

**ALTERACIONES DE LA DECLINACIÓN Y EL CAMPO TONAL
EN EL HABLA DE PERSONAS CON SÍNDROME DE DOWN**

**SPEECH DIFFICULTIES OF PEOPLE WITH DOWN SYNDROME:
CHANGES IN DECLINATION AND TONAL REGISTER**

LOURDES AGUILAR

*Universitat Autònoma de Barcelona
(Espanya)*

Lourdes.aguilar@uab.cat

JOAN BORRÀS-COMES

*Universitat Autònoma de Barcelona
(Espanya)*

joanborrascomes@gmail.com

Artículo recibido el día: 02/04/2020

Artículo aceptado definitivamente el día: 16/06/2020

Estudios de Fonética Experimental, ISSN 1575-5533, XXIX, 2020, pp. 11-33

ABSTRACT

The study examines the phonetic modelling of declination, a measure that relates the F0 contour with the temporal properties, in a corpus of speech of young adults with Down Syndrome (DS) in order to provide reasonable interpretations of the subjective evaluation of monotonous speech that is usually attributed to this group. The acoustic analysis of speech fundamental frequency (F0) during the utterance in a set of declarative sentences pronounced by speakers with DS and its comparison with the same set of sentences spoken by non-disabled speakers has verified the tendency of F0 to decline gradually during utterances in both groups. Nevertheless, both the slope and the difference between high tones and low tones throughout the sentence is lower, a decrease that may explain, in part, the monotonal effect of speech of people with DS.

Keywords: *Down Syndrome, intonation, declination, acoustics, Spanish.*

RESUMEN

El presente estudio examina la manifestación fonética de la declinación, medida que relaciona el contorno de F0 con las propiedades temporales, en un corpus de habla de jóvenes adultos con Síndrome de Down (SD) con el fin de aportar interpretaciones razonables sobre la evaluación subjetiva de monotonía que suele atribuirse a este colectivo. El análisis acústico de los valores de F0 a lo largo de la frase en un conjunto de frases enunciativas pronunciadas por hablantes con SD y su comparación con el mismo conjunto de frases pronunciadas por hablantes sin discapacidad ha permitido comprobar la tendencia del F0 a disminuir gradualmente hasta el final del enunciado en el habla de personas con SD, pero que tanto la pendiente como la diferencia entre tonos altos y tonos bajos a lo largo del enunciado es menor, lo que puede explicar, en parte, el efecto de monotonía en el habla de personas con SD.

Palabras clave: *Síndrome de Down, entonación, declinación, acústica, español.*

1. INTRODUCCIÓN

Podemos convenir que el alto grado de variación y opcionalidad que hallamos en la prosodia dificulta en gran medida su estudio sistemático: los patrones melódicos son muy variados y a menudo idiosincrásicos. Sin embargo, como ha demostrado la fonología de la entonación, no todas las variaciones son igualmente relevantes, y es posible definir categorías prosódicas que deben concebirse como parte de las gramáticas de las lenguas (Ladd, 2008; Pierrehumbert, 1980). En este marco conceptual, nuestro principal interés reside en explorar las posibilidades de aplicar los conocimientos sobre prosodia y entonación en la educación especial.

Formulamos la hipótesis global de que el dominio de la estructura prosódica y su uso pragmáticamente adecuado influyen decisivamente en el éxito comunicativo. Primero, al estructurar el flujo del habla, ayuda a identificar los límites de las palabras y las oraciones, lo que contribuye de manera vital a la comprensión del habla. Segundo, mediante señales prosódicas, los hablantes no solo pueden producir el tipo de oración deseado (pregunta o aseveración), sino también resaltar la información más importante en una frase. Por último, las características individuales de las emociones, actitudes o intenciones de los hablantes también están codificadas mediante pistas que se apoyan en los rasgos llamados suprasegmentales (la frecuencia fundamental (F0), duración e intensidad) (Ladd, 2008).

El hecho de que la prosodia sea de vital importancia para el éxito comunicativo contrasta con el poco interés que ha suscitado este componente en el estudio del habla de las personas con Síndrome de Down (SD). Existe una bibliografía amplia sobre el desarrollo del vocabulario y la gramática en niños con SD, pero en cambio la información disponible sobre el componente fonético y fonológico del habla de jóvenes y adultos, y, en particular, sobre la prosodia es escasa. Se ha observado que los niños con SD son menos capaces de usar la prosodia para expresar funciones pragmáticas del discurso que los niños de desarrollo típico (Zampini *et al.*, 2016); también que muestran dificultades para producir y entender preguntas y emociones, o para señalar el final de turno (Stojanovik, 2011), además de limitaciones para percibir, imitar y producir de forma espontánea una correcta entonación (Roberts *et al.*, 2007; Kent y Vorperian, 2013). En el dominio del ritmo, se describe el habla de personas con SD con interrupciones y tartamudeos frecuentes (Willcox, 1988; Martin *et al.*, 2009; Eggers y van Eerdenbrugh, 2018). En cuanto al registro tonal, se han detectado valores más agudos en los hablantes con SD comparados con otras poblaciones sin trastornos de habla (Lee *et al.*, 2009; Albertini, 2010; Rochet-Capellan y Dohen, 2015).

Junto a la escasez de estudios en el dominio prosódico, en muchas ocasiones, los términos usados para describir la prosodia de las voces patológicas suelen ser imprecisos y sin una relación directa con los rasgos suprasegmentales (voz “monótona”, “robótica”, o “que parezca no fluida”, “con demasiados cortes”, entre otras) y el foco de la intervención clínica se ha centrado en la entonación emocional (Peppé, 2009).

Además, como demuestran Rodríguez *et al.* (2012), la entonación es difícil de entrenar en una población con trastornos de la voz, entre otros motivos, en el caso de los hablantes con Síndrome de Down, colectivo estudiado en el presente estudio, porque es difícil definir un comportamiento lingüístico único. Por un lado, las personas con SD presentan una miríada de perfiles de desarrollo que dependen del coeficiente intelectual, las características faciales y físicas, etc. (Kumin, 2002). Por otro, el SD, como causa genética más común conocida de discapacidad intelectual, causa problemas variados en el lenguaje y el habla (Katz y Lazcano-Ponce, 2008; Dodd y Thompson, 2001).

Aún desconocemos por qué la prosodia del habla de personas con SD es como es y qué factores fonéticos y fonológicos determinan que dicha prosodia se aparte de los modelos representativos del habla general, convirtiéndose en la denominada *prosodia atípica*. Tampoco se ha determinado hasta qué punto la heterogeneidad en los perfiles lingüísticos de los hablantes con SD permite identificar unos patrones prosódicos comunes y así abordar de forma conjunta la intervención clínica.

Para avanzar en esa dirección, se hace preciso adquirir mayor conocimiento sobre la estructura prosódica (acento, organización prosódica y entonación) de los hablantes con SD, no solo desde la perspectiva del desarrollo del lenguaje, sino desde el punto de vista fonético y fonológico en el habla de jóvenes y adultos. Como primera aproximación, nos formulamos preguntas sobre la relación entre el contorno de entonación y las características temporales en la modalidad declarativa neutra en español en el habla de personas con SD con el fin de identificar las variables prosódicas que inciden en la identificación de una voz como discurso atípico.

En el marco teórico de la fonología entonativa (Ladd, 2008), el presente trabajo tiene como objetivo analizar una de las propiedades más extendidas, y ampliamente estudiadas, de la frecuencia fundamental del habla (F0): el descenso gradual de la frecuencia fundamental a lo largo de un enunciado. A esta tendencia se le ha dado una variedad de nombres, de los cuales el más conocido probablemente sea el de declinación. En general, se acepta que la declinación es un efecto fonético (debido a un mecanismo fisiológico y semiautomático de producción) de carácter universal,

al menos para las oraciones declarativas, aunque también se han señalado algunos efectos de la situación (frases aisladas con la introducción del tema nuevo al principio de la frase y sin énfasis) que favorecen dicho efecto mecánico (Umeda, 1982). A diferencia de la declinación (categoría fonética), el escalonamiento tonal descendente (*phonological downstep*) es una categoría fonológica: un declive tonal debido a un proceso fonológico controlado lingüísticamente por el hablante (Pierrehumbert y Liberman, 1982; Beckman y Pierrehumbert, 1986).

Dado que una de las características señaladas repetidamente en el habla de personas con Síndrome de Down es la falta de fluidez (Willcox, 1988; Martin *et al.*, 2009; Kent y Vorperian, 2013; Eggers y van Eerdenbrugh, 2018), nos proponemos examinar la manifestación fonética de la declinación, medida que relaciona el contorno de entonación con las propiedades temporales, con el fin de aportar interpretaciones razonables sobre la discontinuidad de su discurso. Se formula la hipótesis de que, en ausencia de disfemia (dificultad a la hora de comenzar el discurso o repeticiones derivadas de contracciones musculares involuntarias), la producción de enunciados declarativos por hablantes con Síndrome de Down muestra declinación, como efecto fonético, si bien el grado de cambio en los valores de F0 a lo largo del enunciado va a ser menor en comparación con un habla no patológica, lo que a su vez incide en el campo tonal de los hablantes con SD.

Planteamos las siguientes preguntas de investigación:

- PI1: ¿Podemos describir el efecto fonético de la declinación tonal en frases enunciativas pronunciadas por personas con SD?
- PI2: ¿Se pueden establecer patrones comunes de realización fonética de la declinación en el habla de personas con SD?
- PI3: ¿La manifestación de la declinación tonal en enunciativas pronunciadas por personas con SD presenta características diferenciadas con respecto a una población típica?
- PI4: En caso de que exista declinación, ¿pueden describirse diferencias en el campo tonal con respecto a una población típica?

2. PROCEDIMIENTO EXPERIMENTAL

2.1. Informantes

Las muestras de habla proceden del corpus Down recogido en el marco del proyecto Recercaixa-ACUP 2013 “Juguem a comunicar millor!”

(<http://prado.uab.cat/recercaixa/es/index.html>). Para este experimento, se han seleccionado seis hablantes de lengua materna castellana (3 hombres y 3 mujeres) de entre 13 y 22 años, con capacidades cognitivas y de aprendizaje similares. Para obtener medidas de los niveles de desarrollo verbal y cognitivo, se administró a todos los participantes la prueba de vocabulario en imágenes Peabody (Dunn *et al.*, 2006) y la prueba de matrices progresivas de Raven (Raven *et al.*, 1993). Según la prueba Peabody, la edad mental verbal de los participantes osciló entre 5 y 8,25, con un promedio de 6,6 años. Las puntuaciones de la prueba de Raven variaron entre 10 y 22, con un promedio de 15,7. El rango de puntuaciones tanto a nivel cognitivo verbal como no verbal es coherente con el que se describe para personas con SD (Chapman y Hesketh, 2000; Ruiz, 2001), de lo que se infiere que, a pesar del tamaño pequeño de la muestra, el grupo de estudio es representativo.

Las medidas obtenidas se compararán con una base de datos de referencia de hablantes sin discapacidad. El grupo de control consta de seis adultos (3 hombres y 3 mujeres, con edades entre 34 y 46 años), representativos del español peninsular septentrional, que desempeñan una labor profesional en el doblaje.

2.2. Corpus

El corpus Down completo incluye frases de diferente longitud y complejidad, organizadas en grupos según la modalidad entonativa (oraciones declarativas, interrogativas y exclamativas) y la expresión de emociones. En cuanto al modo de producción, consta de lectura de oraciones y de habla guiada, en la que se obtiene una respuesta desencadenada por un contexto comunicativo descrito por la entrevistadora. Por último, los informantes leen en voz alta un cuento corto representando el papel de uno de los personajes (Leyenda de Esopo: El viento del norte y el sol).

Este estudio se centra en la producción de las frases enunciativas: consta de 14 oraciones de estructura sintáctica simple (SVO) y con significado claro orientado hacia una actitud expresiva neutra, por ejemplo, *Iremos por autopista*.

2.3. Procedimiento de grabación

Antes de las sesiones de grabación, los tutores firmaron un acuerdo de cesión de uso de registros de voz con fines únicos de investigación para los centros participantes en el proyecto (Universidad Autónoma de Barcelona, Universidad de Valladolid, Universidad Nacional a Distancia), que se comprometen a no divulgar la identidad de los locutores.

Cada sesión constó de una primera parte en la que se obtuvo habla espontánea, con un encuentro informal entre el/la informante y la entrevistadora (logopeda de formación), con el fin de que el joven se familiarizara con las condiciones de grabación y los requisitos técnicos; y una segunda parte basada en la lectura y la producción guiada, de donde proceden las frases analizadas en el presente estudio.

Todas las sesiones de grabación tuvieron lugar en una sala acondicionada acústicamente del Servicio de Comunicación de la Universidad Pompeu Fabra con el programa Sony Vegas que se ejecutaba en un PC con una tarjeta de sonido RME Hammerfall HDSP 9652 y un mezclador Yamaha 02R96 con tarjetas ADAT MY16AT, a una frecuencia de muestreo de 48 KHz, y con un micrófono inalámbrico de diadema (Senheisser EW100-G2). Dados los retos que plantea obtener un corpus de voz con personas con SD (dificultades de atención y de procesamiento de instrucciones, entre otras emocionales), se les ofreció apoyo visual con una pantalla situada al lado del informante en la que aparecía la frase escrita con una ilustración representativa de su significado. También se realizaron los descansos necesarios para que los jóvenes no se sintieran presionados o incómodos.

2.4. Extracción de datos acústicos y establecimiento de la base de datos

Para constituir la base de datos de análisis, se usó el programa Praat (v. 6.1.04) (Boersma y Weenink, 2016), con el que se etiquetaron las muestras de habla y se extrajeron los valores relativos a los tiempos de inicio y fin (en segundos) de cada uno de los intervalos de interés. A continuación, se obtuvieron todos los puntos de F0 en Hz de cada una de las frases. Cada valor de F0 venía acompañado del nombre del archivo del que había sido obtenido, así como del momento exacto de su producción, con lo que se consiguió una tabla como la de la Figura 1. Todos los datos se importaron después al paquete estadístico R.

A partir de los valores de F0 en Hz, se detectaron —y posteriormente se eliminaron— los valores atípicos. Este procedimiento se realizó teniendo en cuenta la variación de cada hablante, de manera recursiva mientras se detectaran nuevos valores atípicos. El cribado se realizó mediante la función *Boxplot* del paquete *car* (recordemos que cada paquete es una colección de funciones diseñadas para atender una tarea específica en R). En concreto, se eliminaron 1477 casos de la tabla de datos (que constaba de 14710), lo que representó un 10,04% de los casos iniciales.

	Group	File	Speaker	Item	Time	Hz
1	Control	Eva_1	Eva	1	0.200531	228.3763
2	Control	Eva_1	Eva	1	0.210531	234.2452
3	Control	Eva_1	Eva	1	0.220531	234.4478
4	Control	Eva_1	Eva	1	0.230531	233.5192
5	Control	Eva_1	Eva	1	0.240531	234.0516
6	Control	Eva_1	Eva	1	0.250531	234.9520
7	Control	Eva_1	Eva	1	0.260531	234.6129
8	Control	Eva_1	Eva	1	0.270531	233.4350
9	Control	Eva_1	Eva	1	0.280531	229.8831
10	Control	Eva_1	Eva	1	0.290531	218.9879

Figura 1. Extracto de la tabla con los valores de F0 (Hz) y la información asociada del grupo (Group), fichero de voz (File), hablante (Speaker: nombre ficticio), palabra (Item) y tiempo (Time).

Con la ayuda del paquete *hqmisc*, a partir de los valores en Hz se calcularon los valores en semitonos en base 100 a partir de la función $f2st$, que fueron añadidos como nueva columna a la tabla (*st*).

Puesto que cada frase tuvo, lógicamente, una duración distinta, se añadieron dos columnas más a la tabla para indicar los tiempos en los que se produjeron el primero y el último de los puntos de F0 extraídos (todo ello después de haber eliminado los valores atípicos). Esos dos valores pasaron a ser considerados como el momento de inicio y final de cada frase, y a partir de ellos se calculó el tiempo relativo de producción de cada uno de los registros individuales de F0 (*Time_rel*). En este sentido, nuestro estudio de la declinación tonal se fundamenta en los tiempos relativos al inicio y fin de cada frase, para que de esta forma no influya en su descripción las diferencias de duración de las frases.

A continuación, para que el timbre de cada hablante no interfiriera a la hora de comparar los dos grupos de hablantes, todos los registros de F0 (en semitonos) se centraron con respecto a la media de F0 de cada hablante. Dicho de otro modo, a cada uno de los registros individuales de F0 se le restó la media de F0 obtenida para cada locutor, y el resultado de esas restas se guardó como una variable aparte (*st_scaled*). En este proceso de centrado, sin embargo, los valores de F0 no fueron estandarizados. Una estandarización habría dispuesto todos los valores de F0 en una misma escala y habría eliminado los semitonos como unidad de medida del F0 (es decir, perderíamos la distancia real en semitonos entre los tonos más graves y los

más agudos producidos por cada hablante), lo que imposibilitaría comparar los dos grupos de hablantes por lo que respecta a diferencias de campo tonal.

Para el estudio del campo tonal, era necesario obtener los puntos más graves y más agudos de cada frase, sin que en su detección pudiera afectar la declinación tonal. Cabe suponer que tanto al inicio como al final de una frase se pueden producir tonos graves y agudos, pero su realización estará por defecto afectada por la declinación. Para su obtención, primero se calculó, para cada frase, una regresión *loess* de amplitud 0,5 entre los registros de F0 centrados (*st_scaled*) y la línea de tiempo relativo al inicio y final de cada frase (*Time_rel*). Los valores predichos se añadieron como una variable aparte en la base de datos (*fitted*), a los que se restaron los valores reales de F0, con lo que se obtuvo el conjunto de valores residuales (*residuals*). Entre esos valores residuales, se tomaron aquellos puntos que se encontraban en el 20% más grave (*Q20*) y en el 20% más agudo (*Q80*), y, según este criterio estrictamente fonético, esos registros de F0 fueron etiquetados, respectivamente, como tonos graves (*Low*) y agudos (*High*).

La Figura 2 muestra el procedimiento de detección de tonos graves y agudos para una de las frases de nuestro corpus. Todos los puntos del gráfico representan los valores reales de F0 en semitonos, dispuestos en la línea de tiempo relativo al inicio y final de la frase. A partir de ellos se calculó la regresión *loess*, representada con una línea de color negro. A partir de la distancia entre cada punto y el valor que tomaba la regresión en ese mismo tiempo relativo, se identificaron aquellos puntos que ocupaban el 20% más grave (marcados en color rojo) y los que ocupaban el 20% más agudo (en color azul). Dos líneas adicionales en esos dos mismos colores (rojo y azul) ilustran sendas regresiones *loess* aplicadas a los registros identificados como graves y agudos. En el caso representado en la Figura 2, por ejemplo, se muestra cómo la declinación afecta lo que en cada momento puede ser identificado como un punto grave o agudo, mientras que, a su vez, la distancia entre puntos graves y agudos no parece verse alterada a medida que transcurre la frase; es decir, en la frase de ejemplo se aprecia una variación importante del F0 medio a lo largo de la frase pero un campo tonal relativamente estable a lo largo de la misma.

La Figura 3 muestra un ejemplo de la base de datos en ese punto: lo que llamamos tabla *pitch*. En la tabla se pueden leer las columnas relativas al fichero de voz (*File*), el hablante (*Speaker*), la frase en particular (*Item*), el grupo al que pertenece el hablante (*Group*), el tiempo absoluto de cada registro (*Time*), los valores de F0 en semitonos (*st*), el tiempo relativo de cada registro con respecto al primer y último punto de F0 registrado para cada frase (*Time_rel*), los valores de F0 en hercios (*Hz*), los valores de F0 en semitonos escalados con respecto a la media de F0 de cada

hablante (*st_scaled*), los valores de F0 predichos por la regresión *loess* (*fitted*), los valores residuales de esa misma regresión (*residuals*), el valor residual que ejercía de frontera del 20% más grave de los valores residuales para cada frase (*Q20*), el que ejercía de 20% más agudo (*Q80*), y los valores de *st_scaled* correspondientes al 20% más grave (*Low*) y al 20% más agudo (*High*).

Posteriormente, la base de datos se reestructuró para identificar los valores más graves y los más agudos en una única columna (*Tone*), como se muestra en la Figura 4: lo que llamamos tabla *pitch_tones*.

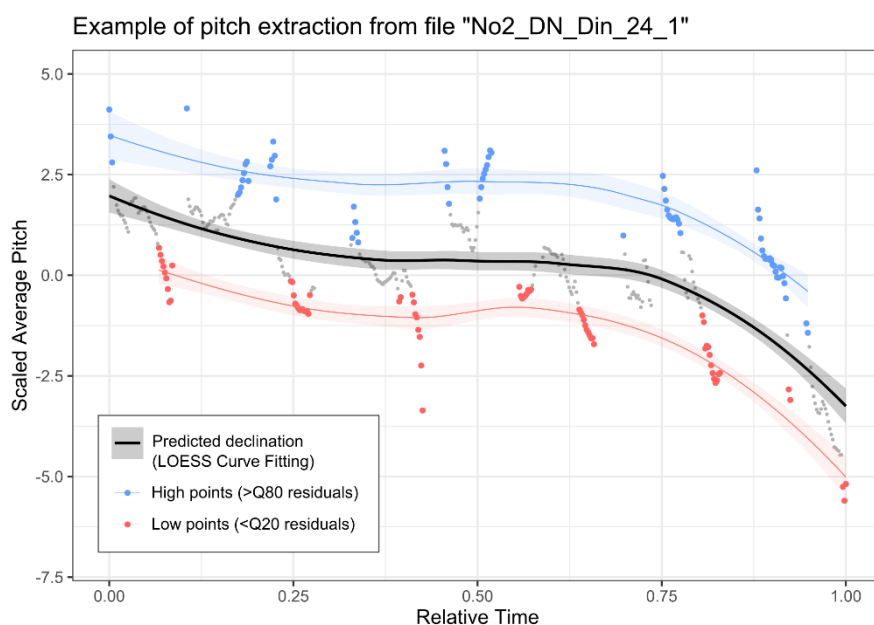


Figura 2. Extracto de la base de datos después de la extracción de registros más graves y agudos con respecto a una regresión *loess* aplicada a cada frase.

File	Speaker	Item	Group	Time	st	Time_rel	Hz	st_scaled	fitted	residuals	Q20	Q80	Low	High
279	AI1_DN_02_1	AI1	DN_02_1	Down	6.671791	10.022853	0.842443730	178.4151	0.271548383	0.77621979	-0.504671403	-0.6389273	0.7206359	NA
280	AI1_DN_02_1	AI1	DN_02_1	Down	6.681791	9.676188	0.844051447	176.9100	0.124882733	0.76246703	-0.637584293	-0.6389273	0.7206359	NA
281	AI1_DN_02_1	AI1	DN_02_1	Down	6.691791	9.604595	0.845659164	174.1563	-0.146709924	0.74827777	-0.894987695	-0.6389273	0.7206359	-0.14670992
282	AI1_DN_02_1	AI1	DN_02_1	Down	6.701791	9.352920	0.847266881	171.6429	-0.398385400	0.73362869	-1.132014090	-0.6389273	0.7206359	-0.39838540
283	AI1_DN_02_1	AI1	DN_02_1	Down	6.711791	9.186751	0.848874598	170.0033	-0.564553976	0.71849646	-1.283050437	-0.6389273	0.7206359	-0.56455398
284	AI1_DN_02_1	AI1	DN_02_1	Down	6.721791	9.135315	0.850482315	169.4989	-0.615989533	0.70285776	-1.318847289	-0.6389273	0.7206359	-0.61598953
285	AI1_DN_02_1	AI1	DN_02_1	Down	6.731791	8.994972	0.852090032	168.1304	-0.756332756	0.68666925	-1.443022005	-0.6389273	0.7206359	-0.75633276
286	AI1_DN_02_1	AI1	DN_02_1	Down	6.741791	8.891557	0.853697749	165.2095	-1.059748323	0.67010617	-1.729854495	-0.6389273	0.7206359	-1.05974832
287	AI1_DN_02_1	AI1	DN_02_1	Down	6.751791	8.092151	0.855305466	159.5873	-1.659153718	0.65323906	-2.312392774	-0.6389273	0.7206359	-1.65915372
288	AI1_DN_02_1	AI1	DN_02_1	Down	6.761791	7.026949	0.856913183	150.0641	-2.724355503	0.63608752	-3.360443025	-0.6389273	0.7206359	-2.72435550
289	AI1_DN_02_1	AI1	DN_02_1	Down	7.101791	10.678212	0.911579563	185.2985	0.926907421	-0.11898584	1.045893263	-0.6389273	0.7206359	0.9269074
290	AI1_DN_02_1	AI1	DN_02_1	Down	7.111791	10.503928	0.913183280	183.4424	0.752623402	-0.14632960	0.899953001	-0.6389273	0.7206359	0.7526234
291	AI1_DN_02_1	AI1	DN_02_1	Down	7.121791	10.506216	0.914790997	183.4667	0.754911462	-0.17397135	0.928882813	-0.6389273	0.7206359	0.7549115
292	AI1_DN_02_1	AI1	DN_02_1	Down	7.131791	10.525194	0.916398714	183.6679	0.773888831	-0.20191148	0.975000308	-0.6389273	0.7206359	0.7738888
293	AI1_DN_02_1	AI1	DN_02_1	Down	7.141791	10.543061	0.918006431	183.8575	0.791755910	-0.23015035	1.021906262	-0.6389273	0.7206359	0.7917559
294	AI1_DN_02_1	AI1	DN_02_1	Down	7.151791	10.628803	0.919614148	184.7704	0.877497713	-0.25868836	1.136186068	-0.6389273	0.7206359	0.8774977

Figura 3. Extracto de la base de datos después de la extracción de registros más graves y agudos con respecto a una regresión loess aplicada a cada frase.

File	Speaker	Item	Group	Time	st	Time_rel	Hz	fitted	residuals	Tone	st_scaled	
1	AI1_DN_01_1	AI1	DN_01_1	Down	2.132234	10.711342	0.000000000	185.6534	1.970098944	-1.0100617	Low	0.96003728
2	AI1_DN_01_1	AI1	DN_01_1	Down	2.262234	11.078604	0.036111111	189.6339	2.093320644	-0.7660218	Low	1.32729884
3	AI1_DN_01_1	AI1	DN_01_1	Down	2.272234	10.709906	0.038888889	185.6300	2.101536799	-1.1429358	Low	0.95860097
4	AI1_DN_01_1	AI1	DN_01_1	Down	2.282234	10.405003	0.041666667	182.3900	2.109578297	-1.4558003	Low	0.65377796
5	AI1_DN_01_1	AI1	DN_01_1	Down	3.292234	12.669013	0.322222222	207.8800	2.005681269	0.9120265	High	2.91770778
6	AI1_DN_01_1	AI1	DN_01_1	Down	3.302234	12.791282	0.325000000	209.3534	1.996453456	1.0435237	High	3.03997716
7	AI1_DN_01_1	AI1	DN_01_1	Down	3.312234	12.821445	0.327777778	209.7184	1.986644437	1.0834060	High	3.07014042
8	AI1_DN_01_1	AI1	DN_01_1	Down	3.322234	12.779447	0.330555556	209.2103	1.976098138	1.0520440	High	3.02814211
9	AI1_DN_01_1	AI1	DN_01_1	Down	3.332234	12.728433	0.333333333	208.5947	1.965037180	1.0120905	High	2.97712764
10	AI1_DN_01_1	AI1	DN_01_1	Down	3.342234	12.725730	0.336111111	208.5622	1.953739751	1.0206854	High	2.97442519
11	AI1_DN_01_1	AI1	DN_01_1	Down	3.352234	12.877550	0.338888889	210.3992	1.942133127	1.1841117	High	3.12624405
12	AI1_DN_01_1	AI1	DN_01_1	Down	3.362234	12.987710	0.341666667	211.7423	1.930144586	1.3062609	High	3.23640553
13	AI1_DN_01_1	AI1	DN_01_1	Down	3.372234	12.773470	0.344444444	209.1381	1.917701405	1.1044632	High	3.02216456
14	AI1_DN_01_1	AI1	DN_01_1	Down	3.422234	10.903021	0.358333333	187.7204	1.845992212	-0.6942758	Low	1.15171645
15	AI1_DN_01_1	AI1	DN_01_1	Down	3.432234	9.740815	0.361111111	175.5321	1.828994205	-1.8394840	Low	-0.01048980
16	AI1_DN_01_1	AI1	DN_01_1	Down	3.442234	8.518220	0.363888889	163.5636	1.810989177	-3.0440737	Low	-1.23308451

Figura 4. Tabla pitch_tones. Base de datos reestructurada que contiene únicamente aquellos valores de F0 asociados a los valores residuales más graves y más agudos con respecto a una regresión loess que representa la declinación de cada frase.

Los datos consignados en la tabla *pitch* se utilizaron para analizar la declinación general de las frases (sección 3), mientras que los datos de la tabla *pitch_tones* sirvieron para analizar los aspectos relacionados con el campo tonal (sección 4).

2.5. Procesado estadístico

Para obtener medidas descriptivas y aplicar pruebas inferenciales que permitan evaluar las relaciones entre las variables se emplearon modelos lineales mixtos utilizando el paquete estadístico *glmmTMB* en el entorno de R.

3. ANÁLISIS DE LA DECLINACIÓN TONAL

Los datos contenidos en la tabla *pitch* (v. Figura 3) permiten comparar las frases de los grupos Control y Down con el fin de observar, por un lado, si en cada grupo la pendiente de declinación tonal es significativamente inclinada, y, por otro lado, si las pendientes de ambos grupos difieren entre ellas en grado significativo, de modo que podamos hablar de dos poblaciones diferenciadas con respecto al uso de la declinación tonal. Para ello, se calculó un modelo lineal mixto utilizando el paquete *glmmTMB* en el entorno estadístico R. Los valores de significación se obtuvieron mediante el uso del paquete *afex*.

La Figura 5 ilustra la diferencia en el grado de pendiente entre ambos grupos. En cada uno de los paneles, los diferentes puntos representan todos los valores individuales de F0 de todas las frases, y las líneas permiten visualizar el resultado de aplicar una regresión *loess* a la totalidad de esos puntos. Si bien es cierto que los dos grupos muestran declive tonal a lo largo del enunciado, este es más pronunciado en las muestras de habla del grupo Control. Nuestro análisis de la declinación tonal tiene en cuenta la totalidad de las frases producidas por cada grupo; sin embargo, los modelos estadísticos que aplicamos controlan la variación asociada a cada hablante y a cada producción verbal, con lo que las pendientes tonales observadas en cada enunciado se tienen en cuenta individualmente para que la descripción estadística de la declinación no quede sesgada por unas pocas producciones y, en consecuencia, se centre en las propiedades comunes.

Para el análisis estadístico de la pendiente de declinación, se estableció el valor de F0 en semitonos escalado por hablante (*st_scaled*) como variable dependiente. Para una interpretación más adecuada de los datos estadísticos, el tiempo relativo de la frase (*Time_rel*, inicialmente en una escala del 0 al 1) se dispuso en una escala del -1 al +1 (*Time_rel_cen*), y esta nueva variable se estableció como factor fijo. También se añadió como factor fijo el Grupo, tomando como valor de referencia el

grupo Down, así como la interacción entre el grupo y el tiempo relativo centrado. Finalmente, se modelaron una interceptación aleatoria para el hablante (*Speaker*) y otra para la frase en particular (*Item*). La tabla 1 muestra los coeficientes fijados del modelo.

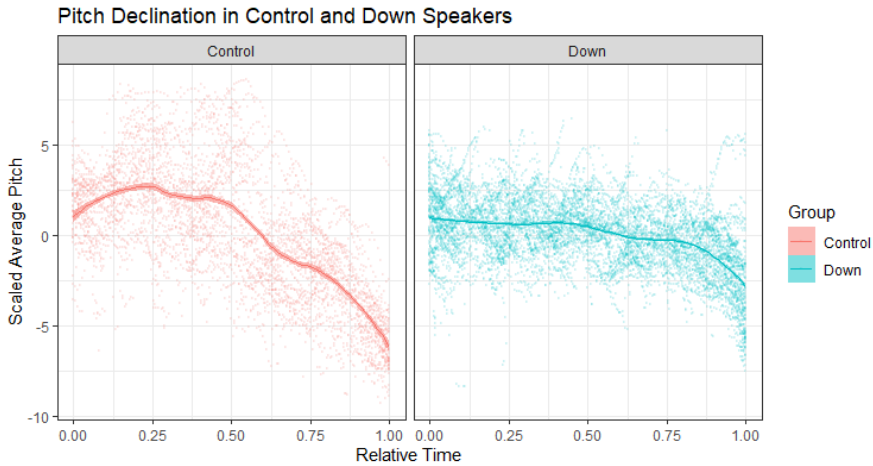


Figura 5. Curva de declinación en el grupo de hablantes de referencia (Control) y en el grupo de hablantes con SD (Down). Las abscisas muestran el tiempo relativo al comienzo y final de cada frase, y las ordenadas los valores de F0 (en semitonos) de todas las frases de todos los hablantes del experimento centrados de acuerdo con la media de cada hablante. La línea muestra en cada grupo la curva estimada de declinación.

	β	SE	z	p
Intercept (GroupDown)	0,262	0,164	1,60	0,110
GroupControl	-0,006	0,283	-0,02	0,982
Time	-1,262	0,034	-36,95	<2E-16
GroupControl: Time	-2,588	0,061	-42,38	<2E-16

Tabla 1. Coeficientes fijados del modelo de análisis de la pendiente de declinación.

En resumen, el efecto temporal sobre el F0 fue hallado significativo para los datos del grupo Down ($\beta = -1,262$), y para el grupo Control se halló incluso

significativamente menor con respecto a los primeros ($\beta = -2,588$ menos; es decir, una pendiente total de $\beta = -3,850$ para el grupo Control).

La lectura de los resultados del modelo como análisis de la varianza (*omnibus test*) indica que los dos grupos no difieren significativamente por lo que respecta a los datos que aportan si no se tiene en cuenta su distribución temporal ($\chi^2(1) = 0,0012$, $p = 0,972$). Es decir, si no tenemos en cuenta la disposición temporal de los registros centrados de F0, no se puede establecer que haya un grupo de hablantes que presente unos registros más agudos o más graves con respecto al otro. Además, tomando los registros de todos los hablantes a la vez, se recoge un efecto principal de la distribución temporal del F0 ($\chi^2(1) = 5349,649$, $p < 0,001$). Es decir, si no tenemos en cuenta que cada uno de los hablantes pertenece a un grupo u otro y analizamos todas las producciones conjuntamente, se describe una variación significativa del F0 a lo largo del tiempo. Finalmente, la interacción entre el grupo y el tiempo relativo también resulta significativa ($\chi^2(1) = 1795,902$, $p < 0,001$). Todo ello indica que los dos grupos presentan declinación, pero que, como ya se observaba en la Figura 5, la declinación del grupo Control es (significativamente) más pronunciada que la del grupo Down.

4. ANÁLISIS DEL CAMPO TONAL

A partir de la tabla *pitch_tones* (v. Figura 4), se quiso analizar si las frases de los grupos Control y Down diferían por lo que respecta al campo tonal a lo largo de la frase. En las Figuras 6 y 7 se incluyen aquellos puntos que constituyen el 20% más grave (*Low*) y el 20% más agudo (*High*) con independencia de la declinación tonal (v. más arriba la explicación de la Figura 2, en la sección 2.4). Para cada tipo de tono y grupo de hablantes, además, se representa el resultado de aplicar una regresión *loess* a cada grupo de registros de F0. La Figura 6 representa los registros en su posición original con respecto a la declinación tonal de las frases; la Figura 7, por su parte, registra la diferencia entre tonos graves y agudos a lo largo del enunciado una vez eliminado el efecto de la declinación tonal.

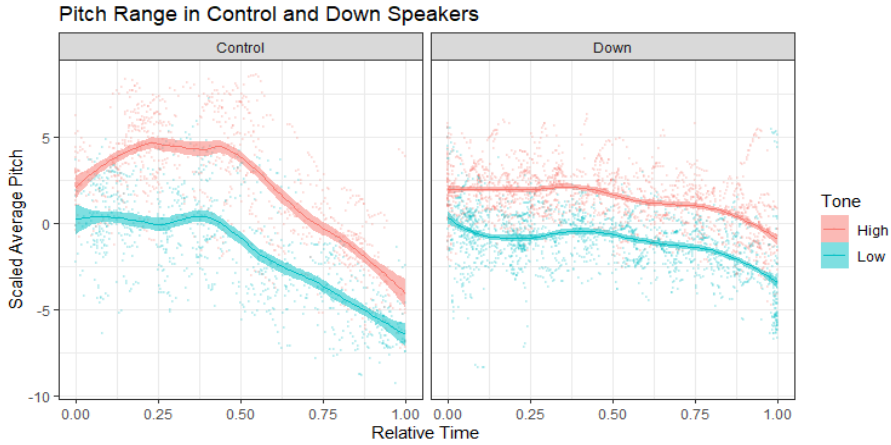


Figura 6. Diferencia entre tonos graves (*Low*) y agudos (*High*) a lo largo del enunciado en el grupo de hablantes de referencia (*Control*) y el grupo de hablantes con *SD* (*Down*).

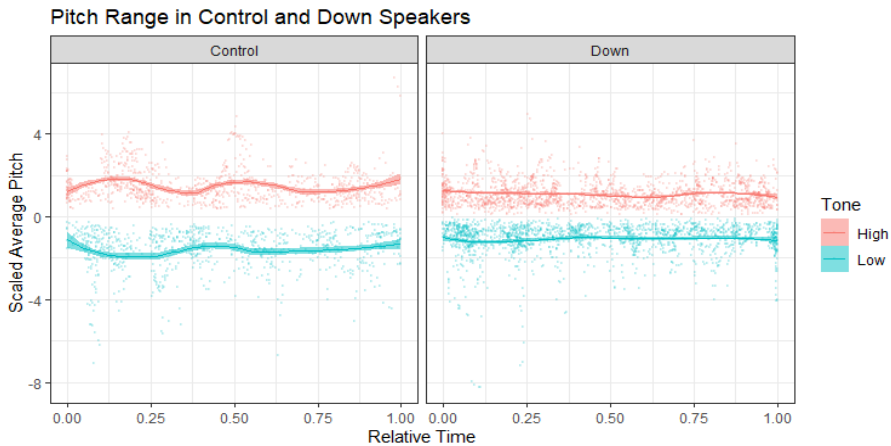


Figura 7. Diferencia entre tonos graves (*Low*) y agudos (*High*) a lo largo del enunciado en el grupo de hablantes de referencia (*Control*) y el grupo de hablantes con *SD* (*Down*), una vez eliminado el efecto de la declinación tonal.

Para determinar si los dos grupos diferían con respecto al campo tonal utilizado, se realizó otro modelo lineal mixto. Se estableció como variable dependiente la columna que contenía los puntos de F0 más graves y los más agudos (*st_scaled*). En este caso, se modelaron como efectos fijos tres variables (dos de ellas categóricas y una de ellas continua), así como todas sus posibles interacciones: el grupo al que pertenecían los hablantes (*Group*), la variable que contenía la clasificación entre tonos graves y agudos de los diferentes registros de F0 (*Tone*) y el tiempo relativo al primer y último punto de F0 de la frase —centrado, como en el modelo anterior— (*Time_rel_cen*). Se añadieron al modelo, también, una interceptación aleatoria tanto para el hablante (*Speaker*) y otra para la frase en particular (*Item*). La tabla 2 muestra los coeficientes fijados de este modelo.

	β	SE	z	p
Intercept (GroupDown, ToneLow)	-0,892	0,205	-4,36	1E-05
GroupControl	-0,820	0,329	-2,49	0,013
ToneHigh	2,361	0,273	8,66	<2E-16
Time	-0,999	0,070	-14,37	<2E-16
GroupControl: ToneHigh	1,408	0,390	3,61	3E-04
GroupControl: Time	-2,473	0,124	-19,99	<2E-16
ToneHigh: Time	-0,028	0,099	-0,28	0,778
GroupControl: ToneHigh: Time	-0,076	0,176	-0,43	0,664

Tabla 2. Coeficientes fijados del modelo de análisis del campo tonal.

Podemos interpretar los resultados de la manera siguiente. Por lo que respecta a los efectos que incluyen *Time*, obtenemos que para los tonos graves del grupo Down se produce una declinación significativamente inclinada ($\beta = -0,999$), que es incluso significativamente menor para los tonos graves del grupo Control ($\beta = -3,472$). Es decir, existe declinación en la producción de los tonos graves de los dos grupos, pero es más prominente en las producciones del grupo Control. Los tonos agudos definen también una curva significativamente descendente con respecto a la declinación, pero no significativamente diferente de la descrita para los tonos graves: tanto para el grupo Down ($\beta = -1,027$) como para el grupo Control ($\beta = -3,576$). Esto implica, en resumen, que tanto tonos graves como agudos por separado definen un patrón de declinación similar y que este se puede describir como más acusado en el caso del grupo Control comparado con el grupo Down.

La lectura de los resultados del modelo como análisis de la varianza indica un efecto principal del tono, en tanto que los tonos agudos contenían valores más elevados que

los graves cuando se tienen en cuenta los datos de los dos grupos conjuntamente ($\chi^2(1) = 245,501, p < 0,001$), y que los dos grupos no difieren significativamente por lo que respecta a los datos que aportan si no se tiene en cuenta su distribución temporal ($\chi^2(1) = 0,0001, p = 0,993$).

La interacción entre tono y grupo, a su vez, aporta un resultado significativo ($\chi^2(1) = 13,138, p < 0,001$), que puede concretarse según los resultados de los contrastes por pares siguientes. Por una parte, la diferencia entre tonos graves y agudos es significativa para los dos grupos, pero es mayor para el grupo Control (*Cohen's d* = 2,109, $p < 0,001$) que para el grupo Down ($d = 1,320, p < 0,001$). Por otra parte, los registros agudos del grupo Control son más elevados que los del grupo Down ($d = 0,432, p = 0,030$), mientras que por lo que se refiere a los registros graves los dos grupos no difieren significativamente pese que, también en estos registros, es el grupo Control el que muestra una tendencia a presentar valores más reducidos ($d = 0,357, p = 0,052$), lo cual es otra forma de mostrar que el campo tonal empleado por el grupo Down es menor que el empleado por el grupo Control.

En resumen, los resultados de la sección 3 indicaban que ambos grupos presentan declinación tonal, pero que esta es más pronunciada en las producciones del grupo Control (v. Figura 5). En la sección 4, además, se ha podido constatar que el campo tonal empleado por el grupo Control es mayor al empleado por el grupo Down (v. Figura 7).

5. DISCUSIÓN DE LOS RESULTADOS

Los resultados permiten alcanzar unas conclusiones que dan respuesta a las preguntas de investigación planteadas en la introducción:

- La declinación es un efecto fonético de declive tonal a lo largo de la oración que se manifiesta en el habla de personas con SD del mismo modo que en el habla de personas sin patologías, de manera que podemos hablar de patrones comunes.
 - Ahora bien, la manifestación fonética de la declinación de los hablantes con SD presenta algunas características que la diferencian de la prosodia típica y que la constituye como una variable prosódica que incide en la identificación de una voz como atípica. La declinación presenta un grado menor de cambio en la curva de F0 de personas con SD: tanto la pendiente como la distancia entre tonos altos y tonos bajos a lo largo del enunciado son menores. Dado que estudios previos relacionan la declinación tonal con la duración real de los enunciados (Swerts *et al.*, 1996, Kocharov *et al.*, 2014), para una interpretación adecuada de nuestras conclusiones conviene
-

tener en cuenta que se fundamentan en el análisis de la declinación con respecto al tiempo relativo de inicio y fin de cada frase. Sería interesante analizar en un futuro las diferencias asociadas con esta doble dimensión del tiempo en la producción de la declinación y si se manifiestan de forma diferente en poblaciones típicas y en poblaciones con trastornos del habla.

- También la variación en el campo tonal se considera un fenómeno de variación fonética, no fonológica. Pero hallamos nuevamente que se trata de un parámetro que incide en la identificación de una voz como atípica y que puede relacionarse con la percepción de monotonía.

Aunque no se dispone de estudios previos sobre la manifestación de la declinación tonal en el habla de las personas con SD, los resultados sobre la diferencia entre tonos altos y bajos están en línea con las descripciones de Albertini *et al.* (2010) y Rochet-Capellan y Dohen (2015), que señalan valores de F0 más altos para los hablantes adultos con SD, italianos y franceses respectivamente. Asimismo, el trabajo de Lee *et al.* (2009) observa para sus informantes británicos, además de valores medios más altos de F0, una reducción del campo tonal.

En las últimas décadas se ha planteado una discusión sobre si los problemas de habla en las personas con SD eran la consecuencia del retraso intelectual (*delay*) o si se trata de algo más complejo (*disorder*). Existen evidencias que permiten argumentar que es un trastorno de habla que tiene que ver con la especial configuración anatómica de su aparato fonador: hipotonía muscular, cavidad oral más pequeña, conductos más estrechos (Rupela *et al.*, 2016; Cleland *et al.*, 2010); además de sus problemas de audición en edad temprana: conductos auditivos más estrechos, tendencia a las infecciones (Laws y Hall, 2014; Borsel, 1996). El hecho de poder describir con parámetros comunes a los de una población típica categorías fonéticas como la declinación o el campo tonal permiten apoyar la hipótesis de un trastorno de habla. Investigaciones adicionales deberán ocuparse de las categorías fonológicas para avanzar un paso más allá en la descripción de la gramática entonativa de los hablantes con SD.

6. CONCLUSIONES

El análisis experimental de un conjunto de frases pronunciadas por hablantes con Síndrome de Down (SD) demuestra la existencia de declinación tonal en las producciones de dichos hablantes. Sin embargo, la comparación con un corpus homólogo de frases pronunciadas por hablantes sin discapacidad permite determinar que aunque hay un descenso gradual de la frecuencia fundamental a lo largo de los enunciados, el grado de cambio es menor en el colectivo de hablantes con SD,

comparado con los registros de personas con habla prototípica. Además, el campo tonal utilizado—es decir, la diferencia de F0 entre tonos graves y agudos— es también más reducido que el que se obtiene en el corpus de habla sin discapacidad. Esa disminución fonética en dos características globales de la frase podría explicar en parte el efecto de monotonía descrito para el habla de personas con SD.

AGRADECIMIENTOS: El corpus de hablantes con Síndrome de Down pudo compilarse gracias al proyecto Recercaixa 2013 “¡Juguemos a comunicar mejor! La mejora de la competencia prosódica como vía de integración educativa e inclusión social del alumnado con Necesidades Específicas de Apoyo Educativo (NEAE)” impulsado por ACUP-Obra Social La Caixa. Agradecemos a todos los miembros del equipo su dedicación, y especialmente a Anna Nogueras su empatía con los jóvenes con SD. Una primera versión de estos resultados fue presentada en el XLIX Simposio Internacional de la Sociedad Española de Lingüística (Tarragona): agradecemos a los asistentes su interés y comentarios.

7. REFERENCIAS

- ALBERTINI, G.; S. BONASSI, V. DALL'ARMI, I. GIACHETTI, S. GIAQUINTO and M. MIGNANO (2010): «Spectral analysis of the voice in Down syndrome», *Research in developmental disabilities*, 31(5), 995-1001.
- BECKMAN, M. E. y J. B. PIERREHUMBERT (1986): «Intonational structure in Japanese and English», *Phonology*, 3, pp. 255-309.
- BOERSMA, P. y D. WEENINK (2016): *Praat v.6.0.19*, www.praat.org [14/06/2016].
- BORSEL, J. V. (1996): «Articulation in Down's syndrome adolescents and adults», *European Journal of Disorders of Communication*, 31(4), pp. 415-444.
- CHAPMAN, R.S. y L. HESKETH (2000): «The behavioral phenotype of Down syndrome», *Mental Retardation and Developmental Disabilities Research Review*, 6, pp. 84-95.
- CLELAND, J.; S. WOOD, W. HARDCASTLE, J. WISHART y C. TIMMINS (2010): «Relationship between speech, oromotor, language and cognitive abilities in children with Down's syndrome», *International journal of language & communication disorders*, 45(1), pp. 83-95.
- DODD, B. y L. THOMPSON (2001): «Speech disorder in children with Down's syndrome», *Journal of Intellectual Disability Research*, 45(4), pp. 308-316.

-
- DUNN, L.; L. DUNN y D. ARRIBAS (2006): *Test de vocabulario en imágenes Peabody*, Madrid, TEA.
- EGGERS, K. y S. VAN EERDENBRUGH (2018): «Speech disfluencies in children with Down Syndrome», *Journal of Communication Disorders*, 71, pp. 72-84.
- KATZ, G. y E. LAZCANO-PONCE (2008): «Discapacidad intelectual: definición, factores etiológicos, clasificación, diagnóstico, tratamiento y prognosis», *Salud Pública de México*, 50, s132-s141.
- KENT, R. D. y H. K. VORPERIAN (2013): «Speech impairment in Down syndrome: A review», *Journal of Speech, Language, and Hearing Research*, 56(1), pp. 178-210.
- KOCHAROV, D., P. SKRELIN y N. VOLSKAYA (2014): «F0 declination patterns in russian», en *International Conference on Speech and Computer*, Novi Sad, Springer, pp. 217-226.
- KUMIN, L. (2002): «Inteligibilidad del habla en las personas con Síndrome de Down: Un marco para señalar factores específicos útiles en la evaluación y tratamiento», *Revista Síndrome de Down*, 19(1), pp. 14-23.
- LADD, D. R. (2008): *Intonational phonology*, Nueva York, Cambridge University Press.
- LAWS, G. y A. HALL (2014): «Early hearing loss and language abilities in children with Down syndrome», *International journal of language & communication disorders*, 49(3), pp. 333-342.
- LEE, M. T.; J. THORPE y J. VERHOEVEN (2009): «Intonation and phonation in young adults with Down syndrome», *Journal of Voice*, 23(1), pp. 82-87.
- MARTIN, G. E.; J. KLUSEK, B. ESTIGARRIBIA y J. E. ROBERTS (2009): «Language characteristics of individuals with Down syndrome», *Topics in Language Disorders*, 29(2), pp. 112-132.
- PEPPÉ, S. J. (2009): «Why is prosody in speech-language pathology so difficult?», *International Journal of Speech-Language Pathology*, 11(4), pp. 258-271.
-

- PIERREHUMBERT, J. y M. LIBERMAN (1982): «Modeling the fundamental frequency of the voice», *Psycritiques*, 27(9), pp. 690-692.
- PIERREHUMBERT, J. B. (1980): *The Phonetics and Phonology of English Intonation*, tesis doctoral, Massachusetts Institute of Technology.
- RAVEN, J.; J.C. RAVEN Y J. COURT (1993): *Test de matrices progresivas: Manual*, Buenos Aires, Paidós.
- ROBERTS, J.E.; J. PRICE y C. MALKIN (2007): «Language and communication development in Down syndrome», *Mental retardation and developmental disabilities research reviews*, 13(1), pp. 26-35.
- ROCHET-CAPELLAN, A. y M. DOHEN (2015): «Acoustic characterisation of vowel production by young adults with Down syndrome», *Proceedings of the XVIII International Congress of Phonetic Sciences (ICPhS)*, Glasgow, Paper 975.
- RODRÍGUEZ, W. R.; O. SAZ y E. LLEIDA (2012): «A prelingual tool for the education of altered voices», *Speech Communication*, 54(5), pp. 583-600.
- RUIZ, E. (2001): «Evaluación de la capacidad intelectual en personas con síndrome de Down», *Revista Síndrome Down*, 21, pp. 134-149.
- RUPELA, V.; S. L. VELLEMAN y M.V. ANDRIANOPOULOS, (2016): «Motor speech skills in children with Down syndrome: A descriptive study», *International journal of speech-language pathology*, 18(5), pp. 483-492.
- STOJANOVIK, V. (2011): «Prosodic deficits in children with Down syndrome», *Journal of Neurolinguistics*, 24(2), pp. 145-155.
- SWERTS, M., E. STRANGERT y M. HELDNER (1996): «F0 declination in read-aloud and spontaneous speech», en T. BUNNELL y W. IDSARDI (eds.): *Proceeding of Fourth International Conference on Spoken Language Processing ICSLP'96*, Filadelfia, IEEE, pp. 1501-1504.
- UMEDA, N. (1982): «“F0 declination” is situation dependent», *Journal of Phonetics*, 10(3), pp. 279-290.

WILLCOX, A. (1988): «An investigation into non-fluency in Down's syndrome», *International Journal of Language & Communication Disorders*, 23(2), pp. 153-170.

ZAMPINI, L.; M. FASOLO, M. SPINELLI, P. ZANCHI, C. SUTTORA y N. SALERNI (2016): «Prosodic skills in children with Down syndrome and in typically developing children», *International Journal of Language & Communication Disorders*, 51(1), pp. 74-83.