

Característiques tonals i espectrals de la parla dirigida a infants en català

Wendy Elvira-García - Universitat de Barcelona
wendyelvira@ub.edu

Georgina Bargalló Osuna - Universitat Autònoma de Barcelona
georgina.bargallo@uab.cat

Rebut / Received: 06-02-23

Acceptat / Accepted: 21-06-23

Resum. Característiques tonals i espectrals de la parla dirigida a infants en català. Aquest estudi preliminar analitza com interactuen els educadors amb els nens de 3-6 anys per caracteritzar la parla dirigida a nens (IDS) en català posant una atenció especial a les característiques fonètiques del to i la intel·ligibilitat. Per tal de fer-ho, comparem les veus de 8 mestres parlant amb un adult i amb nens. Per cada participant i condició, hem anotat en Praat entre 2-6 minuts de conversa i hem extret el to mitjà, la variabilitat (utilitzant la desviació estàndard i el Pitch Dynamism Quotient) i el pendent espectral. Després utilitzem models mixtos generalitzats a R per provar l'efecte de la condició (IDS). Els resultats estan en línia amb els estudis previs per a altres llengües i mostren que l'IDS és més agut (to mitjà més alt), de to més variable (desviació estàndard i PDQ més alts) i més intel·ligible (amb un pendent espectral menys pronunciat i formant de locutor).

Paraules clau: parla dirigida a infants, entonació, pendent espectral, Pitch Dynamism Quotient.

Abstract: Tone and spectral features of child-directed speech in Catalan. This preliminary work analyzes how Catalan educators interact with children aged from 3-6 in order to characterize child-directed speech (IDS) phonetically paying special attention to pitch and intelligibility features. In order to do it, we compared the voices of 8 female teachers speaking with an adult and with children. For each participant and condition, we have annotated in Praat between 2 and 6 minutes of conversation and extracted its mean pitch, pitch variability (using both standard deviation and Pitch Dynamism Quotient) and spectral slope, then we used Generalized Mixed Models in R to test the effect of the child-directed speech condition. The results are in line with previous studies for other languages and show that CDS higher in pitch (higher pitch mean),

more variable (higher standard deviation and PDQ) and more intelligible (a less steep spectral slope and speaker's formant).

Keywords: child directed speech, intonation, spectral slope, Pitch Dynamism Quotient.

1. Introducció

Els adults s'adrecen als nens de manera diferent que als seus iguals. És el que se'n diu "parla dirigida a infants" (*child directed speech*) IDS, també anomenada tradicionalment *maternal linguistic input* o *motherese* i més tard *fatherese* (Shute & Wheldall, 1999).

Aquest comportament o canvi de registre és inconscient, reflexiu i instintiu en presència de qualsevol infant (Kitamura et al., 2001) i obté una reacció positiva pel que fa a la seva atenció i, longitudinalment, a la seva adquisició lingüística.

Se sap des dels primers estudis que la parla dirigida a infants inclou fenòmens sintàctics, fonològics i lèxics (Ferguson, 1964) i que aquests són comuns en diferents llengües (Fernald et al. 1989).

Fonèticament, s'ha constatat que el IDS està associat amb l'alentiment del ritme, l'allargament sil·làbic i l'entonació. I d'aquests, l'entonació és el factor principal en l'adquisició de la primera llengua en infants (Tomasello, 2003) i un element primordial per a la transmissió d'informació a infants preverbals, ja que permet la identificació de la intenció del parlant (Fernald, 1989), les seves emocions i, a més a més, facilita captar l'atenció de l'infant (Trainor et al., 2000). Específicament, s'ha demostrat que els nens prefereixen les veus més agudes i rangs tonals més amplis (Fernald & Kuhl, 1987; Frank, 2020) i que aquesta preferència es va atenuant a mesura que l'edat del nen incrementa.

La influència de l'entonació en el desenvolupament lingüístic es confirma amb un estudi amb infants amb desenvolupament tardà de la parla que va trobar que les mares d'aquests infants usaven un f_0 més baix i un contorn d' f_0 menor que les mares dels infants amb un desenvolupament típic del llenguatge (D'Odorico & Jacob, 2006). El to i el rang tonal tenen també un paper essencial en el manteniment de l'atenció (Fernald & Kuhl, 1987). Segons els estudis de Fernald i Kuhl (1987), el rang entonatiu de la parla entre adults oscil·la entre els 90 i els 300 Hz, mentre que en el cas de la parla dirigida a infants es troba entre 90 i 800 Hz, tot i que pot variar. Per aquest motiu, un f_0 més alt en combinació amb un contorn monòton no captaria l'atenció dels nens tant com ho faria un f_0 més baix amb moviments entonatius de més rang.

Els estudis acostumen a centrar-se en l'anàlisi de la parla dirigida a nens preverbals, perquè com hem dit, més tard, les característiques del IDS s'atenuen, però sabem que hi ha una continuïtat en aquesta modificació prosòdica que continua sent significativa (Díez-Itza, 1993), ja que els adults modifiquen la parla segons les capacitats cognitives

dels nens per facilitar la seva adquisició lèxica, gràcies a l'efecte de la prosòdia sobre el reconeixement de paraules en la parla (Singh et al., 2009).

Els estudis esmentats se centren sobretot en l'anglès, encara que també hi ha mostres d'estudis interlingüístics. En el cas del català, els pocs estudis existents sobre l'IDS se centren en aspectes prosòdics generals. Per exemple, Feijóo i Hilferty (2013), van provar que l'IDS ajuda a segmentar paraules, però no es van analitzar quines pistes acústiques concretes ajudaven més a la segmentació. També s'ha analitzat el nivell d'isosincronia sil·làbica (més alta en IDS) i l'allargament vocàlic (Payne et al., 2010, 2015). A més Payne et al. (2015) proven que no hi ha diferències prosòdiques significatives en la manera d'adreçar-se a nens de 2, 4 i 6 anys.

Aquest treball és un estudi preliminar que, amb una mostra de 8 participants, fa un pas més a l'estudi de les característiques prosòdiques de l'IDS en català. El treball fa un estudi acústic de l'entonació i la intel·ligibilitat de la parla i a més analitza una població normalment negligida als estudis, els professionals de l'educació. Per tant, el nostre objectiu serà el de caracteritzar entonativament la parla dirigida a infants en català.

2. Metodologia

Aquest estudi compara la veu d'un grup d'educadors (logopedes i mestres) en els registres adult i infantil, els següents paràgrafs en detallen el procés.

2.1. Participants

Els participants d'aquest estudi són adults professionals del món educatiu. S'ha escollit aquesta mostra per dos motius principals. En primer lloc, els treballs amb aquest tipus de població són menys habituals i, en cas del català, inexistents. I en segon, perquè tot i que resulta previsible que els educadors facin ús d'aquest registre amb els infants donada la naturalesa de la seva professió, el seu ús del IDS podria diferir dels proposats fins ara, donat que no formen part del cercle més íntim dels infants i, per tant, hi tenen menys vincles afectius, la qual cosa podria afectar a la tria de registre.

S'han analitzat 8 parlants sense patologies conegudes que viuen i exerceixen a Catalunya i tenen el català com a primera llengua, excloent-hi parlants bilingües no dominants de català (Hoff et al., 2020). Tots vuit són de sexe femení i professionals del món de l'educació, dividits en 2 logopedes i 6 mestres d'infantil. Caldria destacar que no s'han comparat ambdós grups perquè el que es compara és el registre infantil amb l'adult, sense diferenciar professió, perquè totes dues ja queden dins de l'àmbit educatiu. Per últim, els infants a qui es dirigeixen els informants per a l'estudi comprenen el rang d'edat de 3 a 6 anys. El rang d'edat de les participants va dels 27 als 54 anys amb una experiència de 5 a 27. Per a una informació més detallada dels informants, es pot consultar la taula 1.

TAULA I. DETALLS DELS PARTICIPANTS

Participant	Professió	Edat	Experiència docent
05	Logopeda	27	5
06	Logopeda	38	18
07	Mestra	30	7
09	Mestra	54	20
10	Mestra	40	19
11	Mestra	37	16
12	Mestra	49	27
13	Mestra	31	10

2.2. Procediment

Les gravacions es van realitzar als centres on les professionals exerceixen (un centre d'educació en el que hi treballen les logopedes i un centre escolar concertat de la localitat de Premià de Mar, Barcelona).

Al centre de logopèdia els enregistraments es van dur a terme en un mateix dia a la tarda, tant pel registre amb adult com per al dirigit als infants i, en el cas del centre escolar, en dos matins consecutius. Pel que fa a la durada de les sessions per a l'obtenció de la veu en registre adult-adult, s'han enregistrat uns 15-20 minuts de parla no estructurada amb l'enregistrador (Taula 2). En el cas de les gravacions per al registre adult-infant s'han enregistrat 30 minuts d'una sessió habitual, fos en visita logopèdica o en impartició de classe a l'aula de l'escola. En la primera hi havia la logopeda i l'infant i, en la segona, la mestra i entre 15 i 22 alumnes.

TAULA 2. TAULA D'INFORMANTS

Participant	Professió	Edat infants	Durada de parla neta		N. total de paraules	
			IDS*	ADS**	IDS	ADS
05	Logopeda	5	4:26	4:39	634	923
06	Logopeda	5	3:41	2:03	713	386
07	Mestra	3-4	5:34	3:49	817	874
09	Mestra	5-6	5:05	3:19	905	518
10	Mestra	4-5	2:49	3:19	485	686
11	Mestra	4-5	2:36	3:10	432	753
12	Mestra	5-6	3:36	3:01	447	543
13	Mestra	5-6	5:40	3:00	946	720

*IDS: *Infant Directed Speech*

**ADS: *Adult Directed Speech*

Pel que fa les característiques de les mostres de parla, els temps de la parla amb adult oscil·la al voltant dels 4 minuts amb una mitjana de paraules de 672 i per a la parla amb infants al voltant dels 3 minuts amb una mitjana de paraules de 675. Així, la comparabilitat de les mostres recau en la similitud de la gravació analitzada, en la comparació d'un mateix individu en ambdós registres (mostres aparellades) i en l'edat dels infants per al registre infantil. L'únic que els diferencia és la persona a qui es dirigeixen en cada registre, diferència amb que treballa les hipòtesis aquest estudi.

2.3. Instruments

La recollida de dades s'ha realitzat amb una gravadora portàtil Zoom, model H4n Pro combinada amb dos micròfons. En una de les sessions, es va fer servir un micròfon de diadema Shure. Per facilitar la comoditat dels participants, en la resta de sessions s'ha fet servir un micròfon Røde portàtil Lavalier Go. Totes les pistes s'han guardat en estèreo i en format PCM (wav), a 44,1 kHz i 16 bits. Les professionals del centre escolar duien la gravadora a la butxaca de la bata i el micròfon enganxat a la solapa; les logopedes duien el micròfon posat i la gravadora jeia prop seu.

2.4. Anotació de les dades

Per a la selecció d'enunciats s'han segmentat els àudios amb el programa GoldWave (Goldwave Inc., 2015), creant arxius diferents per a cada informat i anomenant-los segons fos parla dirigida a adults o parla dirigida a infants. S'han eliminat aquells segments que incloïen veus infantils en solapament o de manera aïllada. La durada dels arxius (Taula 2), s'ha reduït a un mínim de 2 i un màxim de 6 minuts de parla neta. Posteriorment, s'han anotat els àudios amb Praat (Boersma & Weenink, 2021), segmentant per oracions i transcrivint ortogràficament als TextGrids el seu contingut.

2.5. Extracció de paràmetres

Els TextGrids resultants de l'anotació i els seus wavs s'han fet servir posteriorment per poder fer l'extracció de les dades (F0 i espectre) amb el programa R (Team R Core, 2021).

Específicament, per tal de calcular l'f0 de cadascun dels fragments d'àudio s'ha fet servir l'algorisme de detecció de periodicitat de K. Schaefer-Vincent (time-domain) i s'ha extret un valor per cada 5 ms. Aquests valors s'han convertit a semitons de base 100, per tal de poder restar cada valor en semitons la mediana del parlant.

L'error més freqüent dels algorismes de detecció de F0, s'ha determinat que el valor de F0 és una octava més alt o més baix que el valor real. Això, que en Praat s'acostuma a corregir manualment entrant a l'objecte Pitch i corregint, aquí s'ha fet computacionalment. Per cada locutor, s'han calculat clústers de densitat i s'ha identificat el clúster més nombrós com aquell que contenia valors reals. A la resta de clústers se'ls ha sumat o restat una

octava per tal que quedessin inclosos dins del clúster més freqüent. Un cop acabada la normalització del contorn de pitch, s'han tornat a transformar els valors a Hertz.

L'espectre del senyal s'ha extret també en R creant un espectre mig per parlant que s'ha fet servir per calcular el pendent.

2.6. Paràmetres analitzats i anàlisi estadística

Amb els 16 arxius d'àudio i els 16 *TextGrids* s'ha elaborat l'anàlisi estadística amb el programa R (Team R Core, 2021). Com s'avançava a la introducció, aquest treball se centra en les següents mesures del to com a variables: freqüència fonamental, rang entonatiu, variabilitat del to i intel·ligibilitat (a partir de les dades del pendent espectral) per cada participant en parla dirigida a adult i infant. Amb aquestes dades com a variables dependents s'han realitzat models lineals mixtos generalitzats pels quals s'aporten les dades del model i la mida d'efecte a partir de la d de Cohen, tal com s'expressa al paquet glmm (Knudson, 2022).

2.6.1. El to o freqüència fonamental

Aquest paràmetre és el més comprovat en una gran varietat de llengües i de variables (Fernald et al., 1989; Garnica, 1977; Niebuhr & Skarnitzl, 2019; Remick, 1973; Shute & Wheldall, 1999) perquè representa el punt de partida des d'on sorgeixen els altres paràmetres que complementen la caracterització de la parla dirigida als infants. Així doncs, el to d'un so, com explica de Boysson-Bardies (2010), ve determinat per la freqüència de la vibració de les cordes vocals, la qual s'anomena freqüència fonamental, f_0 o *pitch*. Com més ràpid vibren les cordes vocals, més alt serà el to, l' f_0 .

El to mig s'ha calculat a partir de la mediana de tots els punts de f_0 extrets de cada frase dels locutors i a partir d'aquí s'ha fet la mitjana aritmètica de totes les frases per parlant (Taula 4). L'anàlisi estadística implementa una anàlisi lineal d'efectes mixtos de la relació entre l' f_0 i el registre o condició (Bates et al., 2011). Com a efecte fix s'ha introduït la condició (sense terme d'interacció) i com a efectes aleatoris, els pendents aleatoris de cada informant per a la condició estudiada.

2.6.2. Rang entonatiu

En segon lloc, i també com a derivat de l'anàlisi del to fonamental, s'ha treballat amb el rang de l' f_0 , és a dir, amb l'espai entre el to més baix i el to més alt, que conforma l'àrea en què es mouen la freqüència fonamental o f_0 i la variabilitat del parlant. El càlcul del rang s'ha fet mitjançant la subtracció del to més alt menys el to més baix.

També s'inclou un model lineal mixt generalitzat de la relació del rang tonal amb la condició de parla amb adults o amb infants (Brooks et al., 2017). Com a efectes fixos, s'ha inclòs la interacció del rang amb la condició estudiada. I com a efectes aleatoris, s'ha introduït la interacció de la condició per parlant i per frase. Per tal de saber si la variable estudiada era significativa pel model s'ha fet una relació de semblança (anova dels dos models contrastats) del model amb i sense la variable estudiada.

2.6.3. Variabilitat

La variabilitat de l' f_0 és un altre dels paràmetres que més resultats ha donat en estudis previs del registre adreçat als infants (Fernald & Kuhl, 1987; Quigley et al., 2019; Shute & Wheldall, 1999) com un paràmetre essencial en diversos estudis (Fernald et al., 1989). És el terme que tracta la variació en l'entonació, és a dir, la variació de la freqüència fonamental.

En aquest treball s'han fet servir dos indicadors de variabilitat la desviació estàndard del contorn entonatiu en cada parlant i condició i el *Pitch Dynamism Quotient* (PDQ) (Hincks, 2004). La variació estàndard dels valors de f_0 filtrats s'han extret per registre.

El PDQ és una mesura de variabilitat normalitzada, que es calcula a partir de la desviació estàndard del f_0 :

$$PQD = \frac{\sigma}{\text{mitjana}}$$

2.6.4. Intel·ligibilitat

La intel·ligibilitat és un paràmetre perceptiu de claredat de la parla que no té un correlat acústic únic. S'han ressaltat com a paràmetres importants la velocitat de parla, la intensitat, l'estat estacionari de les vocals, i l'espectre entre d'altres (Hargus i Quené, 2014). Per aquest treball s'ha optat fer una aproximació a partir dues mesures espectrals: el pendent espectral i l'existència del formant de locutor o de cantant (Bele, 2006).

Pel fet que el sistema auditiu infantil està ajustat per ser més sensible a les diferències de pendent espectral en el rang de les que es troben habitualment en la parla i la música (Tsang & Trainor, 2002), s'ha considerat un element important a observar en aquest estudi, ja que es pot relacionar amb el formant de locutor en tant que és una habilitat oratòria.

El formant de locutor és un artefacte en forma de pic a l'espectre que apareix a normalment a les veus professionals (locutors, cantants o actors) quan fan servir un tipus de veu projectada per tal d'augmentar la seva qualitat. Articulatoriament s'aconsegueix tensant el tub epilaringi eixamplant la laringe i fent més estreta la part frontal de la cavitat oral (Leino et al, 2011).

El pic a l'espectre es troba normalment a les freqüències que hi hauria entre l' F_3 i F_4 .

En quant al pendent espectral, podem dir que les parles més intel·ligibles serien aquelles que tenen un pendent espectral menys acusat, amb més intensitat a freqüències altes, és a dir, els formants F_3 , F_4 , F_5 i F_6 (Gurlekian et al., 1996).

La determinació de si existeix formant de locutor o no en una gravació no és fàcil perquè sempre hi haurà un pic associat a les freqüències de l' F_3 i F_4 , però a les nostres dades comptar amb gravacions de la mateixa persona en dues condicions possibilita comparar-les directament de manera gràfica tal i com es veu a la Figura 1.

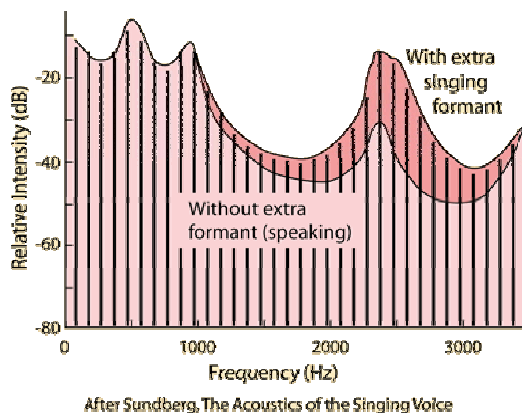


FIGURA 1. FORMANT DEL LOCUTOR, ADAPTADA DE NAVE (2001)

En quant al pendent espectral, s'ha extret a partir d'una regressió lineal que indica quants hertz baixa la intensitat, i s'ha fet la comparació a través de la mitjana d'hertz per interval, per parlant i per condició. Després, s'ha comprovat si aquesta era significativa fent una relació de semblança (anova dels dos models contrastats) de un model que incloïa la variable estudiada i sense la variable.

3. Resultats

3.1. Freqüència fonamental

Les mesures del to per a l' f_0 que es poden veure a la taula 4 mostren una mitjana de 189 Hz quan els receptors són infants i de 243 Hz quan són adults. Així, es pot confirmar el primer punt de la hipòtesi, ja que la freqüència fonamental és més alta amb infants. La mitjana de la freqüència fonamental es pot veure a la taula 3, on s'observa que la condició d'IDS implica una pujada del to de 65.56 Hz (amb un error de +/- 10.55). A la Figura 2 es pot observar la diferència de la mitjana del to segons la condició del registre adult i de l'infant. A la Figura 3 es contempen aquestes mateixes dades de manera desglossada per parlant.

TAULA 3. FREQUÈNCIA FONAMENTAL

	Estimate Std.	Error
(intercepte per ADS)	190.54 Hz	+/- 11.39
Condició IDS	+65.56 Hz	+/- 10.55

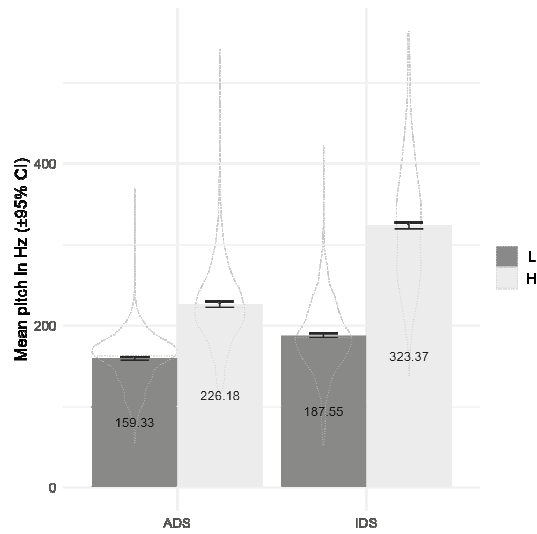


FIGURA 1. MITJANA DEL TO SEGONS CONDICIÓ

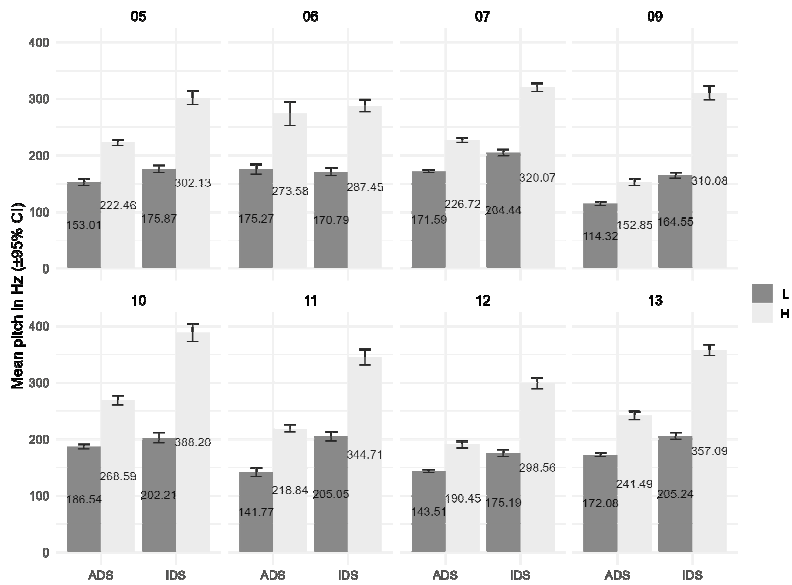


FIGURA 2. MITJANA DEL TO SEGONS EN MÍNIMS I MÀXIMS SEGONS CONDICIÓ PER PARLANT

3.2. Expansió del rang

Tal com il·lustren les figures 4 i 5, el rang en IDS és més ampli que no pas en ADS ($\chi^2= 41.759$, $p<0.001$). Més específicament, el rang en l'IDS és de 489 Hz mentre que amb adults és de 312 Hz, com es pot observar en les dades de la taula 4, restant els mínims dels màxims. La mida de l'efecte d'IDS és $d = -1.403$, $p<0.001$. Aquestes dades discerneixen de les de Fernald et al. (1989) que troben un registre molt més ample mentre que les nostres dades tal com mostra la taula 3, el rang es mou entre 104 i 593 Hz. El rang adult també discerneix i se situa entre els 62 i els 374 Hz.

TAULA 4. VALORS DE L'FO EN HZ

Participant	Hz mínim		Hz màxim	
	Infant	Adult	Infant	Adult
05	50,79	50,52	592,35	385,70
06	50,75	106,78	595,95	569,67
07	53,09	105,13	583,65	353,51
09	83,75	78,67	583,59	282,11
10	71,87	105,83	595,95	483,53
11	120,02	54,99	595,60	363,49
12	103,21	103,04	445,82	322,18
13	50,63	123,61	594,04	411,32
Mitjana	104,08	62,4799	593,20	374,60

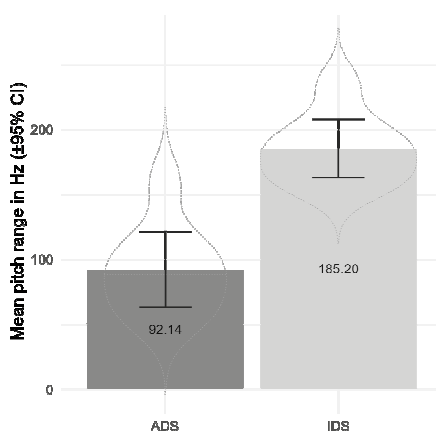


FIGURA 4. MITJANA DEL RANG SEGONS CONDICIÓ

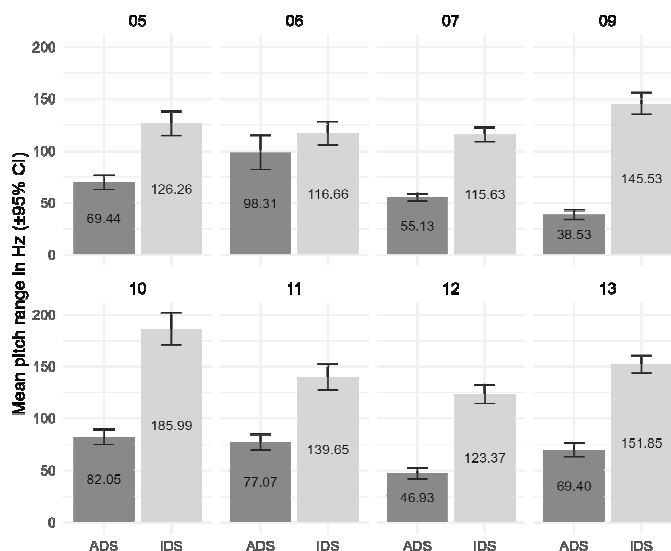


FIGURA 5. MITJANA DEL RANG SEGONS CONDICIÓ I PARLANT

3.3. Variabilitat

Pel que fa a la variabilitat, s’observa a la taula 5 una desviació estàndard de 74 amb infants i de 47 amb adults, diferència que queda corroborada amb la PDQ de 0,29 amb el primer grup i 0,24 amb el segon a la taula 5. A més a més, es confirma que els resultats de variabilitat del to amb PDQ són significatius perquè el resultat del test t per a dades parellades (dues condicions per cada informant) és $t(18,7) = 3.9785, p = 0.005334$.

TAULA 5. MITJANES DELS PARÀMETRES EN HZ

	F0	Rang	Var. Ds*	Var. PDQ**	Pendent
Infant	189	545	74	0,29	
Adult	255	519	47	0,24	

*Variació segons la desviació estàndard

** Variació segons la *Pitch Dinamism Quotient*

En els punts de la figura 6 es pot observar com la variabilitat, segons la PDQ, és més elevada en parla adreçada a infants (IDS), destacada en el color vermell, que la parla adreçada a adults (ADS), en color blau ($\chi^2=36.491, p<0.001$) i l’efecte de la condició

($d = 0.941$, $p < 0.001$) és significativa. Així, queda confirmat el tercer punt de la hipòtesi, pel que fa a la intel·ligibilitat.

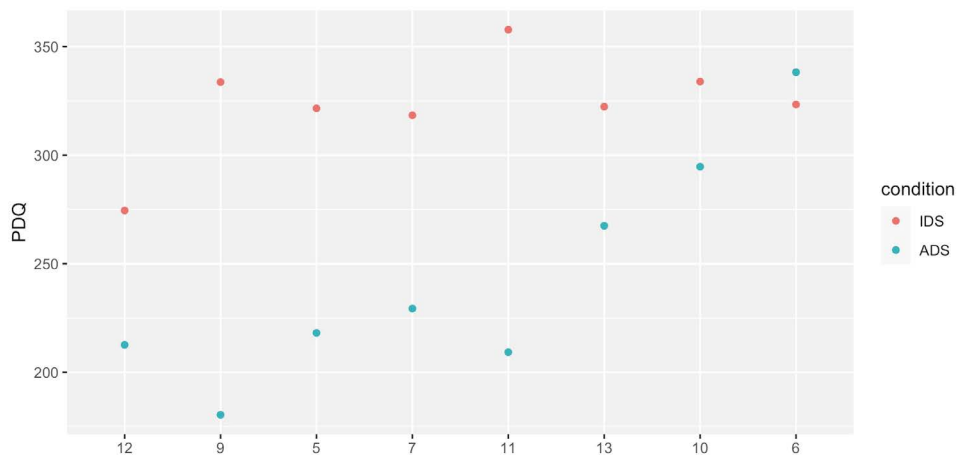


FIGURA 6. REPRESENTACIÓ DELS VALORS DE PITCH DYNAMISM QUOTIENT (EIX VERTICAL) PER INFORMANT (EIX HORIZONTAL).

1.4. Mesures espectrals

Aquest paràmetre s'ha usat com a una mesura per a la intel·ligibilitat. A la taula 6 es pot interpretar que la condició d'IDS implica una pujada del pendent de 1.0113Hz (± 0.36). La informació que s'extreu de les dades és que el pendent és menys inclinat en ADS que en IDS, com mostra la figura 7.

TAULA 6. RESUM DE DADES DE REGRESSIÓ DEL PENDENT ESPECTRAL

	Estimate Std.	Error
(intercepte per ADS)	-0.5938Hz	+/- 0.1751
Condició IDS	+1.0113Hz	+/- 0.3653

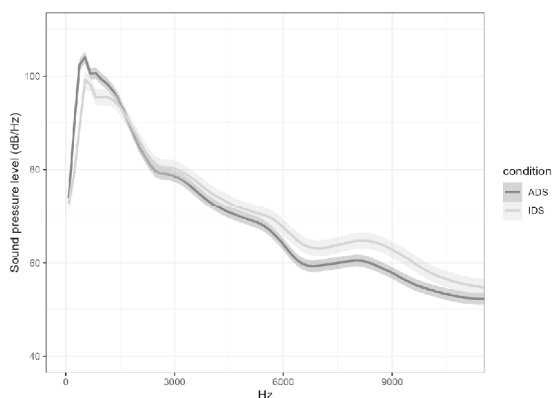


FIGURA 6. PENDENT ESPECTRAL PER CONDICIÓ

A la figura 7 es pot veure la comparativa del pendent espectral per cada participant i en ambdues modalitats de registre en què es pot observar que el pendent és més gran en ADS (línia de sota excepte al participant 12) ($\chi^2=7.664$, $p=0.006$) amb una mida d'efecte de $d=-1.596$, $p=0.006$. A més es pot veure de manera gràfica que per 6 de 8 informants apareix un pic a les freqüències altes que coincidiria amb el formant de locutor. Aquest és especialment clar pels informants 07 i 09.

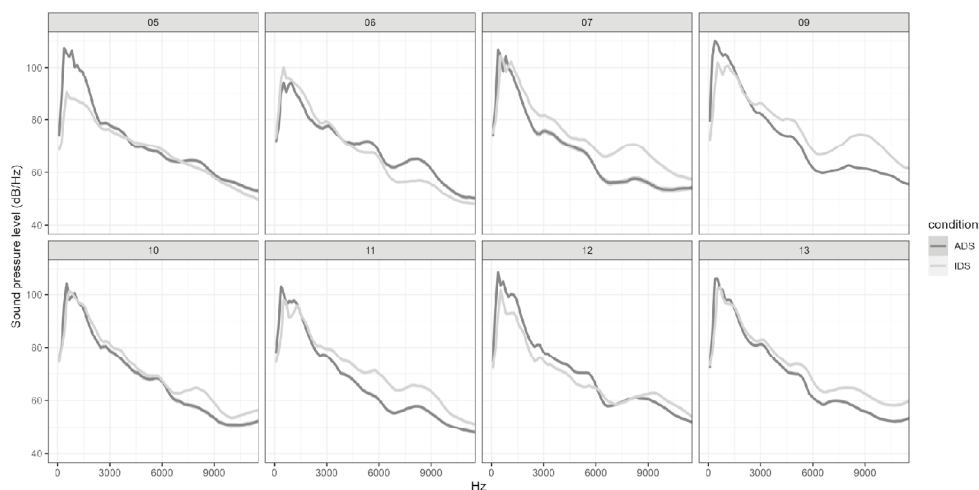


FIGURA 7. PENDENT ESPECTRAL PER PARTICIPANT QUE MOSTRA LA FREQUÈNCIA DELS HARMÒNICS (EIX HORIZONTAL) I LA INTENSITAT (EIX VERTICAL) PER LES DUES CONDICIONS ESTUDIADAES.

4. Discussió

Els resultats d'aquest treball van en línia del que ja s'ha descrit per altres llengües. Altres estudis interlingüístics, com el de Payne et al (2010), també han pogut veure que les dades fonètiques van en la mateixa direcció.

En particular, quant a les dades d'alçada tonal existeix un tret biològic i és que el sistema auditiu madura primer per percebre freqüències altes (Schneider & Trehub, 1992). Això podria explicar que els infants atenguin més a aquestes freqüències més elevades perquè les senten millor.

Consegüentment, els adults adaptarien la seva parla per aconseguir aquesta atenció per part dels seus interlocutors. D'altra banda, i arran de les converses amb les mestres que han participat en l'experiment, l'elevació del to dels adults també pot ser un intent tal de reforçar una relació d'igual a igual, com afirmen altres estudis (Sudartinah, 2016).

En el cas del pendent espectral, Piazza et al. (2017) trobaven en un estudi de 10 llengües que justament la freqüència fonamental no era un tret interlingüísticament comparable, però sí que ho era el timbre de veu com a empremta vocal, és a dir, el pendent espectral seria el que ajuda els infants a diferenciar les veus.

Hi ha diversos autors que, més enllà de la justificació de l'adquisició de la llengua, han buscat una explicació evolutiva del perquè d'aquest canvi de prosòdia envers els infants: d'on ve, com va sorgir i per què. Una possible teoria és que l'expressió vocal sigui filogenèticament més antiga que l'ús de la llengua (Trainor & Desjard, 2002) i que, per conveniència, aquesta expressió vocal derivés funcionalment cap a la discriminació vocàlica amb l'objectiu de transmetre als nens uns coneixements lingüístics, és a dir, usar una parla més intel·ligible perquè els infants entenguin les estructures de la llengua.

En general, i després d'examinar dades de llengües diferents, semblaria que no és un tret o un paràmetre acústic concret, sinó la positivitats i l'afecte del parlant adult (Kitamura et al., 2001) i això s'expressa normalment en la modulació del to.

5. Conclusions

Aquest treball mostra que, també en el cas dels cuidadors professionals com mestres i logopedes catalans, la parla adreçada a infants presenta notables diferències fonètiques de la parla adreçada a adults en els paràmetres d'alçada tonal, el rang tonal, la variabilitat del to i la intel·ligibilitat de la parla.

En el nostre estudi hem pogut observar una pujada del to en tots els participants quan parlen a infants respecte de quan parlen a altres adults. A més a més, s'ha afirmat la diferència de rang tonal, expandit en parla a nens. Dins d'aquest rang més elevat i més ample s'ha verificat l'existència d'una variabilitat o modulació del to més present en la parla adreçada a infants, tant a través de la desviació estàndard i del *Pitch Dinamism Quotient* (PDQ). A més, també hem pogut comprovar que es modifiquen paràmetres

espectrals, probablement per millorar la intel·ligibilitat, cosa que provoca un pendent spectral més petit en IDS i, de vegades, l'aparició del formant de locutor.

En definitiva, els professionals de l'educació fan servir un rang, un to i característiques espectrals a la veu que faciliten la percepció dels infants tal com passa en altres col·lectius (cuidadors i familiars més propers, com els pares) i en altres llengües.

Referències

- Bates, D., Mächler, M., & Dai, B. (2011). lme4: Linear Mixed-Effects Models Using Eigen and S4 Classes. In *R Package Version 0.999375-32: Vols. 0.999375-33*.
- Bele, I. V. (2006). The speaker's formant. *Journal of Voice*, 20(4), 555-578.
- Boersma, P., & Weenink, D. (2021). *Praat: doing phonetics by computer* (6.2.03). University of Amsterdam.
- Broesch, T. L., & Bryant, G. A. (2015). Prosody in Infant-Directed Speech Is Similar Across Western and Traditional Cultures. *Journal of Cognition and Development*, 16(1), 31-43. <https://doi.org/10.1080/15248372.2013.833923>
- Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler, M., & Bolker, B. M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, 9(2), 378-400. <https://journal.r-project.org/archive/2017/RJ-2017-066/index.html>
- de Boysson-Bardies, B. (2010). How Language Comes to Children: From Birth to Two Years. In *Cognitive Linguistics Bibliography (CogBib)*. De Gruyter Mouton. https://www.degruyter.com/database/COGBIB/entry/cogbib_1540/html
- Díez-Itza, E. (1993). Variaciones tonales en el habla a los niños y adquisición del lenguaje. *Estudios de Psicología*, 14(50), 33-47. <https://doi.org/10.1080/02109395.1993.10821193>
- D'Odorico, L., & Jacob, V. (2006). Prosodic and lexical aspects of maternal linguistic input to late-talking toddlers. *International Journal of Language and Communication Disorders*, 41(3), 293-311. <https://doi.org/10.1080/13682820500342976>
- Feijóo, S., & Hilferty, J. (2013). The availability of cues for word segmentation and vocabulary acquisition in Catalan child-directed speech. *RLA. Revista de Lingüística Teórica y Aplicada*, 51(2), 13-27. <https://doi.org/10.4067/S0718-48832013000200002>
- Ferguson, C. A. (1964). Baby Talk in Six Languages. *American Anthropologist*, 66(6), 103-114. https://doi.org/10.1525/aa.1964.66.suppl_3.02a00060
- Fernald, A. (1989). *Intonation and Communicative Intent in Mothers' Speech to Infants: Is the Melody the Message?* 60(6), 1497-1510.
- Fernald, A., & Kuhl, P. (1987). Acoustic Determinants of Infant Preference for Motherese Speech. *Infant Behavior and Development*, 10(3), 279-293. [https://doi.org/https://doi.org/10.1016/0163-6383\(87\)90017-8](https://doi.org/https://doi.org/10.1016/0163-6383(87)90017-8)

- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16(3), 477–501. <https://doi.org/10.1017/S0305000900010679>
- Frank, M. C. (2020). Quantifying Sources of Variability in Infancy Research Using the Infant-Directed-Speech Preference. *Advances in Methods and Practices in Psychological Science*, 3(1), 24–52. <https://doi.org/10.1177/2515245919900809>
- Garnica, O. K. (1977). Some prosodic characteristics of speech to young children a dissertation submitted to the committee on linguistics and the committee on graduate studies. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 63-88). Cambridge University Press.
- Goldwave Inc. (2015). *GoldWave* (6.15). www.goldwave.com.
- Gurlekian, J., Dumm, N., & Hongay, M. (1996). Aplicación de la tecnología informática al estudio de la voz del locutor profesional. *Ecos Fonoaudiológicos*, 1(3).
- Han, M., de Jong, N. H., & Kager, R. (2018). Lexical tones in Mandarin Chinese infant-directed speech: Age-related changes in the second year of life. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.00434>
- Hargus Ferguson, S., & Quené, H. (2014). Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 135(6), 3570-3584.
- Hincks, R. (2004). Processing the prosody of oral presentations. *InSTIL/ICALL Symposium*, 63-69.
- Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2001). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development*, 24(4), 372-392. [https://doi.org/10.1016/S0163-6383\(02\)00086-3](https://doi.org/10.1016/S0163-6383(02)00086-3)
- Knudson, C. (2022). *Glmm: Generalized Linear Mixed Models via Monte Carlo Likelihood Approximation* (1.4.4). <https://CRAN.R-project.org/package=glmm>
- Nave, C. R. (2001). *Sung and spoken sounds*. <http://hyperphysics.phy-astr.gsu.edu/hbase/Music/imgmus/sfor2.gif>
- Leino, T., Laukkanen, A. M., & Radolf, V. (2011). Formation of the actor's/speaker's formant: a study applying spectrum analysis and computer modeling. *Journal of voice*, 25(2), 150-158.
- Niebuhr, O., & Skarnitzl, R. (2019). Measuring a speaker's acoustic correlates of pitch – but which? A contrastive analysis based on perceived speaker charisma. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 1774-1778). Australasian Speech Science and Technology Association Inc.
- Payne, E., Post, B., Astruc, L., Prieto, P., & Vanrell, M. del M. (2010). A cross-linguistic study of prosodic lengthening in child-directed speech. *Speech Prosody*, 5.

- Payne, E., Post, B., Astruc, L., Prieto, P., & Vanrell, M. del M. (2015). Rhythmic modification in child directed speech. In M. Russo (Ed.), *Prosodic Universals. Comparative Studies in Rhythmic Modeling and Rhythmic Typology* (pp. 147-183). Aracne.
- Piazza, E. A., Jordan, M. C., & Lew-Williams, C. (2017). Mothers Consistently Alter Their Unique Vocal Fingerprints When Communicating with Infants [Article]. *Current Biology*, 27(20), 3162-3167. <https://doi.org/10.1016/j.cub.2017.08.074>
- Quigley, J., Nixon, E., & Lawson, S. (2019). Exploring the association of infant receptive language and pitch variability in fathers' infant-directed speech. *Journal of Child Language*, 46(4), 800-811. <https://doi.org/10.1017/S0305000919000175>
- Remick, H. (1973). *Maternal Speech to Children During Language Acquisition*. University of California.
- Schneider, B. A., & Trehub, S. E. (1992). Sources of developmental change in auditory sensitivity. In *Developmental psychoacoustics*. (pp. 3–46). American Psychological Association. <https://doi.org/10.1037/10119-001>
- Shute, B., & Wheldall, K. (1999). Fundamental frequency and temporal modifications in the speech of British fathers to their children. *Educational Psychology*, 19(2), 221–233. <https://doi.org/https://doi.org/10.1080/0144341990190208>
- Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy*, 14(6), 654-666. <https://doi.org/10.1080/15250000903263973>
- Sudartinah, T. (2016). The Role of Parentese in First Language Acquisition: A Psycholinguistic Study. *Journal of English and Education*, 2, 54–66. <https://doi.org/10.20885/jee.vol2.iss1.art6>
- Team R Core. (2021). *R: A Language and Environment for Statistical Computing*. <https://www.r-project.org/>
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is Infant-Directed Speech Prosody a Result of the Vocal Expression of Emotion? *Psychological Science*, 11(3), 188–195. <https://doi.org/https://doi.org/10.3758/BF03196290>
- Trainor, L. J., & Desjard. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels LAUREL. *Psychonomic Bulletin & Review*, 9(2), 335–340. <https://doi.org/10.3758/BF03196290>
- Tsang, C. D., & Trainor, L. J. (2002). Spectral slope discrimination in infancy: Sensitivity to socially important timbres. *Infant Behavior & Development*, 25, 183–194. [https://doi.org/10.1016/S0163-6383\(02\)00120-0](https://doi.org/10.1016/S0163-6383(02)00120-0)