.3 Tabellen visar valens och arousal enligt Russell (1980).

2.4 Hertz och toner

Tidigare forskning av Jee, Jeong, Kim och Kobayashi (2010) visar att robotljud inom frekvensområdet 100-1500 Hz efterliknar den mänskliga ton som används vid tal och upplevs som mest naturlig för människor (Jee et al., 2010).



Figur.4 Frekvensomfång (Hz) för män och kvinnor vid tal, sång och icke-verbala ljud (Anikin, 2023).

Anikin et al. (2023) visar i sin att naturligt mänskligt tal består av en grundton (f_0) och sedan övertoner som ligger ovanför grundtonen. Övertonerna befinner sig i snitt 2 oktaver högre upp än grundtonen. Grundtonen är den mest framträdande frekvensen i ett ljud och uppstår genom stämbandets vibrationer. Vid naturligt tal ligger medianvärdet för grundtonen hos män på 129 Hz och 215 Hz för kvinnor. Både sång och icke-verbala ljud ockuperar ett mycket högre frekvensomfång (Hz) hos båda (Anikin et al., 2023).

2.5 Star Wars och Ben Burtt

I samband med produktionen av *Star Wars: Episode IV – A New Hope* fick Ben Burtt uppdraget att skapa ljudlandskapet till ett helt universum. På den tiden var det ovanligt att någon anställdes enbart för att utveckla specifika ljudeffekter till en film. När filmen hade premiär 1977 ansågs dess ljuddesign av många som banbrytande och Burtt's arbete med robotkaraktärer som särskilt innovativt (Whittington, 2007). BD-1, R2-D2 och BB-8 är några av de icke-verbala robotar som existerar i *Star Wars* universum.

I en intervju berättar Ben Burtt att R2-D2 var den svåraste karaktären han hade arbetat med vid den tidpunkten (Star Wars, 2014). Dels för att R2-D2 saknar ansikte, mun och ögon men

också för han var en robot som skulle ha en betydande roll i filmerna. Burttts ursprungliga idé till R2-D2 var att skapa ett maskinspråk, så att lyssnaren direkt skulle förstå att ljudet kom från en robot.

“I learned to combine my voice with electronic sounds which I play on a keyboard of a synthesizer. Combining the two, I was able to get a sense of performance out of the character” (Star Wars, 2014).

För att producera R2-D2s beeps och visslingar använde Burtt en ARP 2600-synthesizer tillsammans med sin röst (Yoder & Bartmann, 2017). Jee et al. (2010) menar att R2-D2s mänskliga egenskaper framträder genom metalliska övertoner och en variation mellan höga och låga toner, vilket gör att ljudet uppfattas som emotionellt uttrycksfullt snarare än robotaktigt.

3 Problemformulering

I *Star Wars* universum kommunicerar många robotar med icke-verbala ljud, vilket är centralt för deras karaktär och hur deras emotioner uppfattas av publiken. Trots detta finns det begränsad forskning om hur dessa ljud faktiskt förmedlar emotion, valens och arousal. Problemformuleringen för denna studie är därför: *Hur förmedlar robotkaraktärer i Star Wars emotion, valens och arousal genom icke-verbala ljud? Vilka skillnader och likheter finns det i kommunikationen mellan olika robotkaraktärer i Star Wars?*

Tidigare studier har visat att akustiska parametrar, såsom grundton (-fo), frekvensomfång (Hz) och ljudintensitet (dB), kan användas för att förmedla emotioner i icke-verbala robotar (Crumpton & Bethel, 2016). Trots detta saknas detaljerade analyser av hur dessa parametrar används i *Star Wars* franchise.

En central fråga är vilka frekvensområden de icke-verbala robotarnas ljud rör sig inom och huruvida de ligger i samma register som människor. Denna studie syftar därför till att analysera de akustiska parametrarna i robotarnas ljud för att förstå hur emotion, valens och arousal förmedlas utan tal. Analysen baseras på forskarnas erfarenhet inom ljuddesign, vilket möjliggör identifiering och tolkning av subtila akustiska mönster.

3.1 Metodbeskrivning

I denna studie tillämpas en kvalitativ forskningsansats där tidigare forskning och teorier kring ämnen granskas. Utgångspunkten är en teoribaserad analys där tidigare forskning om icke-verbala kommunikation, vokala mönster och akustiska parametrar utgör det teoretiska ramverket. Med stöd från det teoretiska ramverket analyseras ljudklipp med olika icke-verbala robotar. Genom att kombinera kvalitativ tolkning av kontext och auditiva intryck med kvantitativa mätningar av ljudens egenskaper möjliggörs en mer nyanserad analys av hur emotion, valens och arousal uttrycks utan verbalt språk.

3.2 Metodproblem och begränsningar

Urvalet i denna studie är begränsat till tjugosex ljudklipp från tre olika robotkaraktärer i *Star Wars*, vilket innebär att resultatet nödvändigtvis inte kan tillämpas till andra medier som innehåller icke-verbala robotar. Materialet hämtades från filmer, serier och tv-spel, vilket kan medföra variation i

ljuddesign och återanvändning av ljud. Speciellt ljudklipp från serier och spel kan ha återanvänts, vilket kan påverka materialets variation och därmed analysens bredd. Det kan få vissa mönster att framstå som mer framträdande än vad de faktiskt är.

Urvalet baseras på forskarnas egna val av scener, vilket innebär en risk för subjektivitet. Ett annat urval av ljudklipp skulle potentiellt kunna ge andra resultat eller identifiera andra mönster. Trots dessa begränsningar möjliggör det avgränsade urvalet en fokuserad analys av icke-verbala robotar i *Star Wars*. Genom att analysera välkända robotkaraktärer kan forskarna granska ljudklippen mer noggrant, vilket ökar studiens analytiska djup och bidrar till en djupare förståelse för hur icke-verbala robotar förmedlar emotion.

3.3 Urval

Urvalet består av tjugosex scener från *Star Wars* där robotar kommunicerar icke-verbalt. Scenerna kommer från serien *Star Wars: The Clone Wars* (säsong 5, avsnitt 11), filmen *Star Wars: The Force Awakens* samt spelen *Star Wars: Jedi Fallen Order* och *Star Wars: Jedi Survivor*. Avsnittet från *The Clone Wars* är särskilt intressant då det innehåller flera sekvenser där robotarna QT-KT, U9-C4, M5-BZ och R2-D2 kommunicerar med varandra. Att flera icke-verbala robotar har en konversation mellan varandra är sällsynt i *Star Wars* och det ger en möjlighet att analysera hur kommunikationen sker mellan dem.

Fokus ligger på tre centrala robotkaraktärer: R2-D2, BD-1 och BB-8. De representerar olika generationer inom *Star Wars* och förekommer i olika medier, såsom filmer, spel och serier. Urvalet motiveras av att robotkaraktärerna saknar talat språk men ändå kommunicerar genom ljud.

3.4 Analysmetod

Ljudklippen analyseras utefter en analystabell som bygger på flera perspektiv. Valens och arousal, Anikin et al. (2023) forskning om Hz och frekvensvariation i den mänskliga rösten samt Pierrehumbert (1980) forskning om intonation i det engelska språket. Med stöd av Crumpton & Bethel (2016) samt Bethel & Murphy (2008) underlättas analysen av vokala mönster, tonalitet etc. Analystabell (figur.8) hittas i Appendix A.

Analystabellen är uppbyggd i tre delar.

1. Visuellt/kontextuell beskrivning: Ljudklippen beskrivs utifrån vad som händer i scenen och visuellt sammanhang.

2. Auditiv observation: Ljudklippen kategoriseras utifrån Russells (1980) tvådimensionella modell för valens och arousal. Även Russells 28 affektord används för att identifieras och klassificera varje ljud. Under kategorin "säkerhet" bedömer forskarna hur väl ljudet överensstämmer med valens, arousal och affektord, baserat på deras egen bedömning (se Appendix A).
3. Frekvensanalys: Ljudklippen undersöks i ett spektrogram, med fokus på frekvenser, grundton (-fo) och ljudintensitet (dB). Valet av spektrogram motiveras med dess förmåga att visuellt visa ljudets struktur och egenskaper. Frekvenser, grundton (-fo), spektrogram och ljudintensitet (dB) är centralt för syftet med denna studie. I analysen tillämpas Pierrehumberts (1980) modell för höga (H) och låga toner (L), som ett analytiskt verktyg för att tolka tonala mönster i robotljuden.

Denna metod möjliggör en systematisk analys av hur robotar i *Star Wars* kommunicerar emotion, valens och arousal genom icke-verbala ljud.

3.5 Avgränsning

Den här studien avgränsas till robotar som kommunicerar enbart med icke-verbala ljud, vilket möjliggör en fokuserad analys av deras akustiska parametrar och hur de används för att förmedla emotion. Karaktärer som C-3PO och I-11, som använder tal eller kombinerar tal och kroppsliga uttryck, exkluderas ur denna studie. Denna avgränsning skapar förutsättningar för en systematisk jämförelse mellan icke-verbala robotar.

4 Genomförande/Analys

4.1 Introduktion av materialet

Inledningsvis genomfördes en informationssökning inom *Star Wars* franchisen för att identifiera centrala robotkaraktärer. Forskarna gick igenom diverse *Star Wars* material inom tv-spel, serie och film för att hitta robotkaraktärer som passade studien. Robotkaraktärerna QT-KT, U9-C4, M5 var först inkluderade i urvalet men valdes bort då de hade för lite skärmtid i serien *The Clone Wars*, därför valdes robotarna R2-D2, BD-1 och BB-8 ut på grund av deras centrala del i *Star Wars* franchisen. När urvalet av robotar var klart genomfördes en insamling av tillhörande ljudklipp, vilket organiserades och sparades med korrekta namn och tidskoder för att möjliggöra noggrann dokumentation och analys. Totalt analyserades tjugosex olika ljudklipp från filmen *Star Wars: The Force Awakens*, tv-serien *Star Wars: The Clone Wars* (S5E11) samt spelen *Star Wars Jedi Fallen Order* och *Star Wars Jedi Survivor*. Varje observation är kopplad till en specifik scen med tillhörande tidskod som beskriver visuell kontext, ljudtyp, varaktighet samt typ av interaktion.

4.2 Datainsamling och bearbetning

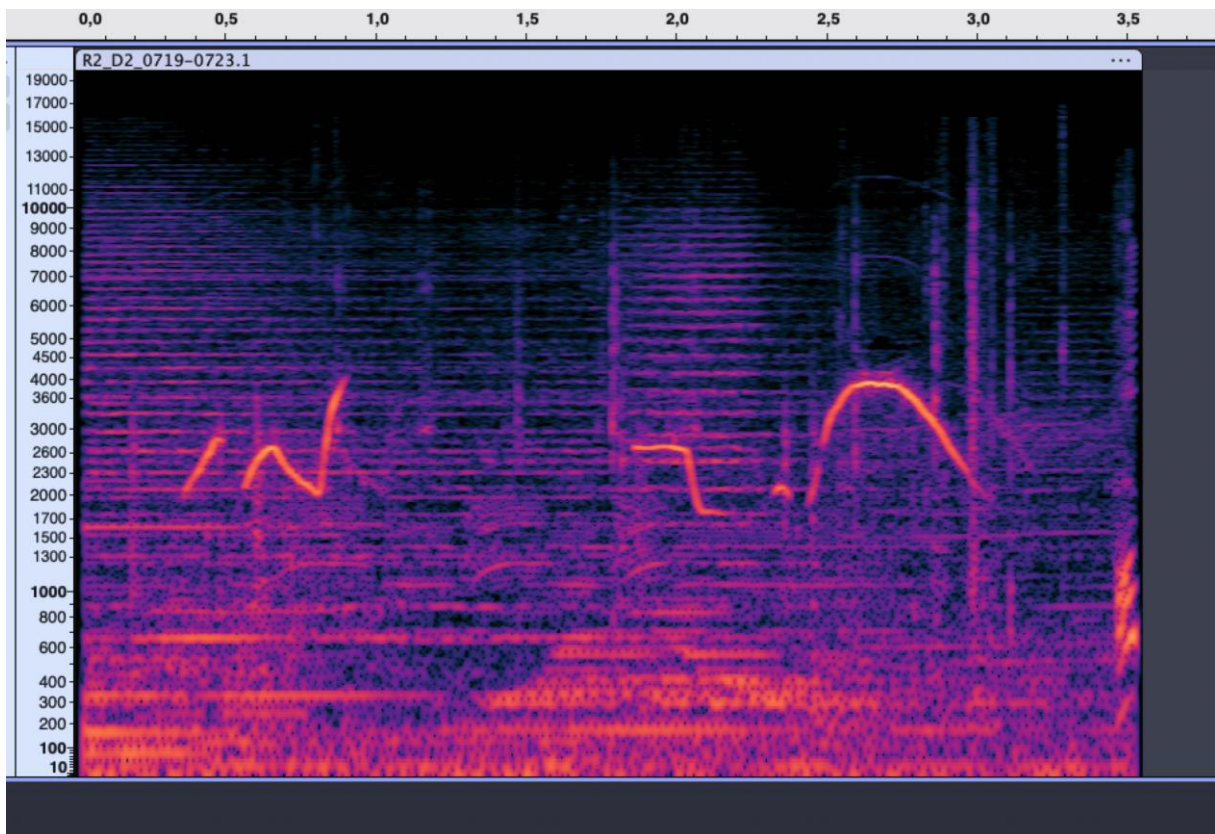
Ljudklippen analyserades systematiskt med fokus på både visuell och auditiv information. Den visuella kontexten noterades för att förstå ljudets funktion i scenen, medan den auditiva delen bedömde valens och arousal enligt Russells (1980) tvådimensionella modell. För varje ljud valdes ett av de 28 affektord som används för att identifiera emotionella uttryck. För att bedöma forskarnas säkerhet i observationerna gjordes en uppskattning på en skala från låg till hög, vilket indikerar hur väl ljudet överensstämde med affektkodningen.

Spektrogram användes för att analysera ljudklippen, vilket gjorde det möjligt att mäta frekvensomfång (Hz), grundton (-fo), ljudintensitet (dB) samt H/L mönster enligt Pierrehumberts (1980) modell. Analystabellen fylldes i av varje forskare individuellt vid separata tillfällen, för att säkerställa dubbelgranskning och minimera risken att påverka varandra. Därefter genomfördes en sammanställning av materialet, där ljudklippen granskades på nytt och diskussioner fördes om hur varje flik i tabellen skulle koda och tolkas.

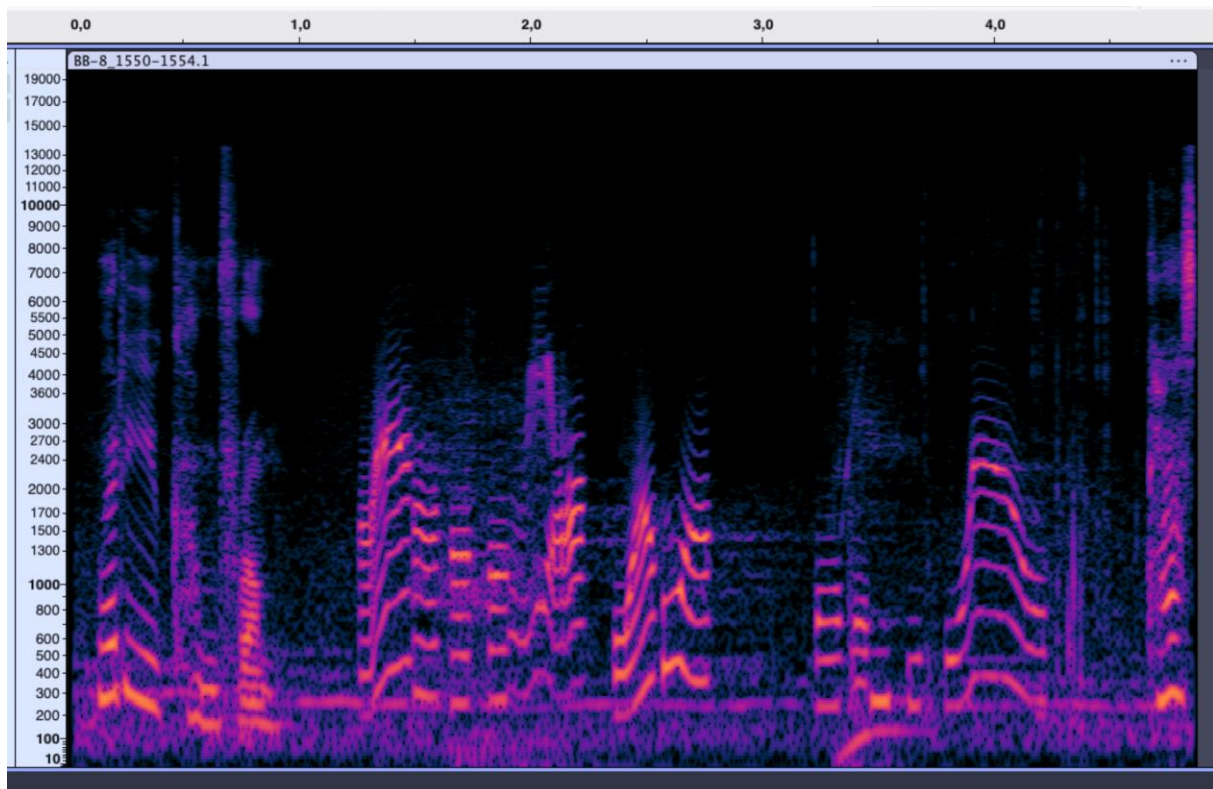
4.3 Analys

4.3.1 Övergripande mönster

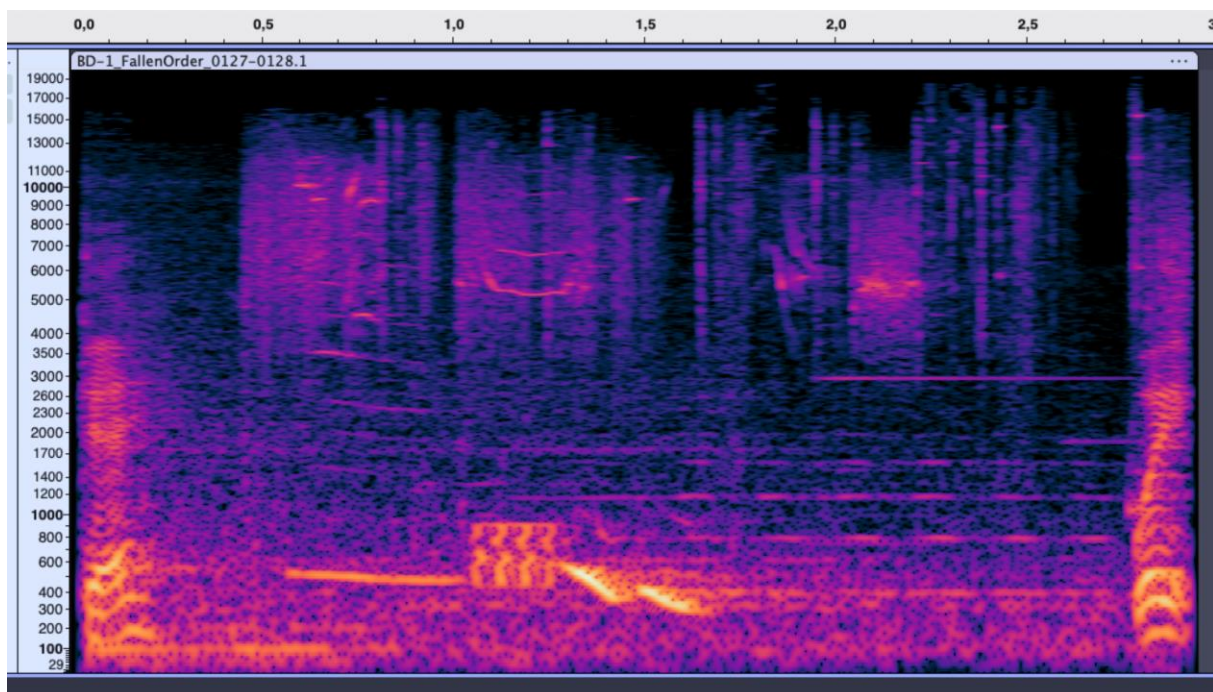
Analysen visar ett tydligt mönster mellan arousal och ljudintensiteten (dB) i robotarnas vokala mönster. Ljudintensitet (dB) syftar på robotens ljudintensitet och inte den totala ljudnivån i scenen. Resultatet visar att i cirka 80,8% av fallen korrelerar hög arousal med hög ljudintensitet och låg arousal med låg ljudintensitet (dB). Detta kan observeras i spektrogram (se Figur 8 & 9) där hög arousal sammanfaller med tydligare intensitet i ljudet och större variation i frekvensomfång.



Figur 8. R2D2, S5E11 (*The Clone Wars*), 00:07:19-00:07:23, ljudklipp nr 04. En tydlig och intensiv signal av R2-D2 med hög ljudintensitet (dB) och ett brett frekvensomfång. Snabba variationer i tonhöjd, vilket indikerar hög arousal och uppmärksamhetssignal.



Figur 9. BB-8, *Star Wars: The Force Awakens*, 00:15:50-00:15:54, ljudklipp nr 14. Ett varierande, dynamiskt ljud från BB-8 med ett relativt brett frekvensomfång. Variationen i frekvens och tydliga förändringar i grundton tyder på hög arousal och ett lekfullt eller socialt uttryck.



Figur 10. BD-1, *Jedi Fallen Order*, 01:27 - 01:28, ljudklipp nr 19. Ett subtilt och lågintensivt ljud från BD-1 med ett begränsat frekvensomfång. Ljudet har mindre variation i tonhöjd vilket motiverar låg arousal.

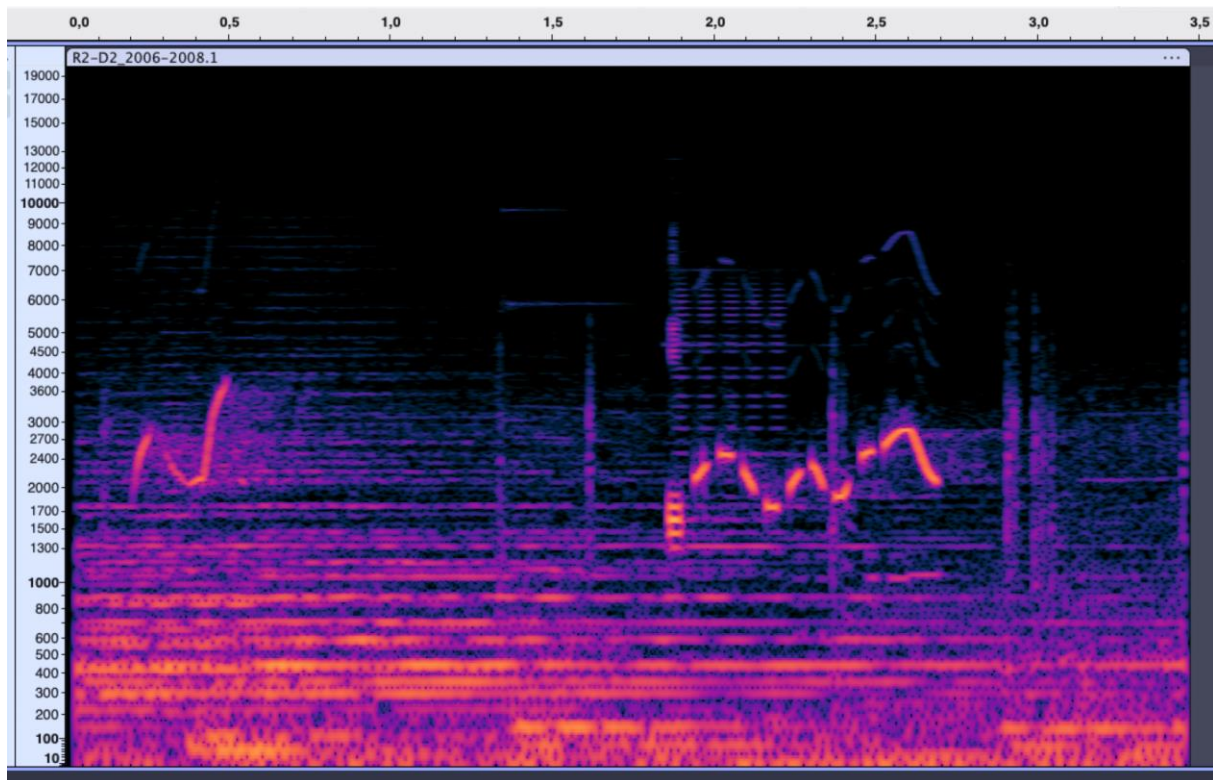
I de återstående 19.2% av observationerna följer inte arousal och ljudintensitet (dB) samma mönster. Ett exempel på detta går att se i Appendix B, ljudklipp nr 03. Resultatet indikerar istället att grundton (-fo) och frekvensomfång (Hz) spelar en större roll. Högre frekvenser eller större variationer i tonhöjd kan göra att ett ljud uppfattas som mer dominant, vilket innebär att ljudet kan upplevas som kraftfullt även när själva ljudintensiteten är låg.

R2-D2 visar ofta upp skillnader i arousal och ljudintensitet (dB), särskilt i gruppkommunikation där han måste använda sig av ljud för att dra till sig uppmärksamhet. BB-8 uppvisar exempel på hög arousal och hög ljudintensitet (dB) i sociala och kooperativa situationer. Medan BD-1 främst producerar inåtriktade och subtila ljud, alltså låg arousal och låg ljudintensitet (dB). Tidigare forskning av Jee et al. (2010) visar att ljudintensitet (dB), ton och tempo kan användas för att förmedla emotion hos sociala robotar, vilket kan relatera till variationerna i R2-D2s ljudintensitet. Skillnaderna mellan robotarna blir tydligare i spektrogrammet (se figur 8, figur 9 & figur 10). R2-D2 och BB-8 har bredare frekvensvariationer och högre ljudintensitet (dB) jämfört med BD-1 som är begränsat och koncentrerat.

4.3.2 Närvarosignalering

En tydlig egenskap hos alla robotkaraktärer är deras användning av ljud för att signalera närvaro i rummet. Michel Chion (2019) introducerar begreppet *acousmatic sound*, vilket syftar på ljud vars källa inte syns i bild men som ändå upplevs som en del av filmens värld. Chion (2019) förklarar att *acousmatic sound* har flera viktiga funktioner i film. En av dessa är att fungera som en närvarosignal. Genom att låta en karaktär höras utan att synas etableras karaktärens existens i scenen. På så sätt förstår tittaren att något eller någon är närvarande, även utan visuell bekräftelse, vilket gör ljudet till ett effektivt verktyg för att introducera eller upprätthålla element i berättelsen.

Ett tydligt exempel är när R2-D2 (se Appendix A, ljudklipp nr 09) deltar i en scen så hörs en auditiv signal, en slags presentation om att han är med och deltar. Ibland hörs han utan att synas i bild men med hjälp av ljud signalerar han sin närvaro i scenen. Samma fenomen gäller för de andra robotkaraktärerna, där varje robot har en unik ljudsignal. Det gör det enklare att särskilja dem vid gruppkonversationer.



Figur 5. R2D2, S5E11 (*The Clone Wars*), 00:20:06-00:20:08, ljudklipp nr 09. R2-D2 ger en kort men tydlig signal och med varierande frekvensomfång. Ljudet visar att ett dynamiskt tonmönster som snabbt drar till sig uppmärksamhet vilket kan tolkas som en aktiv närvarosignal. Med medelhög arousal och neutral till svag positiv valens.

4.3.3 Uppmärksamhetssignaler

Uppmärksamhetssignaler, även benämnda som uppmärksamhetskallande (se Appendix A), används av robotarna när de vill dra till sig uppmärksamhet från andra karaktärer. Detta förekommer som ett återkommande mönster i materialet. Analysen visar att med dessa signaler ofta kännetecknas av hög ljudintensitet (dB) tillsammans med hög arousal, vilket gör att de framstår som mer framträdande i ljudbilden.

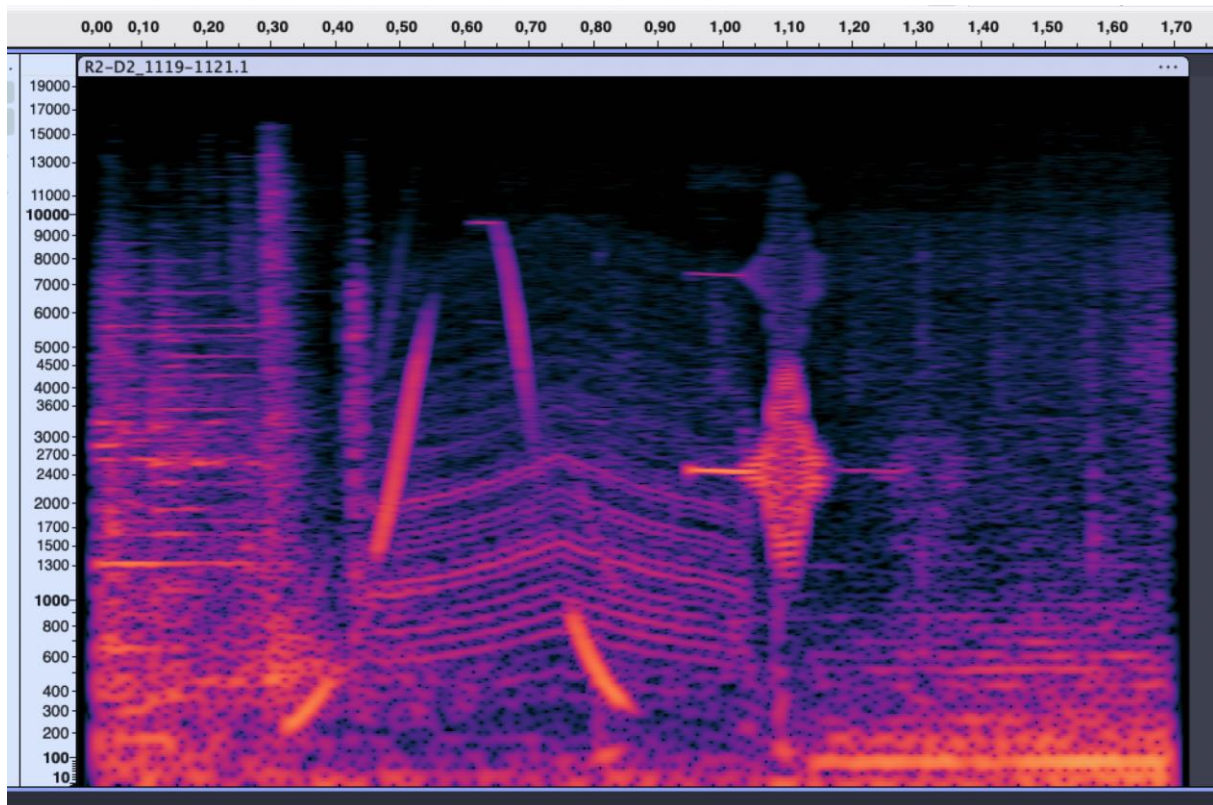
I S5E11 i *The Clone Wars* har de fyra olika robotarna distinkta ljud, som vi kallar Beeps & Boops, vilket kan hjälpa lyssnaren att urskilja de olika robotarna. Det finns olika robot modeller i *Star Wars* universumet och i detta avsnittet är två av de fyra av samma modell. R2-D2 och QT-KT är en klass-2 droid som tillhör *R2-Series Astromechs*. När de kommunicerar har de liknande Beeps & Boops och det blir lätt förvirrande för lyssnaren om vem som faktiskt pratar om fokuset inte är på roboten som talar.

BD-1 har svagare och kortare signaler (se figur 10) jämfört med R2-D2 (se figur 11) och

använder sig av uppmärksamhetssignaler allt mer sällan. Trots detta kan lyssnaren förstå vad BD-1 vill förmedla, vilket tyder på att även subtila uttryck kan kommuniceras effektivt.

R2-D2 är den som använder sig mest av uppmärksamhetssignaler i de utvalda scenerna, både med hög arousal (se Appendix A) och hög ljudintensitet (dB) (se Appendix A). På grund av den höga ljudintensiteten (dB) och arousal så uppfattas R2-D2s signaler som mer akuta och riktade mot omgivningen. R2-D2 rör sig också brett i frekvensområdet (Hz), vilket bidrar till att hans icke-verbala ljud får en mer akut och fokus centrerad känsla.

BB-8 uppvisar ett liknande mönster i sociala sammanhang, där hög arousal och hög ljudintensitet (dB) används för att locka till sig uppmärksamhet. Ett brett frekvensomfång och snabba förändringar i ton anses enligt Crumpton & Bethel (2016) underlätta emotionell kommunikation, vilket överensstämmer med observationerna av både R2-D2 och BB-8.



Figur 11. R2D2, S5E11 (*The Clone Wars*), 00:11:19-00:11:21, ljudklipp nr 06. En tydlig signal med hög ljudintensitet (dB) samt ett brett frekvensomfång. R2-D2 visar snabba variationer i tonhöjd vilket bidrar till hans uppmärksamhetskallande uttryck. Spektrogrammet stödjer observationen att R2-D2 använder sig av starkare och mer riktade signaler jämfört med BD-1.

4.3.4 Valens

I denna studie används begreppet valens utifrån Russells (1980) modell. Valens beskriver om ett känslotillstånd är positivt eller negativt. Vid situationer av gruppkommunikation mellan BB-8 och andra karaktärer kan vi se att BB-8 ofta har en positiv valens. Det kan vi inte se lika tydligt i icke-sociala situationer med BB-8. Det kan bero på att BB-8 finner ett större behov att visa högre energi och positiv valens för att fånga andras uppmärksamhet (se figur 9).

Positiv valens förekommer nästan alltid i sociala sammanhang. I analystabellen (se Appendix A) framgår det att positiv valens uppträder i cirka 70.8% av de sociala situationerna. Negativ valens är vanligare när robotarna är ensamma, när de uttrycker inåtriktad affekt eller reagerar på miljön de befinner sig i (icke-social). Negativ valens uppstår i icke sociala situationer 75% av fallen. Kommunikation mellan de olika robotarna eller mellan andra karaktärer i *Star Wars* har större chans att ha en positiv valens, det är vanligare att en positiv emotion uttrycks då. Medan negativ emotion är kopplad till ensamhet eller introspektiva situationer.

Ett återkommande mönster är att positiv valens ofta sammanfaller med ett bredare frekvensomfång (Hz). Exempelvis har BB-8 vid positiv valens ett genomsnittligt frekvensomfång (Hz) på $\approx 186 - 5420$ Hz, jämfört med $\approx 178 - 4655$ Hz vid negativ valens. BD-1 visar samma trend, med $\approx 250 - 2891$ Hz för positiv valens och $\approx 278 - 1194$ Hz för negativ valens. Större variation i frekvensomfång (Hz) kan tyda på högre emotionell uttrycksfullhet, samt att robotarna är mer dynamiska vid positiv valens.

4.3.5 Arousal

I denna studie används begreppet arousal som är baserat på Russells (1980) modell som beskriver hur stark eller intensiv en känsla är, från låg (nedstämd eller oro) till hög (upprymd eller entusiasm). I analysen tillämpas detta på robotarnas vokala ljud, arousal bedöms av hur intensiv eller dynamisk robotens ljud är. Hög arousal förknippas oftast med höga energetiska ljud och låg arousal till mer subtila och lågmälda uttryck (Russell, 1980).

Genom att relatera Russells (1980) modell till det analyserade materialet framgår att robotarnas vokala signaler uppvisar varierande nivåer av emotionell aktivering. Resultatet visar att arousal tydligt syns i de akustiska parametrarna. Detta överensstämmer med tidigare forskning som visar att parametrar som intensitet, frekvens och variation i tonhöjd används för att förmedla emotion (Zhang & Fitters, 2023).

Observationer från analystabellen (se Appendix A) indikerar att negativa emotionella tillstånd

såsom nedstämdhet och uppgivenhet, ofta sammanfaller med låg arousal (se figur 10). Motsatsen förekommer vid positivt tillstånd, exempelvis glädje eller beslutsamhet, ofta sammanfaller med hög arousal. Hög arousal förekommer också vid sociala interaktioner (se figur 9), vilket tyder på att robotarnas ljud blir mer intensiva i situationer där de kommunicerar med andra.

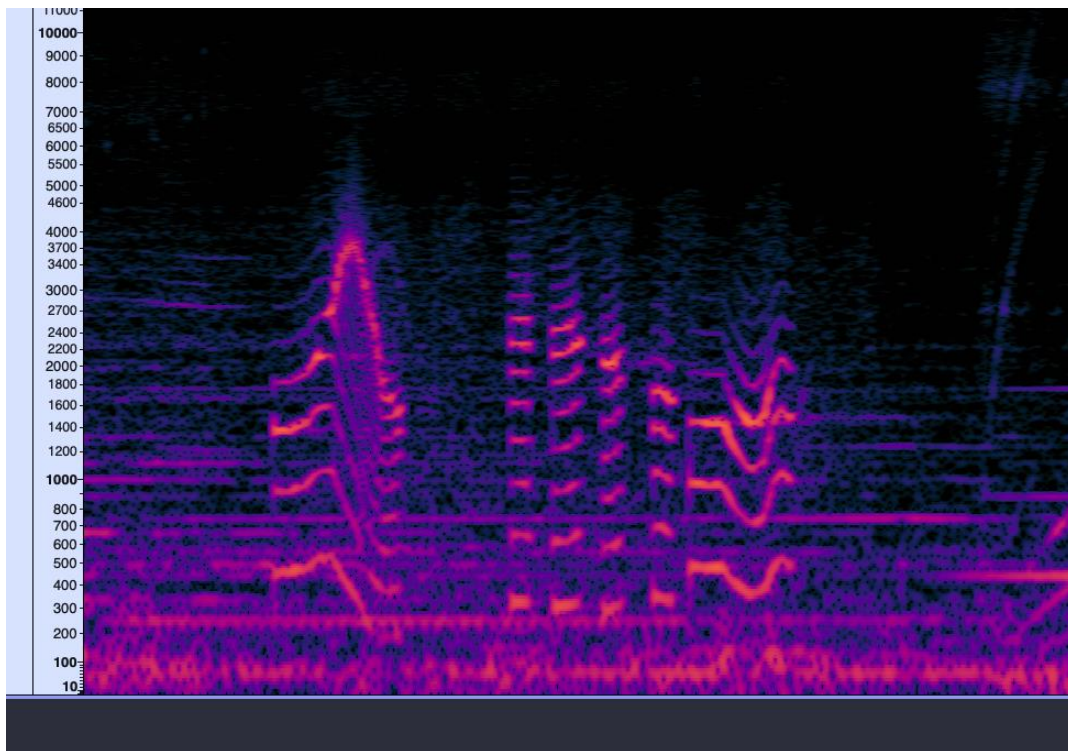
R2-D2 visar flera typer av kommunikation med hög arousal (se figur 11), där hans ljud varierar mer jämfört med andra robotar. Han använder starka och uppmärksammande ljud när han kommunicerar i grupp eller reagerar på andra varelser. R2-D2s ljud fungerar som reaktiva och kommunikativa signaler i många situationer.

BD-1 visar i 62.5% av fallen i hög arousal medan resterande 37.5%, låg arousal. Det kan tyda på att BD-1 använder sig av mer lågmälda och svaga ljud för att uttrycka sitt emotionella stadiet. I samtliga fall där låg arousal förekommer används också inåtriktade uttryck, såsom nedstämdhet, uppgivenhet eller undran (se Appendix A).

4.3.6 H/L Mönster

Ett återkommande mönster i H/L analysen är stigande eller bågformat mönster, exempelvis L-H eller L-H-L (se Appendix A). Alltså börjar robotljuden på en låg frekvens för att sedan stiga till en högre frekvens och ibland går ljuden från hög till låg frekvens. Se exempel i figur 7 på ett L-H och sedan ett bågformat L-H-L mönster. De här tonförändringarna bidrar till att skapa emotionella och dynamiska ljud. Om vi ser till Pierrehumbert (1980) modell för människans talmönster kan vi se likheter hos icke-verbala robotar, som också använder variation i höga och låga toner för att skapa dynamik i sina uttryck. Deras vokala signaler kan liknas med det engelska språket, vilket kan bidra till ökad igenkänning, trots avsaknaden av ett konventionellt språk.

Vidare kan dessa mönster kopplas till variationer i grundton (-fo). Där stigande frekvenser, från lågt till högt, är en stigande grundton (-fo) som bidrar till ett mer dynamiskt uttryck. Detta kan särskilt observeras i ljud med hög arousal, där variationer får en större påverkan på grundtonen (-fo) och det bidrar till ett starkare uttryck.



Figur.7 BB-8, *Star Wars: The Force Awakens*, 01:23:40-01:23:42, ljudklipp nr 17. BB-8 visar ett bågformat mönster (L-H-L) där frekvensen stiger för att sedan falla igen. Tydligt variation i grundton (-fo) vilket bidrar till ett mer dynamiskt och uttrycksfullt ljud.

Vid auditiv observation uppfattas figur.7 som en fråga. När det adderas till visuell kontext från scenen stöds detta. I den aktuella scenen introducerades BB-8 till R2-D2. Spektrogrammet i figur 7 visar ett mönster som liknar det i figur 1, där Pierrehumbert (1980, s. 265) beskriver hur frågor ofta avslutas med en stigande ton. Eftersom mönstret överensstämmer kan det indikera att BB-8:s ljud i klippet har en frågande karaktär.

4.4 Slutsats

Studien har undersökt icke-verbala robotar och hur de används för att uttrycka emotion i film, serier och tv-spel. Robotarna som analyserades var R2-D2, BB-8 och BD-1. Studien har genom analys, undersökt ljudintensitet (dB), grundton (-fo), frekvensomfång (Hz) och vokala mönster. Studiens resultat visar att flera olika vokala mönster kan ses när de icke-verbala robotar analyseras. Liknande mönster hittas i tidigare forskning om den mänskliga rösten.

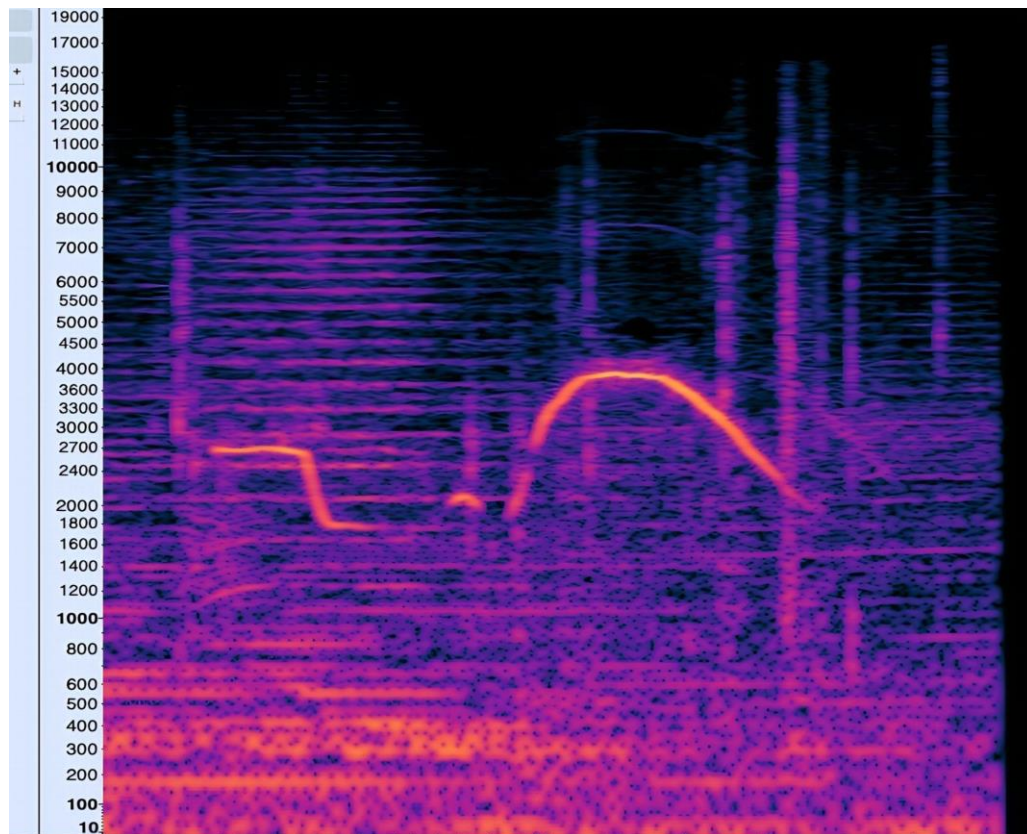
Resultatet tyder på att ljudintensitet (dB) är starkt kopplat till hur arousal upplevs. Detta framhävs extra tydligt i spektrogrammet (se figur 8-11) där hög arousal allt mer sammanfaller med hög ljudintensitet (dB) och mer variationer i frekvens. I majoriteten av observationerna så sammanfaller hög arousal med hög ljudintensitet. Medan låg arousal ofta sammanfaller med

låg ljudintensitet (dB). Det tyder på att ljudintensitet (dB) är en viktig indikation vid emotionellt uttryck i icke-verbala robotar.

Vidare kan vi se samband mellan hur tonförändring har en stor inverkan i hur ett ljud upplevs. Observera figur 6 och 7 där stigande och bågformade tonmönster (L-H, L-H-L) framhäver att variation i grundton (-fo) skapar ett mer dynamiskt och kommunikativt uttryck. Det överensstämmer med tidigare forskning på ämnet (Pierrehumbert, 1980, s 265), vilket visar att om en mening som avslutas med en hög eller låg ton är avgörande för hur vi människor uppfattar intentionen bakom det någon säger. Vi har sett en tendens till samma företeelse hos de robotar som granskats. Exempelvis så avslutas vissa ljud med tydligt hög ton, vilket överensstämmer med Pierrehumbert (1980) forskning om intention. När en mening avslutas med hög ton ställs troligtvis en fråga. Crumpton och Bethel (2016) och Bethel och Murphy (2008) syftar på att variation i tonläge påverkar hur emotion uppfattas.

Studien visar även hur robotarnas ljud varierar beroende på social kontext. Positiv valens tyder ofta på att roboten interagerar med andra karaktärer, medan negativ valens uppstår mer frekvent i icke-sociala situationer när roboten är ensam. Studien visar också att bredare frekvensomfång (Hz) tillkommer oftare vid positiv valens och större känslouttryck. De tre robotarna demonstrerar tydligt olika sätt att uttrycka sig på. R2-D2 använder ofta starka, uppmärksamhetskallande signaler med stort frekvensomfång (Hz), medan BB-8 visar lägre intensitet (dB) i sociala sammanhang. BD-1 använder generellt mer subtila och lågintensiva ljud.

Slutligen visar analysen att robotljud ofta uppvisar likheter med mänskligt intonationsmönster, exempelvis stigande eller bågformade tonsekvenser. Detta kan hjälpa oss människor att tolka robotarnas emotionella uttryck fast de inte kommunicerar med ord. Analysen visar också att robotarnas vokala uttryck härmar människans vokala mönster. Samtidigt indikerar studiens analys att robotarnas emotionella uttryck varierar, de är överdrivna och simplificerade för lyssnarens skull. Deras emotionella tillstånd förstärks genom tydligare variationer i ljudparametrar. Exempelvis kan högre ljudintensitet (dB) tydligt signalera hög arousal, vilket gör att graden av emotionell aktivering uppfattas snabbare och mer direkt.



Figur 6. R2D2, S5E11 (*The Clone Wars*), 00:06:32-00:06:40, ljudklipp nr 02. Stigande tonmönster (L-H), frekvensen rör sig från låg till hög. Detta indikerar variation i grundton (-fo), bidrar till ett mer dynamiskt och uttrycksfullt ljud.

5 Sammanfattning och diskussion

5.1 Sammanfattning

I denna studie undersöks hur icke-verbala robotar visar emotion genom vokala parametrar såsom grundton (-fo), frekvensomfång (Hz), ljudintensitet (dB) och arousal. Med utgångspunkt i tidigare forskning om robotars icke-verbala kommunikation, arousal och valens samt vokala mönster analyserades tre olika robotar från tre genrer i *Star Wars*. Studien kategoriserar olika emotionella tillstånd, social kontext och akustiska egenskaper. Resultaten visar tydliga samband mellan arousal och ljudintensitet (dB), mönster i frekvenser, till exempel bågformade eller stigande. Analysen indikerar också att ljuddesign ofta förenklar emotionella uttryck för att lyssnaren lättare ska förstå indikationen av roboten. Denna studie bidrar till ökad förståelse om icke-verbal kommunikation, hur emotion kan förmedlas och den kan tillämpas för framtida arbeten inom ämnet.

5.2 Diskussion

Som Wrench, Punyanunt-Carter och Thweatt (2025) konstaterar omfattar icke-verbal kommunikation mycket mer än bara ord. Det inkluderar bland annat ögonkontakt, kroppsspråk och tonläge. Utan studiens visuella material som kontext är det möjligt att resultatet hade sett annorlunda ut. Den visuella informationen bidrog till att tydliggöra situationen och typen av interaktion i varje ljudklipp (se Appendix A). Utan denna kontext hade troligtvis den deskriptiva delen av analysen sett annorlunda ut, vilket i sin tur hade kunnat leda till alternativa tolkningar av samma ljud.

Samtidigt visar tidigare forskning att emotionella uttryck kan uppfattas även utan visuella signaler. Murphys och Bethels (2008) fann i sin studie att människor kunde identifiera emotionella uttryck hos robotar trots att de saknade ansikte eller kroppsrörelser. Det tyder på att auditiva signaler kan bära betydande emotionell information. Samtidigt är det rimligt att anta att resultatet kan påverkas om analysen enbart baseras på auditivt material eller en kombination av auditivt och visuellt material.

Studien visar att negativa emotionella uttryck, exempelvis ledsamhet och uppgivenhet ofta indikerar låg arousal. Medan hög arousal visar emotion som glädje eller beslutsamhet. Enligt Russell (1980) kan hög och låg arousal ske vid både positiva och negativa emotioner. Därför används de tjugoåtta affektorden som en tydligare indikation på vilken emotionella karaktär

det specifika ljudet har. Detta kan betyda att när ljuddesigners skapar ljud till icke-verbala robotar så förenklas de emotionella uttrycken för tittarens skull, där låg arousal ger negativ valens och hög arousal förknippas med hög energi, högre ljudintensitet (dB) eller positiv valens.

Detta kan kopplas till Anikin et al. (2023) artikel om vokala mönster, där de menar att variation i grundton (-fo) och frekvensomfång (Hz) används för att förmedla emotion. På bilden (figur.4) så framgår det att emotionella ljud har ett bredare frekvensomfång (Hz), vilket stödjer denna studiens resultat att med högre arousal så framgår även större dynamik i ljuden och med lägre arousal så är det mindre dynamik. Robotljuden använder sig av tydliga variationer för att kommunicera emotion.

5.3 Samhälleliga och etiska aspekter

Icke-verbal kommunikation utgör en central del för hur människor tolkar emotionella uttryck i sociala interaktioner. Enligt Wrench, Punyanunt-Carter och Thweatt (2025) är icke-verbal kommunikation inte ett universellt språk, utan dess betydelse kan variera mellan olika kulturella sammanhang. Detta innebär att samma ljud eller uttryck kan tolkas på olika sätt beroende på mottagarens kulturella bakgrund och erfarenheter. Utifrån studiens perspektiv kan tolkningen av robotarnas icke-verbala ljud potentiellt variera beroende på mottagarens kulturella bakgrund.

Ur ett etiskt perspektiv har ljuddesigners ett viktigt ansvar i hur akustiska parametrar används, särskilt i områden där människor samspekar med teknik, till exempel sociala robotar. Eftersom ljud kan användas för att dra till sig uppmärksamhet eller förstärka emotionella signaler finns det en risk att användare kan påverkas på ett oavsiktligt sätt. Ett exempel är om en robot använder sig av hög ljudintensitet (dB) och snabba förändringar i tonhöjd enbart för att framstå som uttrycksfull kan det påverka hur mottagaren tolkar situationen. Då blir frågan om var gränsen mellan kommunikation och möjligtvis manipulation i ljuddesign går. Detta blir särskilt relevant i utvecklingen av sociala robotar, där förtroendet och upplevelsen för användare måste prioriteras och ljuddesignern påverkas direkt i och med det.

I denna uppsats analyseras ljudklipp från *Star Wars* med hjälp av spektrogram och bilder används som illustrativt stöd. Detta väcker frågor kring upphovsrätt och etisk användning av skyddat material i akademiska sammanhang. *Star Wars* är ett kommersiellt verk med starkt rättighetsskydd, vilket innebär att både ljud och bild omfattas av upphovsrätt. Samtidigt finns det undantag som möjliggör användning av material i forsknings- och utbildningssyfte. Detta

förutsätter att användningen är motiverad, begränsad och korrekt källhänvisad. Materialet används med stöd av citaträtten (Lag (1960:729), 22 §).

I denna studie används materialet enbart i analytiskt syfte, där ljudklippen inte reproduceras i sin helhet utan studeras och beskrivs inom ramen för en vetenskaplig analys. Bilderna inkluderas för att stödja resonemanget och öka förståelsen för det som analyseras. Ur ett etiskt perspektiv är det centralt att respektera upphovspersonernas rättigheter samt att tydligt redovisa källor. Samtidigt belyser användningen av materialet en bredare samhällsfråga om balansen mellan skydd av kreativa verk och möjligheten att bedriva forskning och kunskapsutveckling.

5.4 Framtida arbete

För framtida arbeten finns flera möjliga vägar att vidareutveckla analysen av icke-verbala robotljud i *Star Wars*. Ett nytt urval av material skulle kunna begränsas till långfilmer, där ljuden ofta har en mer unik och konsekvent design för varje scen. Tidigare observationer visar att ljud från serier och spel (t.ex. R2-D2 och BD-1) återanvänds vid olika tidpunkter, vilket innebär att ljuden inte alltid är nya eller specifika. Genom att analysera endast långfilmer blir det lättare att identifiera mönster i hur akustiska parametrar används för att uttrycka emotion, eftersom samma ljuddesignsprinciper och arbetsprocesser används konsekvent över scenerna.

Utöver R2-D2, BD-1 och BB-8 kan framtida studier inkludera mindre framträdande robotar, såsom QT-KT, U9-C4 och M5-BZ. Även om dessa karaktärer är mindre frekventa i filmer och serier, använder de liknande akustiska parametrar för att kommunicera emotion. Såsom variationer i tonhöjd, frekvensomfång och ljudintensitet. Att jämföra robotar med olika nivåer av komplexitet och social interaktion kan ge insikter i hur principerna för icke-verbal kommunikation appliceras på olika karaktärer inom samma universum.

En annan möjlig utveckling är att inkludera fler deltagare i analysen. Denna studie genomfördes av två forskare med erfarenhet av ljuddesign, men framtida studier kan involvera personer med varierande bakgrund och erfarenhet. Allt från professionella ljuddesigners till personer utan någon tidigare erfarenhet av ljudproduktion, samt deltagare i olika åldrar och från olika kulturella miljöer. Genom att låta ett sådant bredare urval analysera samma robotljud med samma metodik kan man undersöka om upplevelsen av valens, arousal och emotion är konsekvent mellan olika grupper. Om resultaten visar liknande mönster hos olika deltagare ökar studiens generaliserbarhet, och det blir möjligt att dra slutsatser om hur akustiska parametrar uppfattas i bredare sammanhang.

Samtidigt kan variationer i tolkning ge nya insikter om hur emotion och akustiska signaler samspelar beroende på individuella skillnader, kulturell bakgrund eller tidigare erfarenhet av robotar och ljud. Att testa upptäckterna på en bredare publik skulle därför inte bara bekräfta tidigare resultat, utan även kunna bidra till nya perspektiv på hur icke-verbala robotar upplevs och hur deras ljuddesign påverkar emotionell tolkning.

Sammanfattningsvis öppnar framtida studier upp för mer konsekvent materialurval, bredare deltagargrupper och inkludering av fler robotkaraktärer. Detta kan bidra till en mer mångsidig och generaliserbar förståelse av hur akustiska parametrar används för att uttrycka emotion i icke-verbala robotar, samt hur dessa uppfattas av olika typer av observatörer.

Referenser

- Anikin, Andrey, Lima, Carlos F., Mothes, Oliver, Ramezanpour, Hamed & Persson, Tomas (2023). *Beyond speech: Exploring diversity in the human voice*. *iScience*, 26(11), 108204. doi:10.1016/j.isci.2023.108204
- Bethel, Cindy L. & Murphy, Robin R. (2008). *Survey of non-facial/non-verbal affective expressions for appearance-constrained robots*. *IEEE Transactions on Systems, Man, and Cybernetics: Part C*, 38(1), s. 83–92. doi:10.1109/TSMCC.2007.905845
- Chion, Michel (2019). *Audio-Vision: Sound on Screen* (2. uppl.). New York: Columbia University Press.
- Crumpton, Joe & Bethel, Cindy L. (2016). *A survey of using vocal prosody to convey emotion in robot speech*. *International Journal of Social Robotics*, 8(2), s. 271–285. doi:10.1007/s12369-015-0329-4
- Jee, Eun-Sook, Jeong, Yong-Jeon, Kim, Chong Hui & Kobayashi, Hisato (2010). *Sound design for emotion and intention expression of socially interactive robots*. *Intelligent Service Robotics*, 3(3), s. 199–206. doi:10.1007/s11370-010-0070-7
- Jørgensen, Jonas & Christiansen, Mads Bering (2021). *The sounds of softness: designing sound for human-soft robot interaction*. *Frontiers in Robotics and AI*, 8, 674121. doi:10.3389/frobt.2021.674121
- Lag (1960:729) om upphovsrätt till litterära och konstnärliga verk. (1960) Tillgänglig på: <https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/lag-1960729-om-upphovsratt-till-litterara-sfs-1960-729/> (Hämtad: 28 mars 2026).
- Pierrehumbert, Janet Breckenridge (1980). *The phonology and phonetics of English intonation*. PhD thesis. Massachusetts Institute of Technology. Tillgänglig på: <https://dspace.mit.edu/bitstream/handle/1721.1/16065/07492108-MIT.pdf> (Hämtad: 18 februari 2026).
- Russell, James A. (1980). *A circumplex model of affect*. *Journal of Personality and Social Psychology*, 39(6), s. 1161–1178. doi:10.1037/h0077714
- Star Wars, (2014). *Ben Burtt Interview: R2-D2* [video]. YouTube. Tillgänglig på: <https://www.youtube.com/watch?v=eUwnFYBPMIU> (Hämtad: 26 januari 2026).
- Taylor, Paul Alexander (1992). *A phonetic model of English intonation*. PhD thesis. University of Edinburgh. Tillgänglig på: <https://era.ed.ac.uk/items/2d8790cd-47dc-4obe-b3dc-0c860246e694> (Hämtad: 18 februari 2026).

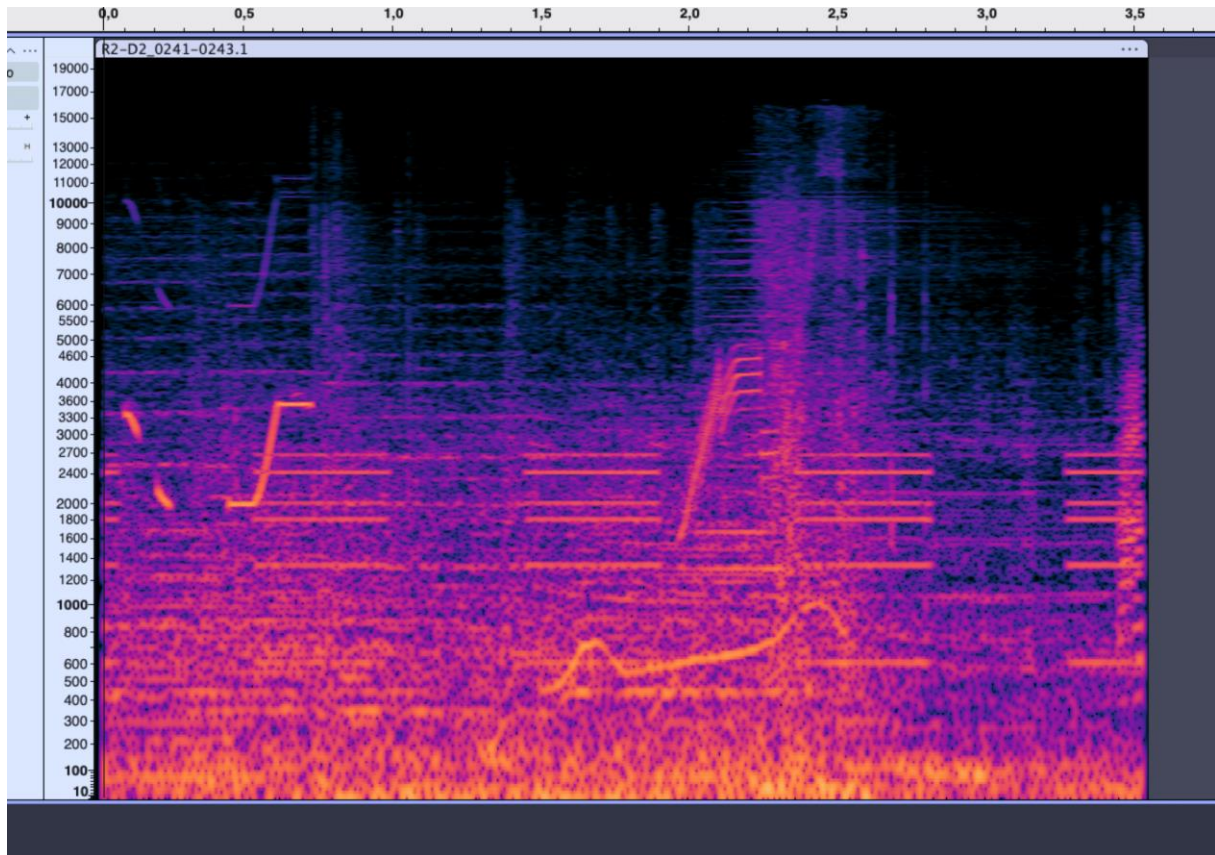
- Watanabe, Kenta, Kawashima, Shun, Inamura, Tetsunari & Yamada, Shinya (2025). *Semiotic analysis of robot sounds in films: implications for sound design in social robotics*. *International Journal of Social Robotics*, 17. doi:10.1007/s12369-024-01186-2
- Yoder, Christina & Bartmann, Martin (2017). A post-workshop interview about making musical instruments talk. I: Gisa Jähnichen (red.), *Studia Instrumentorum Musicae Popularis V (New Series)*. Münster: MV-Wissenschaft, s. 465–482.
- Zhang, Brian J. & Fitter, Naomi T. (2023). *Nonverbal sound in human-robot interaction: a systematic review*. *ACM Transactions on Human-Robot Interaction*, 12(4), doi:10.1145/3583743
- Whittington, William (2007). *Sound Design and Science Fiction*. Austin: University of Texas Press. doi:10.7560/714304
- Wrench, Jason S., Punyanunt-Carter, Narissra M. & Thweatt, Katherine S. (2025). *Interpersonal Communication: A Mindful Approach to Relationships* (2nd ed.). Milne Open Textbooks. Kapitel 5: Nonverbal Communication.

Appendix A – Analystabell

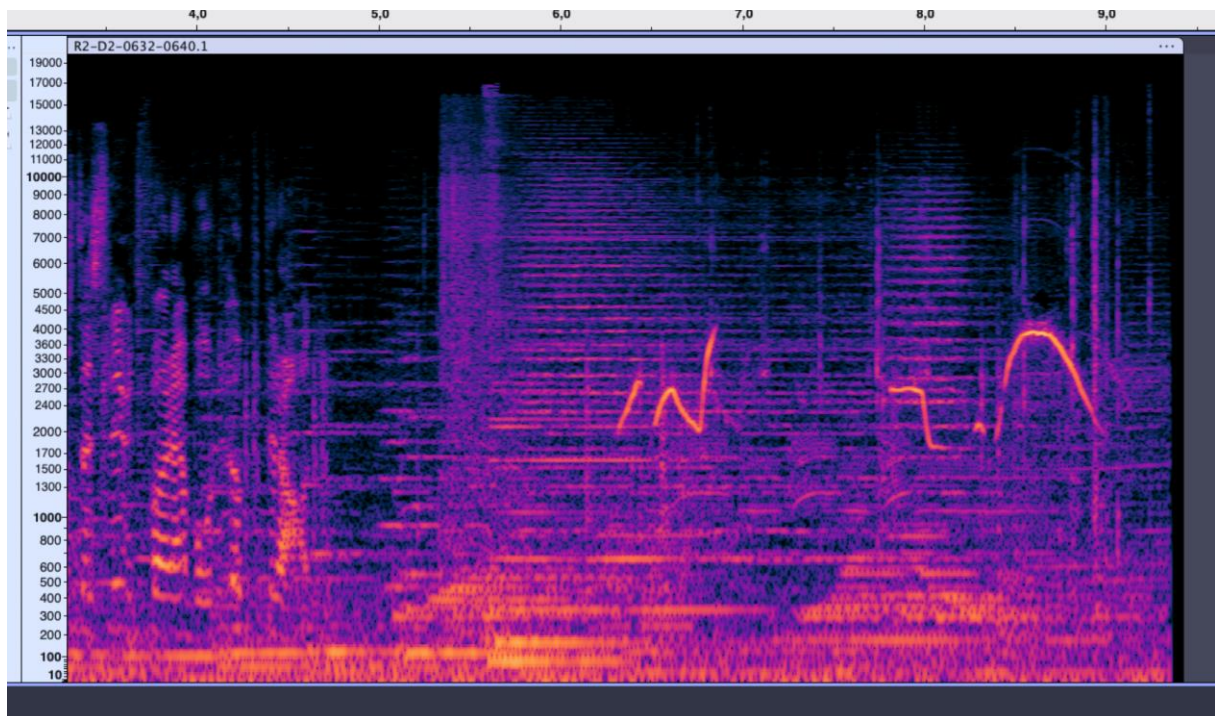
Del 1 – Kontext och auditiv observation										Del 2 – Afekt- och intentionskodning										Del 3 – Frekvens- och akustisk analys													
Robot	Filmavsätt	Titelkod	Visuell kontext	Ljudtyp	Varaklighet	Interaktion	Väns	Arousal	Emotionell kategori	Tolkad intention	Säkerhet	Frekvensområde (Hz)	Grundton (L)	Intensitet (dB)	HL-H	Nr 01	Robot	Filmavsätt	Titelkod	Visuell kontext	Ljudtyp	Varaklighet	Interaktion	Väns	Arousal	Emotionell kategori	Tolkad intention	Säkerhet	Frekvensområde (Hz)	Grundton (L)	Intensitet (dB)	HL-H	Nr 10
R2D2	SSE11 (The Core Wars)	00:02:41-00:02:43	Gruppkommunikation, yndislopp	Totala sakens	Medel (0.5-2.3)	Inlerande	Positiv	Hög	Bräddsammet	Uppmärksamhetsbärande	Hög	1801 - 4660Hz	1801Hz	Hög	L-H-L	Nr 01	BB4	The Force Awakens	00:02:45-00:02:47	Oermälla, ensam, natt	Totala sakens	Medel (0.5-2.3)	Inlerande	Negativ	Hög	Uppmärksamhet	Afektuellt (inaktiskt)	Hög	120-2000Hz	280Hz	Låg	L-H-L	Nr 10
R2D2	SSE11 (The Core Wars)	00:08:32-00:08:40	Oermälla, gruppkommunikation	Totala sakens	Lång (0.2-3)	Inlerande	Negativ	Hög	Uppmärksamhet	Uppmärksamhetsbärande	Hög	250 - 5000Hz	2600Hz	Hög	L-H-L	Nr 11	BB4	The Force Awakens	00:03:22-00:03:27	Inlerande, gruppkommunikation	Bräps & Boops	Medel (0.5-2.3)	Inlerande	Negativ	Hög	Uppmärksamhet	Uppmärksamhetsbärande	Hög	130-1110Hz	2600Hz	Hög	L-H-L	Nr 11
R2D2	SSE11 (The Core Wars)	00:07:14-00:07:16	Oermälla, gruppkommunikation	Totala sakens	Medel (0.5-2.3)	Repons	Positiv	Låg	Stress	Oskärmt / inoppsatt	Hög	250-4500Hz	5500Hz	Hög	L-H-L	Nr 12	BB4	The Force Awakens	00:04:26-00:04:30	Oermälla, yndislopp, natt	Totala sakens	Medel (0.5-2.3)	Inlerande	Negativ	Hög	Stress	Oskärmt / inoppsatt	Hög	250-4500Hz	5500Hz	Hög	L-H-L	Nr 12
R2D2	SSE11 (The Core Wars)	00:07:19-00:07:23	Oermälla, gruppkommunikation	Totala sakens	Lång (0.2-3)	Repons	Positiv	Hög	Engagemang	Socialt kooperativt	Hög	180-2500 Hz	1000Hz	Hög	L-H-L	Nr 13	BB4	The Force Awakens	00:05:54-00:05:54	Oermälla, ensam, natt	Totala sakens	Medel (0.5-2.3)	Repons	Negativ	Hög	Engagemang	Socialt kooperativt	Hög	240-340Hz	1000Hz	Hög	L-H-L	Nr 13
R2D2	SSE11 (The Core Wars)	00:09:02-00:09:05	Oermälla, gruppkommunikation	Pjywhite-sakens	Lång (0.2-3)	Repons	Positiv	Hög	Glädje	Socialt kooperativt	Hög	180-2500 Hz	4200Hz	Hög	L-H-L	Nr 14	BB4	The Force Awakens	00:16:29-00:16:33	Oermälla, gruppkommunikation	Totala sakens	Medel (0.5-2.3)	Inlerande	Positiv	Hög	Glädje	Socialt kooperativt	Hög	180-2500 Hz	4200Hz	Hög	L-H-L	Nr 14
R2D2	SSE11 (The Core Wars)	00:11:19-00:11:21	Oermälla, gruppkommunikation	Totala sakens	Medel (0.5-2.3)	Inlerande	Negativ	Hög	Enustasim	Socialt kooperativt	Hög	150-330Hz	350Hz	Hög	L-H-L	Nr 15	BB4	The Force Awakens	00:16:27-00:16:39	Oermälla, gruppkommunikation	Totala sakens	Medel (0.5-2.3)	Inlerande	Positiv	Hög	Enustasim	Socialt kooperativt	Hög	150-330Hz	350Hz	Hög	L-H-L	Nr 15
R2D2	SSE11 (The Core Wars)	00:11:39-00:11:41	Oermälla, gruppkommunikation	Totala sakens	Medel (0.5-2.3)	Inlerande	Negativ	Hög	Intresse	Uppmärksamhetsbärande	Hög	180-2500 Hz	4000Hz	Hög	L-H-L	Nr 16	BB4	The Force Awakens	01:23:40-01:23:42	Inlerande, yndislopp	Totala sakens	Medel (0.5-2.3)	Övertappande	Positiv	Hög	Enustasim	Socialt kooperativt	Hög	180-2500 Hz	4000Hz	Hög	L-H-L	Nr 16
R2D2	SSE11 (The Core Wars)	00:12:30-00:12:34	Oermälla, mjölprens	Totala sakens	Lång (0.2-3)	Inlerande	Positiv	Hög	Enustasim	Socialt kooperativt	Hög	180-2500 Hz	1600Hz	Hög	L-H-L	Nr 17	BB4	The Force Awakens	02:01:18-02:01:23	Gruppkommunikation, yndislopp	Totala sakens	Lång (0.2-3)	Övertappande	Positiv	Hög	Enustasim	Socialt kooperativt	Hög	180-2500 Hz	1600Hz	Hög	L-H-L	Nr 17
R2D2	SSE11 (The Core Wars)	00:20:08-00:20:08	Oermälla, gruppkommunikation	Pjywhite-sakens	Medel (0.5-2.3)	Inlerande	Positiv	Hög	Enustasim	Socialt kooperativt	Hög	180-2500 Hz	1600Hz	Hög	L-H-L	Nr 18	BB4	The Force Awakens	02:01:18-02:01:23	Gruppkommunikation, yndislopp	Totala sakens	Lång (0.2-3)	Övertappande	Positiv	Hög	Enustasim	Socialt kooperativt	Hög	180-2500 Hz	1600Hz	Hög	L-H-L	Nr 18
BD-1	Jas Fallen Order	01:27 - 01:28	Gruppkommunikation, loamjög	Totala sakens	Kort (0.5-1.5 s)	Repons	Negativ	Låg	Nedslämdhet	Afektuellt (inaktiskt)	Hög	215 - 947Hz	550Hz	Låg	L	Nr 19	BD-1	Jas Fallen Order	02:07 - 02:08	Gruppkommunikation, loamjög	Bräps & Boops	Medel (0.5-1.5 s)	Inlerande	Positiv	Hög	Enustasim	Socialt kooperativt	Hög	812 - 2200Hz	1220Hz	Hög	L-H-L	Nr 20
BD-1	Jas Fallen Order	02:30 - 02:32	Gruppkommunikation, grottmjög	Bräps & Boops	Kort (0.5-1.5 s)	Inlerande	Positiv	Hög	Uppmärksamhet	Uppmärksamhetsbärande	Hög	88 - 1000Hz	1420Hz	Hög	L-H-L	Nr 21	BD-1	Jas Fallen Order	06:04 - 06:06	Gruppkommunikation, grottmjög	Bräps & Boops	Lång (0.2-3)	Repons	Positiv	Hög	Engagemang	Socialt kooperativt	Hög	129 - 2850Hz	602Hz	Låg	L-H-L	Nr 21
BD-1	Jas Survivor	03:20 - 03:27	X	Bräps & Boops	Lång (0.2-3)	Självständigt	Positiv	Hög	Glädje	Lekfull / humoristisk	Hög	215 - 2918Hz	1248Hz	Hög	L-H	Nr 23	BD-1	Jas Survivor	07:51 - 07:51	Gruppkommunikation, yndislopp	Totala sakens	Medel (0.5-1.5 s)	Repons	Negativ	Låg	Uppgivenhet	Socialt kooperativt	Hög	340 - 1440Hz	520Hz	Låg	L-H	Nr 23
BD-1	Jas Survivor	09:25 - 09:26	Gruppkommunikation, yndislopp	Bräps & Boops	Medel (0.5-2.3)	Repons	Positiv	Låg	Bräddsammet	Socialt kooperativt	Hög	215 - 3445Hz	979Hz	Hög	L-H-L	Nr 24	BD-1	Jas Survivor	09:25 - 09:26	Gruppkommunikation, yndislopp	Bräps & Boops	Medel (0.5-2.3)	Repons	Positiv	Låg	Bräddsammet	Socialt kooperativt	Hög	215 - 3445Hz	979Hz	Hög	L-H-L	Nr 24
BD-1	Jas Survivor	09:28 - 09:34	Gruppkommunikation, grottmjög	Bräps & Boops	Lång (0.2-3)	Repons	Positiv	Hög	Ungämsamhet	Socialt kooperativt	Hög	43 - 4000Hz	654Hz	Låg	L-H-L	Nr 25	BD-1	Jas Survivor	09:28 - 09:34	Gruppkommunikation, grottmjög	Bräps & Boops	Lång (0.2-3)	Repons	Positiv	Hög	Ungämsamhet	Socialt kooperativt	Hög	43 - 4000Hz	654Hz	Låg	L-H-L	Nr 25

Figur.8 Analystabell (höger sida) samt översikt över kodningskategorier (vänster sida). Den vänstra delen visar de kodningsalternativ som användes vid analysen av ljudklippen.

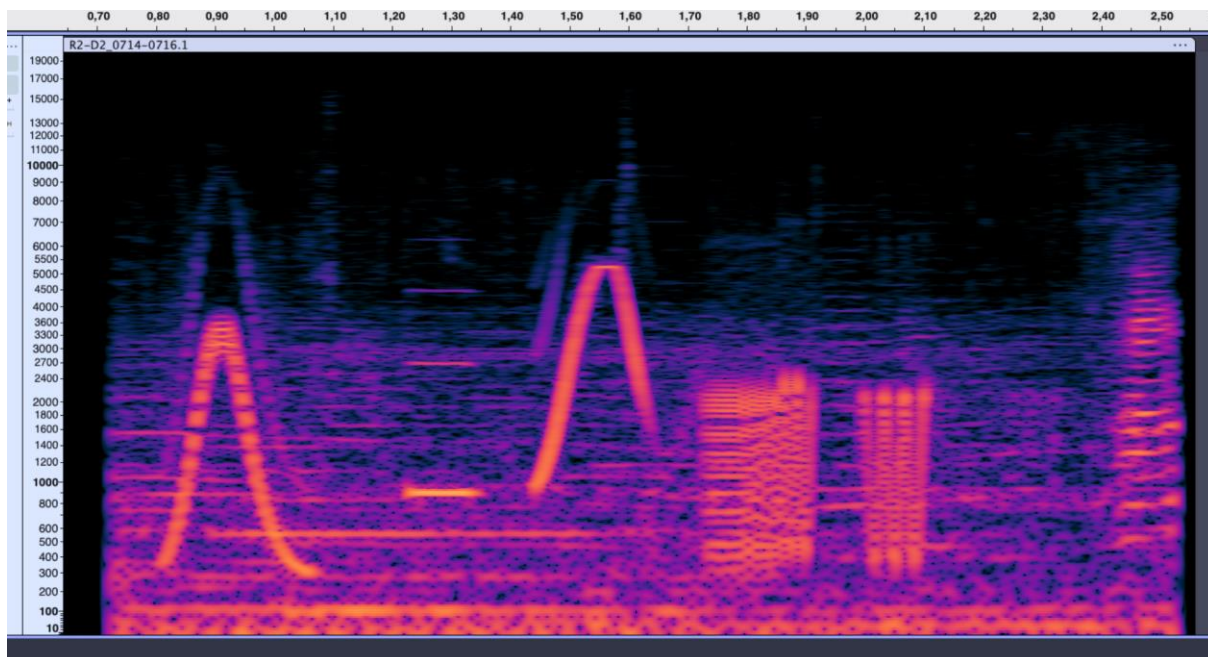
Appendix B - Spektrogram



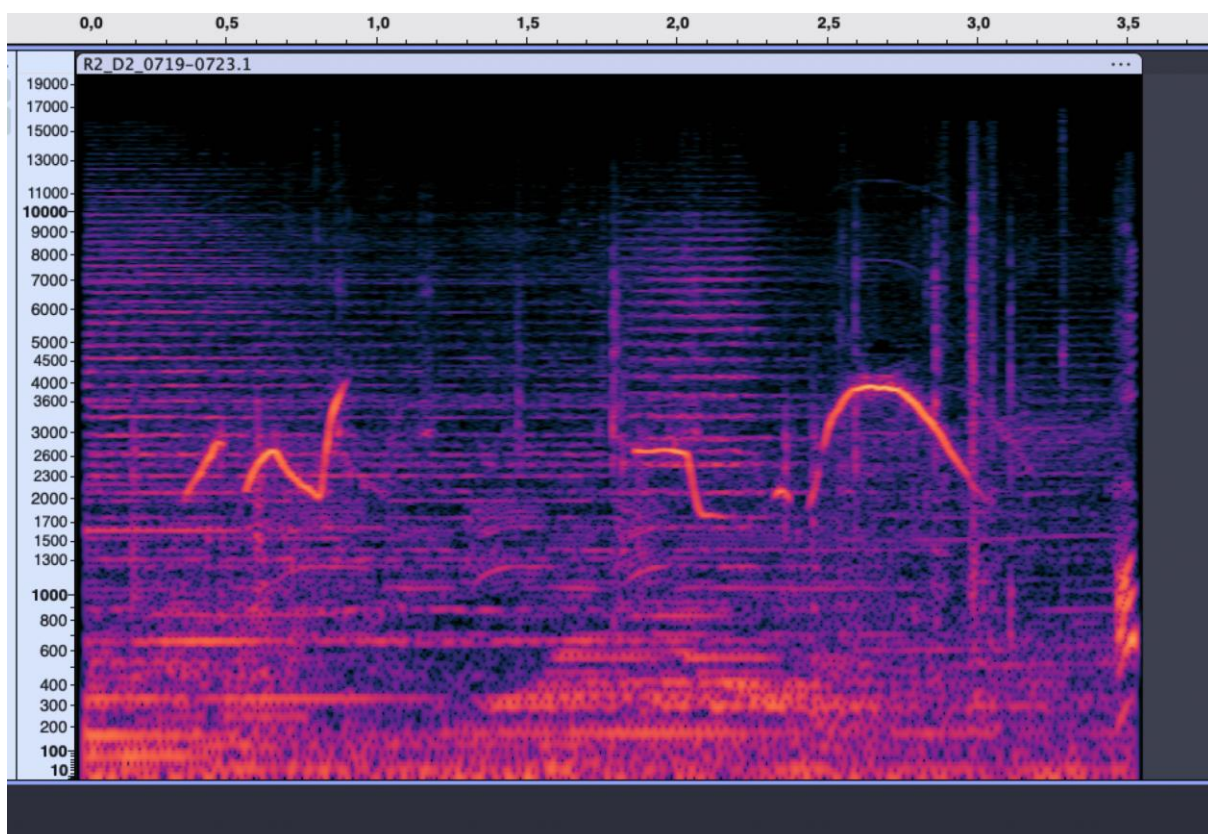
R2D2, S5E11 (*The Clone Wars*), 00:02:41-00:02:43, ljudklipp nr 01.



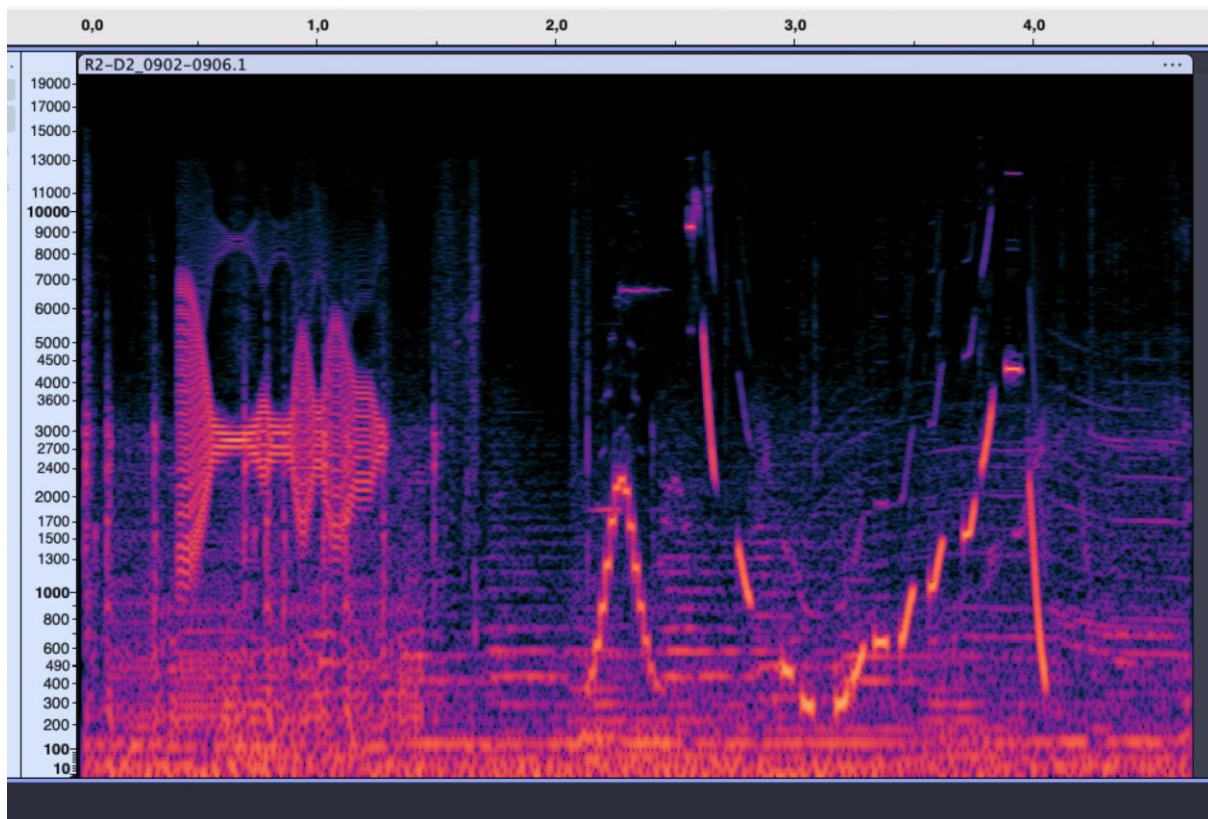
R2D2, S5E11 (*The Clone Wars*), 00:06:32-00:06:40, ljudklipp nr 02.



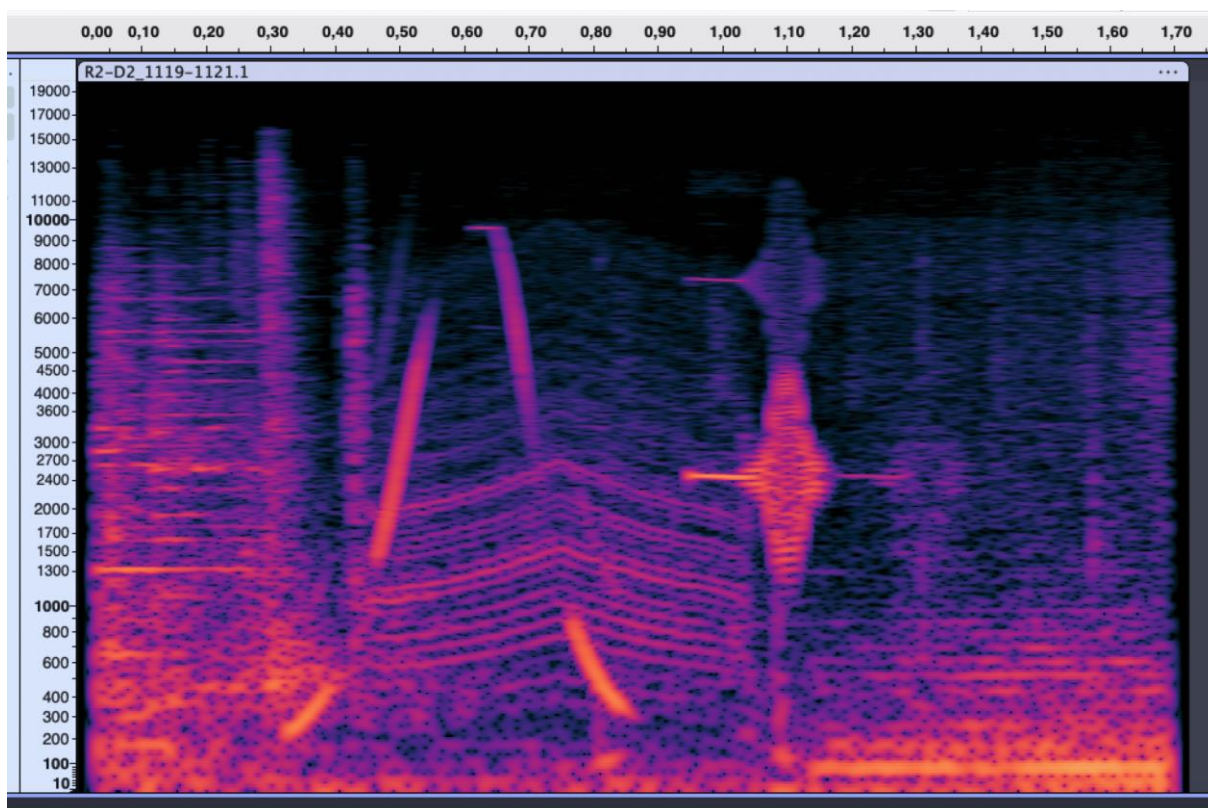
R2D2, S5E11 (*The Clone Wars*), 00:07:14-00:07:16, ljudklipp nr 03.



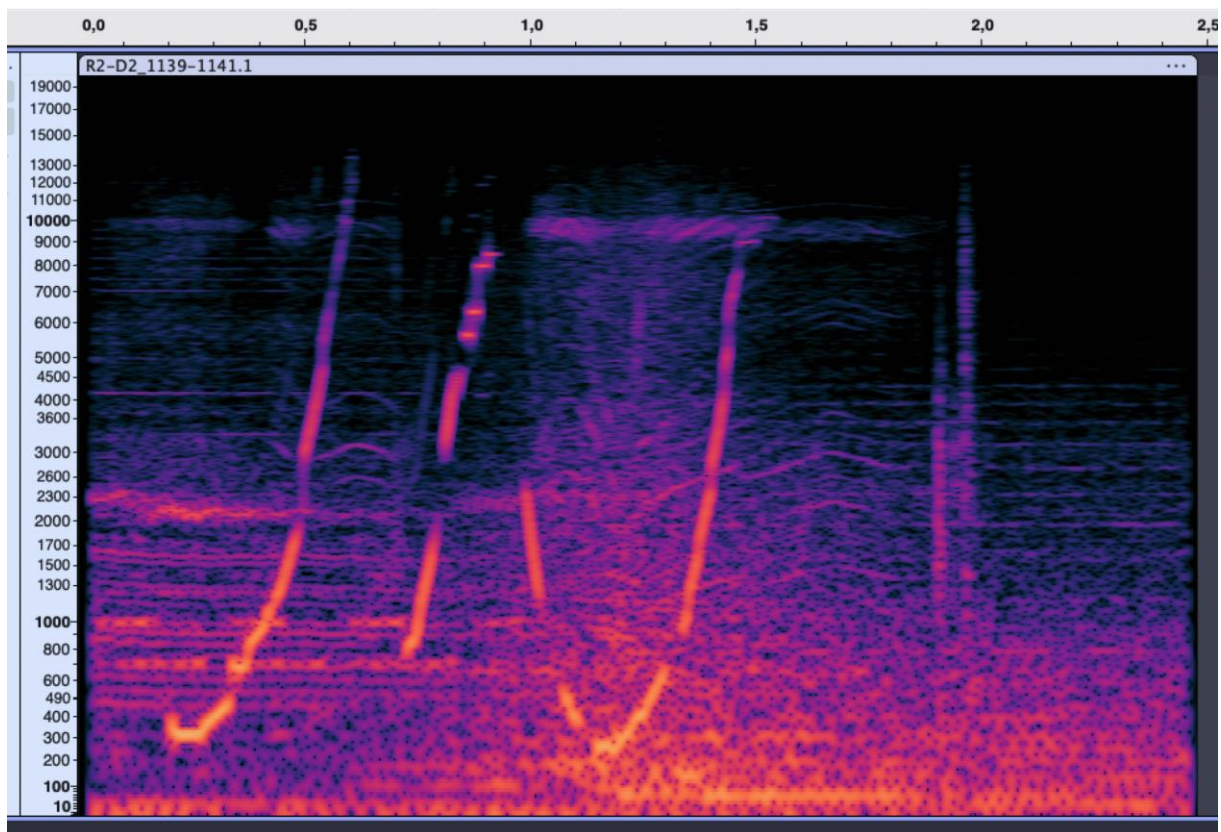
R2D2, S5E11 (*The Clone Wars*), 00:07:19-00:07:23, ljudklipp nr 04.



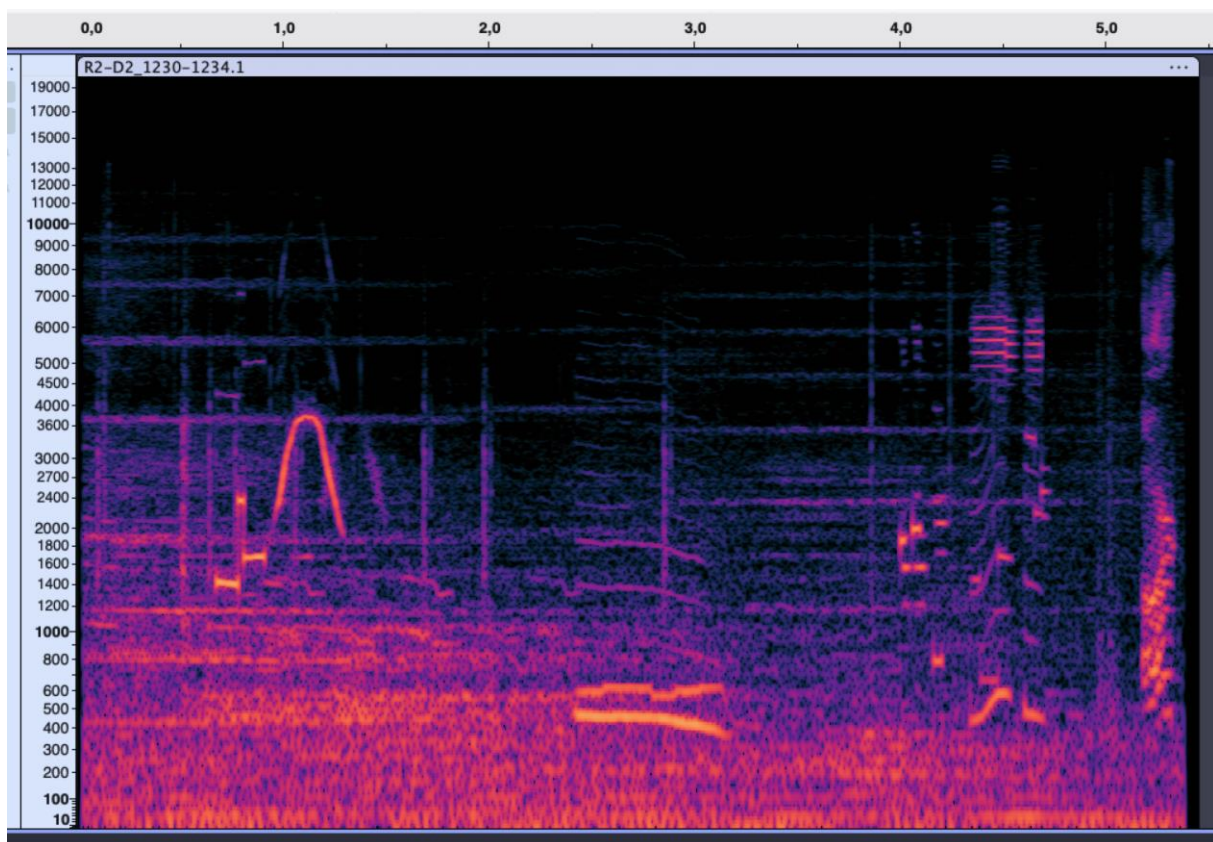
R2D2, S5E11 (*The Clone Wars*), 00:09:02-00:09:06, ljudklipp nr 05.



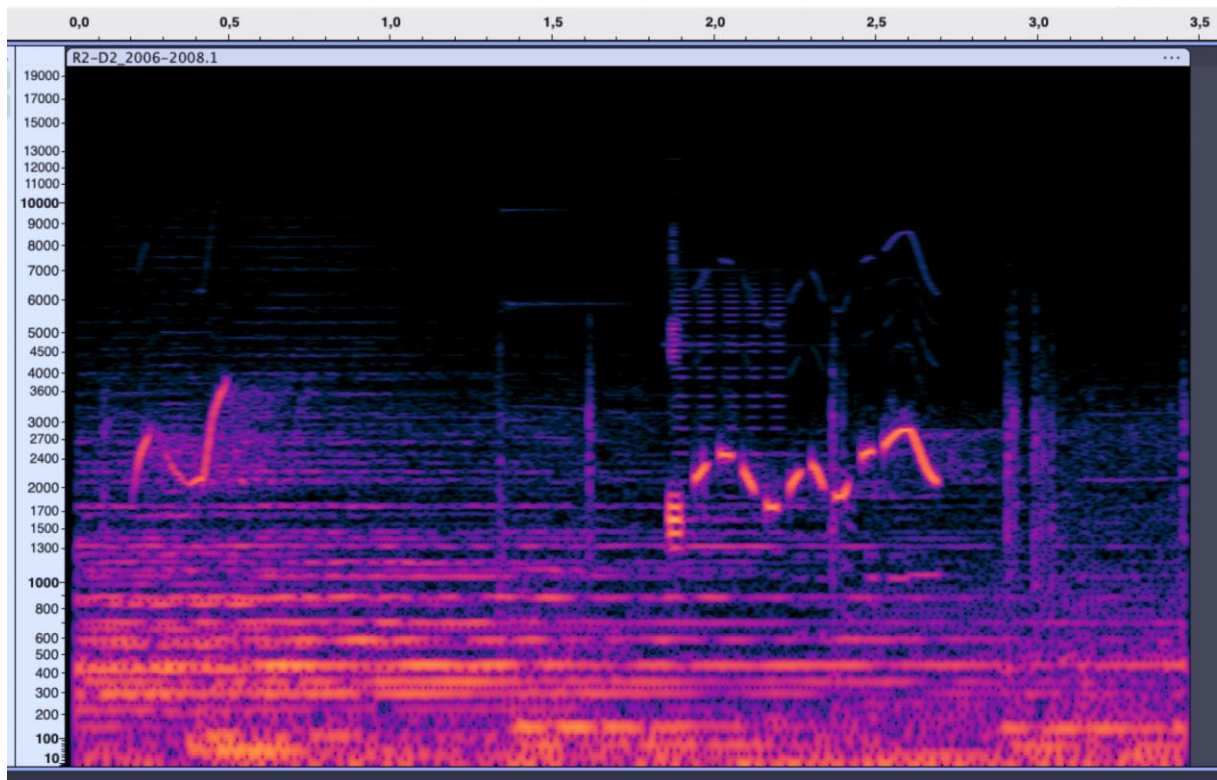
R2D2, S5E11 (*The Clone Wars*), 00:11:19-00:11:21, ljudklipp nr 06.



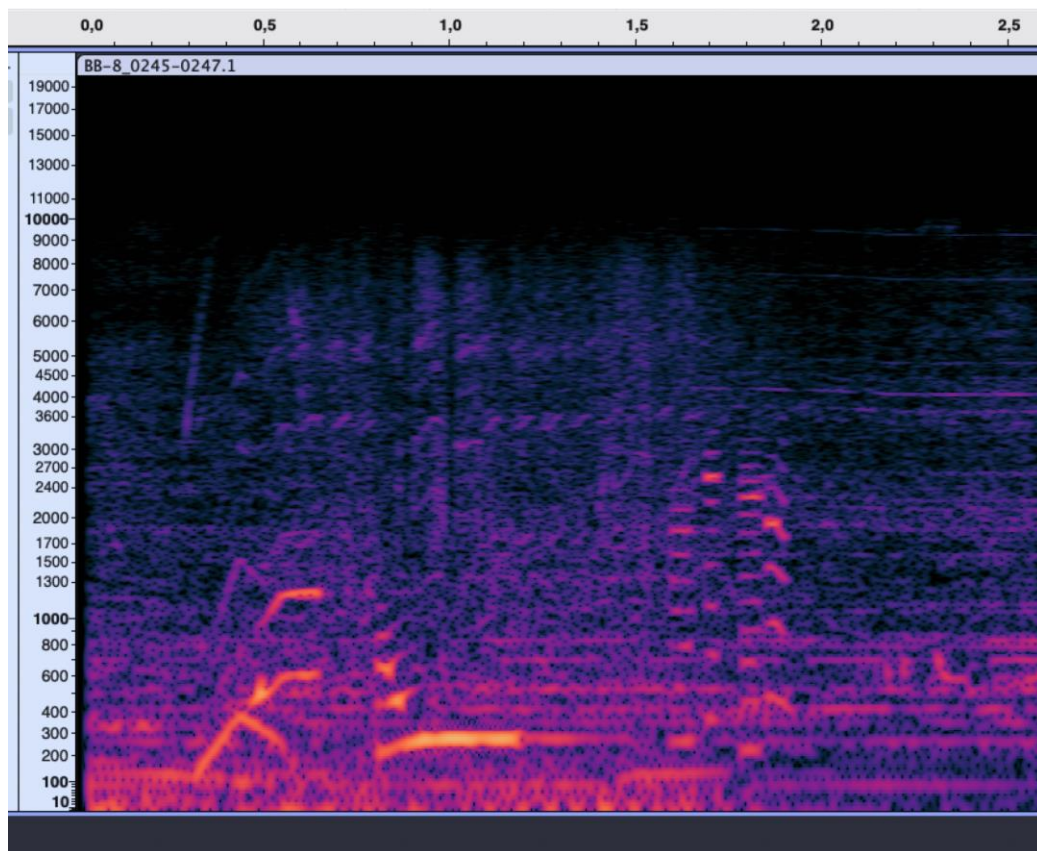
R2D2, S5E11 (*The Clone Wars*), 00:11:39-00:11:41, ljudklipp nr 07.



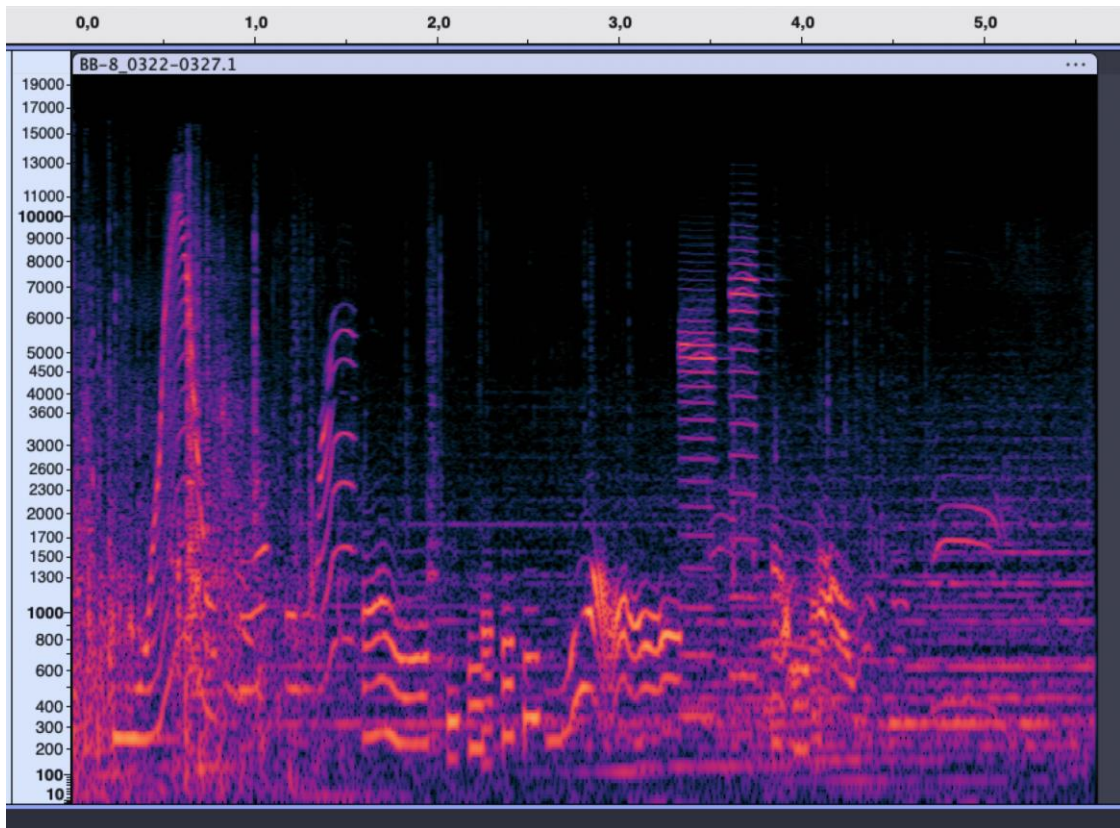
R2D2, S5E11 (*The Clone Wars*), 00:12:30-00:12:34, ljudklipp nr 08.



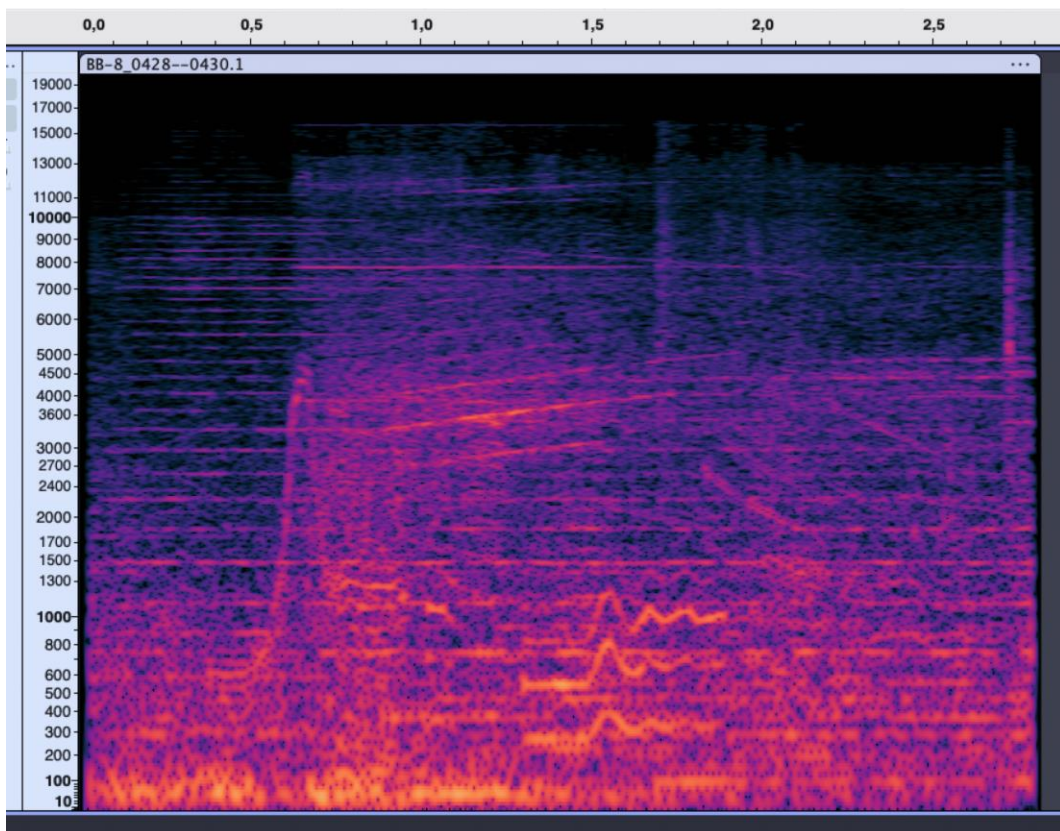
R2D2, S5E11 (*The Clone Wars*), 00:20:06-00:20:08, ljudklipp nr 09.



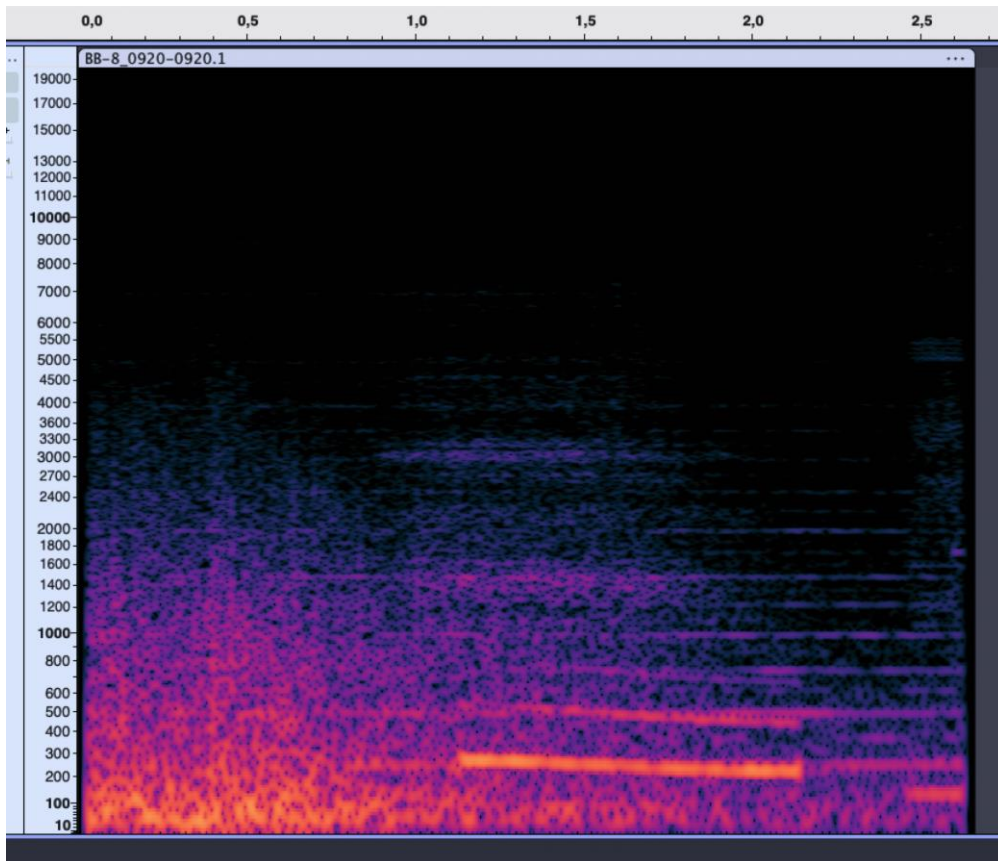
BB-8, *Star Wars: The Force Awakens*, 00:02:45-00:02:47, ljudklipp nr 10.



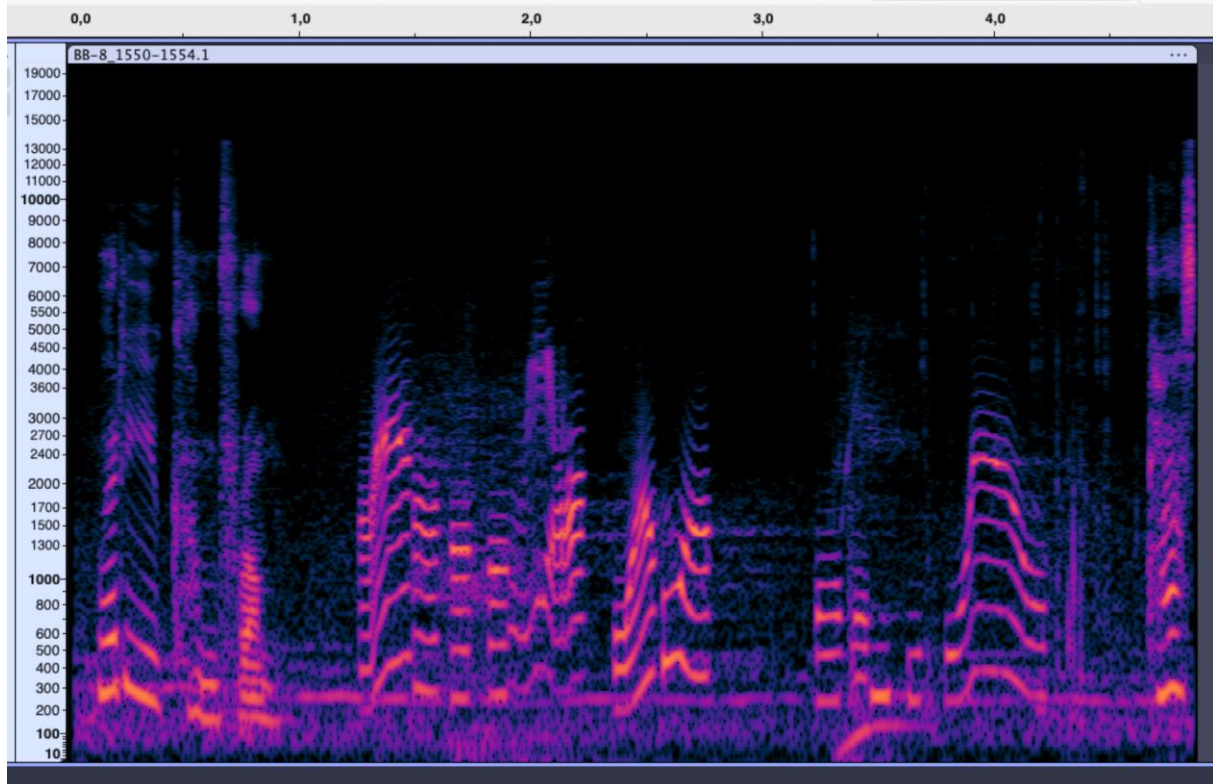
BB-8, *Star Wars: The Force Awakens*, 00:03:22-00:03:27, ljudklipp nr 11.



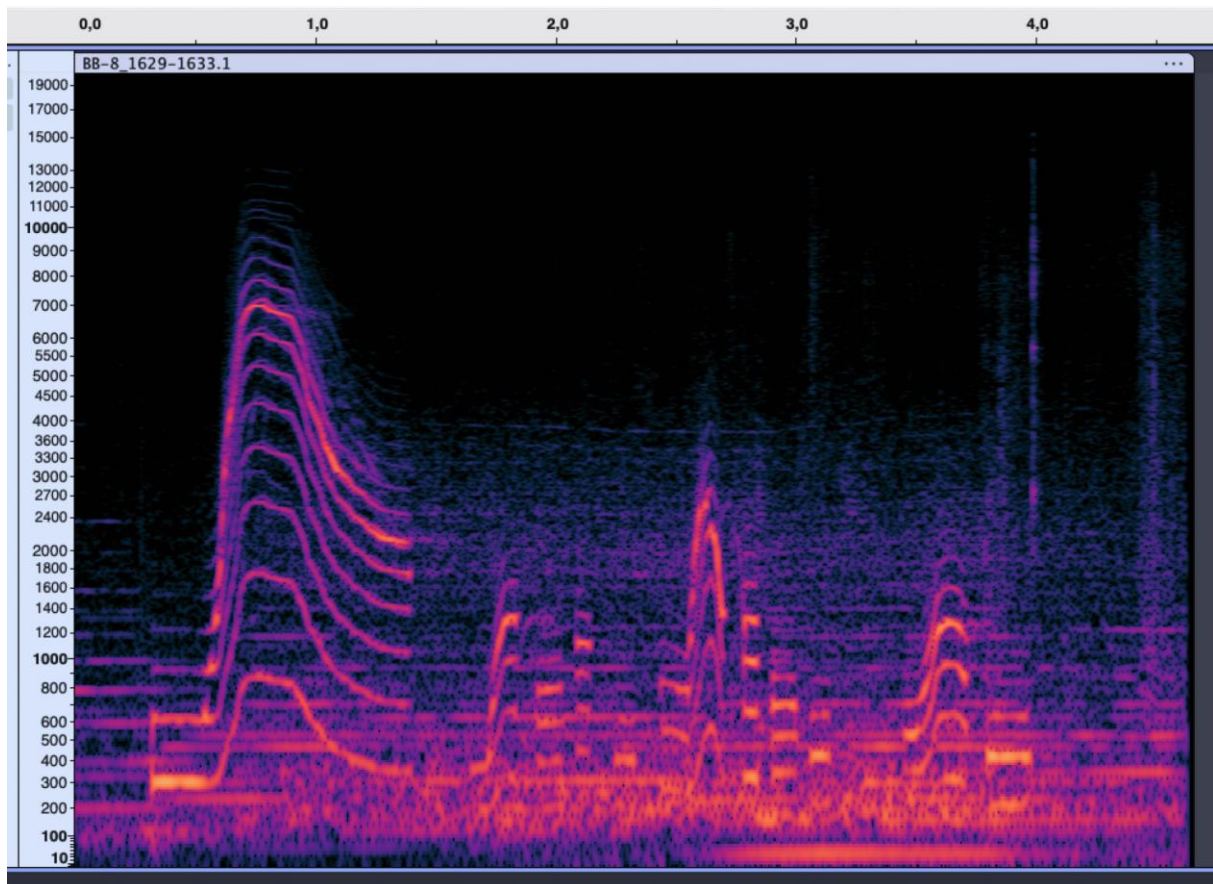
BB-8, *Star Wars: The Force Awakens*, 00:04:28-00:04:30, ljudklipp nr 12.



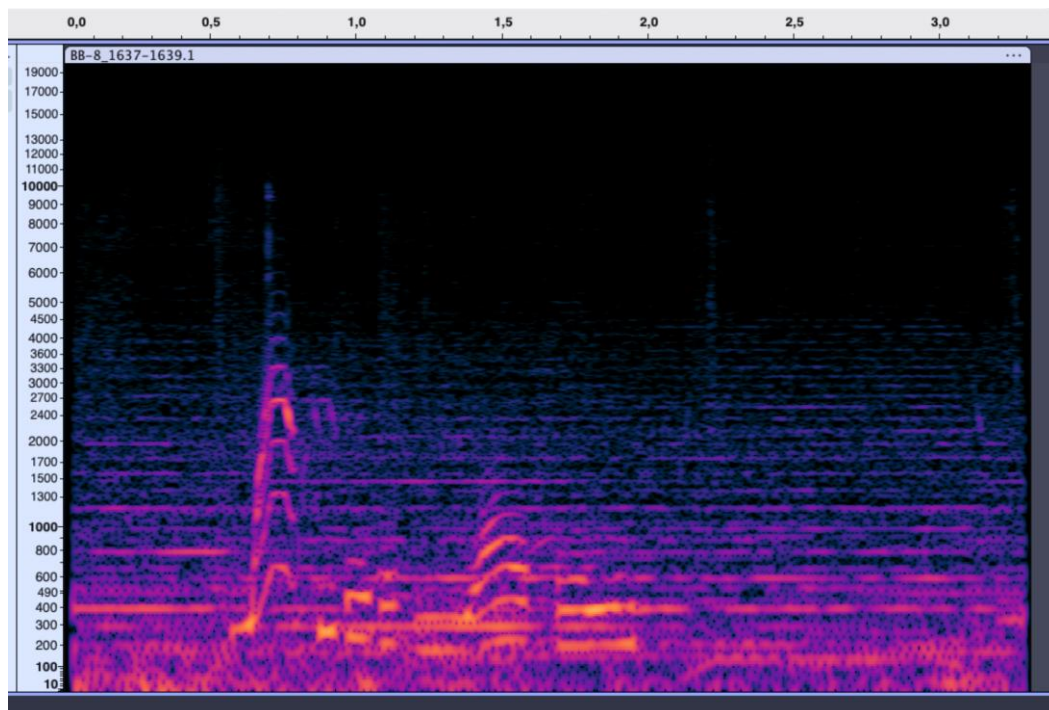
BB-8, *Star Wars: The Force Awakens*, 00:09:20-00:09:20, ljudklipp nr 13.



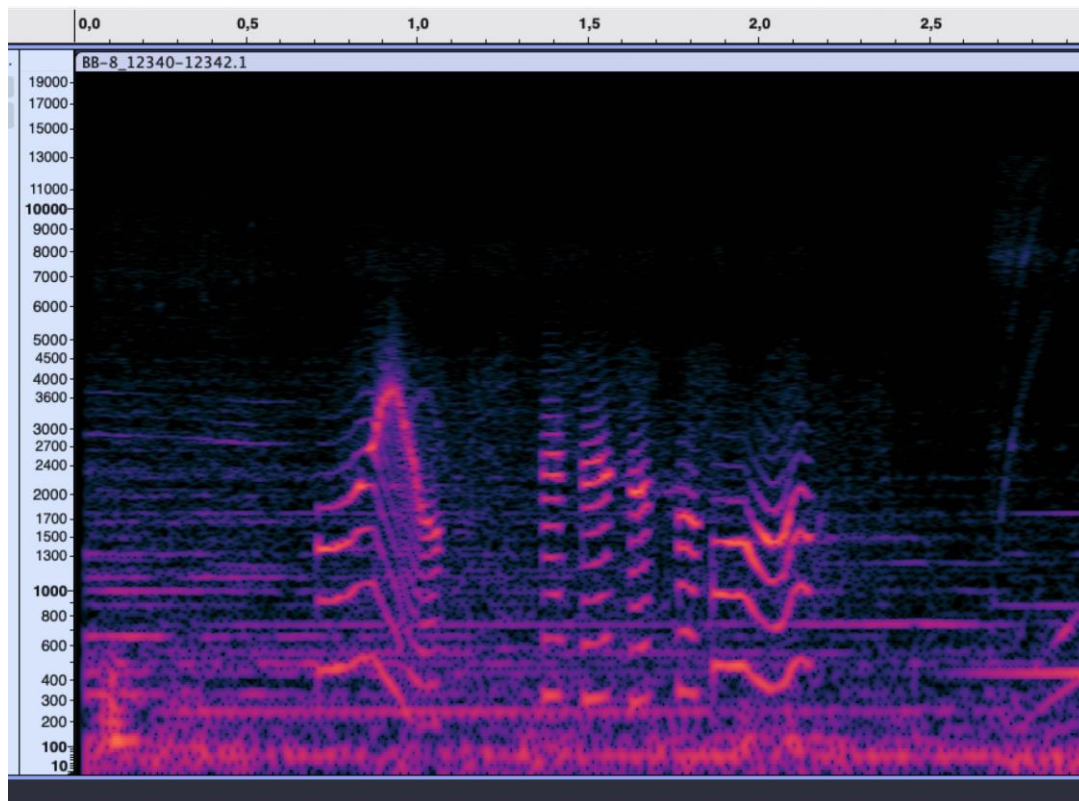
BB-8, *Star Wars: The Force Awakens*, 00:15:50-00:15:54, ljudklipp nr 14.



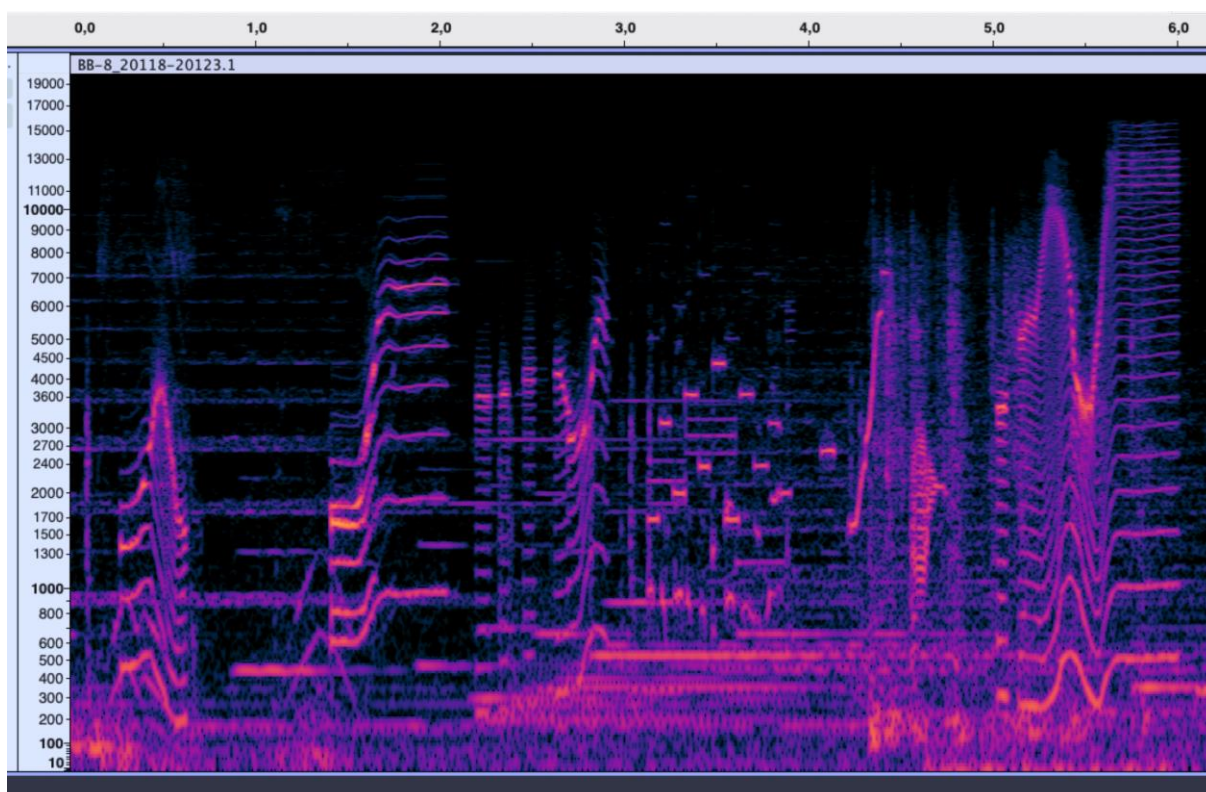
BB-8, *Star Wars: The Force Awakens*, 02:01:18-02:01:23, ljudklipp nr 15.



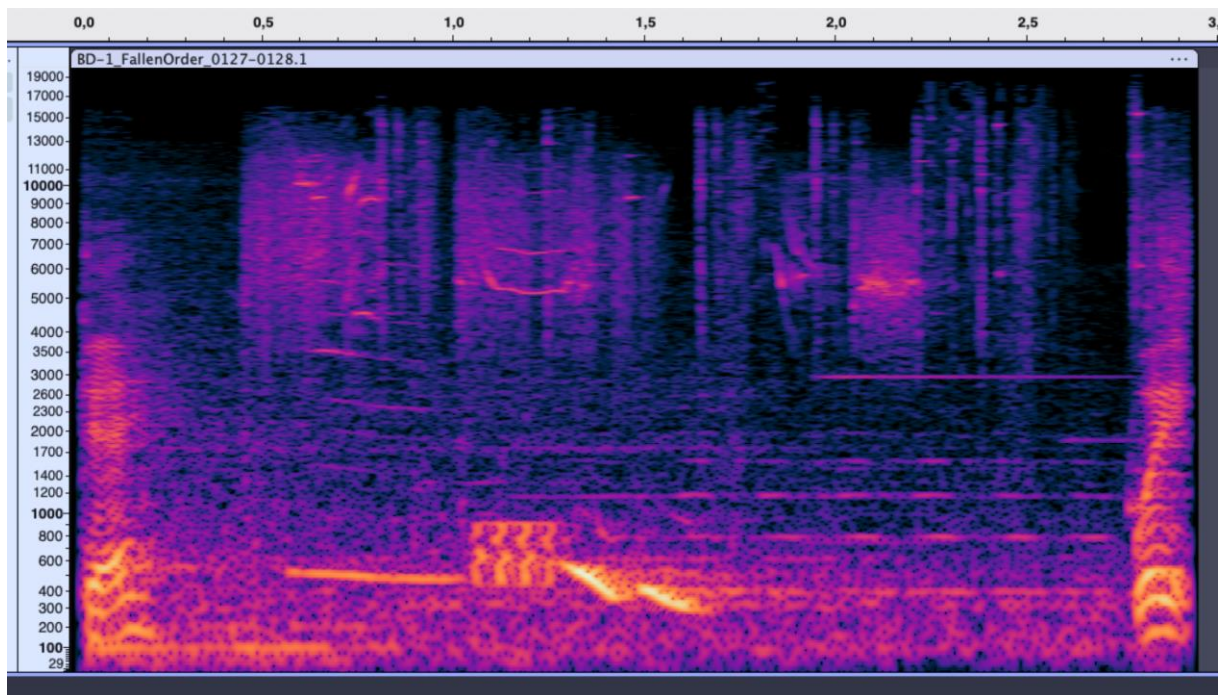
BB-8, *Star Wars: The Force Awakens*, 00:16:37-00:16:39, ljudklipp nr 16.



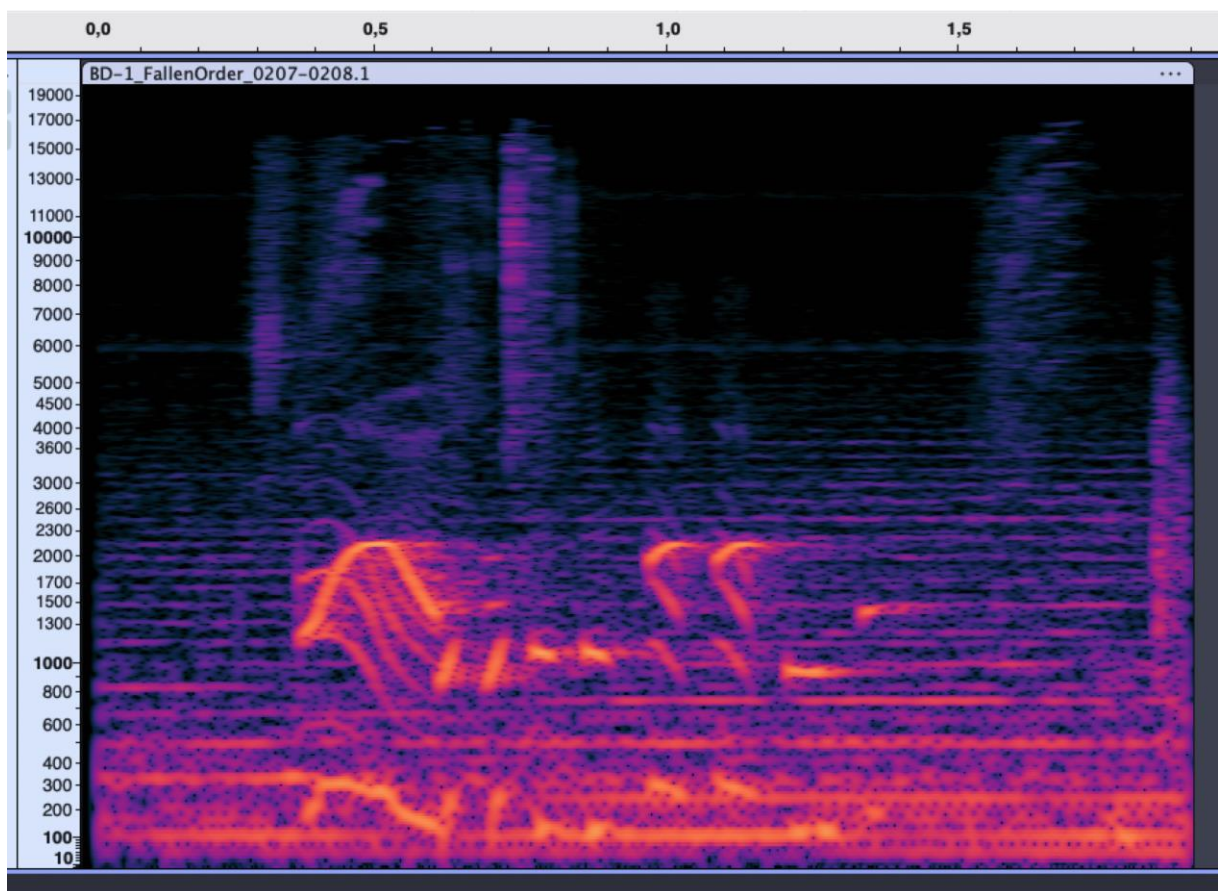
BB-8, *Star Wars: The Force Awakens*, 01:23:40-01:23:42, ljudklipp nr 17.



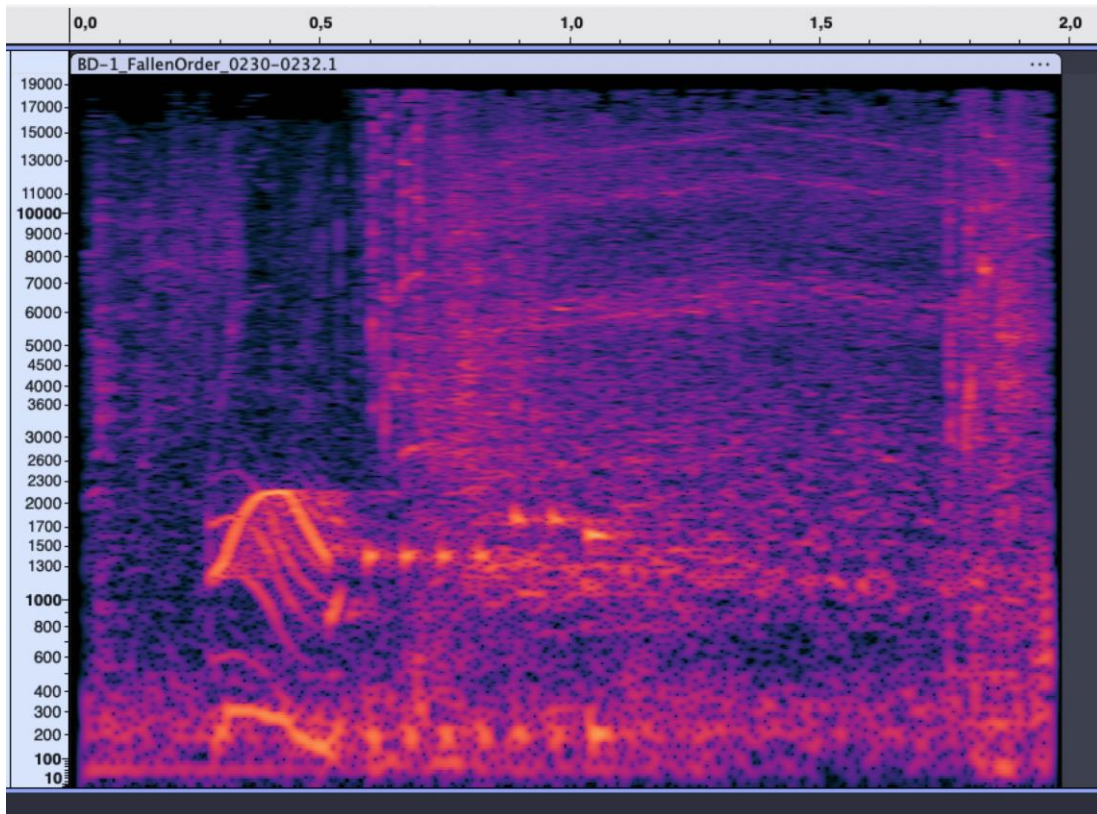
BB-8, *Star Wars: The Force Awakens*, 02:01:18-02:01:23, ljudklipp nr 18.



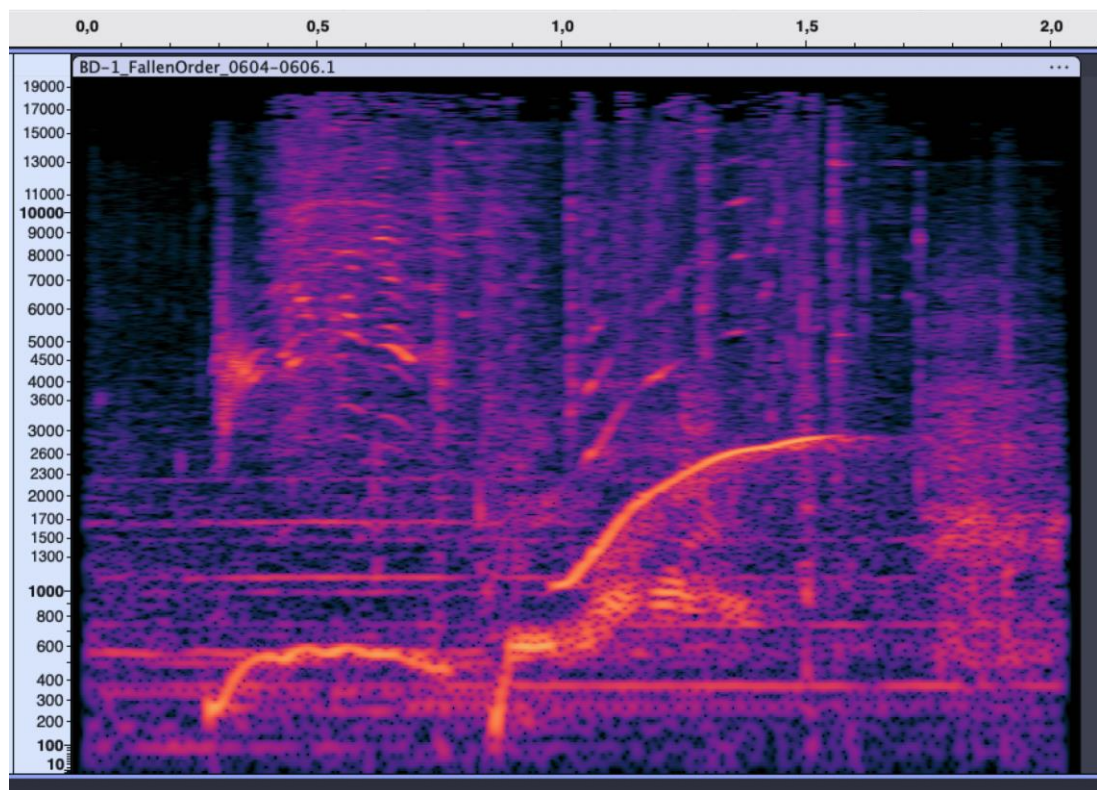
BD-1, *Jedi Fallen Order*, 01:27 - 01:28, ljudklipp nr 19.



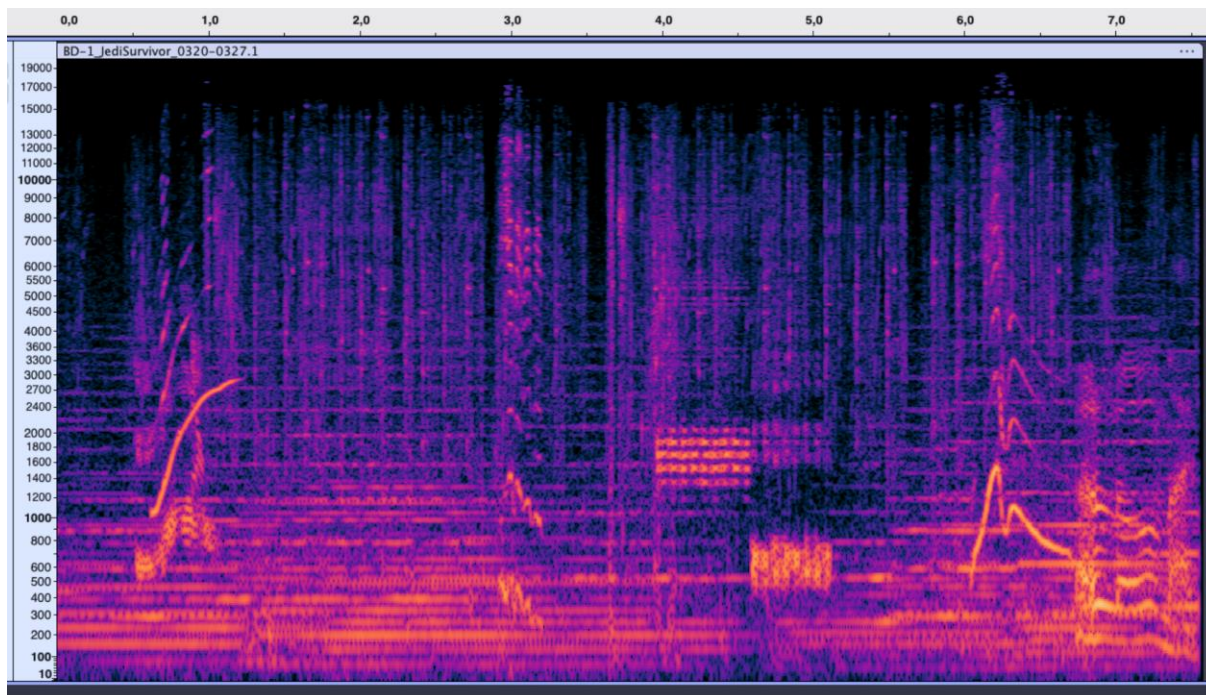
BD-1, *Jedi Fallen Order*, 02:07 - 02:08, ljudklipp nr 20.



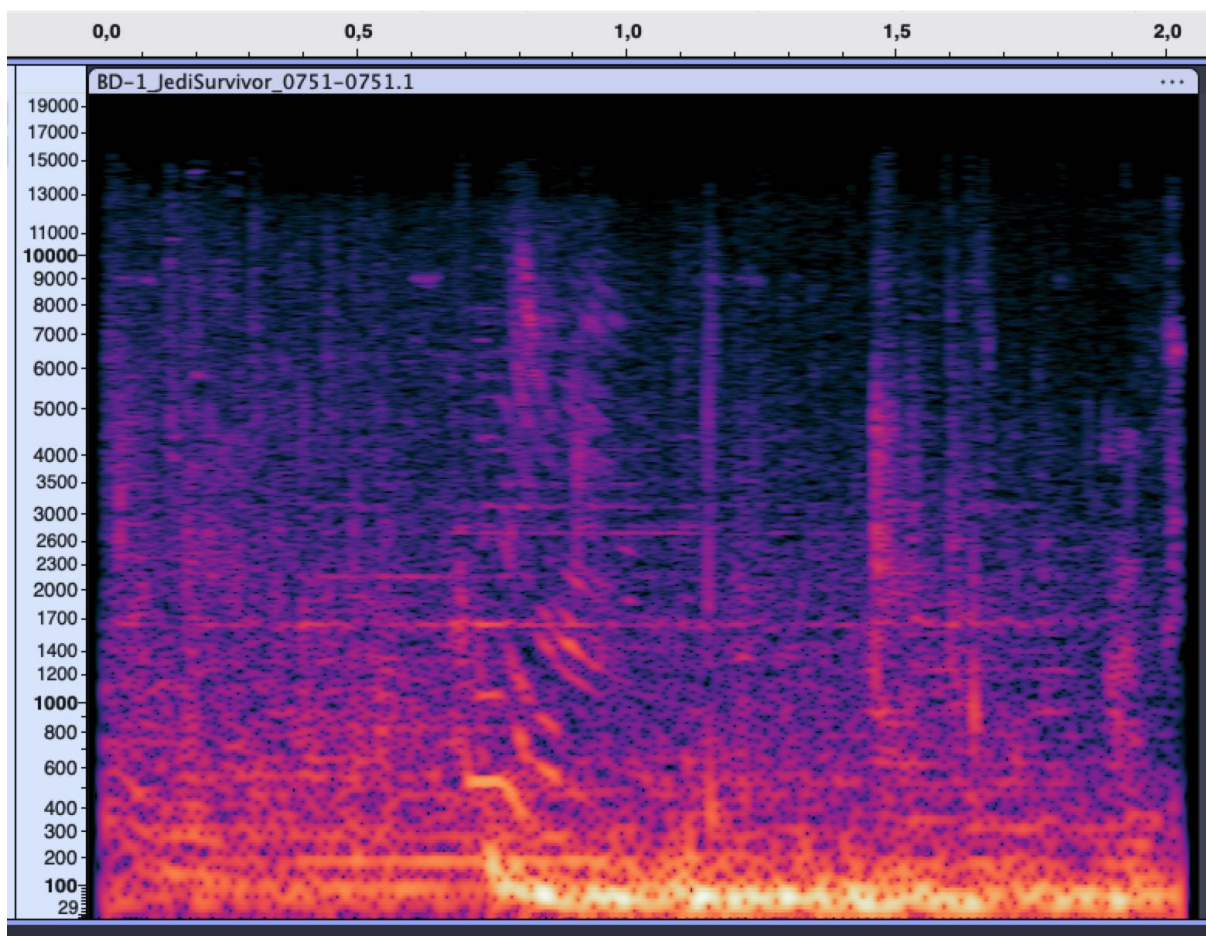
BD-1, *Jedi Fallen Order*, 02:30 - 02:32, ljudklipp nr 21.



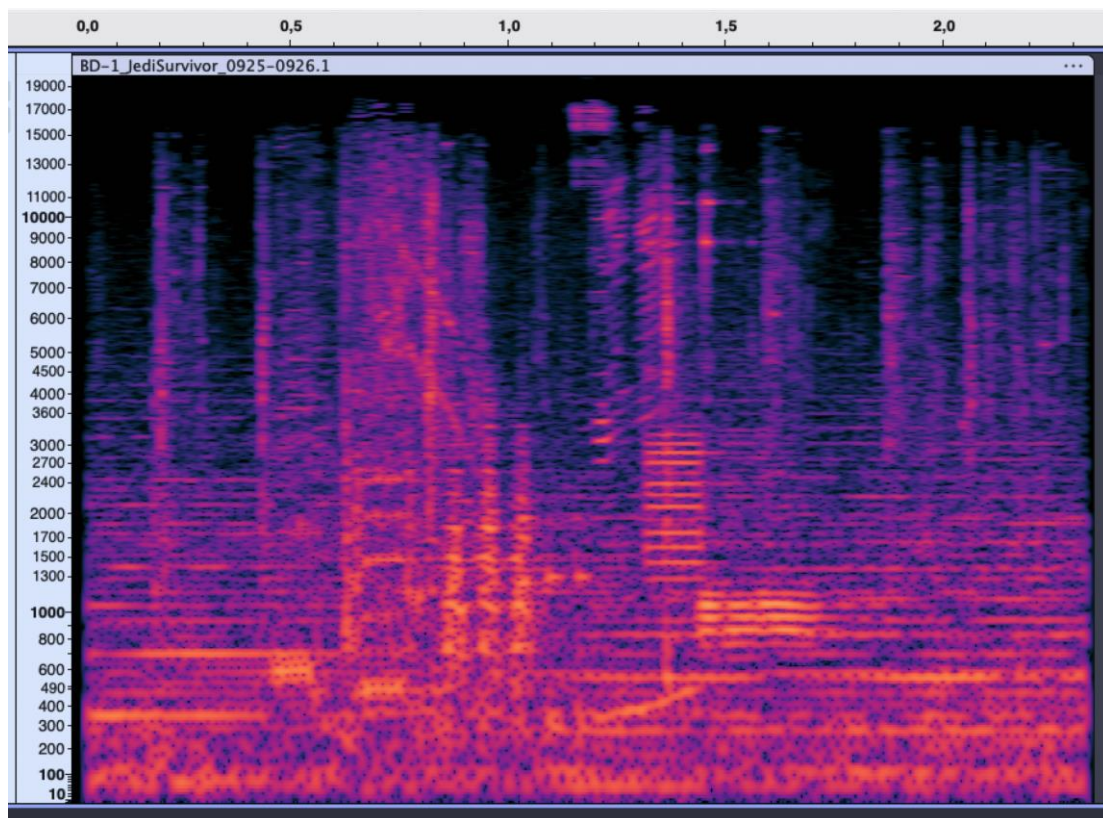
BD-1, *Jedi Fallen Order*, 06:04 - 06:06, ljudklipp nr 22.



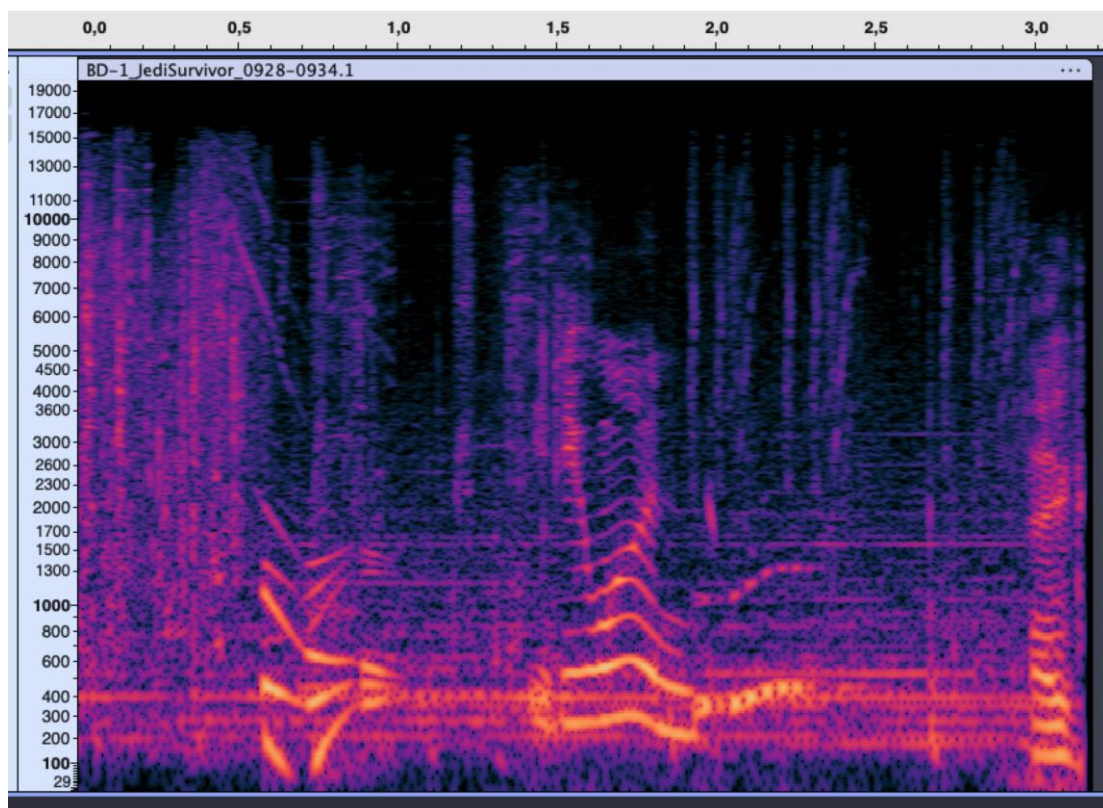
BD-1, *Jedi Survivor*, 03:20 - 03:27, ljudklipp nr 23.



BD-1, *Jedi Survivor*, 07:51 - 07:51, ljudklipp nr 24.



BD-1, *Jedi Survivor*, 09:25 - 09:26, ljudklipp nr 25.



BD-1, *Jedi Survivor*, 09:28 - 09:34, ljudklipp nr 26.

Appendix C – Videomaterial

Video 1: *BD-1 Adorable Scenes*. YouTube. Tillgänglig på: <https://www.youtube.com/watch?v=V7UUJm-8Ba8> (Hämtad: 26 mars 2026).

Video 2: *Star Wars Droid Sound Compilation*. YouTube. Tillgänglig på: <https://www.youtube.com/watch?v=MO3VmTutkcU> (Hämtad: 26 mars 2026).