



## Parálisis supranuclear progresiva: estudio longitudinal a partir del análisis acústico del habla

Eulàlia Pell<sup>1</sup>

Recibido: 20 de marzo de 2018 / Revisado: 30 de abril de 2018 / Aceptado 13 de julio de 2018

**Resumen.** El objetivo de este estudio es medir de forma objetiva y cuantitativa los distintos parámetros del habla en un caso de Parálisis Supranuclear Progresiva durante un periodo de cuatro años. Se realizó un análisis acústico de la vocal sostenida /a/, del automatismo “1,2,3,4” y de la repetición rápida de las sílabas /pa/, /ta/, /ka/ y /pataka/. Se consideraron diversos parámetros: la frecuencia fundamental (F0), el tiempo máximo fonatorio (TMF), la intensidad, los formantes, el shimmer, el jitter, el ruido en relación con los armónicos (NHR), la duración y el número de sílabas, el “Voice Onset Time” (VOT), la tasa de habla, la tasa de silencios y los aspectos prosódicos del habla. Las alteraciones en la voz y en el habla evolucionan hacia la imprecisión articulatoria, una pobre calidad vocal, reducido TMF, reducida tasa del habla, disminución de la variabilidad de la entonación y pausas prolongadas en las diadococinesias. El análisis reveló la coincidencia en el tiempo entre el deterioro progresivo del habla en general y el elevado porcentaje de cierres incompletos en el VOT, con la disfagia

**Palabras clave:** Parálisis Supranuclear Progresiva; disartria; disfagia; análisis acústico; diadococinesias orales

## [en] Progressive supranuclear palsy: longitudinal study by acoustical analysis of speech

**Abstract.** The aim of this study is to measure objectively and quantitatively different speech parameters in PSP for four years. The present investigation performed an acoustic analysis of sustained vowel phonation /a/, maximum phonatory time, numbers sentence (1,2,3,4) and rapid syllable repetition (/pa/, /ta/, /ka/ and /pataka/). Various parameters were considered: F0, intensity, formants, shimmer, jitter, harmonic-to-noise ratio (NHR), the mean number of syllables, the median syllable duration, voice onset time (VOT), speech rate, silence rate and prosodic elements. Alterations in voice and speech evolve towards articulatory imprecision, poor vocal quality, reduced TMF, reduced speech rate, decreased intonation variability and prolonged pauses in diadochokinesia. This analysis revealed a coincidence in time between progressive global speech impairment and highly percentage of incomplete closures in VOT, with progressive dysphagia.

**Key words:** Progressive Supranuclear Palsy; Dysarthria; Dysphagia; Acoustical analysis; oral diadochokinesia

**Cómo citar:** Pell, E. (2018). Parálisis supranuclear progresiva: estudio longitudinal a partir del análisis acústico del habla. *Revista de Investigación en Logopedia*, 8(2), 115-128.

<sup>1</sup> Hospital Sociosanitari Francolí (Tarragona). España.  
epell.gipss@gencat.cat

## Introducción

La Parálisis Supranuclear Progresiva (PSP), también denominada Síndrome de Steele Richardson Olszewski, se trata de un parkinsonismo primario degenerativo de etiología desconocida que evoluciona de forma progresiva. Las principales manifestaciones clínicas (Colosimo, Bak, Bologna et al. 2014) son: síndrome parkinsoniano (rigidez extrapiramidal de predominio axial -cuello y tronco superior-, bradicinesia con caídas frecuentes hacia atrás, rigidez y desequilibrios), hipomimia facial, oftalmoplejía supranuclear (parálisis de la mirada conjugada en plano vertical) y síndrome pseudobulbar (disfagia, disartria y labilidad emocional). El debut de los síntomas es entre los 55-70 años y de predominio masculino (Rehman, 2000). Dos años después del inicio de la enfermedad el 41% de los pacientes presentan disartria; la disfagia está presente a partir de los 42-48 meses (Colosimo et al. 2014). La lectura resulta alterada por el empeoramiento de la función oculomotora. Según el estudio de Kluin (1993), la comunicación se ve comprometida por la palilalia, el tartamudeo adquirido, la disminución de la habilidad para la interacción y la ecolalia. El tratamiento clínico es sintomático y su diagnóstico diferencial es con la enfermedad de Parkinson, la atrofia sistémica múltiple, la demencia o la degeneración corticobasal (Kim & McCann, 2015). La supervivencia después del inicio de los síntomas se sitúa hacia los 7 años (Golbe, 2014). La causa más frecuente de muerte en la PSP es la neumonía secundaria a la disfagia (Litvan, Sastry & Sonies, 1997).

Fisiológicamente, la rigidez asociada a la PSP conlleva una reducción en el recorrido muscular y en la velocidad de los movimientos de la musculatura del velo del paladar, de la lengua, de los labios y de la laringe (Golbe, 2014) que progresivamente compromete la inteligibilidad del habla hasta llegar a la anartria en las últimas fases de la enfermedad (Gómez-Haro, Espert, Gadea & Navarro, 1999). La disartria presente es fundamentalmente de tipo mixta con una combinación de componentes espásticos, hipocinéticos y atáxicos. En los casos más evolucionados, la disartria es tan severa que se caracteriza por una emisión vocal en forma de gruñido ininteligible (Colosimo et al. 2014).

Los aspectos alterados en el habla (Golbe, 2014) se centran en: insuficiencia prosódica (habla monotonal y monointensa, acento reducido, frases cortas, tasa del habla variable, imprecisión consonántica y descargas rápidas de habla); incremento de la actividad refleja (hiperreflexia), reducciones del recorrido muscular y de la velocidad del movimiento de la musculatura velopalatina, lingual, labial, mandibular y costodiafragmática que conlleva distorsión vocálica e hipernasalidad. Acústicamente su habla se caracteriza por la reducción de la velocidad articulatoria y un incremento de la tasa de pausas, asociados principalmente a los componentes espásticos de la disartria que predominan por encima de los hipocinéticos (Skodda, Visser & Schlegel, 2011). Las disartrias de tipo espástico, atáxico o mixto muestran correlaciones significativas con los trastornos de la deglución en general (Nishio & Niimi, 2004).

En cuanto a la deglución, las dificultades en la masticación y la disfunción oral y faríngea producen disfagia pero no suele presentarse como síntoma inicial (Jackson-Menaldi, 2002). En los estudios con FEES en la PSP (Litvan, 1997) se evidencia un retraso en la activación del reflejo deglutorio, acumulación de residuos principalmente en vallécula, pobre control del bolo (escape oral), penetraciones laríngeas, tiempo excesivo invertido en las ingestas y aumento de las aspiraciones conforme avanza la afectación cognitiva (demencia subcortical, déficit cognitivo frontal, di-

ficultades en la planificación, perseveración, apraxia). El tratamiento de la disfagia se centra en las maniobras posturales principalmente la de la flexión cervical y en el ajuste dietético hacia texturas homogéneas. En las fases más avanzadas, se debe valorar otras formas alternativas a la nutrición vía oral (Bascañana, 1999).

Existen en la literatura algunos estudios en los que se compara acústicamente los aspectos del habla en la disartria producida por la PSP y la enfermedad de Parkinson como el de Metter & Hanson, (1986). En dicho trabajo comentan también la variabilidad en cuanto a la tasa del habla y el número de pausas entre dos individuos con edades (58 años) y diagnósticos (PSP) idénticos que comparten un mismo grado de afectación de la disartria. Estos mismos autores, añaden que la progresión de la PSP va asociada a cambios en la alteración del habla después de evaluar durante 4 años a un paciente diagnosticado inicialmente de Parkinson (con la progresión de los síntomas se confirmó la PSP). Kuin (1993) estudia los rasgos acústicos del habla de 44 pacientes afectados de PSP concluyendo que la disartria que presentan es una combinación de componentes espásticos, hipocinéticos y atáxicos. Kent (1979) concreta que en algunos individuos con disartria atáxica, la disminución de la velocidad del habla se acompaña de pausas entre sílabas que a menudo afectan a las diadococinesias orales. Otro estudio, (Kent, Kent, Duffy, Thomas, Weismer, & Stuntebeck, 2000), concluye que las afectaciones en el habla en la disartria atáxica difiere en función de la tarea a realizar, observándose que no había mucha variación en la repetición de palabras o frases pero sí en la producción de una vocal sostenida o en la repetición de sílabas.

El objetivo del presente estudio es concretar de forma cuantitativa la evolución de los aspectos del habla en un caso de parálisis supranuclear progresiva. Teniendo en cuenta los estudios previos, se prioriza la tarea de fonación sostenida y de las diadococinesias orales con los fonemas oclusivos sordos. Para obtener datos prosódicos y melódicos del habla, se propone la producción del automatismo numeral (contar del uno al cinco) que a su vez permite observar el fenómeno de la coarticulación con el diptongo producido en “cuatro” y cuantificar la duración de sus producciones vocálicas.

## **Método:**

Se realiza un estudio longitudinal de los datos cuantitativos obtenidos a partir de la grabación y análisis acústico del habla durante un periodo de cuatro años (con el consentimiento y autorización correspondiente), de un paciente varón de 67 años diagnosticado de PSP que inicia el tratamiento logopédico dos años después del diagnóstico, con una frecuencia semanal. Al inicio, ya presentaba disartria mixta por afectación principalmente a nivel articulario y disfagia orofaríngea a líquidos por la que requería espesante para todos los líquidos a viscosidad néctar. Los resultados de la valoración psicológica con el Mini Examen Cognitivo de Lobo (MEC) durante los tres primeros años revelaron una adecuada capacidad cognoscitiva. El castellano fue la lengua de las grabaciones al ser su lengua materna. El estudio finalizó a medianos del último año, cuando fue derivado a otro centro debido al empeoramiento de su situación clínica por la severa limitación en la comunicación, la sospecha patológica de deterioro cognitivo (según MEC) con presencia de bradipsiquia y reducción de la

velocidad de procesamiento de la información así como necesidad de vía alternativa de hidratación y nutrición.

Para los registros acústicos se utilizó un micrófono de cabeza con capuchón de goma colocado a una distancia aproximada de 10 cm de la boca del paciente. Se realizaron en la sala de tratamiento con un bajo ruido ambiente, siempre al inicio de la sesión, con una periodicidad semestral el primer año. En los años consecutivos se intentó llevar un registro más frecuente (bimensual) hasta en el último año en el que sólo se obtuvieron 2 registros más prolongados en el tiempo (4 meses de diferencia) debido a la falta de asistencia del paciente como consecuencia del aumento de caídas sufridas. En total se obtienen datos de 14 registros (2 el primer año, 5 el segundo y el tercer año y 2 para el último año).

Las grabaciones consistían en la producción de una /a/ de forma sostenida durante al menos 3 segundos, repitiendo varias veces y haciendo un promedio de las medidas descartando siempre la primera y la última producción. De este modo se obtuvo el valor del TMF que se define como el máximo periodo de tiempo en el que el paciente es capaz de mantener la fonación vocal confortable. Permite estudiar la eficiencia del cierre glótico y del sistema respiratorio en general y sus valores de normalidad se sitúan entorno los 15 s (Jackson-Menaldi, 2002). A partir de la fonación máxima de la vocal /a/ se obtiene también el valor de la frecuencia fundamental que está determinada por la vibración de las cuerdas vocales. Su valor varía en función de la edad y el sexo. La información que se extrajo con el programa PRAAT (Boersma, 2018) de la producción vocálica sostenida fue la siguiente: la frecuencia fundamental (F0), su media y desviación estándar; la intensidad media; los 5 formantes; la razón ruido/armónicos (NHR); el número de pulsos glóticos; los periodos sin voz; la presencia del formante nasal; la tasa del habla; las perturbaciones de la amplitud (shimmer) y de la frecuencia fundamental (jitter). Este último parámetro no debe superar el 1% para poder considerarse normal (Jackson-Menaldi, 2002). Las perturbaciones que afectan a la amplitud de la onda vocal (shimmer) se consideran anormales a partir del 1% (Jackson-Menaldi, 2002). En relación a la intensidad, el rango dinámico normal para los hombres es de 50dB con un mínimo y máximo de 35-95dB (Melle, 2007). El ruido entre armónicos está asociado con el cierre incompleto de las cuerdas vocales, con la variación de la frecuencia jitter y con la amplitud shimmer. Los valores de NHR en la población sana son inferiores a 12 dB. El estudio de las resonancias del conducto vocal que están relacionadas con la inteligibilidad del habla se realiza a partir de los formantes y sus relaciones acústico-articulatorias, principalmente de los dos primeros (F1 y F2). El F1 (250-1000 Hz) se refiere a la primera cavidad de resonancia entre las cuerdas vocales y la parte más elevada de la lengua; F2 (850-2500 Hz) a la segunda cavidad de resonancia entre la lengua y el paladar (Saxena, M., Behari, M., Kumaran, S.S., Goyal, V., Narang, V., 2014). La frecuencia de F1 es un valor inversamente proporcional a la altura de la lengua en la cavidad vocal y el valor de F2 es directamente proporcional a la anteriorización de la lengua (Golbe, 2014). Para la vocal /a/ los valores de referencia de Jackson-Menaldi (2002) son de 740 Hz para F1 y 1350 Hz para F2.

Se solicitó además, la tarea de contar del 1 al 5 aunque sólo se analizó los cuatro primeros elementos suprimiendo el último en cada registro para evitar la presencia de pausas al producir involuntariamente la conjunción “y” al finalizar el enunciado. Del análisis del enunciado numeral se extrajo la curva melódica, la de la intensidad, la duración del diptongo /ua/ en la palabra “cuatro”, la duración de cada palabra

(número), el porcentaje de voz para “1,2,3,4” así como el porcentaje de pausas (relacionado con la velocidad del habla).

Para completar el análisis del habla se solicitó la tarea de producción silábica rápida a través de unidades de tiempo (diadococinesias). Para su estudio se indicó la repetición silábica rápida y repetitiva de /pa/, /ta/, /ka/ y /pataka/; se descartó la primera y la última sílaba de cada grabación. Las repeticiones se centran en las oclusivas sordas /pa/, /ta/, /ka/ por su composición articulatoria y por las posibilidades de extracción del VOT. El paciente debe repetir en cada serie (una para cada oclusiva) todas las sílabas que pueda en una sola inspiración. Se calcula la tasa de habla valorando el número de producciones para cada sílaba en 1 segundo de duración intentando mantener el ritmo de la producción. Se cuantifican también: el total de producciones de la combinación consonante + vocal (siempre la /a/) y se calculan los datos estadísticos para cada sílaba; la duración de cada sílaba (con su respectiva media y desviación típica) y el VOT para cada una de las oclusivas. Se descarta el cálculo de aquellas combinaciones silábicas que no cumplen con el criterio acústico de presencia de barra de explosión. Entendemos por VOT el intervalo temporal entre la explosión (marcador acústico del cierre articulatorio producido por los sonidos oclusivos) y el inicio de la sonoridad de la vocal (el primer pulso glótico) que puede observarse y cuantificarse en un espectrograma (Özsancak, Auzou, Jan & Hannequin, 2001). El VOT permite explorar acústicamente la coordinación entre las conductas laríngeas y supralaríngeas (Jackson-Menaldi, 2002). Las dos características principales del VOT son los valores positivos para las consonantes sordas y negativos para las sonoras, y se incrementa a medida que la oclusión se desplaza hacia la parte posterior del tracto vocal (es mayor en /k/ que en /p/ o /t/). Se siguieron los siguientes criterios de exclusión en la medición del VOT (Özsancak et al. 2001); cuando la explosión no resultaba evidente; la vocal era débil y sin pulsos glóticos observables y si se observaba un continuo vocal durante toda la producción de las oclusivas. En el espectrograma de PRAAT se obtuvieron los tiempos de inicio de la sonoridad de cada oclusiva de forma manual. Se calculó el tiempo transcurrido en milisegundos entre el inicio de la explosión y el primer pulso glótico. Se calcularon el valor medio de las producciones de cada uno de los elementos de la serie realizados con una sola inspiración. Se desestimó su cálculo cuando no se apreciaba la barra de explosión por ausencia de ésta o por no observarse estrías regulares en el espectrograma. En el caso de observarse dos o más barras de explosión, se optó por tomar como medida de inicio de la explosión la última barra.

Se utilizó el procedimiento “Medias” de SPSS®, para comprobar si existían diferencias significativas o no significativas en los valores medios al inicio y al final del estudio de distintos parámetros (la duración de /a/, pa, ta, ka, pataka; la duración de la producción del numeral y sus distintos elementos; la duración de la vocal /a/ en las distintas diadococinesias y el valor del VOT para cada una de ellas). Se compararon los datos estadísticos de tres registros iniciales obtenidos en la primera grabación (del primer año) y en la última (correspondiente al cuarto año).

Se cuantificaron los datos con el programa estadístico SPSS® versión 15.0.

## Resultados

### Tiempo Máximo Fonatorio (TMF)

El valor obtenido para la vocal /a/ presenta variaciones entre los distintos registros pero se observa una clara tendencia a reducirse en la fase más avanzada, que coincide con la presencia de gruñido en la que la inteligibilidad del habla resulta gravemente alterada. La producción final inteligible como /a/ es de 2,7 segundos (s) y dista enormemente de la inicial que es de 8,9 s (Tabla 1).

En la Tabla 2 puede observarse los datos relativos al primer año (fase inicial de la disartria) y al último (disartria severa) de la duración de la vocal durante la producción silábica. Los resultados obtenidos en dicha tabla muestran que existen diferencias significativas entre las medias iniciales y finales en la duración de /a/ para la producción de la diadococinesia “ka” ( $p < 0,001$ ). No se observan diferencias significativas para el resto de variables ( $p > 0,05$ ).

Tabla 1. Duración de los distintos elementos

	Media / DS	Mínimo / Máximo	Inicio <sup>a</sup>	Final <sup>b</sup>	p <sup>c</sup>
/a/	7174,07 / 2401,598	2725,25 / 9823,00	8921,67	2725,25	0,360
/pa/	221,00 / 58,144	157,33 / 330,30	196,11	330,30	0,048
/ta/	295,24 / 62,294	178,50 / 386,70	369,35	386,70	0,449
/ka/	346,13 / 77,000	243,50 / 518,70	518,70	376,60	0,894
/pataka/	361,83 / 151,620	175,20 / 705,00	291,03	202,80	0,681
“1,2,3,4”	2314,46 / 475,909	1362,00 / 3148,30	2325,10	2007,90	0,480
“Uno”	549,88 / 108,488	403,50 / 772,10	587,70	674,40	0,076
“Dos”	538,24 / 1933,255	246,10 / 1020,30	574,90	387,60	0,863
“Tres”	454,89 / 97,985	228,90 / 643,40	475,90	451,10	0,289
“Cuatro”	430,90 / 74,832	309,00 / 531,00	531,00	374,40	0,548
Diptongo “ua” <sup>c</sup>	148,81 / 73,852	127,8 / 253,10	178,90	<sup>d</sup>	0,469

**Nota:** Valores obtenidos con los datos de los 14 registros totales. Unidad: milisegundos (ms). DS: desviación estándar.

<sup>a</sup> Media de la producción en la primera grabación. <sup>b</sup> Media de la producción en la grabación final.

<sup>c</sup> Diptongo extraído de la palabra “cuatro”. <sup>d</sup> No se obtuvo muestra inteligible ni identificable acústicamente. <sup>e</sup> Procedimiento Medias de SPSS® para los valores iniciales y finales de cada parámetro (significación estadística  $p < 0,05$ ).

Tabla 2. Duración de /a/ en las diadococinesias al inicio y al final del estudio

	Media (I / F)	D.S (I / F)	Mínimo / Máximo	p <sup>b</sup>
/a/ PA	167,65 / 305,05	30,675 / 13,647	138,20 / 246,70	0,082
/a/ TA	162,09 / 363,70	28,015 / - <sup>a</sup>	130,10 / 234,10	0,160
/a/ KA	308,25 / 359,54	1,626 / -	307,10 / 309,40	0,000

**Nota:** Valores en milisegundos (ms). (I): corresponde a los datos de la primera grabación del primer año. (F): Se refiere a los datos de la última grabación (4 años después de la inicial).

<sup>a</sup> No existen datos al disponer sólo del valor de una única producción válida en toda la serie registrada. <sup>b</sup> Procedimiento Medias de SPSS® para los valores iniciales y finales de cada parámetro (significación estadística  $p < 0,05$ ).

## Frecuencia Fundamental (F0)

La F0 no muestra cambios significativos y se mantiene en los límites normales con una media de 120 Hz y una desviación estándar (DS) de 9,16. El valor mínimo obtenido es de 103,93 Hz y el máximo de 135,35 Hz. En el primer registro obtiene una F0 de 122,04 Hz y en el último registro 117,34 Hz. Las roturas o cambios repentinos (Keller, Vigneux & Laframboise, 1991) sólo se observaron de forma puntual en dos registros a medianos del segundo y del tercer año. En el jitter (que no tiene en cuenta los cambios de F0) se observa un incremento destacado de su valor a partir del tercer año. La media y la desviación del jitter son 0,88 % y 0,78 respectivamente; con un valor máximo en todos los registros de 2,78 % y un mínimo de 0,22 %. En el primer registro el jitter se sitúa en el 0,26 % y finaliza con 0,92 %. A medianos del tercer año de grabación, se incrementa temporalmente en dos ocasiones hasta el 2,45 % y 2,78 %.

En la Figura 1, se observa como F1 y F2 se mantienen estables durante los dos primeros años, y a partir del tercero, se observa una clara tendencia al incremento de F2 y a presentar cierta inestabilidad en F1. Los respectivos valores que adquiere F1 (media de 663,66 Hz, DS: 95,90) y sobretodo F2 (1323,69 Hz y DS: 334,49) son relativamente similares a los valores de referencia para la normalidad.

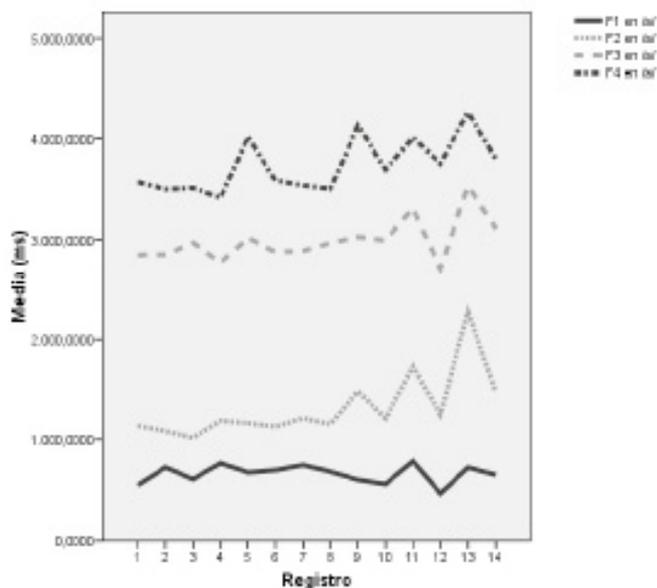


Figura 1. Evolución del valor de los formantes.

## Intensidad, shimmer y elementos de la prosodia

La media obtenida de todos los registros para la intensidad es de 70,34 dB (DS: 4,59). El valor inicial es de 76,08 dB (shimmer: 2,75 %) y en el último año de 76,91 dB (shimmer: 3,60 %). Entre el primer y segundo año disminuyen los valores hasta alcanzar un mínimo de 64 dB (shimmer: 3,48 %). A partir del tercer año, se observa

un aumento progresivo, y por tanto, una alteración del shimmer (valor máximo en el tercer año de 20,95 dB) al igual que del jitter ya comentado anteriormente. La media obtenida de todos los datos del estudio para el shimmer es de 5,94 % con DS de 5,20; el valor máximo registrado es de 20,95 % y el mínimo es 2,52 %.

El acento y la prosodia (en la producción de la cadena “1,2,3,4”) representan la monotonía con una curva de intensidad casi plana. La curva melódica evoluciona con los años en la misma dirección (Figura 2).

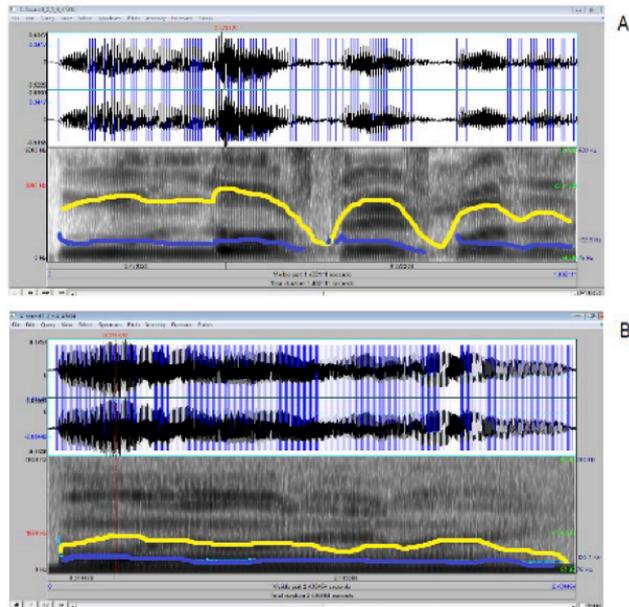


Figura 2. Curva melódica y curva de la intensidad en la producción del enunciado “uno, dos, tres, cuatro”.

A: primera grabación. B: grabación final del último año. La curva más clara corresponde a la intensidad; la más oscura al tono (melodía).

### Ruido en relación con los armónicos (HRN)

El HRN en todas las grabaciones se presenta severamente alterado con valores superiores al normal. Se obtiene una media (para todos los registros) de 17,80 dB, con una desviación estándar de 4,64; el valor máximo y mínimo registrados son 24,30 dB y 7,07 dB, respectivamente. En la primera grabación se observa un HRN de 24,30 dB y en la última 17,87 dB.

### Resonancia

A través del espectrograma se observa un incremento generalizado del ancho de banda formántica y la presencia ocasional del formante nasal. A finales del segundo año y a principios del tercero, empieza a observarse la presencia del formante a 250 Hz que consigue controlar en las posteriores producciones.

## Diadococinesias

Para la tasa de habla en las diadococinesias no se obtienen valores acústicamente aceptables para la combinación /pataka/ en ninguno de los registros del estudio. El primer año se obtienen los valores máximos de producción: /pa/: 3, /ta/: 2 y /ka/: 1. Las puntuaciones disminuyen de forma progresiva hasta obtener en la última grabación los siguientes resultados: /pa/: 1, /ta/: 1 y /ka/: 0.

La tasa media para cada producción durante un segundo son: /pa/: 2,34 (DS: 0,85), con un máximo de 4 y un mínimo de 1; /ta/: 1,64 (0,50), máx.2 y mín.; y /ka/: 1 (DS: 0,68), máx. 2 y mín 0. A partir del tercer año, disminuyen claramente las tasas para /pa/ y /ka/, manteniéndose relativamente estables las de /ta/.

La duración en milisegundos (ms) de /ta/ y /pa/ se ve incrementada con el tiempo conforme avanza la enfermedad (Tabla 1). Los resultados obtenidos muestran que existen diferencias significativas entre las medias iniciales y finales de “pa” ( $p=0,048$ ). No se observan diferencias significativas para el resto de variables ( $p>0,05$ ). La duración de cada sílaba presenta irregularidades (Figura 3). El cómputo total de producciones para /ka/ en cada registro realizado disminuye drásticamente al no poderse detectar el VOT en la mayor parte de sus ejecuciones, produciendo acústicamente rasgos del fonema oclusivo velar sonoro /b/. Se observó dos y a veces más barras de explosión en /k/ cuando era acústicamente identificable. Además, en el último año analizado, resulta casi imposible diferenciar las producciones silábicas por mostrar en el espectrograma un continuo de sonido.

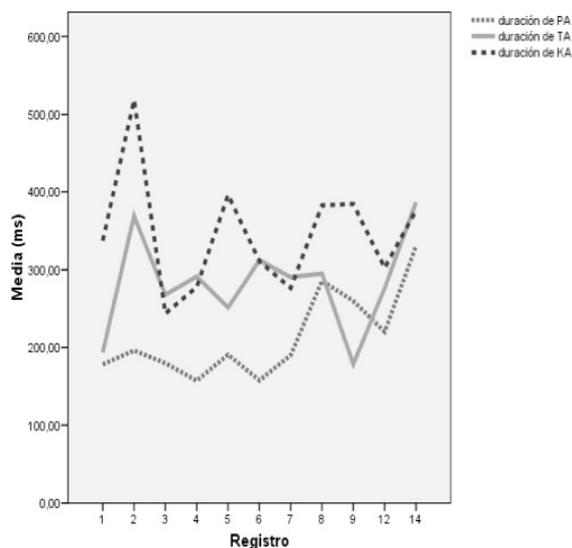


Figura 3. Evolución de la duración de las sílabas “pa”, “ta” y “ka”.

Los registros 10, 11 y 13 se han suprimido del gráfico por la imposibilidad de obtener una producción válida de “ka”.

En la duración (ms) para la producción “1,2,3,4” (Tabla 1) se observa un incremento progresivo que en su registro final es un 12 % mayor que en la producción inicial (1362 ms). Cada uno de los elementos del automatismo aumenta de forma

distinta hasta obtener su valor final: “uno” es un 12,25 % mayor que al inicio (403,50 ms); “dos”, un 6,9 % de los 246,10 ms inicial; “tres”, un 25,58 % de los 228,90 ms del primer año; “cuatro” incrementa sus 329,10 ms solamente un 0,2 % en la última grabación. Con estos resultados se pone de manifiesto la tendencia a la prolongación de las vocales y las claras dificultades asociadas a la disartria en la coarticulación, en este caso concreto para el diptongo /ua/ de “cuatro”. Este diptongo llega a duplicar su valor inicial (127,20 ms) respecto del último año registrado (valor final: 253,10 ms). Cabe tener en cuenta que en las últimas grabaciones resultaba imposible diferenciar las vocales en el espectrograma al mostrar un continuo sonoro.

## Pausas

Para la duración total del enunciado “1,2,3,4” se observa un incremento destacado durante el tercer año llegando a 3089,80 ms (al inicio es de 1362 ms y en el último registro es de 1996,90 ms). La media de porcentajes de emisión de voz en la producción numeral es de 84,49 % con una desviación estándar de 10,69; al inicio es de 88,62 % y al final del 100 %. Para las pausas los valores son: 14,39 % y DS:10,64; siendo en la grabación inicial de 11,38 % y desconocido en el último registro al no poder identificar acústicamente ningún aspecto claro en el continuo de sonido que producía. En el último año, la tendencia al habla encadenada reduce el valor de las pausas de forma significativa. Se progresa hacia un patrón en el que se inicia la producción del automatismo de forma lenta para terminar acelerando al producir la última palabra.

## Voice Onset Time (VOT)

Para /ka/ se observa un alto porcentaje (69,6 %) de imposibilidad para detectar el VOT en el conjunto de registros de este estudio, a diferencia de /pa/ y /ta/ que mantienen un 100 % de éxito en su detección. La pérdida de barra de explosión para /ka/ resulta evidente desde el primer año registrado pero en los registros finales no es posible analizar ningún dato válido para este fonema. Acústicamente, a finales del segundo año e inicio del tercero, empieza a observarse imprecisión a nivel articulatorio con múltiples barras de explosión para /k/, sonorización de /k/ y muestras de nasalidad en /p/ al final de algunas diadococinesias registradas.

Tabla 3. Valores de VOT en las sílabas /pa/, /ta/ y /ka/.

	Media / DS	Mínimo / Máximo	Inicio <sup>a</sup>	Final <sup>b</sup>	<i>p</i> <sup>d</sup>
/pa/	12,38 / 4,827	6,53 / 25,30	10,55	25,30	0,585
/ta/	20,36 / 5,703	11,00 / 31,40	31,40	22,30	0,438
/ka/	24,66 / 15,241	20,15 / 44,42	29,40	- <sup>c</sup>	0,331

**Nota:** Valores obtenidos con los datos de los 14 registros totales. Unidad: milisegundos (ms).

DS: desviación estándar. <sup>a</sup> Media de la producción en la primera grabación. <sup>b</sup> Media de la producción en la grabación final. <sup>c</sup> No se obtuvo muestra inteligible ni identificable acústicamente. <sup>d</sup> Procedimiento Medias de SPSS® para los valores iniciales y finales de cada parámetro (significación estadística  $p < 0,05$ ).

## Discusión

Los resultados del presente estudio longitudinal permiten cuantificar la alteración en el habla en la parálisis supranuclear progresiva. La disartria está presente en la PSP de forma muy inicial con una combinación de componentes hipocinéticos, espásticos y atáxicos.

En el caso presentado, la curva melódica se vuelve plana progresivamente conforme avanza la enfermedad. Kent (1997) evaluando la conversación de pacientes con disartria atáxica detecta el mismo perfil prosódico aplanado. En cuanto a la resonancia, únicamente se aprecian dos episodios aislados de presencia de nasalidad al producir /pa/ al final de alguna serie. Los armónicos y los formantes aparecen cada vez menos delimitados y claros en los distintos espectrogramas.

La estabilidad fonatoria se ve comprometida desde los estados iniciales con valores inferiores a la normalidad para el TMF en la producción sostenida de /a/, coincidiendo con los 13,2 s obtenidos por Rusz (2015) en los 12 individuos estudiados. El progresivo incremento de la duración de la vocal en las diadococinesias así como el aumento del tiempo empleado en la producción de cada sílaba repercuten en la velocidad del habla, volviéndola más pausada, lenta e imprecisa. Las puntuaciones obtenidas por el paciente son considerablemente superiores a las medias obtenidas en el trabajo de Tykalova (2017) con valores para /pa/ de 84 ms, /ta/ 90 ms y /ka/ 89 ms. En su investigación incluyó la producción de la sílaba con tres vocales distintas (a, i, u), aspecto que deben ser tenido en cuenta puesto que la coarticulación y las relaciones fonológicas de cada lengua influye en su ejecución.

Coincidiendo con la literatura, no se aprecian alteraciones significativas en la frecuencia fundamental, en la intensidad ni en F1 y F2, obteniéndose resultados considerados no patológicos. Resultados similares obtiene Saxena (2014) que valora la intensidad de la voz en 19 individuos con PSP en la tarea de repetición silábica obteniendo una medida de 70-71 dB. El mismo autor, no detecta diferencias significativas respecto al grupo control en los valores de los dos primeros formantes en la producción de sílabas. La alteración del jitter, el shimmer y el HRN detectadas en el presente estudio longitudinal coinciden con los presentados por Rusz (2015) en su investigación con 12 individuos con PSP mostrando valores medios de 1,60 %, 8,48 % y 15,03 dB respectivamente. La falta de control del sistema nervioso subyacente a la misma enfermedad puede ser el responsable de estos cambios que se relacionan con la percepción de la ronquera (Jackson-Menaldi, 2002).

La tasa de diadococinesias fue valorada durante el periodo de un segundo, obteniendo valores medios para los 14 registros de 2,34 para /pa/; 1,64 para /ta/; 1 para /ka/ y 0 para /pataka/. Se interpreta que el habla es lenta, con una evidente reducción de la velocidad articulatoria y un incremento del porcentaje de pausas en cada una de las series. Este dato dista de algunos estudios previos como el de Kluin (1993) en el que cuantificó la tasa de repeticiones (realizadas por 44 individuos con PSP) durante 5 segundos: 17 /pa/, 16 /ta/, 13 /ka/ y 6 /pataka/. En todas las investigaciones, incluida la presente, destaca la inestabilidad en las repeticiones silábicas cuyo mecanismo (según Skodda, 2012) puede ser inducido por la combinación de la disfunción en el control motor a nivel de los ganglios basales junto con el componente espástico que conlleva a la imprecisión articulatoria.

Según los datos, la duración de las sílabas con fonemas oclusivos sordos disminuye conforme avanza la enfermedad y la gravedad en la disartria; en cambio

las pausas aumentan en las diadococinesias. Este fenómeno se interpreta como un indicador de la imprecisión y de la falta de estabilidad del control motor del habla (Wang, Kent, Duffy, Thomas, 2005) a la vez que se considera un aspecto clásico a nivel acústico en las disartrias atáxicas (Akerman, Hertrich & Hehr, 1995). Existen estudios previos como los de Skodda (2011) y el de Rusz (2015) en el que se valora el porcentaje de pausas en tareas como la lectura o el monólogo respectivamente, obteniendo valores de 14.42 % y 37,1 %. En el presente trabajo, la tarea a realizar consistía en la producción de un enunciado (automatismo numeral) en el que se detectó de media un 14,39 % que fue disminuyendo en los registros posteriores hasta ser indetectable debido al continuo fonatorio en el que se convirtió la expresión oral.

En relación al VOT, se mantiene la misma proporción que la observada en la literatura: cuanto más posterior es la consonante, más tarde empieza la sonoridad. Así se observa también en el estudio del VOT en español con pacientes que presentan disartria atáxica (Delgado, 2016) aunque realizan la tarea de repetición de palabras y no la de sílabas. El valor del VOT para /ka/ en la grabación inicial de la presente investigación duplica el de /pa/. En el estado final del estudio, el valor para la bilabial es aún mayor que el de /ta/ y podría deberse al aumento progresivo del trismus que limita la movilidad y precisión labial. Estas mismas dificultades fonatorias debidas a los desórdenes en la coordinación de la musculatura, añadido a la rigidez de la misma, repercuten en la apertura de las cuerdas vocales (Hlavnička, J., Tykalová, T., Čmejla, R., Klempir, J., Růžička, E. & Rusz, J., 2017) prolongando o reduciendo el VOT y llegando a un continuo sonoro en las consonantes sordas como se observa en /ka/. Saxena (2014) detecta un incremento del VOT en el caso de la PSP en las palatales y bilabiales. En otros estudios recientes se obtiene un valor medio de todas las diadococinesias sin distinción del punto articulatorio, como en Tykalova (2017), limitando la comparación de los resultados con los aquí presentados.

Las posibilidades de detección del VOT para /ka/ se reducen drásticamente conforme avanza la enfermedad y empeora la calidad vocal debido a la tendencia a la continuación ininterrumpida del sonido invadiendo el espacio temporal de las pausas. El porcentaje de no detección del VOT (69,6 %) discrepan de los presentados en otros estudios ya existentes, como el de Özsancak (2001) que se sitúa en 2,9 % en los 48 casos de disartria espástica estudiados realizando tareas de repetición de palabras y frases, lectura de texto y repetición silábica.

Existe una amplia variabilidad entre pacientes con disartria y en un mismo individuo, como se detecta en el presente estudio con el empeoramiento generalizado de la mayoría de los parámetros acústicos durante el tercer año de grabación. Coincide con el incremento de las dificultades deglutorias: la aparición de episodios de expulsión vía oral del contenido alimenticio durante la preparación y propulsión del bolo; aumento del residuo oral y faríngeo y la disminución de la amplitud bucal, que requirió de adaptación dietética (textura triturada y homogénea). El aumento del trismus observado coincide también en el tiempo con la disminución progresiva de F1 en la producción de /pa/.

Acústicamente, la presencia de disfonía y la pobreza en armónicos en el espectrograma indica una inadecuada aducción de las cuerdas vocales y una pobre presión subglótica (menor esfuerzo vocal), lo que sugiere una insuficiente protección de la laringe durante la deglución. Este mismo planteamiento se puede interpretar teniendo en cuenta la relación directa entre F1 y la afectación de la eficacia deglutoria

por limitación de la apertura bucal. Del mismo modo, las dificultades al producir diadococinesias sin modificar el estado “sordo” de /k/ sugiere el compromiso de la seguridad en la deglución al deteriorarse la función laríngea. Estos aspectos, junto a una ineficacia tusígena (a menudo presente en este tipo de disartrias) y el reducido descenso de la epiglotis (Litvan, 1997) predisponen a la disfagia.

En la parálisis pseudobulbar, la afectación de la voz (progresivamente más tensa) y de la deglución con compromiso añadido de la masticación, evolucionan de forma progresiva hacia una severa limitación en la comunicación y en el acto deglutorio. La relación existente entre estos dos aspectos fundamentales de la vida humana podría llegar a anticiparse o a concretarse usando los medios disponibles para la medición acústica del habla. Así por ejemplo, las relaciones entre F1 y F2 (con la apertura bucal y la posición lingual respectivamente) y el compromiso de la función laríngea (con las dificultades de producir correctamente los oclusivos sordos), deberían ser tenidos muy en cuenta durante la evaluación y la intervención logopédica, como signo previo a una muy probable disfagia futura.

## Referencia bibliográficas

- Ackerman, H., Hertrich, I. & Hehr, T. (1995). Oral diadochokinesis in neurological dysarthrias. *Folia Phoniatrica et Logopaedica*, 47, 15-23
- Asensi, LL., Pórtoles, S. & del Río, A. (1997). Barra de explosión, VOT y frecuencia de las oclusivas sordas del castellano. *Estudios de fonética experimental*, 9, 221-242
- Bascuñana, H. (1999). Características de la disfagia neurológica en las demencias y en las enfermedades neurológicas progresivas que afectan al sistema nervioso central. *Rehabilitación* 33,38-42
- Boersma, Paul & Weenink, David (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.40, retrieved 11 May 2018 from <http://www.praat.org/>
- Colosimo, C., Bak, TH., Bologna M. & Beradelli, A. (2014). Fifty years of progressive supranuclear palsy. *Journal of Neurology, Neurosurgery and Psychiatry*, 95,938-944
- Delgado, J. (2016). Medida de la severidad de la disartria atáxica a través del análisis acústico. *Estudios de Fonética Experimental*, 25, 149-166
- Golbe, L. (2014). Progressive supranuclear palsy. *Seminars in Neurology*, 34,151-159
- Gómez-Haro, C., Espert, R., Gadea, M. & Navarro, J.F. (1999). Parálisis supranuclear progresiva: aspectos neurológicos, neuropatológicos y neuropsicológicos. *Revista de Neurología*, 29 (10):936-956
- Hlavnička, J., Tykalová, T., Čmejla, R., Klempir, J., Růžička, E. & Rusz, J. (2017). Dysprosody Differentiate Between Parkinson's Disease, Progressive Supranuclear Palsy, and Multiple System Atrophy. 1844-1848. 10.21437/Interspeech.2017-762.
- Jackson-Menaldi, M.C. (2002). *La voz patológica*. Ed. Médica Panamericana.
- Keller, E., Vigneux, P., Laframboise, M. (1991). Acoustic analysis of neurologically impaired speech. *British Journal of Disorders of Communication*, 26, 74-94.
- Kent, R.D., Kent, J.F., Rosenbek, J.C., Vorperian, H.K. & Weismer, G. (1997). A speaking task analysis of the dysarthria in cerebellar disease. *Folia Phoniatrica et Logopaedica*, 49, 63-82
- Kent, R.D., Kent, J.F. Duffy, J.R., Thomas, J.E., Weismer, G. & Stuntebeck, S.(2000). Ataxic dysarthria. *Journal of speech, language and hearing research*, 43, 1275-89.

- Kim, J-H. & McCann, C.M. (2015). Communication impairments in people with progressive supranuclear palsy: A tutorial. *Journal of Communication Disorders*, 56, 76-87
- Kluin, K.J., Foster, N.L., Berent, S. & Gilman, S. (1993). Perceptual analysis of speech disorders in progressive supranuclear palsy. *Neurology*, 43, 563-566
- Litvan, I., Sastry, N. & Sonies, B. (1997). Characterizing swallowing abnormalities in progressive supranuclear palsy. *Neurology*, 48, 1654-1662.
- Melle, N. (2007). *Guía de intervención logopédica en la disartria*. Editorial Síntesis.
- Metter, E.J. & Hanson, W.R. (1986). Clinical and acoustical variability in hypokinetic dysarthria. *Journal of Communication Disorders*, 19, 347-366.
- Nishio, M. & Niimi, S. (2004). Relationship between speech and swallowing disorders in patients with neuromuscular disease. *Folia Phoniatrica et Logopaedica*, 56, 291-304
- Özsancak, C., Auzou, P., Jan, M. & Hannequin, D. (2001). Measurement of Voice Onset Time in dysarthric patients: methodological considerations. *Folia Phoniatrica et Logopaedica*, 53, 48-57
- Rehman, H.U. (2000). Progressive supranuclear palsy. *Postgraduate Medical Journal*, 76, 333-336
- Rusz, J., Bonnet, C., Klempíř, J., Tykalová, T., Baborová, E., Novotný, M., Ulseh, A. & Ružička, E. (2015). Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supra nuclear palsy and multiple system atrophy. *Journal of Neurology*, 262, 992.
- Saxena, M., Behari, M., Kumaran, S.S., Goyal, V. & Narang, V. (2014). Assessing speech dysfunction using BOLD and acoustic analysis in parkinsonism, *Parkinsonism and Related Disorders*, 20, 855-861
- Skodda, S., Visser, W. & Schlegel, U. (2011). Acoustical analysis of speech in Progressive Supranuclear Palsy. *Journal of Voice*, 25, 725-731.
- Skodda, S., Grönheit, W. & Schlegel, U. (2012). Instability of syllable repetition in progressive supranuclear palsy. *Journal of Neural Transmission*, 119, 457-462
- Tykalova, T., Rusz, J., Klempir, J., Cmejla, R., & Ruzicka, E. (2017). Distinct patterns of imprecise consonant articulation among Parkinson's disease: Progressive supranuclear palsy and multiple system atrophy. *Brain and Language*, 165,1-9.
- Wang, Y-T., Kent, R.D., Duffy, J.R. & Thomas, J.E. (2005). Dysarthria associated with traumatic brain injury: speaking rate and emphatic stress. *Journal of Communication Disorders*, 38, 231-260