



**UNIVERSIDAD  
DE GRANADA**

**FACULTAD DE TRADUCCIÓN E INTERPRETACIÓN**

Trabajo Fin de Grado

**SESGOS EN EL *SOFTWARE* DE PROCESAMIENTO DEL  
LENGUAJE NATURAL DE LA APLICACIÓN SIRI EN  
DISTINTAS VARIEDADES DEL ESPAÑOL DE ESPAÑA**

Curso 2020/2021

Autores: **Begoña Fernández Martínez y Manuel Torres Rodríguez**

Tutora: **Rocío Díaz Bravo**





UGR

Universidad  
de Granada

### Declaración de Originalidad del TFG

(Este documento debe adjuntarse cuando el TFG sea depositado para su evaluación)

D./Dña. BEGOÑA FERNÁNDEZ MARTÍNEZ, con DNI  
(NIE o pasaporte) 31019668M, declaro que el presente Trabajo de  
Fin de Grado es original, no habiéndose utilizado fuente sin ser citadas debidamente. De  
no cumplir con este compromiso, soy consciente de que, de acuerdo con la Normativa  
de Evaluación y de Calificación de los estudiantes de la Universidad de Granada de 20  
de mayo de 2013, esto *conllevará automáticamente la calificación numérica de cero  
[...]independientemente del resto de las calificaciones que el estudiante hubiera  
obtenido. Esta consecuencia debe entenderse sin perjuicio de las responsabilidades  
disciplinarias en las que pudieran incurrir los estudiantes que plagie.*

Para que conste así lo firmo el 22 / 06 / 2021 (FECHA)

Firma del alumno



ugr

Universidad  
de Granada

### Declaración de Originalidad del TFG

(Este documento debe adjuntarse cuando el TFG sea depositado para su evaluación)

D./Dña. MANUEL TORRES RODRÍGUEZ, con DNI  
(NIE o pasaporte) 47390410Z, declaro que el presente Trabajo de  
Fin de Grado es original, no habiéndose utilizado fuente sin ser citadas debidamente. De  
no cumplir con este compromiso, soy consciente de que, de acuerdo con la [Normativa  
de Evaluación y de Calificación de los estudiantes de la Universidad de Granada](#) de 20  
de mayo de 2013, esto *conllevará automáticamente la calificación numérica de cero  
[...]independientemente del resto de las calificaciones que el estudiante hubiera  
obtenido. Esta consecuencia debe entenderse sin perjuicio de las responsabilidades  
disciplinarias en las que pudieran incurrir los estudiantes que plagie.*

Para que conste así lo firmo el 22/06/2021 (FECHA)

Firma del alumno



# ÍNDICE DE CONTENIDOS

RESUMEN .....	1
ABSTRACT .....	1
1. INTRODUCCIÓN .....	2
2. MOTIVACIÓN Y OBJETIVOS .....	3
3. MARCO TEÓRICO .....	5
3.1 Variedades del español en España .....	5
3.1.1 Las hablas andaluzas .....	5
3.1.2 Las hablas canarias .....	9
3.1.3 El castellano .....	11
3.2 El español de las comunidades autónomas bilingües .....	14
3.2.1 El español de Cataluña, Islas Baleares y Comunidad Valenciana .....	14
3.2.2 El español del País Vasco y el norte de Navarra .....	16
3.2.3 El español de Galicia .....	18
4. ESTADO DE LA CUESTIÓN .....	20
4.1 Definición, objetivo e historia del procesamiento del lenguaje natural .....	20
4.1.1 Definición .....	20
4.1.2 Objetivo .....	21
4.1.3 Historia .....	21
4.2 Siri: historia y desarrollo .....	24
4.3 Comparativa con otros asistentes virtuales .....	28
4.4 Uso de Siri .....	30
5. HIPÓTESIS Y PREGUNTAS DE INVESTIGACIÓN .....	31
5.1 Hipótesis .....	31
5.2 Preguntas de investigación .....	32
6. DESARROLLO DE LA INVESTIGACIÓN: METODOLOGÍA Y DISEÑO DE LA ENCUESTA .....	33
6.1 Metodología .....	33
6.2 Diseño de la encuesta .....	34
7. ANÁLISIS Y DISCUSIÓN DE LAS MUESTRAS .....	37
8. CONCLUSIÓN .....	49
9. REFERENCIAS BIBLIOGRÁFICAS .....	51
ANEXOS .....	56

I. Registro de muestras.....	56
II. Siri y las variedades del español en España .....	59
II.1 Muestras relacionadas con las hablas andaluzas .....	59
II.2 Muestras relacionadas con las hablas canarias .....	69
II.3 Muestras relacionadas con el castellano .....	79
III. Siri y el español de las comunidades autónomas bilingües .....	89
III.1 Muestras relacionadas con el español de Cataluña, Islas Baleares y Comunidad Valenciana.....	89
III.2 Muestras relacionadas con el español del País Vasco y el norte de Navarra..	99
III.3 Muestras relacionadas con el español de Galicia.....	107
IV. Protocolo de encuesta para el Trabajo Fin de Grado .....	117
IV.1 Instrucciones sobre el desarrollo de la encuesta .....	117
IV.2 Preguntas .....	117
V. Informe favorable del Comité de Ética .....	120

## **RESUMEN**

La presencia de los asistentes de voz virtuales en nuestro día a día es cada vez mayor. Es una realidad innegable que su uso se ha incrementado tanto en jóvenes como en adultos y personas mayores. Ahora es momento de plantearse cómo influyen las distintas variaciones (diatópica y diastrática) y cómo puede variar la efectividad en función de la variedad del español que transcriba.

**Palabras clave:** Siri, procesamiento del lenguaje natural, inteligencia artificial, variedades del español, rasgos lingüísticos

## **ABSTRACT**

The presence of virtual voice assistants in everyday life is becoming more and more common. It is undeniable that their use has increased among young people, adults and the elderly. It is now time to consider the influence of the different variations (diatopic and diastratic) and how effectiveness may vary depending on the variety of Spanish being transcribed.

**Keywords:** Siri, natural language processing, artificial intelligence, varieties of Spanish, linguistic features

# **1. INTRODUCCIÓN**

Hoy en día, la población usa de manera generalizada y creciente los asistentes virtuales. Estos asistentes permiten al usuario manejar sus dispositivos electrónicos a través del control por voz, algo que facilita el uso de los mismos para las personas con mayor dificultad para habituarse a las nuevas tecnologías, y para toda la población en general. Aunque existen numerosos ejemplos, uno de los asistentes virtuales más usados en nuestro país es Siri, desarrollado por la empresa Apple. El desarrollo de esta aplicación y de muchas otras ha supuesto un gran avance para la convivencia de los humanos con las máquinas, y ha sido posible gracias a los *software* de procesamiento del lenguaje natural (PLN).

El PLN es un campo de conocimiento dentro de la inteligencia artificial (IA) que se encarga de investigar la manera en la que las personas podemos comunicarnos con las máquinas mediante el uso de alguna lengua. Estrictamente, cualquier lengua puede ser procesada por las máquinas, pero, lógicamente, hay limitaciones tanto económicas como prácticas que provocan que solo las lenguas más habladas por la población tengan cabida en el mundo digital y las aplicaciones.

Si partimos de la base de que el PLN no procesa aún todas las lenguas que se hablan en el mundo, entenderemos que, para el PLN, la variedad dialectal de muchas de estas lenguas también es complicada de procesar. Es por esta pluralidad que los *software* de PLN no pueden ofrecer resultados igual de precisos para todas las lenguas del mundo y todas las variedades dentro de esas lenguas. Esto quiere decir que existe la posibilidad de que algunas variedades se queden atrás en el desarrollo de estos *software* y que muchos hablantes no puedan disfrutar de unos resultados precisos cuando utilicen sus asistentes virtuales.



Esta investigación trata de identificar si existen sesgos en el *software* de PLN de la aplicación Siri de Apple cuando esta es sometida a hablantes de diverso origen social y de distintas variedades del español de España. De esta manera podremos examinar si existe algún tipo de discriminación tecnológica para los hablantes que hablen una variedad del español concreta.

## **2. MOTIVACIÓN Y OBJETIVOS**

Nuestro interés en lingüística computacional, *machine learning* y procesamiento computacional de variedades lingüísticas y dialectos es lo que nos ha motivado a llevar a cabo esta investigación. Son temas de gran importancia en la actualidad, pero encontramos que aún no se han llevado a cabo estudios sobre el español y sus variedades dentro de los *software* de procesamiento natural del lenguaje.

Los objetivos que perseguimos son aportar nuevos focos de investigación y dar visibilidad a la estigmatización que sufren muchas variedades del español, en este y en el resto de ámbitos, y explicar por qué es tan necesaria la mejora del rendimiento de aplicaciones de reconocimiento del habla.

A raíz de la lectura de la investigación «Gender and Dialect Bias in YouTube’s Automatic Captions» de Rachel Tatman (2017) de la Universidad de Washington conocimos y empezamos a interesarnos por el fenómeno denominado «brecha digital». Este consiste en cualquier desigualdad en el acceso, el uso o el impacto de las nuevas tecnologías entre distintos grupos sociales. Estos grupos pueden definirse según varios criterios: de género, geográficos, geopolíticos, culturales... En este caso, la variable a la que prestaremos atención es la variable geográfica, conocida en lingüística como variación diatópica.

En realidad, la brecha digital define al «pobre digital» como persona que, por falta de conocimiento o de ingresos, carece de acceso a las tecnologías digitales. También forma

parte de la brecha digital la escasa habilidad que pueda tener algún sector de la población (por ejemplo, los jubilados) para manejar estas nuevas herramientas. Pero en nuestra investigación, lo que queremos demostrar es que aun teniendo acceso a las nuevas tecnologías y haciendo un correcto uso de ellas, la variable geográfica, en este caso dialectal, hace que el resultado que se obtiene de ellas no siempre sea igual de preciso.

Esta irregularidad ocurre porque la mayoría de los sistemas de identificación lingüística no están diseñados para manejar la diversidad dialectal; hay un gran desequilibrio en el conjunto de datos de entrenamiento o *input*. Para programar un *software* de procesamiento del lenguaje natural, se debe contar con un *input* de muestras que sean representativas de los diferentes dialectos de una misma lengua, sin sesgos. Porque si hay sesgos en la recopilación de muestras, también los habrá en el sistema (Torralba y Efros, 2011; Bock y Shamir, 2015). Esto es algo de grandísima importancia porque pensamos que las diferencias en la precisión del reconocimiento automático del habla que se basan en la identidad social del hablante son una cuestión ética.

Si permitimos que estos sistemas funcionen sistemáticamente peor para los usuarios de los grupos desfavorecidos (como podrían ser los hablantes andaluces) que para los de los grupos privilegiados (por ejemplo, los hablantes del castellano), estaremos contribuyendo a agravar las desigualdades existentes. El objetivo que perseguimos con este trabajo es hacer una llamada de atención para que el procesamiento del lenguaje natural funcione igual de bien para todos los usuarios independientemente de su origen sociolingüístico. Es decir, desarrollar herramientas de procesamiento del lenguaje natural que sean socialmente inclusivas y luchar contra la discriminación digital.

Este es un tema especialmente preocupante dado el creciente uso de estas tecnologías por parte de los españoles. En 2019, los resultados de la 1ª Ola del Estudio General de Medios (EGM) presentados por la Asociación para la Investigación de Medios de Comunicación

(AIMC) mostraban que un 10,7 % de la población española utiliza algún asistente virtual de voz, es decir, unas 4 300 000 personas. Según esta encuesta, el asistente más utilizado, con diferencia, es Siri de Apple, declarado por el 86 % de los encuestados. El uso de los asistentes virtuales es cada vez más común en personas de la tercera edad, que no solamente utilizan sus dispositivos para recordar tareas, programar alarmas o consultar el tiempo, sino también para combatir algo que les afecta a muchos: la soledad. Por esto mismo, no podemos permitir que la identidad social de los hablantes repercuta en el rendimiento de estas nuevas tecnologías que tanto facilitan la vida de los usuarios. Este análisis de errores que compara el rendimiento entre grupos sociolingüísticos pretende ayudar a garantizar que esto deje de ocurrir y poner de manifiesto cualquier desequilibrio que pueda existir.

### **3. MARCO TEÓRICO**

#### **3.1 Variedades del español en España**

##### **3.1.1 LAS HABLAS ANDALUZAS**

###### **3.1.1.1 Introducción**

La importancia de las hablas andaluzas no reside solo en el peso poblacional que Andalucía tiene sobre el total de España (el 17,84 % de los españoles son andaluces), sino que se remonta hasta el siglo XIII, cuando los castellanos desembarcaron en el sur peninsular. Entonces, Sevilla se convirtió en el referente urbano de toda la región, cuya habla, tras crecer con habitantes de diversos orígenes geográficos, avanzó hacia soluciones lingüísticas novedosas (Instituto Nacional de Estadística, 2021). Desde entonces, observada esta circunstancia desde una perspectiva más sociolingüística que histórica, los hablantes andaluces marcaban su distancia de las variedades castellanas del norte peninsular (Moreno-Fernández, 2020, p. 60).

Definir con precisión quiénes son los hablantes de andaluz es más complejo de lo que pueda parecer en un principio, pues las fronteras lingüísticas y las fronteras administrativas, a menudo, son distintas. Es posible establecer una franja septentrional (Sierra de Huelva, comarcas cordobesas del Valle del Guadiato y Los Pedroches, zona oriental de Jaén, las comarcas de Huéscar granadina y Los Vélez almeriense) donde los rasgos de la variedad castellana y las hablas andaluzas se funden y sería complicado determinar qué variedad se habla en cada una de ellas. No obstante, al bajar esa línea, es casi imposible encontrar unas características comunes (Narbona-Jiménez *et al.*, 2011, p. 17).

De acuerdo con Moreno-Fernández (2014, pp. 152-155), podemos establecer una división administrativa general de las hablas andaluzas por la delimitación de las distintas provincias: por un lado, las hablas de Andalucía occidental (Sevilla, Cádiz, Córdoba y Huelva) y, por otro, las de Andalucía oriental (Málaga, Granada, Almería y Jaén).

Hay lingüistas, como el catedrático de la Universidad de Granada, José Mondéjar Cumpián (Álvarez-García, 2004, p. 32), que afirman que el andaluz no es un dialecto, ya que, para ello, tendría que haber una unidad fonética, léxica, morfológica y sintáctica que no existe. Sí admiten la existencia de hablas andaluzas, una variedad dialectal dentro de un territorio geográfico más reducido. Como ejemplo, señalan el uso frecuente en Andalucía occidental del *ustedes* conjugado con un verbo en 2ª persona del plural (*ustedes sabéis*), algo que consideran un vulgarismo, pues en situaciones en las que se requiere un registro más elevado, corregirían el uso del verbo en 2ª persona del plural y emplearían, como dicta la Real Academia Española en su obra *Nueva gramática de la lengua española*, la 3ª persona.

Frente a estos, destacados lingüistas como Antonio Narbona, Rafael Cano y Ramón Morillo sí consideran el andaluz un dialecto (2011, p. 20):

El andaluz es históricamente dialecto del castellano, en cuanto que es una expansión del mismo por conquista, repoblación y colonización [...]. En su caracterización no deben contar únicamente los fenómenos innovadores. Al lado de los *evolutivos* (*revolucionarios*, en algunos casos), se descubren otros *conservadores*, y muchos de los primeros (sobre todo, de pronunciación) no gozan de prestigio.

En medio de unos y otros, no son pocos los lingüistas que coinciden en que procede calificar o no al andaluz como un dialecto. Este término, más que connotaciones lingüísticas, posee connotaciones políticas y económicas, por lo que los conflictos derivados de esta cuestión son, en la mayoría de los casos, estériles. No hay un solo rasgo que sea común en el habla de la totalidad de los andaluces ni un solo rasgo que sea exclusivo de las hablas andaluzas. Lo que sí es cierto es que, hoy por hoy, casi nadie duda que el andaluz es una variedad hablada del español (Narbona-Jiménez *et al.*, 2011, p. 20).

### 3.1.1.2 Plano fónico

En el plano fónico de las hablas andaluzas sobresalen por su frecuencia el **yeísmo generalizado**, consistente en pronunciar el sonido de las palatales /ɲ/ y /j/, el **seseo**, el rasgo más frecuente entre hablantes de Córdoba, zona oeste de Sevilla y zonas centrales de Andalucía, y el **ceceo**, generalmente asociado a las zonas rurales de Andalucía. Los hablantes seseantes articulan una pronunciación predorsal de la /s/ para los grafemas «s», «z», «ce» y «ci», mientras que los ceceantes articulan el fonema /θ/ para dichos fonemas. Según Moreno-Fernández, el «ceceo se considera rural, frente al urbano seseo, y no goza del prestigio abierto entre los hablantes andaluces» (2010, p. 74). En este sentido, otro de los rasgos emergentes es el **heheo**: la aspiración de la /s/ en posición explosiva, tanto al principio de palabra, como en posición intervocálica (Pharies, 2007, p. 198).

También es interesante observar la **aspiración, asimilación o pérdida de /-s/ en posición final de sílaba** (Moreno-Fernández, 2010, p. 75), que, al aspirarse, /s/ se transforma en [h]; la **elisión de la /d/ intervocálica**, sobre todo, en participios; la pérdida de consonantes finales (-r, -l, -d); la **abertura de vocales finales por pérdida o debilitamiento de consonante final** (este rasgo es específico de Andalucía oriental); y la **lenición o pronunciación fricativa de la africada /tʃ/ (/tʃ/ > /ʃ/)**. En este último rasgo puede influir la variación diatópica, ya que es más frecuente en Andalucía occidental (Moreno-Fernández, 2020, 61) y diastrática, pues no se suele dar en el acrolecto, es decir, en la variedad hablada por los ciudadanos con mayor nivel socioeducacional (Muñoz-Basols y Lacorte, 2017, p. 367).

La **velarización de las consonantes nasales -l y -r en posición final de sílaba**, la **aspiración de la fricativa velar sorda /x/** como fricativa laríngea sorda [h] (se exceptúa en la zona oriental de Andalucía, en especial, Granada, Jaén y Almería (Muñoz-Basols y Lacorte, 2017, p. 366)), **abertura de vocales finales** y el debilitamiento o **pérdida de las consonantes finales** alveolares -r, -l, -n y -d se consideran también rasgos significativos de las hablas andaluzas (Moreno-Fernández, 2020, p. 61).

### 3.1.1.3 Plano gramatical o morfosintáctico

En relación con la gramática de las hablas andaluzas, destaca el **tuteo**, como en el resto de España, si bien es cierto que en buena parte de Andalucía occidental hay **alternancia de vosotros y ustedes en el sistema de tratamiento y uso de se con valor de segunda persona del plural** (*Se vais a enterar* en lugar de *Os vais a enterar*). Se puede escuchar por gran parte del territorio el **uso de infinitivo con valor de imperativo**.

En el basilecto, o variedad popular empleada por el estrato inferior o con menos recursos de una comunidad (Muñoz-Basols y Lacorte, 2017, p. 346), son rasgos frecuentes el

**cambio de género en algunos sustantivos** (*el tele* en lugar de *la tele*) y el **uso del artículo definido ante nombres propios de persona** (*la Tere*). A diferencia del castellano, en Andalucía destaca la **ausencia de leísmo de cosa, de láismo y de loísmo**. No obstante, «como aproximación al español castellano podría explicarse la tendencia al leísmo que se ha comenzado a detectar en las generaciones más jóvenes» (Moreno-Fernández, 2020, p. 63).

#### **3.1.1.4 Plano léxico**

En lo que se refiere al léxico de las hablas andaluzas, hay algunos **arcaísmos**, como *casapuerta* ('portal o zaguán') o *escarpín* ('calcetín'), y **arabismos**, como *almatriche* ('acequia o canal de riego') o *zagal* ('muchacho'). Encontramos también **andalucismos léxicos** que se emplean a lo largo de toda la geografía andaluza, como *búcaro* ('botijo') o *malaje* ('persona desagradable'). Hay también interjecciones, como *¡ojú!* o *¡isú!*, propias del habla coloquial para expresar sorpresa (Muñoz-Basols y Lacorte, 2017, pp. 367-368). Algunos de los vocablos se usan también en América, como *cobija* ('ropa de abrigo para la cama') o *molestoso* ('molesto') (Moreno-Fernández, 2020, p. 63).

### **3.1.2 LAS HABLAS CANARIAS**

#### **3.1.2.1 Introducción**

Repartido a lo largo de todo el territorio insular, a unos 2000 km de distancia del resto del país, las hablas canarias han actuado como depositarias de elementos singulares de ida y vuelta, es decir, que afectan a varios niveles, desde su léxico particular hasta su entonación característica. Estas hablas, históricamente, han surgido como una mezcla de las hablas de los castellanos que desembarcaron allí y de las hablas de todos los que llegaban allí de paso: dominicanos en el siglo XVIII, puertorriqueños en el siglo XIX y cubanos y venezolanos en el pasado siglo XX (Muñoz-Basols y Lacorte, 2017, p. 64).

Desde la Península, la mayoría de los habitantes que se asentaron allí eran procedentes de Andalucía, muchos de ellos nobles andaluces que tomaron posesión de las islas, en las que alcanzaban acuerdos comerciales y tomaban ventaja de su posición estratégica en el mapa. No es de extrañar, por tanto, los estrechos lazos que unen a Andalucía con el archipiélago canario, desde el plano histórico hasta el lingüístico, como vamos a observar a continuación.

### 3.1.2.2 Plano fónico

En el plano fónico, las hablas canarias han hecho suyos muchos rasgos analizados en las hablas andaluzas: el **yeísmo generalizado** (los fonemas palatales /ʎ/ y /j/ se pronuncian igual), la **aspiración, asimilación o pérdida de /s/ en posición final de sílaba** (Moreno-Fernández 2010, 75), la **aspiración de la fricativa velar sorda /x/** como fricativa laríngea sorda [h] y la **elisión de la /-d-/ intervocálica**, aunque este fenómeno es menos frecuente por el contraste entre habla rural (elisión) y urbana (conservación) (Etxebarria-Aróstegui, 2002, pp. 203-204).

A diferencia de las hablas andaluzas y el castellano, el **seseo** (*s* dorsodental) en las hablas canarias está muy generalizado y es empleado por casi la totalidad de la población. Otro de los rasgos que da personalidad a las hablas canarias es el sonido palatal correspondiente a la *che*, denominado ***che* adherente**, pues la lengua se adhiere a la superficie del cielo de la boca y la pronunciación es similar a [tj] o [j]. De la misma manera, la *che* puede ser **sonorizada** (Moreno-Fernández, 2020, p. 65). Al contrario de lo que sucede en el norte peninsular, en Canarias se observa una **entonación menos ascendente** para las oraciones interrogativas absolutas y más coincidente con el dialecto caribeño que con las hablas del resto de España (Muñoz-Basols y Lacorte, 2017, p. 370).



### 3.1.2.3 Plano gramatical o morfosintáctico

En lo que se refiere al plano gramatical de las hablas canarias, cabe destacar el **tuteo** para mostrar cercanía y familiaridad entre los hablantes, así como el **uso de *ustedes***, si bien en La Gomera, El Hierro y La Palma se alterna con el uso de *vosotros* como muestra de confianza (Calderón-Campos, 2010, p. 229). El **uso del pretérito perfecto simple** es preferente frente al pretérito perfecto simple y se dan ciertos giros sintácticos muy característicos en estas hablas, como **art. + *más/menos que* + verbo**. Por último, en las hablas canarias se hace un **uso etimológico de los pronombres personales átonos** (*la, le y lo*), pese a que algunos hablantes pueden mostrar leísmo de persona y leísmo de cortesía (Muñoz-Basols y Lacorte, 2017, p. 370).

### 3.1.2.4 Plano léxico

Las hablas canarias mantienen en este plano sus semejanzas con el dialecto caribeño. En las voces características de estas hablas podemos encontrar orígenes muy diversos: **guanchismos** (*jameo* ‘cueva’ o *tabaiba* ‘un tipo de árbol’), **americanismos** (*guagua* ‘autobús’ o *flis* ‘insecticida’), **arabismos** (*guayete* ‘niño’ o *arife* ‘calor sofocante’), **portuguesismos** (*enchumbar* ‘empaparse de agua’) y **arcaísmos** (*demorarse* ‘retrasarse’). Gran parte de estas voces se pueden oír también en Venezuela, país del que llegaron numerosos habitantes a Canarias en el siglo XVII. Es por ello que en Canarias se conoce a Venezuela como *la octava isla* (Moreno-Fernández, 2020).

## 3.1.3 EL CASTELLANO

### 3.1.3.1 Introducción

Pese a que existe una idea muy generalizada, si bien errónea, de que todos los habitantes de España hablamos la modalidad castellana, el castellano es la variedad empleada en la

península ibérica, especialmente en los dos tercios del norte (Moreno-Fernández, 2020, p. 58).

La formación inicial del castellano tuvo lugar en la Edad Media. A partir de entonces, esta variedad fue añadiendo elementos de otras variedades romances que ya se encontraban en la Península. Pese a su apariencia uniforme y regulada, el castellano se fue amoldando a las condiciones sociales de los lugares geográficos en los que se instaló (Moreno-Fernández, 2020, pp. 58-59).

Dentro de la variedad castellana, se pueden identificar dos grandes variedades del castellano en la península ibérica: por un lado, el castellano manchego (Castilla-La Mancha) y, por otro, el castellano norteño (al norte de Madrid), que incluye el norteño occidental (Castilla y León) y el norteño oriental (Aragón). En relación con las hablas andaluzas, el castellano comparte dos zonas limítrofes en las que estas dos variedades se funden y forman las llamadas «hablas de tránsito»: el oeste extremeño y el este murciano (Muñoz-Basols y Lacorte, 2017, p. 360).

### 3.1.3.2 Plano fónico

En el plano fónico del castellano, al igual que sucede en las hablas andaluzas y canarias, destaca un **yeísmo generalizado**. Suele llamar la atención la **pronunciación apicoalveolar de la /s/**, pues, para ello, la punta de la lengua roza los alveolos superiores. A colación de este último rasgo, cabe destacar que en la variedad castellana se produce la **distinción entre los fonemas /s/ y /θ/** (Moreno-Fernández, 2020, p. 59).

Asimismo, existe una tendencia a la **conservación de -r, -l, -n y -s como consonantes finales**, bien de sílaba o bien de palabra, algo que contrasta con la **elisión de la /-d-/ intervocálica** en el habla coloquial o espontánea y la **elisión del elemento oclusivo en la combinación /ks/ para el grafema «x»** (*tasi* por *taxi*). Este último no es el único

fenómeno que acontece en el plano fónico y que está relacionado con **el fonema fricativo velar sordo /x/**, que **se pronuncia de manera tensa** en el castellano, a la inversa de lo que sucede en la mayoría de variedades del español, en las que se articula de manera relajada o suave (Muñoz-Basols y Lacorte, 2017, p. 362).

### 3.1.3.3 Plano gramatical o morfosintáctico

En el plano gramatical, hay que señalar por su extensión y generalización el **tuteo**, el **uso de vosotros** y, sobre todo, el **leísmo** de cosa, del que se hace un uso intenso. El **laísmo** (empleo del pronombre de objeto directo *la* en lugar del pronombre de objeto directo *le*) y el **loísmo** (empleo del pronombre de objeto directo *lo* en lugar del pronombre de objeto directo *le*) son rasgos morfosintácticos singulares del castellano del centro peninsular y del occidente peninsular (Muñoz-Basols y Lacorte, 2017, p. 363).

A excepción de algunas zonas del norte peninsular, el **uso del pretérito perfecto compuesto** para acciones del pasado reciente es predominante, así como el empleo del **infinitivo con valor de imperativo**. En las últimas décadas, se viene observando un uso proporcionalmente mayor de las **formas en -se del en los pretéritos imperfecto y pluscuamperfecto de subjuntivo** en comparación con otras variedades del español (Moreno-Fernández, 2020, p. 59).

### 3.1.3.4 Plano léxico

En lo que se refiere al léxico del castellano, destaca el uso de **castellanismos** y **españolismos**, como *chaval*, *-a* para referirse a ‘chico, -a’, *droguería* como ‘establecimiento en el que se venden productos de limpieza y pintura’ o *fiambra* al hacer alusión a un ‘táper’ (Moreno-Fernández, 2020, p. 59).

## **3.2 El español de las comunidades autónomas bilingües**

### **3.2.1 EL ESPAÑOL DE CATALUÑA, ISLAS BALEARES Y COMUNIDAD VALENCIANA**

#### **3.2.1.1 Introducción**

Pese a que abundan los estudios sobre la influencia del español sobre la lengua catalana y viceversa, es sorprendente el escaso interés que suscita el estudio de la historia del español en las regiones bilingües peninsulares e insulares. Sin embargo, desde hace poco tiempo para acá numerosos lingüistas que han dirigido su atención al estudio de los aspectos sociológicos e ideológicos que cobran importancia en el uso que se hace del español en Cataluña (Blas-Arroyo, 2004, p. 1065).

Del mismo modo, en este período han visto la luz algunos estudios con orientación sociolingüística en los que se investigan las matrices lingüísticas y sociales que fomentan o restringen el uso del español en las comunidades del ámbito lingüístico catalán. En dichos trabajos, expertos en lingüística como José Ramón Gómez Molina o Monserrat Casanovas, analizan muestras de habla para obtener una idea más exacta de la relevancia social que alcanzan los fenómenos interferenciales del español dentro de estas regiones (Blas-Arroyo, 2004, p. 1066).

#### **3.2.1.2 Plano fónico**

Existen diversos fenómenos fonéticos que caracterizan el español de estas regiones, como la **confusión entre los fonemas interdental /θ/ y la sibilante alveolar /s/** (que alienta el seseo entre algunos de los miembros del basilecto), la **reducción de las vocales [e] y [a] átonas a una vocal central** o el **cierre de [o] átona en [u]**. No obstante, sí hacen distinción los hablantes del catalán entre las vocales [e] y [o] tónicas, ya que en la lengua catalana ambas poseen valor fonemático (Blas-Arroyo, 2004, p. 1071).

Algunos fenómenos del plano fónico surgen a raíz de la **sustitución** de la variante estándar por la variante *extranjera*. Entre ellos, destacan el **ensordecimiento de la oclusiva alveolar sonora /-d/ a final de palabra**, donde se convierte habitualmente en /t/, la **velarización de la alveolar /l/** y la **sonorización de la alveolar sorda /-s-/** por fonética sintáctica en algunos contextos intervocálicos (Blas-Arroyo, 2004, p. 1075).

### 3.2.1.3 Plano gramatical o morfosintáctico

En lo que se refiere al plano gramatical del castellano de Cataluña, Baleares y Comunidad Valenciana, es necesario reseñar el **uso predominante de las variantes flexivas verbales** (*llamaré*) en detrimento de las demás posibilidades expresivas que encontramos en la lengua española (*voy a llamar, llamo*). Dentro del plano verbal se puede destacar el uso de la perífrasis obligativa **haber de + inf.** en lugar de *tener que + inf.* Respecto a los elementos deícticos, como los demostrativos (*este, ese, aquel*) o los adverbios de lugar (*aquí, ahí, allí*), en el español de las comunidades de ámbito lingüístico catalán se observa una drástica **reducción de los términos intermedios en el habla**: solo se usan *este* o *aquel* y *aquí* o *allí* (Blas-Arroyo, 2004, p. 1072).

Los verbos *llevar* y *traer* se usan en estas comunidades de forma indistinta y el verbo **hacer** tiene unos usos más específicos que en el resto de comunidades autónomas: *hacer una película* se entiende como *emitir o proyectar una película*. El pronombre **que átono** se emplea sin **valor interrogativo** en oraciones interrogativas (*¿Qué quieres cenar?* Con un *Sí* o *No* como posibles respuestas) y la preposición **sin** también puede adquirir **funciones adverbiales** (*¿Traes la comida? – No, vine sin*) (Moreno-Fernández, 2020, p. 37).

Tras hacer un repaso a los fenómenos predominantes del castellano en contacto con el catalán, también hay que subrayar que existen **fenómenos sustitutorios**, como el que

sufren algunos sustantivos, cuyas formas femeninas se clasifican normalmente como vulgarismos en el español estándar: *¡Qué calor más mala hace hoy!* (Blas-Arroyo, 2004, p. 1076).

#### **3.2.1.4 Plano léxico**

Algunos de los términos que destacan del plano léxico del español de las comunidades de ámbito lingüístico catalán son *paleta* para hacer referencia a un ‘albañil’, *rachola* para mencionar una ‘baldosa’, *prestache* para aludir a un ‘estante’, *nen* para nombrar a un ‘niño’ o *plegar del trabajo* cuando alguien acaba de ‘salir del trabajo’. Esta última no es la única perífrasis que se usa por un fenómeno de **sustitución léxica**: hay otras como *dejar de pasar boniato* (‘sorprender’), *encontrar a falta* (‘echar de menos’), *no padecer* (‘no preocuparse’) o *querer decir* (‘estar seguro’).

### **3.2.2 EL ESPAÑOL DEL PAÍS VASCO Y EL NORTE DE NAVARRA**

#### **3.2.2.1 Introducción**

En la obra *Lenguas y dialectos de España* (García-Mouton, 1994), la autora destaca que el euskera es la única lengua prerromana peninsular, es decir, la única cuyo origen es anterior a la llegada del latín a la Península. Junto al finés, el lapón, el estonio, el carelio y el húngaro, el vasco es una de las seis lenguas no indoeuropeas que ha sobrevivido hasta la actualidad, soportando eso sí los embistes del régimen franquista, que vetó su uso, y demás convulsiones que han acabado fragmentándola (del Moral-Aguilera, 2014).

El uso del vasco no se limita a las tres provincias de Euskadi, pues se habla de igual modo en las tierras noroccidentales de Navarra. Fuera del territorio español, su uso también está extendido en el departamento francés de los Pirineos Atlánticos. En cuanto a su estructura, el vasco funciona por la suma de elementos aislados, como prefijos y sufijos. Es decir, es una lengua aglutinante (García- Mouton, 1994, p. 51).

En relación con la convergencia del español y el vasco en las tierras donde conviven ambas lenguas, existen escasos recursos que analicen este fenómeno.

### 3.2.2.2 Plano fónico

Entre las características fónicas compartidas por las lenguas vasca y española, destaca la neutralización de la palatal central y la palatal lateral, lo que se conoce también como el **yeísmo**. Un rasgo que se ha expandido desde el vasco al español ha sido el **ensordecimiento de obstruyentes finales**, pero, sin duda, el rasgo fónico más reconocible de los hablantes de español en el País Vasco y en Navarra es la **elisión de la dental sonora**, acompañada por el **cierre de /o/ en /u/**, en las terminaciones en *-ado*. Es también un rasgo notable del español de los hablantes vascos la intensidad de la vibrante, que presenta un número de oclusiones mayor que en el resto de España (Oñederra-Olaizola, 1995, pp. 1104-1107).

De la misma manera, el español hablado en el País Vasco y en Navarra se distingue por su inconfundible entonación y por la tensión articulatoria, por lo que, a menudo, el español de los hablantes cultos de esa zona resulta mucho más normativo que el de otros lugares de la Península. En los hablantes del vasco con bajo nivel sociocultural se aprecia un **seseo con /s/ apical**: *voy a haser poca siesta* (García-Mouton, 1994, p. 53).

### 3.2.2.3 Plano gramatical o morfosintáctico

En el plano gramatical o morfosintáctico, se puede señalar el **uso predominante de *estar* en lugar de *haber*** (*no están papeles* en lugar de ‘no hay papeles’) (Moreno-Fernández, 2020, p. 37). En cambio, existe un contraste entre dos rasgos opuestos: la **omisión de los clíticos de tercera persona** (*pues eché mechas en el pelo*) y el **uso redundante del pronombre átono** (*La llamé a María*). En relación con estos dos últimos rasgos, se sitúa el **leísmo femenino** (*Le avisé a Carmen*) (García-Mouton, 1994, p. 53).

En zonas del entorno vasco castellanizadas, es imposible pasar por alto un rasgo que ha llegado hasta ahí por influencia de la lengua vasca: la **sustitución del pretérito imperfecto de subjuntivo por el condicional** (*Si yo podría asistir a la reunión, iría*) (García-Mouton, 1994, p. 53).

#### **3.2.2.4 Plano léxico**

Son muchos más los **préstamos que la lengua vasca** toma del español que los que el español toma de la lengua vasca. Sin embargo, el número de estos últimos no es nada desdeñable: *aita* ‘padre’, *ama* ‘madre’, *anaia* ‘hermano’, *morrosko* ‘fuerte’, *belari* ‘reunión’, *bertzale* ‘patriota’, *pelotari* ‘jugador de pelota’ (Moreno-Fernández, 2020, p. 37).

### **3.2.3 EL ESPAÑOL DE GALICIA**

#### **3.2.3.1 Introducción**

Frente a lo acontecido en Cataluña o Euskadi, Galicia es una comunidad autónoma en la que la emigración ha sido permanente a lo largo de los últimos siglos y la inmigración, por el contrario, insignificante. La ausencia de una nobleza gallega y de una clase media fuerte y acomodada, desde el reinado de Fernando III, aceleró el abandono de la lengua gallega y la adopción del español como lengua de uso en el día a día para progresar de manera gradual en la sociedad.

Sin embargo, en los últimos años de la dictadura franquista comenzó a fraguarse un proceso rupturista del código diglósico: las clases bajas y medias percibían que ascender socialmente iba siendo cada vez más posible y, al mismo tiempo, los jóvenes con mayor nivel socioeducacional comenzaron a defender el uso del gallego. Como consecuencia de tales hechos y de la cascada de movimientos políticos que comenzaron a sucederse desde



1978, la situación del gallego ha mejorado en las últimas décadas (Rojo-Sánchez, 2004, pp. 1088-1089).

### 3.2.3.2 Plano fónico

Uno de los rasgos más llamativos del español hablado en Galicia es, sin duda, la **entonación** de los hablantes. Sin embargo, desde el plano científico e investigador, no pasa desapercibido su **sistema vocálico**, como el claro resultado de la integración del sistema de la lengua gallega, con siete vocales y cuatro grados de apertura, en el de la lengua española, con cinco vocales y tres grados de apertura. Un rasgo fónico en común con el resto de hablas analizadas es la **pérdida, debilitamiento o elisión de consonantes finales** ante palabras que comienzan por vocal (Rojo-Sánchez, 2004, pp. 1093-1094).

### 3.2.3.3 Plano gramatical o morfosintáctico

Entre los rasgos gramaticales que destacan del español de Galicia se encuentran el uso frecuente de los **diminutivos en -iño/a**, el uso del **pretérito perfecto simple vinculado al presente** y el uso del verbo **tener + participio** con valor reiterativo: *Lo tengo visto miles de veces en esta tienda* (Moreno-Fernández, 2020, p. 37). Otra de las construcciones perifrásticas propias es **haber de + infinitivo** con valor de posterioridad o de obligatoriedad: *Si me echas una mano, he de invitarte a algo*. También **llevar + participio** es mucho más frecuente en las hablas de esta comunidad que en el resto de España: *Ese equipaje lleva hechos muchos golpes* (Rojo-Sánchez, 2004, p. 1096).

En algunos casos, se aprecia una **transferencia del género** del equivalente gallego al español, como *el vispera* o *el sal*. En el español hablado en Galicia, también se dan casos en los que los hablantes hacen uso del **pretérito imperfecto de subjuntivo con valor de indicativo**: *Cuando nos dimos cuenta, ella ya se fuera* (Rojo-Sánchez, 2004, p. 1095).

### **3.2.3.4 Plano léxico**

En este apartado, podemos hacer distinción entre dos grupos distintos: términos que existen en las lenguas gallega y española, aunque con distinto significado, y palabras gallegas que se han integrado en el español de Galicia sin tener en cuenta su carácter diferencial. En el primer grupo, encontramos términos como *sentir* ‘oír’, *librar* ‘vaciar’, *pota* ‘olla’, *parvo* ‘bobo’ o *toro* ‘rodaja (de pescado, por ejemplo)’. En el segundo grupo, hallamos términos como *cheirar* ‘oler mal’, *lambón* y *larpeiro* ‘goloso’, *esmagar* ‘aplastar’, *carallada* ‘cosa sin importancia’ o *rosmar* ‘refunfuñar’ (Rojo-Sánchez, 2004, pp. 1098-1099).

## **4. ESTADO DE LA CUESTIÓN**

### **4.1 Definición, objetivo e historia del procesamiento del lenguaje natural**

#### **4.1.1 DEFINICIÓN**

El procesamiento del lenguaje natural (PLN) es el enfoque informático del análisis de textos que se basa tanto en un conjunto de teorías como en un conjunto de tecnologías. Al ser un área muy activa de investigación y desarrollo, no existe una única definición consensuada. En el artículo «Natural Language Processing» (2001) escrito por Elizabeth D. Liddy, se define el PLN como el conjunto de técnicas computacionales teóricamente motivadas para analizar y representar textos naturales en uno o más niveles de análisis lingüístico con el fin de lograr un procesamiento del lenguaje similar al humano para una serie de tareas o aplicaciones.

Los «textos naturales» pueden ser de cualquier lengua, modo o género; así como orales o escritos. El único requisito es que estén en una lengua utilizada por los seres humanos para comunicarse entre sí. Además, el texto que se analice no debe construirse

específicamente para el propósito del análisis, sino que debe recogerse del uso real (Liddy, 2001, p. 3).

Aunque depende de otras disciplinas, es apropiado considerar el PLN como una disciplina dentro de la inteligencia artificial, ya que se esfuerza por conseguir un rendimiento similar al humano. Entre las principales disciplinas de las que depende encontramos: la lingüística, que se centra en los modelos formales y estructurales del lenguaje y en el descubrimiento de los universales del lenguaje; la informática, que se ocupa de desarrollar representaciones internas de los datos y de procesar eficazmente estas estructuras; y la psicología cognitiva, que estudia el uso del lenguaje como una ventana a los procesos cognitivos humanos y cuyo objetivo es modelar el uso del lenguaje de una forma psicológicamente plausible (Liddy, 2001, pp. 4-6).

#### **4.1.2 OBJETIVO**

El objetivo del procesamiento del lenguaje natural de los asistentes de voz es lograr un procesamiento del lenguaje similar al humano. La palabra *procesamiento* no debería sustituirse por *comprensión*, ya que hoy en día se está de acuerdo en que, aunque el objetivo del PLN es la verdadera comprensión del lenguaje natural, y esta aún no se ha logrado. Si este fuese el caso, el PLN sería capaz de parafrasear un texto de entrada, traducir el texto a otro idioma, responder a preguntas sobre el contenido del texto y extraer conclusiones del texto. Aunque sí se ha avanzado mucho en la consecución de los primeros tres objetivos, el PLN no puede aún, por sí mismo, extraer inferencias del texto (Lizzy, 2001, p. 3).

#### **4.1.3 HISTORIA**

El desarrollo del procesamiento del lenguaje natural comenzó en los años cuarenta. La traducción automática (TA), también llamada MT por sus siglas en inglés, fue la

primera aplicación informática relacionada con el lenguaje natural. Aunque Warren Weaver y Andrew Booth iniciaron uno de los primeros proyectos de TA en 1946, no fue hasta 1949 cuando Weaver sugirió utilizar las ideas de la criptografía y la teoría de la información para la traducción de idiomas. La investigación comenzó en varias instituciones de investigación de Estados Unidos en pocos años (Johnson, 2017, p. 71).

Los primeros trabajos de TA adoptaron la visión simplista de que las únicas diferencias entre las lenguas residían en sus vocabularios y en los órdenes de palabras permitidos. Los sistemas desarrollados, desde esta perspectiva, se limitaban a buscar en el diccionario las palabras apropiadas para la traducción y a reordenarlas después para que se ajustaran a las reglas de la lengua de destino, sin tener en cuenta la ambigüedad léxica inherente al lenguaje natural. El resultado fue nefasto, lo que hizo que los investigadores se dieran cuenta de que la tarea era mucho más difícil y que era necesaria una teoría del lenguaje más adecuada (Johri *et al.*, 2021, pp. 365-375).

Durante este periodo, empezaron a surgir otras áreas de aplicación del PLN, como el reconocimiento del habla. La comunidad del procesamiento del lenguaje y la comunidad del habla se dividieron entonces en dos bandos: la comunidad del procesamiento del lenguaje dominada por la perspectiva teórica de la gramática generativa y la comunidad del habla dominada por la teoría de la información estadística. Debido a los desarrollos de la teoría sintáctica del lenguaje y de los algoritmos de análisis sintáctico, en la década de los cincuenta hubo un entusiasmo excesivo que llevó a la gente a creer que los sistemas de traducción de alta calidad totalmente automáticos serían capaces de producir resultados idénticos a los de los traductores humanos, algo bastante irreal dados los conocimientos lingüísticos y los sistemas informáticos disponibles por aquel entonces (Johri *et al.*, 2021, pp. 365-375).

Las insuficiencias de los sistemas existentes en aquel momento dieron lugar al informe ALPAC (*Automatic Language Processing Advisory Committee of the National Academy of Science - National Research Council*) de 1966, que concluía que la TA no era realizable de forma inmediata y recomendaba que no se financiara. Aunque el trabajo en PLN disminuyó considerablemente durante los años posteriores al informe, se produjeron algunos avances significativos, tanto en cuestiones teóricas como en la construcción de prototipos de sistemas. El trabajo teórico de finales de los años sesenta y principios de los setenta se centró en la cuestión de cómo representar el significado y en el desarrollo de soluciones computacionalmente manejables que las teorías gramaticales existentes no eran capaces de producir. En 1965, el lingüista Noam Chomsky introdujo el modelo transformacional de la competencia lingüística. Sin embargo, las gramáticas generativas transformacionales estaban demasiado orientadas a la sintaxis para dar cabida a los aspectos semánticos (Lizzy, 2001, p. 5).

Paralelamente al desarrollo teórico, se desarrollaron muchos prototipos de sistemas para demostrar la eficacia de determinados principios. Uno de los prototipos fue el ELIZA de Weizenbaum, que se construyó para replicar la conversación entre un psicólogo y un paciente, simplemente permutando o haciéndose eco de la entrada del usuario. Otro ejemplo podría ser el SHRDLU de Winograd, que simulaba un robot que manipulaba bloques sobre una mesa. Este, a pesar de sus limitaciones, demostró que la comprensión del lenguaje natural era realmente posible para el ordenador (Lizzy, 2001, p. 5).

A finales de la década de los setenta, la atención se desplazó hacia las cuestiones semánticas, los fenómenos del discurso y los objetivos y planes comunicativos. En este periodo también se trabajó mucho en la generación de lenguaje natural. El planificador del discurso de McKeown, llamado TEXT, y el generador de respuestas de McDonald, llamado MUMMBLE, utilizaban predicados retóricos para producir descripciones

declarativas en forma de textos cortos, normalmente párrafos. La capacidad de TEXT para generar respuestas coherentes en línea se consideró un gran logro (Lizzy, 2001, p. 6).

A principios de la década de los ochenta, motivados por la disponibilidad de recursos computacionales críticos, la creciente conciencia dentro de cada comunidad de las limitaciones de las soluciones aisladas a los problemas del PLN y un impulso general hacia las aplicaciones que trabajaban con el lenguaje en un contexto amplio del mundo real, los investigadores comenzaron a reexaminar los enfoques no simbólicos que habían perdido popularidad en los primeros tiempos. A finales de la década de los ochenta, los enfoques simbólicos se habían utilizado para abordar muchos problemas significativos en PLN y los enfoques estadísticos demostraron ser complementarios en muchos aspectos a los enfoques simbólicos (Lizzy, 2001, p. 6).

En los últimos diez años del milenio, este campo ha crecido rápidamente. Esto puede atribuirse a la mayor disponibilidad de grandes cantidades de texto electrónico, a la disponibilidad de ordenadores con mayor velocidad y memoria y a la llegada de Internet. Los enfoques estadísticos lograron resolver muchos problemas genéricos de la lingüística computacional, como la identificación de partes del habla o la desambiguación del sentido de las palabras, y se han convertido en un estándar en todo el PLN. Los investigadores están desarrollando ahora sistemas de PLN de nueva generación que tratan razonablemente bien el texto general y tienen en cuenta una buena parte de la variabilidad y la ambigüedad del lenguaje (Lizzy, 2001, p. 6).

## **4.2 Siri: historia y desarrollo**

Siri es una aplicación desarrollada por iOS que cumple las funciones de asistente personal. Su funcionamiento interno es toda una incógnita. Desconocemos la existencia de estudios sobre este aspecto, si bien nuestra investigación se ha centrado más en sus

avances a lo largo de los últimos lustros. La herramienta ha sido diseñada para contar con personalidad propia y utiliza el procesamiento del lenguaje natural para responder a preguntas planteadas por los usuarios, dar recomendaciones y llevar a cabo otras acciones mediante la delegación de solicitudes hacia un conjunto de servicios en línea que se ha ido agrandando con el paso de los años (Barrios *et al.*, 2018, p. 94).

En diciembre de 2007, Dag Kittlaus, Adam Cheyer, Tom Gruber y Norman Winarsky, miembros del grupo SRI Venture, crearon este asistente de voz. Este grupo internacional, especializado en inteligencia artificial, nació en la década de los 40 del siglo pasado. La Universidad de Stanford, situada en California (Estados Unidos), fue la entidad encargada de impulsar el proyecto y dar alas a este grupo como una organización sin ánimo de lucro (Rosas-Pérez y Torres-Castillo, s. f., p. 2).

Como hemos mencionado anteriormente, entre sus creadores figura Adam Cheyer, actual vicepresidente de I+D+i en el grupo Samsung. La trayectoria de este ingeniero informático es su mejor aval, ya que cuenta con numerosos reconocimientos, bien por haber creado aplicaciones relacionadas con la inteligencia artificial o bien por haber impulsado proyectos que habían sido abandonados, como el caso de Bixby, asistente de voz de los dispositivos Samsung (Rosas-Pérez y Torres-Castillo, s. f., p. 2).

Pese a su esfuerzo y su empeño en sacar adelante este trabajo, el grupo SRI Venture se vio obligado por problemas de viabilidad económica a traspasar el proyecto a Apple Inc. en abril de 2010. Esta empresa hegemónica en el mundo de la telefonía móvil hizo la puesta a punto de esta herramienta en el dispositivo iPhone 4S, que salió al mercado el 4 de octubre de 2011. En sus nuevas versiones, Siri se ha ido integrando en iOS hasta el punto de ofrecer ayuda con otras aplicaciones, como WhatsApp, los recordatorios y la agenda de contactos.

Siri tiene soporte de idiomas para el alemán (de Alemania, Austria y Suiza), el árabe, el chino (cantonés y mandarín), el coreano, el danés, el español (de Chile, España, Estados Unidos y México), el finés, el francés (de Bélgica, Canadá, Francia y Suiza), el hebreo, el inglés (de Australia, Canadá, Estados Unidos, la India, Irlanda, Nueva Zelanda, el Reino Unido, Singapur y Sudáfrica), el italiano (de Italia y Suiza), el japonés, el malayo, el neerlandés (de Bélgica y Países Bajos), el noruego, el portugués de Brasil, el ruso, el sueco, el tailandés y el turco. Es decir, la aplicación está disponible en 20 idiomas y en 41 variedades distintas de dichos idiomas.

Hasta 2012, solo contaba con soporte para tres idiomas: el inglés, el alemán y el francés. No fue hasta el 7 de marzo de 2012, día en el que se publicó la actualización al iOS 5.1, cuando se hizo disponible para el japonés, el español, el italiano, el chino y el coreano (Jiménez-Martín y Sánchez-Allende, 2015, p. 22).

En septiembre de ese mismo año, Siri ofrecía ya la compatibilidad con las tabletas desarrolladas por Apple, los iPads, y con los coches por medio de Bluetooth, gracias al iOS 6 que fue presentado al público en dicho mes. No obstante, no fue hasta el lanzamiento un año más tarde del iOS 7 cuando Siri cumplió su «mayoría de edad». Siri dejó atrás entonces su condición de beta, o sea, la fase de desarrollo de la aplicación en la que se prueba y se proporciona una retroalimentación a sus desarrolladores para que puedan mejorarla (Jiménez-Martín y Sánchez-Allende, 2015, p. 22).

En esa actualización, Siri presentó una nueva interfaz, contestaba más rápido a las preguntas planteadas por los usuarios y presentaba nuevas funcionalidades, como devolver llamadas, ajustar el brillo de la pantalla automáticamente si así lo solicitaba el usuario y leer los mensajes del buzón de voz (Jiménez-Martín y Sánchez-Allende, 2015, p. 22).



En 2014, un especialista que trabajaba en el equipo de Siri, Álex Acero, confirmó que el asistente de voz empezaría a utilizar una red neuronal a partir de ese mismo año. Para ello, el equipo de *hardware* se encargó de hacer un uso de los datos prestados por los usuarios en los ajustes del teléfono y el equipo de *software* optimizó para que fuese posible que la herramienta funcionase adecuadamente en el nuevo *hardware* (Woods, 2018, p. 335).

Desde entonces hasta hoy, el objetivo de este asistente virtual ha sido siempre ser lo más parecido posible a cualquier humano. Por un sistema parecido al de *cookies* que tenemos que aceptar cuando visitamos un sitio web, Siri ha desarrollado un carácter intuitivo. Si antes era necesario pedirle una canción en concreto o, en cambio, reproducía aleatoriamente cualquiera de las que figuraba en nuestro álbum, ahora Siri es capaz de responder con acierto a una petición como «Siri, ponme algo de música que me guste».

Ahora bien, la trayectoria de este asistente no ha sido solo un camino de luces, ha habido polémicas que han ensombrecido y apagado sus logros, al menos, de manera temporal. En agosto de 2019, se supo que los empleados de Apple Inc. tenían acceso a las grabaciones de Siri que hacían los usuarios. Por ello, la empresa se vio en el compromiso de reconocer este error y asegurar por activa y por pasiva que nunca más se iba a repetir.

En abril de 2021, una guía publicada por el Centro Criptológico Nacional, órgano dependiente del Centro Nacional de Inteligencia, recogía textualmente la siguiente afirmación:

Debido a que la información procesada por Siri es enviada a Apple [...], y a que, a lo largo de la historia, se han identificado numerosas vulnerabilidades en este servicio, la recomendación de seguridad y privacidad es evitar el uso de Siri en la medida de lo posible.

### 4.3 Comparativa con otros asistentes virtuales

En este apartado procederemos a realizar un análisis comparativo de los tres *software* de procesamiento de lenguaje natural principales. Como ya hemos comentado, en 2011 se dio la bienvenida a Siri, creado por Apple, uno de los asistentes virtuales más exitosos. Durante los años siguientes se unieron Google Assistant, creado por Google, y Alexa, creado por Amazon (Saavedra-Montejo, 2020).

Siri es una herramienta que se basa en la inteligencia artificial disponible para todos los dispositivos de Apple. Destaca de entre el resto porque utiliza el PLN para responder preguntas, hacer recomendaciones y llevar a cabo acciones. Además, este *software* se adapta progresivamente a las preferencias individuales de cada usuario, pudiendo incluso mandar mensajes de texto, cantar y contar historias (Saavedra-Montejo, 2020).

Se trata de uno de los *software* de PLN más demandados por la privacidad que ofrece, ya que los datos no tienen que asociarse a ningún perfil personal, por lo que la información de los usuarios no se puede utilizar para hacerles llegar publicidad. Además de contar con una interfaz bastante sencilla y de responder rápidamente y con una voz muy natural, la gran ventaja de este *software* es que está disponible en 23 idiomas diferentes, lo que convierte a Siri en el asistente virtual que ofrece acceso a un mayor número de personas (Saavedra-Montejo, 2020).

Google Assistant fue desarrollado por Google y está implementado en altavoces inteligentes, en todos los *smartphones* de Android y en un gran número de dispositivos domésticos. Este *software* permite a los usuarios tener una interacción por medio de la voz natural. Responde preguntas abstractas y complejas y puede mantener conversaciones bidireccionales. Además de realizar búsquedas en internet, programar alarmas o eventos

y configurar *hardwares*, este asistente virtual destaca por su uso en la domotización de los hogares (Saavedra-Montejo, 2020).

Google Assistant quizás sea uno de los más inteligentes, ya que es capaz de reconocer comandos muy complejos e incluso el tono de voz. Se puede programar en 20 idiomas distintos, lo que facilita el acceso a una gran parte de la población mundial.

En 2014, Amazon creó Alexa, un asistente virtual doméstico capaz de controlar altavoces, cámaras de video, focos, interruptores, enchufes, etc. Usando simples comandos de voz, los usuarios pueden crear rutinas de automatización con todos sus dispositivos inteligentes. Este *software* cuenta con el número más amplio de *gadgets*, ya que incluso los dispositivos que no pertenecen a Amazon pueden ser controlados a través de Alexa. Aunque sí se pueden personalizar los comandos para mejorar su efectividad, Alexa está disponible en tan solo 5 idiomas.

Pese a que estos asistentes de voz se encuentran en constante actualización e incluyen novedades y nuevos idiomas a su configuración, aún no integran las variedades dialectales de cada idioma. Tampoco existen investigaciones acerca de esto, por lo que este tema es algo totalmente nuevo en la ciencia de la inteligencia artificial. En los estudios sobre el funcionamiento y el rendimiento de estos *software* no se presta atención a los errores cometidos por las variedades dialectales. Tan solo encontramos referencias a los dialectos dentro del PLN de los asistentes de voz para el inglés (Tait y Wilks, 2019, pp. 405-418), pero no existen estudios que evalúen la efectividad de las aplicaciones cuando son sometidas a diferentes variedades del español. Por ello, merece una especial mención este **hueco investigador**.

#### **4.4 Uso de Siri**

Según el informe emitido en 2020 por la consultora experta en transformación de redes sociales Ditrendia, Siri es el asistente virtual más reconocido, con un 74 %, seguido de Alexa (59 %) y el asistente de Google (52 %). Lo más reseñable de estos datos es el incremento en casi un 50 % de reconocimiento de Alexa, que en 2019 solo era identificada por el 9,8 % de las personas encuestadas (Rivero, 2020, p. 38).

Tras la llegada de la pandemia de la COVID-19, ciertas entidades bancarias han promovido el uso de asistentes virtuales para el uso de sus aplicaciones en línea, como es el caso del BBVA, que ha firmado un convenio con Apple para que sus clientes puedan enviar dinero y realizar gestiones con la ayuda de Siri (Rivero, 2020, p. 39).

Solo dos años antes, en 2018, este mismo informe señalaba que la revolución de los asistentes virtuales y de voz había llegado a Europa. Sin embargo, esta estaba yendo a más velocidad en Estados Unidos, donde el 64 % de la población reconocía sentirse cómoda utilizando asistentes de voz delante de otras personas. Casi veinte puntos por debajo se situaba el Reino Unido (46 %), que sacaba ventaja a Alemania (44 %) y Francia (38 %) (Rivero, 2019, p. 24).

A veces, la utilidad de estas figuras del mercado electrónico se mide en la capacidad de dar con la respuesta a aquello que le preguntamos. En el caso de Siri, en 2018, tenía una ratio de respuestas del 40 %, de las que un 80 % eran correctas. En cambio, el mejor desempeño correspondía al asistente de Google, que era capaz de responder en un 77 % de las ocasiones, con un índice de acierto superior al 95 % (Rivero, 2019, p. 27).

Si aumentamos la perspectiva, el pasado año 2019, Microsoft realizó un estudio mundial sobre el uso de los asistentes virtuales y de voz. Aunque Alexa es ampliamente conocida y publicitada en Estados Unidos, puede resultar sorprendente que no sea la más conocida.

Siri y el asistente de Google quedaron empatados en el primer puesto, con un 36 % de los encuestados que reconocieron utilizar cada uno de ellos. Alexa quedó en tercer lugar, con un 25 % de encuestados que lo utilizaban, seguida de cerca por Cortana, con un 19 %.

La popularidad de Siri y Google se debe probablemente al gran número de teléfonos móviles (aproximadamente 250 millones en Estados Unidos) en comparación con los altavoces inteligentes (más de 50 millones en Estados Unidos). Tanto Siri como Google están más vinculados a los teléfonos móviles, mientras que Alexa está más vinculada a los altavoces inteligentes (aunque Siri y Google están empezando a infiltrarse en los altavoces inteligentes, mientras que Alexa también puede utilizarse en teléfonos móviles).

Al poner en perspectiva estos datos, es posible tomar conciencia de lo mucho que se ha avanzado en este aspecto. Una encuesta llevada a cabo en Estados Unidos en el año 2012 reveló que el 87 % de los usuarios que tenían como móvil un iPhone 4S usaba algunas de las funciones de Siri, al menos, una vez al mes. Las funciones más recurrentes eran hacer llamadas telefónicas, enviar mensajes de texto y buscar información (Echauri-Soto, 2015, p. 3).

## **5. HIPÓTESIS Y PREGUNTAS DE INVESTIGACIÓN**

### **5.1 Hipótesis**

Creemos firmemente que va a existir una desigualdad en el resultado de las transcripciones de la aplicación Siri al ser esta expuesta a diferentes variedades del español de España. Pensamos que Siri no va a ser capaz de proporcionar resultados igual de precisos para todas las variedades a las que la exponemos, ya que no será capaz de comprender todos los rasgos lingüísticos de cada una de ellas.

Las variedades que ofrecerán peores resultados podrían ser el andaluz y el canario. Creemos que rasgos dialectales propios de estas variedades, tales como la pérdida o debilitamiento de sonidos finales, pueden confundir al *software* de procesamiento de lenguaje natural (PLN) de esta aplicación, haciendo que los resultados sean considerablemente peores.

Por el contrario, pensamos que otras variedades como el castellano tendrán resultados óptimos ya que el *software* de PLN de Siri podría estar programado con un *input* procedente de hablantes de estas zonas y no tendrá problema al transcribirlas. Creemos que podría estar programada con ese tipo de *input* porque el castellano es la variedad del español que, de manera popular y errónea, se considera normativa y más prestigiosa.

## **5.2 Preguntas de investigación**

Antes de dar comienzo al diseño de la encuesta y a la recolección de muestras para evaluar la efectividad de Siri, es necesario reflexionar sobre los aspectos que queremos analizar y poder contribuir con nuestro granito de arena a dar respuesta a las siguientes preguntas.

En primer lugar, ¿cuáles son los rasgos que más dificultad presentan para ser identificados en herramientas que procesan el lenguaje natural, como es el caso de Siri? Del mismo modo, cabe plantearse si influye positiva o negativamente nuestro lugar de origen y determinadas variables sociales, como el nivel de estudios, al utilizar aplicaciones con *software* de PLN.

Por último, tras pasado este límite, en el plano pragmático no es nada descabellado plantearse si es útil Siri al procesar el lenguaje natural, dejando atrás la barrera territorial. Es decir, ¿su índice de acierto es tal que es recomendable usarla? ¿Merece la pena dictar por nuestra voz lo que podemos hacer a golpe de clic?

## **6. DESARROLLO DE LA INVESTIGACIÓN: METODOLOGÍA Y DISEÑO DE LA ENCUESTA**

### **6.1 Metodología**

El estudio se divide en tres fases bien diferenciadas:

- a. La revisión bibliográfica de las variedades del español en el marco teórico y la investigación sobre el procesamiento del lenguaje natural y Siri. Esta primera fase, plenamente teórica, se caracteriza por el repaso de contenidos que ya hemos estudiado en otras asignaturas del grado, como ‘Lengua A Nivel 3 Español’, y la búsqueda bibliográfica, principalmente artículos de profesores universitarios, acerca del procesamiento del lenguaje natural y de Siri.
- b. El diseño de una encuesta y la recopilación de muestras de cada una de las variedades. En primer lugar, hemos diseñado una encuesta en la que se recogen aspectos relevantes fónicos, gramaticales o morfosintácticos y, en última instancia, léxicos. La encuesta consta de una primera parte con preguntas sobre el lugar de origen del participante, sus estudios y la manera de contactar con su familia si alguna vez ha vivido fuera. La segunda parte de la encuesta contiene tres imágenes: una de un cazador acompañado de un perro, otra de una casa muy infantil y una última de un carro de la compra con varios productos en su interior. Como se puede comprobar en el punto IV de los anexos, cada pregunta tiene un fin lingüístico relacionado con los tres planos mencionados anteriormente (fónico, gramatical y léxico).

Hemos redactado una solicitud al Comité de Ética en Investigación de la Universidad de Granada, antes de proceder a la recopilación de las muestras. Una vez que esta Comisión emitió un informe favorable a nuestra investigación

(Herrera-Viedma y O'Valle-Ravassa, 2021), disponible en el punto V de los anexos, comenzamos a recibir las pruebas de los voluntarios que decidan aportar su granito de arena a este proyecto. Pese a que en un principio tan solo perseguíamos que este fuese un estudio cualitativo, también es cuantitativo, pues hemos contado con un total de 42 voluntarios, 7 por cada una de las variedades analizadas: las hablas andaluzas, las hablas canarias, el castellano, el castellano en contacto con el catalán, el castellano en contacto con el euskera y el castellano en contacto con el gallego.

- c. La discusión y el análisis de las muestras recibidas. Tras compendiar la transcripción realizada por Siri de todas y cada una de las muestras, llega el momento de analizar los aciertos y los errores. La discusión no solo debe ceñirse a esos dos aspectos, sino además profundizar en los rasgos más repetidos en los planos fónico y gramatical y en los vocablos usados en el plano léxico.

La mayoría de los voluntarios que se han prestado a participar en nuestro Trabajo Fin de Grado son jóvenes, cursan aún estudios universitarios o los han terminado recientemente. De los 42 participantes, 25 son mujeres y 17 son hombres. Su procedencia geográfica y el número exacto de la muestra de cada uno de los participantes está disponible en el punto I de los anexos. El siguiente punto recoge las muestras relacionadas con las hablas andaluzas, canarias y el castellano y el punto III contiene las muestras del castellano en contacto con el catalán, el euskera y el gallego.

## **6.2 Diseño de la encuesta**

Mediante el diseño de la encuesta, pretendemos dar con las respuestas a las preguntas que hemos planteado anteriormente. Por ello, tras darle algunas vueltas y valorar las distintas



opciones, esta es la encuesta que vamos a plantear a los voluntarios que se presten a participar:

¡Hola!

Somos Manuel Torres y Begoña Fernández, estudiantes de cuarto curso de Traducción e Interpretación en la Universidad de Granada. Estamos llevando a cabo esta encuesta para un estudio sobre procesamiento del lenguaje natural que realizaremos en nuestro Trabajo de Fin de Grado.

Estamos buscando a voluntarios para que nos ayuden a realizar este estudio mediante la grabación de un breve audio. Si desea participar, por favor, siga las siguientes instrucciones:

- Busque un lugar tranquilo y sin ruido para grabarse mientras responde las preguntas.
- Puede utilizar la grabadora del teléfono móvil o directamente una nota de voz a través de la aplicación WhatsApp.
- Si prefiere mandarnos su grabación por correo electrónico, envíela a esta dirección: [tfgvariedades21@gmail.com](mailto:tfgvariedades21@gmail.com)

La participación es completamente voluntaria. Al enviar su audio **está dando su consentimiento** para que se utilice con fines científicos y educativos.

**Estructura del audio:**

Empiece la grabación diciendo su nombre y lugar de nacimiento (ciudad/pueblo y país). Es importante que diga «*Me llamo...*» para que todas las encuestas empiecen igual. Por ejemplo, «*Me llamo Ana y soy de Pozoblanco, un pueblo de Córdoba, España*».

A continuación, **responda a las siguientes preguntas**. Intente hacerlo de la manera más exhaustiva posible y, lo más importante, de manera espontánea y natural (para que se asemeje lo más posible a como usted hablaría normalmente en una conversación):

- ¿Podría decirnos cómo era su ciudad/pueblo cuando era pequeño/a?
- ¿Estudia o trabaja en la actualidad? Si trabaja, ¿cuál es su profesión? Si estudia o ha estudiado, ¿qué estudia o ha estudiado? Si estudia o ha estudiado en la universidad, indique cuál, por favor.
- ¿Ha vivido alguna vez fuera de su lugar de nacimiento? ¿Dónde? ¿Cuánto tiempo? ¿Cómo se comunicaba con su familia desde allí?

**Describe las siguientes imágenes**, por favor:

Imagen 1



Imagen 2



Imagen 3



## 7. ANÁLISIS Y DISCUSIÓN DE LAS MUESTRAS

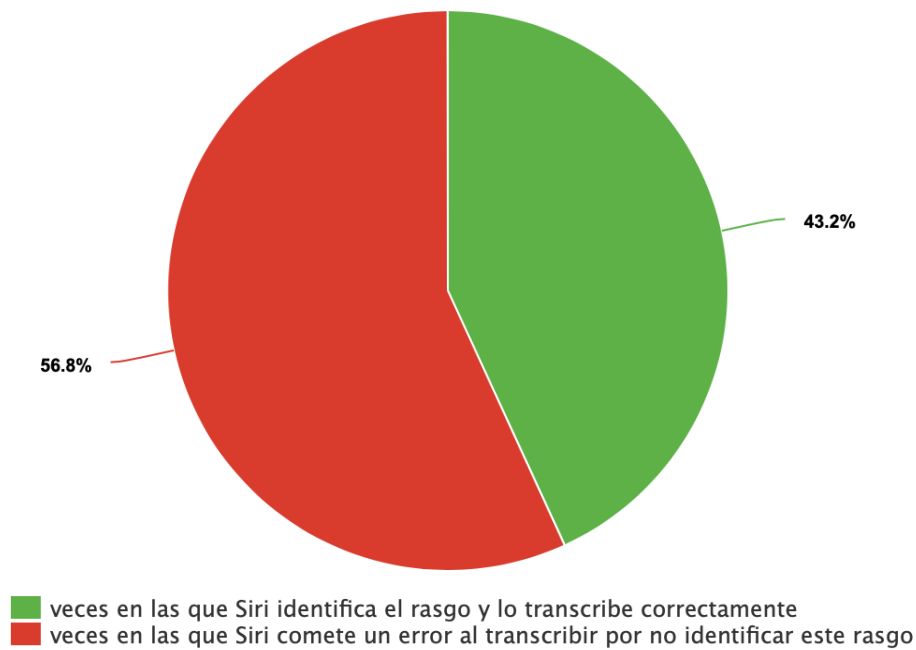
Dedicaremos este apartado a comentar los aspectos más importantes dentro del análisis de los rasgos lingüísticos de las muestras, tanto del plano fónico, como gramatical y sintáctico. Además, ofreceremos datos y gráficos que hemos diseñado con la información sacada de las tablas de análisis.

Comenzaremos con el **plano fónico**, posiblemente el más significativo dentro de las variedades dialectales. Hemos observado que cada variedad tiene sus propios, y muy diferentes, rasgos lingüísticos fonéticos. Lo primero en lo que reparamos fue en dividir las variedades según la articulación de los fonemas /s/ y /θ/. Encontramos muestras en las que los hablantes distinguen de forma generalizada estos dos fonemas dentro de la variedad castellana y de esta en contacto con catalán, con euskera y con gallego. En realidad, estas variedades también se consideran castellano en contacto con otras lenguas en zonas bilingües. Por el contrario, en la variedad canaria encontramos que todos y cada uno de los participantes eran seseantes. La variedad andaluza muestra diferencias en

cuanto a este rasgo dentro de los diferentes participantes ya que dos seseaban, otros dos ceceaban y tres distinguían. Aunque existan muchas diferencias a la hora de pronunciar ambos fonemas, Siri es perfectamente consciente de ello y solo ha cometido 2 errores en relación al seseo, y porque este iba acompañado además de la pérdida de una consonante final.

Esto nos lleva a comentar los rasgos lingüísticos en relación a los sonidos al final de palabra. Es muy común dentro de las hablas andaluzas perder por completo las consonantes *-s*, *-d*, *-r*, *-z* y *-l* en posición implosiva. Las veces en las que hemos registrado la pérdida de alguna de estas ha sido 102, habiendo cometido Siri un error de transcripción 54 de las veces. Algo similar ocurre con las hablas canarias. Los hablantes canarios pierden también la *-d*, por ejemplo, pero en el caso de la *-s*, el sonido no se pierde por completo, pero sí se debilita, dando lugar a un fenómeno que se conoce como aspiración. En la recogida de muestras encontramos ejemplos de aspiración de *-s* en hasta 112 ocasiones, siendo Siri incapaz de identificar 58 de ellas. Esto hace que Siri confunda los plurales de los singulares, transcribiendo «unas patatas» como «una patata» o «Soy de Paradas» como «Soy de Parada». Como vemos, este es el rasgo lingüístico fonético que más errores genera en la transcripción del habla natural, resultando perjudicados, en la inmensa mayoría de los casos, los hablantes andaluces y canarios.

A continuación, mostramos el gráfico 1, en el que observamos los resultados de la pérdida o aspiración de *-s* en las muestras de las hablas andaluzas y canarias:



*Ilustración 1. Porcentaje de acierto/error en la transcripción de la aspiración o pérdida de la /-s/ en las muestras de las hablas andaluzas y canarias.*

Por el contrario, dentro del castellano y el español en las zonas bilingües, los hablantes rara vez pierden el sonido a final de palabra. En todo caso, algunos hablantes pierden la *-d*, habiendo sido el caso en 15 ocasiones dentro de las cuatro variedades anteriormente mencionadas, frente a los 452 ejemplos de ocasiones en las que mantienen ese y otros sonidos. Aunque, al contrario que la pérdida de *-s*, la pérdida de *-d* no ha ocasionado ningún error en la transcripción de Siri.

Curiosamente, algo similar ocurre con la pérdida de consonantes al final de sílaba. Los andaluces y canarios pierden *-s*, *-c*, *-d* o *-p* al final de sílaba, mientras que los hablantes de las otras cuatro variedades las mantienen en la mayoría de los casos. Pero este rasgo es perfectamente procesado por Siri, ya que en las 56 ocasiones en las que esto ha ocurrido, Siri solo ha cometido un error de transcripción en 2 de ellas.

También son destacables las distintas opciones que se dan en las muestras para la *-d-* intervocálica. En todas las variedades encontramos que esta puede perderse o mantenerse;

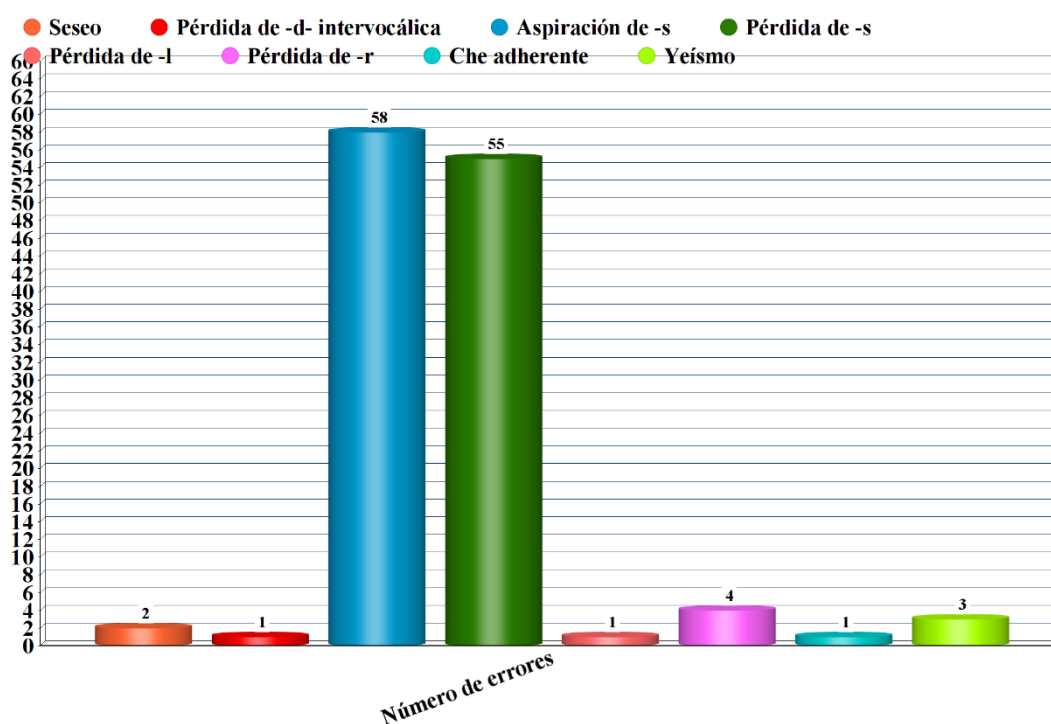
no existe una práctica generalizada. Encontramos 6 ejemplos (de las hablas andaluzas, castellano, castellano en contacto con catalán y castellano en contacto con euskera) en los que se pierde la *-d-* intervocálica. Siri tan solo comete un error. Este se da en una muestra procedente de Navarra en la que la persona pierde la *-d-* al decir «al lado de» y Siri transcribe «a la de». En cambio, encontramos dentro de todas las variedades hasta 28 ejemplos en los que esta *-d-* intervocálica se mantiene. Algo curioso relacionado con esto es la elisión de la dental sonora acompañada por el cierre de /o/ en /u/ en las terminaciones en *-ado* (*estudiau* por *estudiado*), lo cual encontramos en una muestra del castellano en contacto con euskera. Este rasgo no supone ninguna complicación para Siri.

Otro de los rasgos que dividen en dos mitades a los hablantes del español en España es la pronunciación de la fricativa velar sorda /x/. Mientras que en las variedades castellana y de las zonas bilingües se hace una articulación tensa de /x/ (registrado en 74 ocasiones), en las variedades andaluza y canaria se hace una articulación relajada de /x/ (registrado en 32 ocasiones). Tan solo existe una anomalía dentro del andaluz y es que en dos de las muestras, las procedentes de la provincia de Jaén, hacen hasta en 3 ocasiones una articulación tensa de /x/ como es común en ese territorio. Pero estas diferencias no han generado error alguno, por lo que podemos afirmar que Siri reconoce y transcribe este rasgo a la perfección.

Tras este estudio podemos comprobar también que existe un yeísmo generalizado (86 ejemplos entre todas las muestras) en todo el territorio nacional, ya que solo en 14 ocasiones hemos localizado ejemplos de distinción entre los fonemas palatales /ʎ/ y /j/, sobre todo en muestras del castellano en contacto con catalán. Pero Siri no tiene prácticamente ningún problema para identificar este rasgo, habiéndose equivocado solo en 2 ocasiones, transcribiendo «valla» por «vaya» y «ya» por «ella».

Por último, quisiéramos comentar rasgos llamativos y propios de cada variedad relacionados con su fonética. Por ejemplo, dentro de las hablas canarias encontramos la pronunciación de la *che* adherente hasta en 11 ocasiones, habiendo cometido Siri un error solo en 1 de ellas al transcribir «hecho» por «y yo». El castellano en contacto con catalán también ofrece algunos rasgos fonéticos típicos como el ensordecimiento de la oclusiva alveolar sonora /d/ a final de palabra (3 ocasiones) o la reducción de las vocales [e] y [a] átonas a una vocal central (15 ocasiones).

Para finalizar, ofrecemos este gráfico de barras en el que se muestran los rasgos lingüísticos fonéticos que han provocado errores y el número de veces en el que hemos registrado un error:



*Ilustración 2. Número de errores por cada uno de estos rasgos fónicos.*

Dejando a un lado los aspectos más representativos del plano fónico, no está de más recordar que, al haber diseñado una encuesta dentro de un marco claramente definido y

con preguntas muy concretas, es más sencillo observar rasgos léxicos o fónicos que sintáctico-gramaticales. No obstante, como bien se puede apreciar en el anexo el **plano gramatical y sintáctico** también ha arrojado interesantes resultados.

Nada más comenzar con el análisis, uno de los aspectos en los que más rápido reparábamos era el uso etimológico de los pronombres clíticos o, en su defecto, el laísmo, el loísmo o el leísmo. Mientras que de los dos primeros no hemos encontrado ninguna muestra ni ejemplo, sí los ha habido del último. En concreto, nos hemos topado con dos de las siete muestras del gallego donde existía el leísmo, así como un ejemplo en una muestra del castellano y otro en una muestra del castellano en contacto con el euskera. En todos estos casos, la efectividad de Siri ha sido máxima, pues ha reconocido tanto el leísmo como el uso etimológico de los pronombres clíticos sin mayor problema.

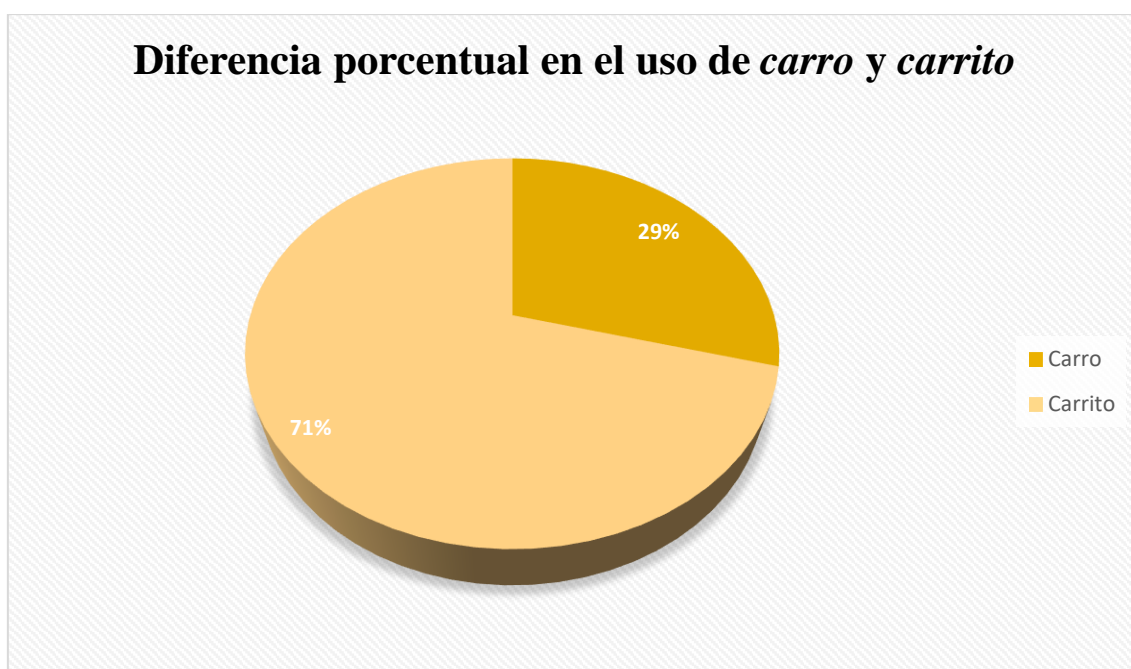
Lamentablemente, esto no ha ocurrido con las sinalefas, rasgos fonético-sintácticos para los que Siri se ha visto en apuros a la hora de identificarlos: «he estudiado» transcrito como «estudiado», «de España» registrado como «España», y así hasta un total de 21 sinalefas que esta herramienta que procesa el lenguaje natural no ha detallado con acierto.

Aunque las sinalefas forman parte del lenguaje común, corriente y que se puede escuchar día a día en las calles, el índice de acierto ha sido menor que con el uso agramatical de la tercera persona del imperfecto de indicativo del verbo *haber*, más estigmatizado y peor considerado: *ya que no habían [...] videojuegos, en el pasado, [...] habían más turistas*. Estos dos ejemplos corresponden a dos muestras distintas de voluntarios canarios que se han prestado voluntariamente a participar y a enriquecer nuestro proyecto, con un resultado positivo en la transcripción, pues Siri ha entendido a la perfección lo dicho en ambos casos. Un caso parecido se puede apreciar en una de las muestras del castellano en contacto con el gallego, en la que la participante dice textualmente «han venido mucha gente», algo que Siri ha sabido transcribir.



Otro de los rasgos más presentes en los hablantes del gallego y del español es el uso del pretérito perfecto simple vinculado al presente, muy bien reconocido también por Siri: *estéticamente no cambió mucho, el dibujo que siempre se hizo, nunca me dio como una sensación de claustrofobia.*

En la tercera imagen, la del carro, o carrito, de la compra, ha ocurrido un fenómeno que nos permite enlazar con el siguiente plano. No nos puede pasar desapercibido ese detalle de *carro* o *carrito* porque hemos puesto a Siri a prueba con el uso de los diminutivos en *-ito*. De nuevo, ha superado con éxito el reto. Mientras que un total de siete participantes, en su mayoría navarros y madrileños, han optado por *carro de la compra*, han sido diecisiete los que se han decantado por *carrito de la compra*, mayoritariamente andaluces, gallegos y los procedentes de las comunidades donde el español y el catalán están en contacto permanente:



*Ilustración 3. Diferencia porcentual en el uso de *carro* y *carrito*.*

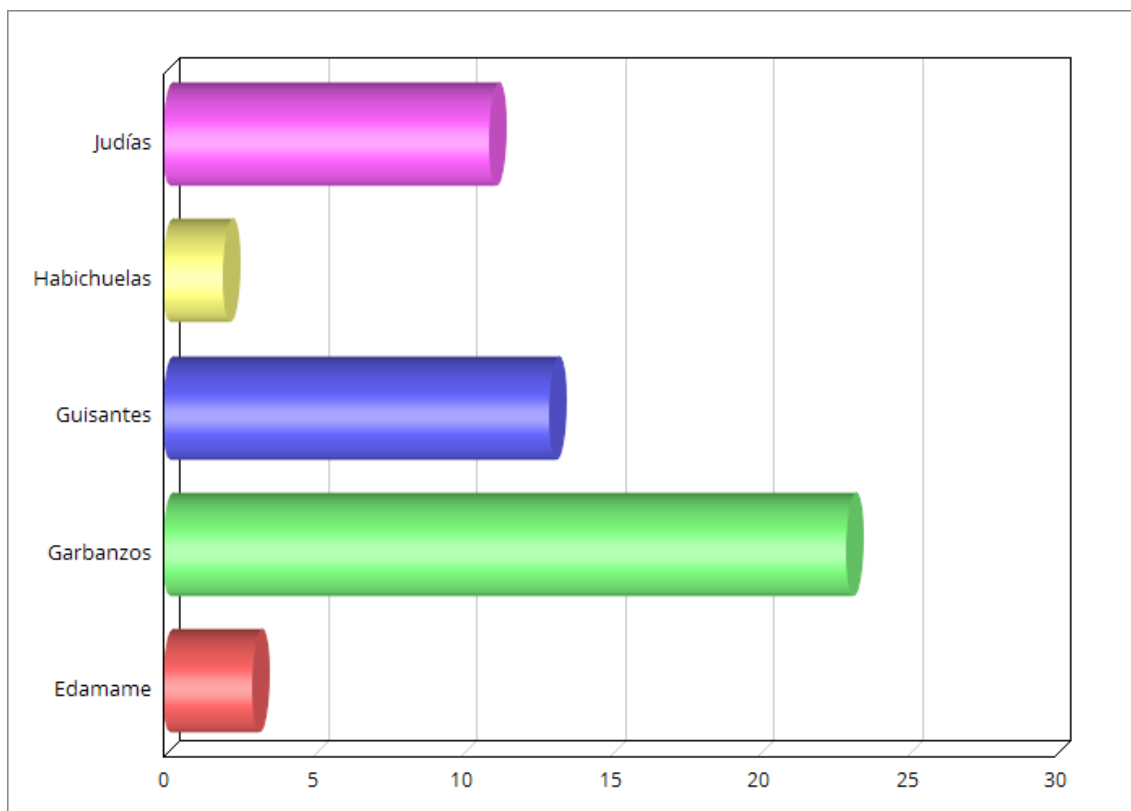
Toca ahora ceñirnos al **plano léxico**, para el que escogimos específicamente las tres imágenes que se ven en la encuesta por el juego que sabíamos que podían dar. Así ha

sido. En la primera imagen, han sido seis los participantes que veían un *lago*, mientras que otros once se referían más bien a un *pantano*. Del mismo modo, en algunos casos, hemos oído que el cazador apuntaba a unos *patos* (ocho veces); al hiperónimo *aves*, que es más general (cuatro ocasiones); y, con menor frecuencia, a unos *pajaritos* o *faisanes* (una vez cada uno).

En la segunda de las ilustraciones, donde observábamos la representación gráfica de la casa idealizada por todo niño pequeño, veíamos un árbol del que colgaban manzanas. En las 42 muestras hemos observado distintas alternativas para ese mismo concepto: *manzano* (en seis muestras distintas); *árbol* (a secas, cuatro ocasiones); *árbol con manzana* o *árbol que tiene manzanas* (tres veces); *árbol frutal* (en dos muestras); y *manzanero*, término propio de las hablas canarias, en una única ocasión.

Por último, el tercer dibujo nos muestra un carro de la compra con productos muy variados. En relación con la comida que aparece dentro del carro de la compra, la riqueza ha quedado de manifiesto también al hacer referencia a las *papas* o *patatas*: en las siete muestras de participantes canarios y en dos de las muestras de andaluces, en concreto, las dos personas sevillanas que han participado, es decir, un total de nueve veces se ha recurrido a *papas*, mientras que en el resto de variedades: el castellano y este mismo en contacto con el catalán, el euskera y el gallego los participantes se han decantado unánimemente por *patatas*.

Por último, con respecto a las verduras y legumbres que aparecen en el carro han sido muchas las alternativas aportadas, como podemos observar en el siguiente gráfico de manera más clara:



*Ilustración 4. Número de veces que se ha repetido cada término para hacer referencia a las legumbres.*

Ahora vamos a proceder a señalar los aspectos más importantes de cada una de las variedades que hemos estudiado durante nuestra investigación. Comenzaremos con las **hablas andaluzas**. Esta es, de entre las seis variedades que hemos estudiado, la variedad con la que Siri presenta los peores resultados.

En primer lugar, casi el 100 % de las veces, los hablantes andaluces pierden la *-s* implosiva, tanto en interior como al final de palabra, algo que dificulta bastante el trabajo de Siri para distinguir los plurales. También es común dentro de esta variedad la pérdida de *-r*, *-d*, *-l* y *-z* al final de palabra, aunque únicamente la pérdida de *-r* ha dado errores en la transcripción realizada por Siri, que no ha sabido identificar este rasgo en cuatro de las nueve ocasiones en las que aparece.

Otro rasgo al que debemos prestar atención en las hablas andaluzas es a la pronunciación de los fonemas /s/ y /θ/. Hemos recogido dos muestras en las que los hablantes cecean, otras dos en las que los hablantes sesean y tres en las que los hablantes distinguen entre un fonema y otro. La efectividad de Siri para reconocer y transcribir correctamente estos rasgos ha sido óptima, tan solo presentando el mismo error en dos ocasiones distintas en las que, en una misma palabra, se han dado tanto seseo como pérdida de *-r* final de palabra («cazador» entendido como «casado»).

Finalmente, otro rasgo típico de esta variedad que también ha presentado un alto porcentaje de error ha sido el acortamiento de palabras como *mu* en lugar de *muy*, por ejemplo. Siri solo ha sido capaz de reconocer y transcribir correctamente este rasgo en el 20 % de las veces en las que aparece.

Después de las hablas andaluzas, pusimos a prueba a Siri con las **hablas canarias**. Los resultados de Siri con esta variedad son tan solo ligeramente más precisos que los resultados de las hablas andaluzas.

Mientras que los hablantes andaluces pierden las *-s* finales de sílaba y de palabra, los hablantes canarios la aspiran. Esto ha provocado que Siri cometa errores en casi el 50 % de las ocasiones, ya que no era capaz de distinguir entre plural y singular. Además de la aspiración de *-s*, los rasgos más distintivos de esta variedad son el seseo, presente en todas las muestras, y la pronunciación de la *che* adherente, presente en más de la mitad de ellas. Siri ha identificado el seseo el 100 % de las ocasiones y tan solo ha cometido un fallo en cuanto a la pronunciación de la *che* adherente.

Además, hemos observado que los hablantes canarios que se han prestado a esta investigación utilizan indistintamente los fonemas palatales /ʎ/ y /j/, pudiendo pronunciar

la palabra *llamada* tanto [ja'maða] como [ʎa'maða]. En lo que sí coinciden todas las muestras es en la pronunciación relajada de la fricativa velar sorda /x/.

A continuación, sometimos a Siri a muestras del **castellano**, con las que mostró una eficacia casi del 100 %, cometiendo muy pocos errores, sobre todo, por sinalefa y por taquilalia, es decir, la rapidez excesiva del habla.

Las personas que hablan esta variedad se caracterizan por el mantenimiento de las consonantes finales, lo cual facilita enormemente el trabajo de Siri, además de por un yeísmo generalizado y por la distinción de los fonemas /s/ y /θ/. Llama la atención que la muestra que más errores ha presentado sea la muestra 01, de una persona de Ciudad Real (Castilla-La Mancha) que lleva cuatro años residiendo en Granada y manteniendo contacto directo con las hablas andaluzas. Este contacto ha influido en su forma de pronunciar algunos sonidos en algunas ocasiones, lo que ha hecho que Siri obtuviese peores resultados al transcribir los fragmentos de esta muestra.

Dicho sea de paso que en las muestras de esta variedad observamos la mayor variación lingüística respecto a los alimentos que encontramos en la tercera imagen. Entre los siete hablantes del castellano que han participado en la investigación han surgido cinco términos distintos para designar el mismo alimento, independientes de la variedad diatópica de cada hablante: edamame, judías, habas, judías verdes y guisantes. Algo que, por supuesto, no le ha causado ningún problema a Siri a la hora de transcribir.

Pasando del castellano a las variedades del español en las comunidades autónomas bilingües, es turno de enfocar nuestra atención en el **castellano en contacto con el catalán**. Es conveniente resaltar que la efectividad de Siri con estas muestras ha sido muy alta, presentando errores tan solo con algunos de los rasgos que vamos a detallar a continuación.

En primer lugar, lo que en otras variedades del español observábamos como el mantenimiento de la *-d*, en esta variedad era más bien el ensordecimiento de la oclusiva alveolar sonora *-d* (*universidad, ciudad*). De la misma manera, hemos detectado tanto algunos casos de yeísmo, como otros de distinción entre los fonemas palatales /*ʎ*/ y /*j*/.

De cualquier modo, Siri ha mostrado su utilidad al haberlos identificado correctamente todas las veces que han aparecido.

Por último, otro elemento distintivo del habla de las personas catalanas y compartido por una amplia mayoría de estas es la reducción de las vocales [e] y [a] átonas a una vocal central: *Farners, Terrassa, Vidreras...* Por lo demás, podemos destacar que, en algunas de las muestras, hemos percibido que los participantes utilizaban el español como su lengua secundaria o subsidiaria, algo normal y muy común en las regiones bilingües. Prueba de esto era el cambio de género para ciertos términos, como *vallo* en lugar de *valla*.

Después del catalán, era turno del **castellano en contacto con el euskera**. Para ello, hemos recurrido tanto a participantes navarros como a participantes vascos. En un primer lugar, Siri ha tenido facilidad para identificar todo lo que decían los participantes, salvo con algunos rasgos muy concretos que han supuesto algún problema más. La elisión de la dental sonora intervocálica *-d-*, acompañada por el cierre de /o/ en /u/, en las terminaciones en *-ado* ha entrañado alguna dificultad para que Siri reconociese lo que se decía: *rodeau*, en lugar de *rodeado*; *estudiau*, en lugar de *estudiado...*

El yeísmo en el español del norte de Navarra y Euskadi se hacía mucho más patente en algunos de los participantes: Siri no ha identificado en una muestra *valla* y, en su lugar, ha interpretado que lo que se decía era *vaya*, palabra que aparecía en las muestras de los hablantes de otras variedades y Siri ha transcrito correctamente. Aparte del yeísmo, esta variedad ha sido, sin duda, con la que Siri ha tenido mayor problema para determinar los

diminutivos en *-ito*, pues no ha distinguido estas desinencias ni en *pueblito*, que lo ha transcrito como *polito*, ni en *vallita*, que lo ha transcrito como *galleta*. Pese a todo lo indicado con anterioridad, el alto rendimiento de Siri ha vuelto a quedar probado en esta variedad.

En última instancia, el **castellano en contacto con el gallego** ha sido la última variedad que hemos explorado en su combinación con Siri. De esta variedad, más allá de lo señalado con anterioridad, como el uso extendido del pretérito perfecto simple vinculado al presente. También ha quedado de manifiesto la influencia del gallego, con algunos galleguismos, como *Rías Baixas* o *a perla das mariñas*. Hemos detectado asimismo un caso de influencia del inglés: *barra de chocolate*, como traducción literal y poco adaptada de *chocolate bar*. También hemos identificado algunos otros diminutivos, como *-ete* o *-illo*: *solete*, *floreциllas*. Alguna construcción, como *llevar + participio*, también ha estado presente en las muestras de los participantes gallegos: *llevo trabajados*.

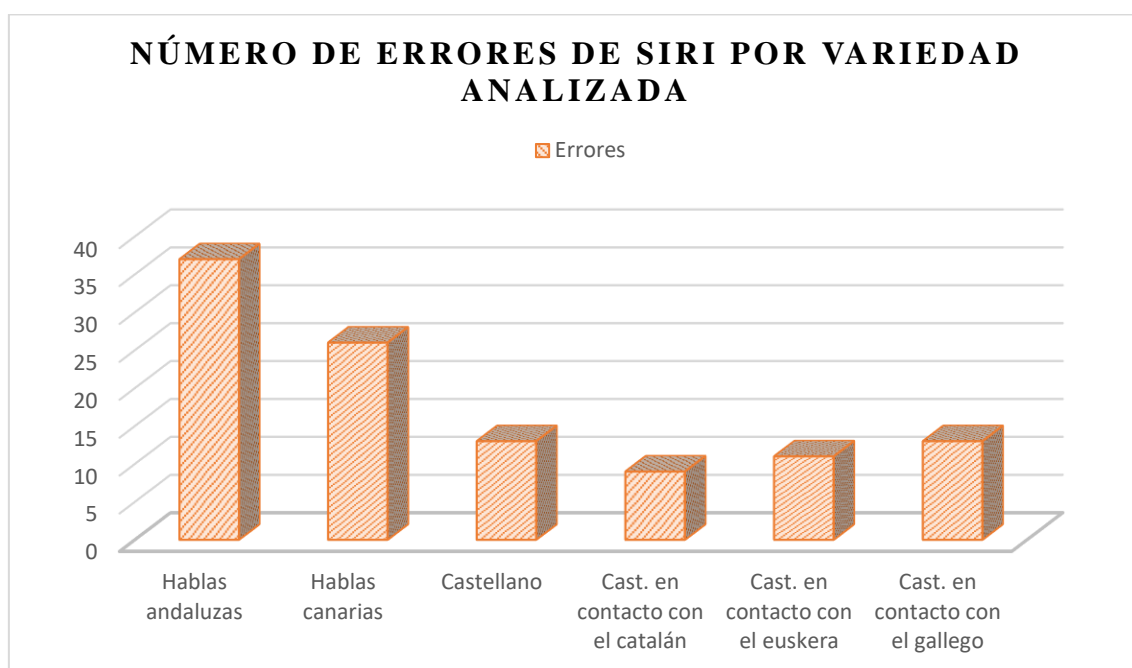
## **8. CONCLUSIÓN**

En este trabajo, hemos pretendido arrojar luz sobre un campo en el que aún queda mucho por ahondar y avanzar. El hueco investigador que existe en esta materia nos ha permitido desarrollar, desde un punto de vista innovador, un estudio que valorase tanto las variedades del español como el procesamiento del lenguaje natural.

La estructura de la encuesta, así como la difusión que le hemos dado, ha dado sus frutos. En un principio, pretendíamos que este fuese un estudio transversal más cualitativo que cuantitativo, si bien la gran participación con la que nos hemos encontrado ha permitido que podamos seguir una metodología mixta (cuantitativa y cualitativa). Se antoja necesaria la publicación de una guía rápida por parte Apple, compañía desarrolladora de esta herramienta, para que el público lego en informática alcance a entender, aunque sea

de manera simple, el uso interno de este recurso. Esto sería una solución a corto plazo. A largo plazo, sería buena idea elaborar una guía de recomendaciones para que mejore la herramienta.

La principal dificultad percibida es la pérdida o aspiración de las consonantes finales, para las que Siri aún debe ser adaptada y actualizada. Al igual que con las sinalefas, Siri omitía parte del contenido o lo reformulaba sin dar con una alternativa acertada.



*Ilustración 5. Número de errores de Siri por variedad analizada.*

Los resultados del estudio son claros y no dejan lugar a dudas: Siri es una gran aplicación para el procesamiento del lenguaje natural, pero debe superar muchas barreras para atender a todas las variedades del español en igualdad y con semejante efectividad. Mientras que las hablas andaluzas y canarias son las que salen peor paradas al dictar a esta herramienta (con una eficacia aceptable, pero mejorable), el castellano (incluido el de las regiones bilingües) cuenta con una eficiencia más que evidente. Dando respuesta así a una de las preguntas que nos formulamos al comenzar la investigación, los resultados



ponen de relieve que el lugar de origen influye positiva o negativamente al utilizar esta herramienta de procesamiento del lenguaje natural.

En conclusión, el uso de Siri es recomendable, sobre todo, para personas con discapacidad visual, dificultad para leer y/o escribir, así como para todos aquellos usuarios que pretenden hacer a viva voz lo que también se puede llevar a cabo a golpe de clic. Las actualizaciones venideras y el desarrollo de una tecnología cada vez más avanzada permitirán dar con soluciones que acaben con los problemas mencionados anteriormente.

## **9. REFERENCIAS BIBLIOGRÁFICAS**

Álvarez-García, M. (2004). Consideración de la modalidad lingüística andaluza en el *Diálogo de la lengua* y en la actualidad. En *Cauce*, 27, pp. 27-34.

[https://idus.us.es/bitstream/handle/11441/31387/Consideraci%  
c3%b3n%20de%20a%20modalidad%20ling%  
c3%bc%c3%adstica%20andaluza.pdf?sequence=1&isAllowed=y](https://idus.us.es/bitstream/handle/11441/31387/Consideraci%c3%b3n%20de%20a%20modalidad%20ling%c3%bc%c3%adstica%20andaluza.pdf?sequence=1&isAllowed=y)

Bachate, R. P., & Sharma, A. (2020). Acquaintance with Natural Language Processing for Building Smart Society. En *E3S Web of Conferences* (Vol. 170, p. 02006). EDP Sciences.

[https://www.e3s-  
conferences.org/articles/e3sconf/pdf/2020/30/e3sconf\\_evf2020\\_02006.pdf](https://www.e3s-conferences.org/articles/e3sconf/pdf/2020/30/e3sconf_evf2020_02006.pdf)

Baeza, B. (1996). Procesamiento del lenguaje natural: Fuentes bibliográficas para su estudio en Lengua Española. En *Procesamiento del Lenguaje Natural*, 18.

<http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/viewFile/3767/2224>

Barrios, K., López, J., Mendieta, S., Benavides, R., & Sáez, Y. (2018). Sistema de reconocimiento de voz: un enlace en la comunicación hombre-máquina. *Revista de*

*Iniciación Científica*, 4, pp. 92-95.

<https://revistas.utp.ac.pa/index.php/ric/article/view/1827/2637>

Blas-Arroyo, J. L. (2004). El español actual en las comunidades del ámbito lingüístico catalán. En *Historia de la lengua española*, pp. 1065-1086. Ariel.

Bock, B., & Shamir, L. (2015). Assessing the efficacy of benchmarks for automatic speech accent recognition. En *Proceedings of the 8th International Conference on Mobile Multimedia Communications*, pp. 133-136.

[https://www.researchgate.net/profile/Lior\\_Shamir/publication/281069316\\_Assessing\\_the\\_efficacy\\_of\\_benchmarks\\_for\\_automatic\\_speech\\_accent\\_recognition/links/55da473108ae9d659491edc0.pdf](https://www.researchgate.net/profile/Lior_Shamir/publication/281069316_Assessing_the_efficacy_of_benchmarks_for_automatic_speech_accent_recognition/links/55da473108ae9d659491edc0.pdf)

Calderón-Campos, M. (2010). Formas de tratamiento. En *La lengua española en América: normas y usos actuales*, pp. 225-236. Valencia: Universitat de València.

<https://www.uv.es/aleza/esp.am.pdf>

Council, J. (2021). Behind Microsoft's Nuance Deal: Natural Language Processing, Explained; From virtual assistants to call center attendants, voice is beginning to rival text as the primary computing interface. En *Wall Street Journal*.

<https://www.proquest.com/newspapers/behind-microsofts-nuance-deal-natural-language/docview/2511961896/se-2?accountid=14542>

Del Moral-Aguilera, R. (2014). *Breve historia de las lenguas*. Castalia.

Echauri-Soto, G. A. (2015). *La interacción de las personas con los asistentes virtuales: Estudio sobre los usos, percepción e impactos del asistente Siri de Apple*.

[http://amic2015.uaq.mx/docs/memorias/GI\\_01\\_PDF/GI\\_01\\_La%20interaccion\\_de\\_las\\_personas.pdf](http://amic2015.uaq.mx/docs/memorias/GI_01_PDF/GI_01_La%20interaccion_de_las_personas.pdf)

- Etxebarría-Aróstegui, M. (2002). *La diversidad de lenguas en España*. Espasa.
- García-Mouton, P. (1994). Vasco o euskera. En *Lenguas y dialectos de España*, pp. 51-53. Arco Libros.
- Herrera-Viedma, E. & O'Valle-Ravassa, F. J. (2021). *Comité de Ética en Investigación de la Universidad de Granada*. Granada, España: Vicerrectorado de Investigación y Transferencia.
- Hilotin, J. (2021, Mar 28). Chatbots: How robots are taking over the world, one chat at a time. En *Gulf News*. <https://www.proquest.com/newspapers/chatbots-how-robots-are-taking-over-world-one/docview/2505967407/se-2?accountid=14542>
- Instituto Nacional de Estadística. (2021). *Cifras oficiales de población resultantes de la revisión del Padrón municipal a 1 de enero. Población por comunidades y ciudades autónomas y tamaño de los municipios*. Madrid: Instituto Nacional de Estadística. <https://www.ine.es/jaxiT3/Tabla.htm?t=2915&L=0>
- Jiménez-Martín, P., & Sánchez-Allende, J. (2015). De Eliza a Siri: La evolución. En *Tecnología y desarrollo*, 13, p. 28. [https://revistas.uax.es/index.php/tec\\_des/article/view/616/572](https://revistas.uax.es/index.php/tec_des/article/view/616/572)
- Johnson, R. G. (2016, May). There and back again—Andrew Booth, a British Computer Pioneer, and his interactions with US and other contemporaries. En *IFIP International Conference on the History of Computing*, pp. 58-70. <https://link.springer.com/content/pdf/10.1007%2F978-3-319-49463-0.pdf>
- Johri, P., Khatri, S. K., Al-Taani, A. T., Sabharwal, M., Suvanov, S., & Kumar, A. (2021). *Natural Language Processing: History, Evolution, Application, and Future Work*.

- En *Proceedings of 3rd International Conference on Computing Informatics and Networks: ICCIN 2020*, pp. 365-375. Springer Singapore.
- Liddy, E. D. (2001). Natural language processing. En *Encyclopedia of Library and Information Science*, 2nd Ed. NY. Marcel Decker, Inc.  
<https://surface.syr.edu/cgi/viewcontent.cgi?article=1043&context=istpub>
- Martínez-Celdrán, E., & Fernández-Planas, A. M. (2007). *Manual de fonética española: Articulaciones y sonidos del español*. Ariel.
- Millstein, F. (2020). *Natural language processing with python: natural language processing using NLTK*. Frank Millstein.
- Moreno-Fernández, F. (2010). *Las variedades de la lengua española y su enseñanza*. Arco Libros.
- Moreno-Fernández, F. (2014). *La lengua española en su geografía: Manual de dialectología hispánica* (2nd ed.). Arco Libros.
- Moreno-Fernández, F. (2020). *Variedades de la lengua española*. Routledge.
- Muñoz-Basols, J. & Lacorte, M. (2017). *Lingüística hispánica actual: guía didáctica y materiales de apoyo*. Routledge.
- Narbona-Jiménez, A., Cano-Aguilar, R., & Morillo-Velarde Pérez, R. (2011). *El español hablado en Andalucía*. Universidad de Sevilla.
- Oñederra-Olaizola, M. L. (2004). El español en contacto con otras lenguas: español-vasco. En *Historia de la lengua española*, pp. 1103-1116. Ariel.
- Pharies, D. A. (2007). *Breve historia de la lengua española*. University of Chicago Press.
- Potowski, K., & Shin, N. L. (2019). *Gramática española: variación social*. Routledge.

Real Academia Española. (2009). *Nueva gramática de la lengua española* (Vol. 2). Espasa Libros.

*Resumen y conclusiones del Informe mobile en España y en el Mundo 2020.* <https://mktefa.ditrendia.es/blog/informe-mobile-2020>

Rivero, F. (2019). Informe ditrendia: Mobile en España y en el Mundo 2018. En *Informe técnico, Ditrendia.* [https://www.amic.media/media/files/file\\_352\\_2531.pdf](https://www.amic.media/media/files/file_352_2531.pdf)

Rivero, F. (2020). Informe ditrendia: Mobile en España y en el Mundo 2020 + Especial COVID-19. En *Informe técnico, Ditrendia.* [https://www.amic.media/media/files/file\\_352\\_2531.pdf](https://www.amic.media/media/files/file_352_2531.pdf)

Rojo-Sánchez, G. (2004). El español de Galicia. En *Historia de la lengua española*, pp. 1087-1102. Ariel.

Rosas-Pérez, M., & Torres-Castillo, J. M. (s. f.). *Modeling Voice Conversation Assistant.* <http://ceur-ws.org/Vol-1784/p2.pdf>

Ryan, K. (1993). The role of natural language in requirements engineering. En *Proceedings of the IEEE International Symposium on Requirements Engineering*, pp. 240-242. IEEE. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=324852>

Saavedra-Montejo, Á. (2020). *Oye Siri, Okey Google, Hey Alexa...: Determinantes de la intención de uso de los asistentes virtuales por voz e influencia de la privacidad percibida.* [https://repositorio.uam.es/bitstream/handle/10486/692318/saavedra\\_montejo\\_alvaro\\_TFM.pdf](https://repositorio.uam.es/bitstream/handle/10486/692318/saavedra_montejo_alvaro_TFM.pdf)

- Tait, J., & Wilks, Y. (2019). Anniversary article: Then and now: 25 years of progress in natural language engineering. En *Natural Language Engineering*, 25(3), pp. 405-418. <https://www.proquest.com/docview/2224981895>
- Tatman, R. (2017). Gender and dialect bias in YouTube's automatic captions. En *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, pp. 53-59. <https://www.aclweb.org/anthology/W17-1606.pdf>
- Torralba, A., & Efros, A. A. (2011). Unbiased look at dataset bias. En *CVPR 2011*, pp. 1521-1528. IEEE. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5995347>
- Woods, H. S. (2018). Asking more of Siri and Alexa: Feminine persona in service of surveillance capitalism. *Critical Studies in Media Communication*, 35(4), p. 334-349.

## **ANEXOS**

### **I. Registro de muestras**

<b>LAS HABLAS ANDALUZAS</b>		
<b>Nombre</b>	<b>Procedencia</b>	<b>Muestra</b>
Marta	Algeciras, Cádiz	1
Andrea	Puente Genil, Córdoba	2
José Ángel	Tabernas, Almería	3
Eduardo	Córdoba	4
Manolo	Paradas, Sevilla	5
Virginia	Paradas, Sevilla	6
María	Baeza, Jaén	7

**LAS HABLAS CANARIAS**

<b>Nombre</b>	<b>Procedencia</b>	<b>Muestra</b>
Daniela	Radazul, Tenerife	8
Esteban	La Orotava, Tenerife	9
Borja	Santa Cruz de Tenerife	10
Cristina	La Orotava, Tenerife	11
Luis	Puerto de la Cruz, Tenerife	12
Natalia	La Orotava, Tenerife	13
Carlos	Puerto de la Cruz, Tenerife	14

**EL CASTELLANO**

<b>Nombre</b>	<b>Procedencia</b>	<b>Muestra</b>
Sofía	Ciudad Real, Castilla-La Mancha	15
Juan	Madrid	16
Esther	Madrid	17
Luis Antonio	Ciudad Real, Castilla-La Mancha	18
Isabel	Madrid	19
Marta	Santander, Cantabria	20
Andrea	Madrid	21

**EL CASTELLANO EN CONTACTO CON CATALÁN**

<b>Nombre</b>	<b>Procedencia</b>	<b>Muestra</b>
Meritxell	Sant Cugat del Vallès, Barcelona, Cataluña	22
Jordi	Sant Cugat del Vallès, Barcelona, Cataluña	23
Júlia	Santañí, Mallorca, Balears	24

Marta	Terrassa, Barcelona, Cataluña	25
Paula	Santa Coloma de Farners, Girona, Cataluña	26
Pau	Vidreres, Girona, Cataluña	27
Helena	Valencia, Comunidad Valenciana	28

#### **EL CASTELLANO EN CONTACTO CON EL EUSKERA**

<b>Nombre</b>	<b>Procedencia</b>	<b>Muestra</b>
Asier	Mutilva, Navarra	29
Miguel	Vitoria-Gasteiz, País Vasco	30
Adrián	Fitero, Navarra	31
Beatriz	Mutilva Alta, Navarra	32
Juanjo	Pamplona, Navarra	33
Javi	Mutilva, Navarra	34
Laura	Guipúzcoa, País Vasco	35

#### **EL CASTELLANO EN CONTACTO CON EL GALLEGO**

<b>Nombre</b>	<b>Procedencia</b>	<b>Muestra</b>
Blanca	La Coruña	36
Iria	Vigo, Pontevedra	37
Rober	Oleiros, La Coruña	38
Paloma	Sada, La Coruña	39
Marta	Monforte de Lemos, Lugo	40
Claudia	Monforte de Lemos, Lugo	41
Sabela	Monforte de Lemos, Lugo	42



## II. Siri y las variedades del español en España

### II.1 MUESTRAS RELACIONADAS CON LAS HABLAS ANDALUZAS

<b>Muestra</b>	1			
<b>Procedencia del hablante</b>	Andaluz (Procedente de Algeciras, Cádiz)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Debilitamiento de la fricativa velar sorda /x/ Pérdida de -s	Me llamo Marta y soy de Algeciras, Cádiz	Me llamo Marta y soy la alegría de Cádiz	Reconoce el yeísmo, el debilitamiento de la fricativa velar sorda /x/ y la pérdida de -z	Error fonético: «Algeciras» entendido como «alegría» por debilitamiento de /x/ y pérdida -s
Pérdida de -d Mantiene la -s Mantiene la -d- intervocálica	Mi ciudad cuando era pequeña pues e.. en realidad e.. era entretenida	Mi ciudad cuando era pequeña en realidad era entretenida	Reconoce la pérdida de -d	
Pérdida de -s y de -d	Porque para los niños de nuestra edad, quedábamos todos en una plaza	Porque para los niños de nuestra edad quedábamos todo en una plaza	Reconoce la pérdida de -d y -s	Error fonético: «Todos» entendido como «todo» por pérdida de -s
Pérdida de -s Pérdida de -d Acortamiento	Pero te lo pasabas bien, pa esa edad, no sé	Pero te lo pasaba bien pasada no sé		Error fonético: «Pasabas» entendido como «pasaba» por pérdida de -s Error fonético-sintáctico: «Pa esa edad»

				entendido como «pasada» por acortamiento y pérdida de <i>-d</i>
Sinalefa	He estudiado Farmacia	Estudiado farmacia		Error fonético-sintáctico: «He estudiado» entendido como «estudiado» por sinalefa
Pérdida de <i>-s</i>	Veo un bosque, con un laguito, hay un cazador con su perro	Pues veo un bosque con un laguito hay un cazador con su perro	Reconoce la pérdida de <i>-s</i>	
Pérdida de <i>-s</i> y de <i>-r</i> Mantiene la <i>-s</i> al final de patos	Está intentando cazar unos patos	Está intentando cazar unos patos	Reconoce la pérdida de <i>-s</i> y <i>-r</i> , así como la <i>-s</i> de «patos»	
Pérdida de <i>-s</i> Preferencia por los diminutivos en <i>-ito</i>	Una casita con su sol, sus arbolitos	Una casita con su sol, su arbolito	Reconoce el uso de los diminutivos en <i>-ito</i>	Error fonético: «Sus arbolitos» entendido como «su arbolito» por pérdida de <i>-s</i>
Pérdida de <i>-s</i>	Con chocolate, plátanos, patatas	Con chocolate, plátano, patatas		Error fonético: «Plátanos» entendido como «plátano» por pérdida de <i>-s</i>

<b>Muestra</b>		2
<b>Procedencia del hablante</b>	Andaluza (Natural de Puente Genil, Córdoba)	
<b>Información social</b>	Estudios universitarios	
<b>Rasgos lingüísticos</b>		

Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Debilitamiento de la fricativa velar sorda /x/ Pérdida de -l	Me llamo Andrea y soy de Puente Genil, un pueblo de Córdoba	Me llamo Andrea y soy de Puente Genil un pueblo de Córdoba	Reconoce el yeísmo, el debilitamiento de la fricativa velar sorda /x/ y la pérdida de -l	
Pérdida de -s y -l Sinalefa	Mi pueblo, cuando era pequeña, era más o menos igual que ahora	En mi pueblo cuando era pequeña era más o menos igual que ahora	Reconoce la pérdida de -s y -l y la sinalefa	
Pérdida de -d	Estoy estudiando Traducción e Interpretación en la Universidad de Granada	Estoy estudiando traducción e interpretación en la Universidad de Granada	Reconoce la pérdida de -d	
Seseo Sinalefa Pérdida de -s y -d	He vivido, un año entero en Portsmouth, que es una ciudad de Inglaterra	He vivido un año entero en Portsmouth que es una ciudad de Inglaterra	Reconoce el seseo y la pérdida de -s y -d	Error fonético-sintáctico: «Que es una» entendido como «que una» por sinalefa
Uso etimológico de los pronombres clíticos Seseo Pérdida de -s	Simplemente, pues los llamaba o hacíamos videollamada	Simplemente pues lo llamaba o hacíamos videollamada	Reconoce el uso etimológico de los pronombres clíticos y el seseo	Error fonético: «Pues los llamaba» entendido como «pues lo llamaba» por pérdida de -s
Pérdida de -s y -r Seseo Debilitamiento de la fricativa velar sorda /x/	Podemos ver como un tipo de bosque y hay un cazador que está sujetando una escopeta,	Podemos ver como un tipo de bosque y hay un casado que está sujetando una escopeta	Reconoce la pérdida de -s y el debilitamiento de la fricativa velar sorda /x/	Error fonético: «Cazador» entendido como «casado» por seseo y pérdida de -s

Seseo Pérdida de -s Variación léxica ( <i>habas, judías, edamames</i> )	Lo que parecen habas o judías o edamames	Lo que parecen habas o judía o edamame	Reconoce el seseo	Error fonético: «Judías o edamames» entendido como «judía o edamame» por pérdida de -s
--	--	--	-------------------	---

<b>Muestra</b>	3			
<b>Procedencia del hablante</b>	Andaluz (Natural de Tabernas, Almería)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Pérdida de -s Sinalefa	Me llamo José Ángel, soy de Tabernas, de un pueblo de Almería, España	Me llamo José ángel soy de taberna de un pueblo Almería España	Reconoce la sinalefa	Error fonético: «Tabernas» entendido como «taberna» por pérdida de -s
Confusión del verbo <i>estar</i> por el verbo <i>ser</i> Sinalefa	En mi pueblo es el único desierto de Europa	En mi pueblo es el único desierto de Europa	Reconoce la sinalefa y la confusión del verbo <i>estar</i> por <i>ser</i>	
Sinalefa Pérdida de -s y -d	Se han rodado una infinidad de películas, sobre todo, de temática del oeste	San rodado una infinidad de películas sobre todo de temática del oeste	Reconoce la pérdida de -s y -d	Error fonético-sintáctico: «Se han rodado» entendido como «san rodado» por sinalefa
Uso etimológico de los pronombres clíticos	Los llamo por teléfono	Lo llamo por teléfono	Reconoce el uso etimológico de los pronombres	Error fonético: «Los llamo» entendido como

Pérdida de -s Yeísmo			clíticos y el yeísmo	«lo llamo» por pérdida de -s
Pronunciación tensa de la fricativa velar sorda /x/ Pérdida de -s y -f Preferencia por los diminutivos en -illo	En la primera imagen vemos a un cazador y a un perrillo cazando perdices	En la primera imagen vemos a un cazador y a un perrillo cazando perdices	Reconoce la pronunciación tensa de la fricativa velar sorda /x/, la pérdida de -s y -f y el diminutivo en -illo	
Pérdida de -s y -c Preferencia por los diminutivos en -ito Variación léxica (patatas, guisantes)	En la tercera imagen es un carrito de la compra con productos necesarios, productos básicos, como son los plátanos, las cebollas, las patatas, los guisantes, los garbanzos...	Y la tercera imagen de un carrito de la compra con un producto necesario producto básico cómo son los plátano la cebolla la patata los guisantes	Reconoce la pérdida de -c y el diminutivo en -ito	Error fonético: «Productos necesarios, productos básicos [...] plátanos, las cebollas, las patatas» entendido como «producto necesario, producto básico [...] plátano, la cebolla, la patata» por pérdida de -s

<b>Muestra</b>	4			
<b>Procedencia del hablante</b>	Andaluz (Natural de Córdoba)			
<b>Información social</b>	Estudios medios - Actualmente trabaja como autónomo			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores

Yeísmo Pérdida de -s	Me llamo Eduardo, soy de Córdoba, España	Me llamo Eduardo soy de Córdoba España	Reconoce el yeísmo y la pérdida de -s	
Seseo Pérdida de -s y -d	Recuerdo, más que la ciudad, mi barrio	recuerdo Mac que la Ciudad mi barrio	Reconoce el seseo y la pérdida de -d	Error fonético: «Más» entendido como «Mac» por pérdida de -s
Pérdida de -s Seseo Sinalefa	Muy pocos coches, muy pocas construcciones, un barrio muy pequeño, muy diferente al que es ahora	Muy poco coche muy pocas construcción de un barrio muy pequeño muy diferente al que es ahora	Reconoce el seseo	Error fonético: «Pocos coches [...] construcciones» entendido como «poco coche [...] construcción de» por pérdida de -s Error fonético- sintáctico: «Al que es ahora» entendido como «al que ahora» por sinalefa
Debilitamiento de la fricativa velar sorda /x/ Pérdida de -z	Mucho juego en la calle y muy feliz	Mucho juego en la calle y muy feliz	Reconoce el debilitamiento de la fricativa velar sorda /x/ y la pérdida de -z	
Pérdida de -s y -p	Vendo gafas para las ópticas	Vendo gafas para la óptica	Reconoce la pérdida de -p	Error fonético: «Las ópticas» entendido como «la óptica» por pérdida de -s
Pérdida de -s Variación léxica (COU)	Estudí en los Hermanos Maristas hasta COU	Estudí Elo hermano Marita hasta Cole		Error léxico: No reconoce el nombre propio del colegio ni

				las siglas 'EGB'
Seseo Pérdida de -s y -r	Aparece un cazador disparándole a unos patos	Aparece un casado disparándole a unos patos	Reconoce la pérdida de -s	Error fonético: «Cazador» entendido como «casado» por seseo y pérdida de -r
Yeísmo Pérdida de -s Debilitamiento de la fricativa velar sorda /x/ Seseo Variación léxica (legumbres, judías)	Con el carrito lleno de legumbres, judías, garbanzos, patatas, plátanos, chocolate...	Con el carrito lleno de legumbres judías garbanzos patatas plátano chocolate	Reconoce el yeísmo, el debilitamiento de la fricativa velar sorda /x/ y el seseo	Error fonético: «Plátanos» entendido como «plátano» por pérdida de -s

<b>Muestra</b>	5			
<b>Procedencia del hablante</b>	Andaluz (Natural de Paradas, Sevilla)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Pérdida de -s Distinción entre los fonemas /s/ y /θ/	Me llamo Manolo, soy de Paradas, un pueblo de Sevilla, España	Me llamo Manolo soy de parada un pueblo de Sevilla España	Reconoce el yeísmo	Error fonético: «Paradas» entendido como «parada» por pérdida de -s
Acortamiento	Mi pueblo era un pueblo muy tranquilo con <i>mu</i> poco tráfico	Mi pueblo era un pueblo muy tranquilo con un poco de tráfico		Error fonético-sintáctico: « <i>Mu</i> poco tráfico» entendido como «un poco de

				tráfico» por acortamiento
Pérdida de -s Ceceo Mantiene la -r Acortamiento Variación léxica ( <i>adoquinar</i> )	Algunas calles estaban sin adoquinar todavía, pero era <i>mu</i> bonito	Alguna calle está sin adoquinar todavía pero era muy bonito	Reconoce el ceceo y el acortamiento de <i>muy</i> por <i>mu</i>	Error fonético: «Algunas calles» entendido como «alguna calle» por pérdida de -s
Pérdida de -s y -d Debilitamiento de la fricativa velar sorda /x/ Ceceo	Ahora estoy trabajando en una entidad bancaria, en CaixaBank	Ahora estoy trabajando en una entidad bancaria en CaixaBank	Reconoce la pérdida de -s y -d, el debilitamiento de la fricativa velar sorda /x/ y el ceceo	
Pérdida de -s Ceceo Acortamiento Debilitamiento de la fricativa velar sorda /x/	La manera de comunicarme con mis familiares estando fuera, a pesar de que estaba <i>mu</i> cerca porque eran 50 kilómetros <i>na</i> más, pues era por el teléfono fijo	La manera de comunicarme con mis familiares estando fuera a pesar de que estábamos <i>muy</i> cerca porque eran 50 km mamá fuera por el teléfono fijo	Reconoce la pérdida de -s, el ceceo y el debilitamiento de la fricativa velar sorda /x/	Error fonético-sintáctico: Omisión de <i>muy</i> por acortamiento y « <i>na</i> más, pues era» entendido como «mamá fuera» por acortamiento
Ceceo Debilitamiento de la fricativa velar sorda /x/ Pérdida de -s y -l Sinalefa	En la segunda imagen, veo una casa de campo, con un jardín, árboles y el sol	En la segunda imagen veo una casa de campo con un jardín árboles y <i>eso</i>	Reconoce el ceceo, el debilitamiento de la fricativa velar sorda /x/ y la pérdida de -s	Error fonético-sintáctico: «El sol» entendido como « <i>eso</i> » por pérdida de -l
Ceceo Pérdida de la -d- intervocálica Pérdida de -s Variación léxica ( <i>papas, guisantes</i> )	En la tercera imagen, un carrito de un supermercado que contiene guisantes, chocolate, garbanzos, papas y plátanos	En la tercera imagen un carrito de un supermercado que contiene guisante chocolate garbanzo papa y plátano	Reconoce el ceceo y la pérdida de la -d- intervocálica	Error fonético: «Guisantes, chocolate, garbanzos, papas y plátanos» entendido como «guisante, chocolate,



				garbanzo, papa y plátano» por pérdida de -s
--	--	--	--	---

<b>Muestra</b>	6			
<b>Procedencia del hablante</b>	Andaluzá (Natural de Paradas, Sevilla)			
<b>Información social</b>	Estudios medios			
<b>Rasgos lingüísticos</b>				
<b>Rasgos</b>	<b>Ejemplo en la grabación</b>	<b>Transcripción de Siri (aciertos / errores)</b>	<b>Tipos de aciertos</b>	<b>Tipos de errores</b>
Yeísmo Ceceo Pérdida de -s	Me llamo Virginia, soy de Paradas, un pueblo de Sevilla	Me llamo Virginia soy de parada un pueblo de Sevilla	Reconoce el yeísmo y el ceceo	Error fonético: «Paradas» entendido como «parada» por pérdida de -s
Pérdida de -s y -d Debilitamiento de la fricativa velar sorda /x/	Mi ciudad era muy diferente a como es ahora y, bueno, jugaba con mis vecinas	Mi ciudad era muy diferente a como es ahora y bueno jugaba con mi vecina	Reconoce la pérdida de -d y el debilitamiento de la fricativa velar sorda /x/	Error fonético: «Mis vecinas» entendido como «mi vecina» por pérdida de -s
Pérdida de -s Debilitamiento de la fricativa velar sorda /x/	Estudio un grado de Moda, de Patronaje	Estudio 1º de moda de patronaje	Reconoce la pérdida de -s y el debilitamiento de la fricativa velar sorda /x/	Error: Interpreta «un grado» como «1º»
Ceceo	Y he vivido en Sevilla un año	Y en vivido en Sevilla un año	Reconoce el ceceo	Error fonético-sintáctico: «He vivido» entendido como «en vivido»
Debilitamiento de la fricativa velar sorda /x/	En la primera imagen, vemos un hombre cazando	En la primera imagen vemos un hombre cazando	Reconoce el debilitamiento de la fricativa velar	

Pérdida de -s			sorda /x/ y la pérdida de -s	
Ceceo Debilitamiento de la fricativa velar sorda /x/ Pérdida de -s	En la segunda imagen, vemos una casa con sus árboles	En la segunda imagen de una casa con su árboles	Reconoce el ceceo y el debilitamiento de la fricativa velar sorda /x/	Error fonético: «Sus árboles» entendido como «su árboles» por pérdida de -s
Debilitamiento de la fricativa velar sorda /x/ Pérdida de -s Variación léxica ( <i>papas</i> )	En la tercera imagen, vemos un carro con comida: chocolate, plátanos, papas...	En la tercera imagen de un carro con comida chocolate plátano papa	Reconoce el ceceo y el debilitamiento de la fricativa velar sorda /x/	Error fonético: «Plátanos, papas» entendido como «plátano, papa» por pérdida de -s

<b>Muestra</b>	7			
<b>Procedencia del hablante</b>	Andaluza (Natural de Baeza, Jaén)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Pronunciación tensa de la fricativa velar sorda /x/ Pérdida de -s	Me llamo María. Soy de Baeza, un pueblo de Jaén, España	Me llamo Maria soy de Baeza un pueblo de Jaén España	Reconoce el yeísmo, la pronunciación tensa de la fricativa velar sorda /x/ y la pérdida de -s	
Pérdida de -s y -d	Estoy estudiando en la Universidad de Granada Administración y Dirección de Empresas	Estoy estudiando en la universidad de Granada administración y dirección de empresas	Reconoce la pérdida de -s y -d	

Mantiene la <i>-d-</i> intervocálica Pérdida de <i>-s</i> y <i>-c</i>	Me he comunicado con mi familia a través de llamadas, videollamadas... para estar en contacto con ellos	Me he comunicado con mi familia a través de llamada vídeo llamada para estar en contacto con ellos	Reconoce la pérdida de <i>-c</i>	Error fonético: «Llamadas, videollamadas» entendido como «llamada, vídeo llamada» por pérdida de <i>-s</i>
Pérdida de <i>-r</i>	Supongo que también será perro cazador	Supongo que también será perro cazador	Reconoce la pérdida de <i>-r</i>	
Pérdida de <i>-s</i> y <i>-d-</i> intervocálica Uso agramatical del adverbio <i>donde</i> Variación léxica ( <i>vainas de guisantes, patatas, garbanzos</i> )	Veo un carrito de la compra, donde se ha comprado chocolate, plátanos, como vainas de guisantes, lo otro patatas y garbanzos	Veo un carrito de la compra dónde se ha comprado chocolate plátano como vainas de guisantes lo otro patata y garbanzo	Reconoce el uso agramatical del adverbio <i>donde</i>	Error fonético: «Plátanos [...] patatas y garbanzos» entendido como «plátano [...] patata y garbanzo» por pérdida de <i>-s</i>

## II.2 MUESTRAS RELACIONADAS CON LAS HABLAS CANARIAS

<b>Muestra</b>	8			
<b>Procedencia del hablante</b>	Canario (Natural de Santa Cruz de Tenerife)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Aspiración de <i>-s</i>	Me llamo Daniela, soy de Tenerife, de Canarias	Me llamo Daniela, soy de Tenerife de Canaria		Error fonético: «Canarias» entendido como «Canaria» por aspiración de <i>-s</i>

Seseo Pérdida de <i>-d</i>	Vivo en un pueblo que se llama Radazul, que es una ciudad dormitorio	Vivo en un pueblo que se llama a Radazul que es una ciudad dormitorio	Reconoce el seseo y la pérdida de <i>-d</i>	Error sintáctico: «Se llama Radazul» entendido como «se llama Radazul»
Yeísmo Aspiración de <i>-s</i>	Mi familia se comunica conmigo a través de videollamadas	Mi familia se comunica conmigo a través de videollamada	Reconoce el yeísmo	Error fonético: «Videollamadas» entendido como «videollamada» por aspiración de <i>-s</i>
Variación léxica ( <i>gorriones, bosque, lago</i> )	Se ve un cazador disparando a gorriones con un perro marrón en el bosque y un lago al fondo	Se ve un cazador disparando a gorriones con un perro marrón en el bosque y un lado al fondo		Error fonético: «Lago» entendido como «lado»
Aspiración de <i>-s</i> Preferencia por los diminutivos en <i>-ito</i> Variación léxica ( <i>guisantes, papas</i> )	Se ve un carrito de supermercado con unas imágenes digitalizadas en el interior, que contienen guisantes, plátanos, chocolate y papas	S.B. un carrito de supermercado con unas imágenes digitalizado en el interior que contienen guisantes plátano chocolate y papa	Reconoce el uso de los diminutivos en <i>-ito</i>	Error fonético: «Plátanos y papas» entendido como «plátano y papa» por aspiración de <i>-s</i>

<b>Muestra</b>	9			
<b>Procedencia del hablante</b>	Canario (Natural de La Orotava, Tenerife)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores

Aspiración de -s	Hola, soy Esteban, de La Orotava, de un pequeño pueblo del norte de Tenerife	Hola soy Esteban trataba de un pequeño pueblo del norte de Tenerife		Error fonético: «La Orotava» entendido como «trataba» por taquilalia
Yeísmo Aspiración de -s Variación léxica ( <i>habituarse</i> )	Muchos turistas de diferentes lugares paseando por las calles en grupos guiados por una persona	Mucho turista de diferentes lugares paseando por las calles en grupo guiado por una persona	Reconoce el yeísmo	Error fonético: «Muchos turistas» y «grupos guiados» entendido como «mucho turista» y «grupo guiado» por aspiración de -s
Pluralización del verbo impersonal <i>haber</i> Aspiración de -s Yeísmo Seseo	También se podían ver a muchos niños jugando en todas las plazas, ya que no habían pues videojuegos que los mantuvieran metidos a ellos en sus casas	También se podían ver muchos niños jugando en toda la plaza ya que no habían pues videojuego que lo mantuviera metido en su casa	Reconoce el seseo, la pluralización del verbo impersonal <i>haber</i> y el yeísmo	Error fonético: «Todas las plazas», «videojuegos» y «los mantuvieran metidos en sus casas» entendido como «toda la plaza», «videojuego» y «lo mantuviera metido en su casa» por aspiración de -s
Debilitamiento de la fricativa velar sorda /x/ Aspiración de -s Sinalefa	Además he hecho un máster de Marketing Digital	Tema hecho el máster de la digital	Reconoce el debilitamiento de la fricativa velar sorda /x/	Error fonético-sintáctico: «He hecho» entendido como «hecho» por sinalefa Error léxico: El término «marketing» no es entendido

Seseo Aspiración de -s Variación léxica ( <i>papas, garbanzo</i> )	Como pueden ser plátanos, chocolate, papas y creo que algún garbanzo	Cómo pueden ser el plátano de chocolate papada y creo que algún garbanzo	Reconoce el seseo	Error fonético: «Plátanos» y «papas» entendido como «plátano» y «papada» por aspiración de -s
---	--	--	-------------------	--

<b>Muestra</b>	10			
<b>Procedencia del hablante</b>	Canario (Natural de Santa Cruz de Tenerife)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Debilitamiento de la fricativa velar sorda /x/ Aspiración de -s	Me llamo Borja, soy de Santa Cruz de Tenerife, en las islas Canarias, España	Me llamo Borja soy de Santa Cruz de Tenerife en las islas Canarias España	Reconoce el yeísmo, el debilitamiento de la fricativa velar sorda /x/ y la aspiración de -s	
Seseo Sinalefa Aspiración de -s Pérdida de -d y -c	Lo cierto es que Santa Cruz de Tenerife, que es la ciudad capital de la isla de Tenerife y capital de provincia, no ha cambiado especialmente respecto a la actualidad	Lo cierto es que Santa Cruz de Tenerife que es la ciudad capital de la isla de Tenerife y capital de provincia no ha cambiado especialmente a la actualidad	Reconoce el seseo, la aspiración de -s y la pérdida de -d y -c	Error fonético-sintáctico: «Que es la» entendido como «que la» por sinalefa
Pronunciación de la <i>che</i> adherente	Solo lo ha hecho de manera ligera	Solo lo ha hecho de manera ligera	Reconoce la pronunciación de la <i>che</i> adherente y el debilitamiento de la fricativa velar	

Debilitamiento de la fricativa velar sorda /x/			sorda /x/	
Seseo Aspiración de -s Yeísmo	Hace 7 años ya	Hace siete años ella	Reconoce el seseo y la aspiración de -s	Error fonético: «Ya» entendido como «ella» por yeísmo
Aspiración de -s Mantiene la -r Seseo Variación léxica ( <i>escopeta, cartuchos, doble cañón</i> )	En la que podemos observar a un cazador con una escopeta de cartuchos de doble cañón	En la que podemos observar a un cazador con una escopeta de cartucho de doble cañón	Reconoce el seseo	Error fonético: «Cartuchos» entendido como «cartucho» por pérdida de -s
Aspiración de -s Debilitamiento de la fricativa velar sorda /x/ Seseo	La casa es estereotípica de tejado de tejas, ventanas y una puerta y parece que es pues muy bonita	La casa es estéreo típica de tejado de teja ventana y una puerta y parece que es pues muy bonita	Reconoce el seseo y el debilitamiento de la fricativa velar sorda /x/	Error fonético: «Tejas, ventanas» entendido como «teja, ventana» por pérdida de -s
Aspiración de -s Seseo Yeísmo Variación léxica ( <i>papas, habichuelas</i> )	Los productos que tienen incluidos son chocolate, plátanos, parecen avellanas, habichuelas y papas	Los productos que tienen incluido son chocolate plátano parecen avellana habichuela y papa	Reconoce el seseo y el yeísmo	Error fonético: «Incluidos», «plátanos» y «avellanas, habichuelas y papas» entendido como «incluido», «plátano» y «avellana, habichuela y papa» por pérdida de -s

<b>Muestra</b>	11
----------------	----

<b>Procedencia del hablante</b>	Canaria (Natural de La Orotava, Tenerife)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Aspiración de -s Pérdida de -z y -d Seseo Yeísmo	Me llamo Cristina Fernández y soy de La Orotava, una ciudad de Tenerife	Me llamo Cristina Fernández y soy de La Orotava una ciudad de Tenerife	Reconoce el seseo, el yeísmo, la aspiración de -s y la pérdida -z y -d	
Aspiración de -s Pérdida de -d Yeísmo	Tenemos fiestas que son como un sello de identidad para todos los que nos hemos criado en La Orotava y lo vivimos con muchísimo orgullo	Tenemos fiesta que son como un sello de identidad para todo lo que nos hemos criado en la Orotava y lo vivimos con muchísimo orgullo	Reconoce el yeísmo y la pérdida de -d	Error fonético: «Fiestas» y «todos los» entendido como «fiesta» y «todo lo» por aspiración de -d
Mantiene la -d- intervocálica Aspiración de -s Pérdida de -d Debilitamiento de la fricativa velar sorda /x/ Uso de anglicismos: <i>online</i> Variación léxica ( <i>cursar</i> )	Soy estudiante de la Universidad <i>Online</i> de La Rioja y curso el máster de Formación del Profesorado	Soy estudiante de la universidad online de La Rioja y curso el máster de formación del profesorado	Reconoce la aspiración de -s, la pérdida -d y el debilitamiento de la fricativa velar sorda /x/	
Aspiración de -s Mantiene la -r Debilitamiento de la fricativa velar sorda /x/ Seseo	Podemos ver a un joven vestido de militar junto a su perro. El joven está apuntando con un rifle hacia el cielo	Podemos ver a un joven vestido de militar junto a su perro el joven está apuntando con un rifle hacia el cielo	Reconoce la aspiración de -s, el debilitamiento de la fricativa velar sorda /x/ y el seseo	



Variación léxica ( <i>rifle</i> )				
Seseo Debilitamiento de la fricativa velar sorda /x/ Aspiración de -s Pérdida de -d	Parece un dibujo pintado con la ingenuidad de un niño, pues por la línea y el trazo	Parece un dibujo pintado con la ingenuidad de un niño pues por la línea y el trazo	Reconoce el seseo, el debilitamiento de la fricativa velar sorda /x/, la aspiración de -s y la pérdida de -d	
Pronunciación tensa de la fricativa velar sorda /x/ ( <i>imagen</i> ) Debilitamiento de la fricativa velar sorda /x/ ( <i>rojo</i> ) Seseo Mantiene la -r Aspiración de -s Variación léxica ( <i>habichuelas, papas</i> )	La última imagen representa un carro de la compra azul y rojo, en el interior podemos observar varios alimentos, como tres habichuelas, unos garbanzos, unos plátanos, unas papas y una tableta de chocolate	La última imagen representa un carro de la compra azul y rojo en el interior podemos observar varios alimentos como tres habichuela uno garbanzo uno plátano una papa y una tableta de chocolate	Reconoce la pronunciación tensa y el debilitamiento de la fricativa velar sorda /x/ y el seseo	Error fonético: «Habichuelas, unos garbanzos, unos plátanos, unas papas» entendido como «habichuela, uno garbanzo, uno plátano, una papa» por aspiración de -s

<b>Muestra</b>	12			
<b>Procedencia del hablante</b>	Canario (Natural de El Puerto de la Cruz, Tenerife)			
<b>Información social</b>	Estudios medios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Seseo Aspiración de -s Pérdida de -d	Me llamo Luis, soy del Puerto de la Cruz, una pequeña ciudad situada en el norte de	Me llamo Luis soy del Puerto de la cruz una pequeña ciudad situada en el norte de Tenerife en las Islas Canarias	Reconoce el yeísmo, el seseo, la aspiración de -s y la pérdida de -d	

	Tenerife, en las Islas Canarias			
Aspiración de -s Mantiene la -d- intervocálica Seseo Pluralización del verbo impersonal <i>haber</i>	En el pasado se hacían más eventos, <i>habían</i> más turistas	En el pasado se hacían más eventos habían más turistas	Reconoce la aspiración de -s, el seseo y la pluralización del verbo impersonal <i>haber</i>	
Aspiración de -s Acortamiento	Ahora mismo, muchos turistas se van <i>pal</i> sur	Ahora mismo muchos turistas se van para el sur	Reconoce la aspiración de -s y el acortamiento de <i>para el</i> por <i>pal</i>	
Aspiración de -s Mantiene la -r Debilitamiento de la fricativa velar sorda /x/	Estudio en un colegio británico y estoy en segundo de bachiller	Estudio en un colegio británico y estoy en segundo de bachiller	Reconoce la aspiración de -s y el debilitamiento de la fricativa velar sorda /x/	
Debilitamiento de la fricativa velar sorda /x/ Seseo Mantiene la -r	En la primera imagen, veo a un cazador con su perro	En la primera imagen veo a un cazador con su perro	Reconoce el seseo y el debilitamiento de la fricativa velar sorda /x/	
Debilitamiento de la fricativa velar sorda /x/ Aspiración de -s Variación léxica ( <i>manzaneros</i> )	Veo una casa, con su jardín. El jardín tiene flores, un arbusto y dos manzaneros	Veo una casa con su jardín el jardín tiene flores un arbusto y dos Manzanero	Reconoce el debilitamiento de la fricativa velar sorda /x/	Error fonético: «Manzaneros» entendido como «manzanero» por aspiración de -s
Aspiración de -s Pronunciación de la <i>che</i> adherente Variación léxica ( <i>papas</i> )	Tiene varios alimentos dentro, como fruta, frutos secos, verduras, papas y chocolate	Tiene varios alimentos dentro como fruta frutos secos verdura papá y chocolate	Reconoce la pronunciación de la <i>che</i> adherente	Error fonético: «Verduras, papas y chocolate» entendido como «verdura, papa y chocolate» por aspiración

				de -s
--	--	--	--	-------

<b>Muestra</b>	13			
<b>Procedencia del hablante</b>	Canaria (Natural de La Orotava, Tenerife)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Sinalefa Aspiración de -s	Me llamo Natalia, soy de La Orotava, que es un pueblo situado en el norte de Tenerife	Me llamo Natalia soy de la Orotava que es un pueblo situado en el norte de Tenerife	Reconoce la sinalefa y la aspiración de -s	
Pronunciación de la <i>che</i> adherente Aspiración de -s	Está hecho sobre cuestras, porque forma parte de un valle	Está y yo sobre cuestra porque forma parte un valle		Error fonético: «Hecho» entendido como «y yo» por pronunciación de la <i>che</i> adherente «Cuestras» entendido como «cuestra» por aspiración de -s
Debilitamiento de la fricativa velar sorda /x/ Aspiración de -s Pérdida de -n Seseo	El sentir de la gente, las tradiciones siguen siendo las mismas	El sentir de la gente las tradiciones sigue siendo la misma	Reconoce el debilitamiento de la fricativa velar sorda /x/	Error fonético: «Siguen» y «las mismas» entendido como «sigue» y «la misma» por aspiración de -s y pérdida de -n
Seseo Aspiración de -s	Yo soy canaria, pero vivo en	Yo soy canaria pero vivo en Andalucía	Reconoce el seseo y la aspiración de -	

	Andalucía desde hace aproximadamente 4 años	desde hace aproximadamente cuatro años	s	
Mantiene la -d- intervocálica Aspiración de -s	En un campo con un lago justo al lado, rodeados de vegetación	En un campo con un lago justo al lado rodeado de vegetación		Error fonético: «Rodeados» entendido como «rodeado» por aspiración de -s
Aspiración de -s Variación léxica ( <i>manilla de plátanos</i> )	Vemos una manilla de plátanos	Vemos una manilla de plátano		Error fonético: «Plátanos» entendido como «plátano» por aspiración de -s
Aspiración de -s Variación léxica ( <i>papas</i> )	Hay unas tres papas	Hay una tres papa		Error fonético: «Unas tres papas» entendido como «una tres papa» por aspiración de -s

<b>Muestra</b>	14			
<b>Procedencia del hablante</b>	Canario (Natural de El Puerto de La Cruz, Tenerife)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Seseo Aspiración de -s	Soy del Puerto de la Cruz, en la provincia de Santa Cruz de Tenerife	Soy del Puerto de la cruz en la provincia de Santa Cruz de Tenerife	Reconoce el seseo y la aspiración de -s	

Seseo Aspiración de -s Preferencia por los diminutivos en <i>-ito</i>	Es una ciudad pues costera, pequeñita	Es una ciudad pues costera <b>pequeñita</b>	Reconoce el seseo, la aspiración de -s y el uso de los diminutivos en <i>-ito</i>	
Aspiración de -s Pérdida de -d Debilitamiento de la fricativa velar sorda /x/	En la actualidad estudio y trabajo, compagino las dos	En la actualidad estudio y trabajo compagino las dos	Reconoce el debilitamiento de la fricativa velar sorda /x/, la aspiración de -s y la pérdida de -d	
Aspiración de -s Pérdida de -d	La universidad en la que estudié fue en Queen Margaret University, en Edimburgo	La universidad en la que estudié fue en Queen Margaret University en Edimburgo	Reconoce la aspiración de -s y la pérdida de -d	
Sinalefa Aspiración de -s	Una casa así un poco rústica	Una casa <b>siun</b> poco rústica	Reconoce la aspiración de -s	Error fonético-sintáctico: «Así un» entendido como «siun» por sinalefa
Pronunciación de la <i>che</i> adherente Aspiración de -s Variación léxica ( <i>guisantes, papas</i> )	Un carrito de la compra con garbanzos, chocolate, guisantes, papas y plátanos	Un carrito de la compra con garbanzos chocolate <b>guisante</b> <b>papa</b> y <b>plátano</b>	Reconoce la pronunciación de la <i>che</i> adherente	Error fonético: «Guisantes, papas y plátanos» entendido como «guisante, papa y plátano» por aspiración de -s

### II.3 MUESTRAS RELACIONADAS CON EL CASTELLANO

<b>Muestra</b>		15
<b>Procedencia del hablante</b>	Castellano-manchega (Natural de Ciudad Real)	
<b>Información social</b>	Estudios universitarios	

Rasgos lingüísticos				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Distinción entre los fonemas /s/ y /θ/ Yeísmo Mantiene la -s y -d	Me llamo Sofía y soy de Ciudad Real, Castilla-La Mancha, España	Me llamo Sofía y soy de Ciudad Real Castilla-La Mancha España	Reconoce el yeísmo	
Preferencia por los diminutivos en -ito Mantiene la -d- intervocálica Pérdida de -s	Está ya casi pegada al pueblecito que tenemos al lado	Estaría casi pegada al pobrecito que tenemos al lado	Reconoce la pérdida de -s y el uso de los diminutivos en -ito	Error léxico: «Pueblecito» entendido como «pobrecito»
Mantiene la -z Pérdida de -s Yeísmo Uso etimológico de los pronombre clíticos	Por teléfono principalmente, los llamo una vez a la semana mas o menos	Por teléfono principalmente lo llamo una vez a la semana más o menos	Reconoce el yeísmo y el uso de los pronombres clíticos	Error fonético: «Los llamo» entendido como «lo llamo» por pérdida de -s
Distinción entre los fonemas /s/ y /θ/ Mantiene la -s Omisión de la preposición a	Apuntando unos pájaros que están volando por el cielo	Apuntando unos pájaros que están volando por el cielo	Reconoce la totalidad del fragmento	
Mantiene la -s Yeísmo Preferencia por los diminutivos en -ito Variación léxica (valla)	Hay florecitas, una valla y un par de árboles	Hay florecitas, una valla y un par de árboles	Reconoce el yeísmo y el uso de los diminutivos en -ito	
Pronunciación tensa de la fricativa velar sorda /x/ Pérdida de -s	Y en la imagen tres veo un carro de la compra	Y en la imagen de un carro de la compra	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	Error fonético: «Tres» entendido como «de» por pérdida de -s

Pérdida de -s Variación léxica ( <i>edamame, patata</i> )	Tiene varios alimentos: edamame, chocolate, garbanzos, unas patatas y tres plátanos	Que tiene varios alimentos edamame chocolate garbanzos en la patatas y tres plátanos		Error fonético: «Unas patatas» entendido como «en la patata» por pérdida de -s
---	---	--	--	---

<b>Muestra</b>	16			
<b>Procedencia del hablante</b>	Nació en el Puerto de Santa María (Cádiz) pero ha vivido la mayor parte de su vida en Madrid. También ha vivido 3 años en Italia y otros 3 en EE.UU.			
<b>Información social</b>	Estudios universitarios. Actualmente trabaja en el Banco Santander.			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -s Pérdida de -d Sinalefa	Me llamo Juan y soy de Madrid de España	Me llamo Juan y soy de Madrid de España	Reconoce el yeísmo, la pronunciación tensa de la fricativa velar sorda /x/	Error fonético-sintáctico: «De España» entendido como «España» por sinalefa
Mantiene la -s y -d Distinción entre los fonemas /s/ y /θ/	Siempre ha sido una ciudad con mucho atasco, con mucho movimiento	Siempre ha sido una ciudad con mucho atasco con mucho movimiento	Reconoce la totalidad del fragmento	
Mantiene la -s y -r Pronunciación tensa de la fricativa velar sorda /x/	Actualmente estoy trabajando en el Banco Santander	Actualmente estoy trabajando en el banco Santander	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Sinalefa Mantiene la -s y -d	He estudiado Administración y Dirección de empresas en inglés	Estudiado administración y dirección de empresas en inglés		Error fonético-sintáctico: «He estudiado»

Mantiene la <i>-d-</i> intervocálica				entendido como «estudiado» por sinalefa
Mantiene la <i>-s</i> Variación léxica ( <i>dando a los patos</i> )	Un hombre cazando con un perro, dando a los patos	Un hombre cazando con un perro dando a los patos	Reconoce la totalidad del fragmento	
Mantiene la <i>-s</i> Pronunciación tensa de la fricativa velar sorda /x/ Variación léxica ( <i>judías</i> )	Con chocolate, avellanas (que soy alérgico), un plátano, patatas y judías	Con chocolate avellanas que soy alérgico un plátano patatas y judías	Reconoce la totalidad del fragmento	

<b>Muestra</b>	17			
<b>Procedencia del hablante</b>	Madrileña (Natural de Madrid)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la <i>-s</i> Pérdida de <i>-d</i>	Me llamo Esther y soy de Madrid	Me llamo Esther y soy de Madrid	Reconoce el yeísmo y la pérdida de <i>-d</i>	
Uso redundante del verbo <i>acabar</i> Distinción entre los fonemas /s/ y /θ/ Pérdida de <i>-d</i>	Acabo de acabar la carrera de Traducción e Interpretación en la Universidad Complutense de Madrid	Acabo de acabar la carrera de traducción e interpretación en la Universidad Complutense de Madrid	Reconoce el uso redundante del verbo <i>acabar</i> , así como la pérdida de <i>-d</i>	
Distinción entre los fonemas /s/ y /θ/ Leísmo	Les ayudo con los deberes, también pues para que	Les ayudo con los deberes también pues para que estén más	Reconoce el leísmo y la sinalefa	



Mantiene la <i>-s</i> y <i>-r</i> Sinalefa	estén más concentrados a la hora de estudiar	concentrados a la hora de estudiar		
Mantiene la <i>-d-</i> intervocálica Mantiene la <i>-s</i> y <i>-r</i> Pronunciación tensa de la fricativa velar sorda /x/	El año pasado viví fuera un año en Bruselas, en Bélgica para hacer el Erasmus	El año pasado viví fuera un año en Bruselas en Bélgica para hacer el Erasmus	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la <i>-s</i> y <i>-r</i> Pronunciación tensa de la fricativa velar sorda /x/ Variación léxica ( <i>pájaros</i> )	Es un cazador que está intentando matar pájaros	Es un cazador que está intentando matar pájaros	Reconoce la totalidad del fragmento	
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la <i>-s</i> Preferencia por los diminutivos en <i>-ito</i> Variación léxica ( <i>cerca</i> )	La típica casa que dibujan los niños, pues con el sol, los arbolitos, unas manzanas y pues una cerca	La típica casa que dibujan los niños pues con el sol los arbolitos unas manzanas y pues una cerca	Reconoce el uso de los diminutivos en <i>-ito</i> , así como la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la <i>-s</i> Variación léxica ( <i>habas</i> )	Es un carro de la compra, en el que se meten las habas, los garbanzos, las patatas, los plátanos y una tableta de chocolate	Es un carro de la compra en el que se meten las habas los garbanzos las patatas los plátanos y una tableta de chocolate	Reconoce la totalidad del fragmento	

<b>Muestra</b>		18
<b>Procedencia del hablante</b>	Castellano-mancheño (Natural de Ciudad Real)	
<b>Información social</b>	Estudios universitarios	

Rasgos lingüísticos				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Distinción entre los fonemas /s/ y /θ/ Mantiene la -s, -d y -l	Me llamo Luis Antonio y soy de Ciudad Real, España	Me llamo Luis Antonio llegue Ciudad Real España	Reconoce el yeísmo	Error fonético: «Y soy de» entendido como «llegue» por taquilalia
Mantiene la -d- intervocálica	Un pueblo meramente agrícola, un poco desordenado en la construcción	Un pueblo meramente agrícola un poco desordenado en la construcción	Reconoce la totalidad del fragmento	
Distinción entre los fonemas /s/ y /θ/ Mantiene la -s y -l	Tenemos una empresa de exportación, una empresa internacional	Tenemos una empresa de exportación una empresa internacional	Reconoce la totalidad del fragmento	
Sinalefa Mantiene la -s Pérdida de la -d- intervocálica (solo en el segundo <i>he estudiado</i> )	También he estudiado en Lérida, he estudiado en Zaragoza	También he estudiado en Lérida he estudiado en Zaragoza	Reconoce la pérdida de la -d- intervocálica	Error fonético-sintáctico: «He estudiado» entendido como «estudiado» por sinalefa
Distinción entre los fonemas /s/ y /θ/ Mantiene la -s y -r Sinalefa Variación léxica ( <i>abatir, patos</i> )	Esperando el disparo del cazador, si abate a alguno de los patos	esperando el disparo del cazador si abate a alguno de los patos		Error fonético-sintáctico: «Abate a alguno» entendido como «abate alguno» por sinalefa
Distinción entre los fonemas /s/ y /θ/ Sinalefa Mantiene la -s	Una casa donde la chimenea parece que está en funcionamiento	Una casa donde la chimenea parece que está en funcionamiento	Reconoce la sinalefa	

Preferencia por los diminutivos en <i>-ito</i> Mantiene la <i>-s</i> Pronunciación tensa de la fricativa velar sorda /x/ Variación léxica ( <i>judías verdes, garbanzos, patatas</i> )	Es un carrito de la compra, donde se ven unas judías verdes, una tableta de chocolate, unos garbanzos, unas patatas y unos plátanos que deben ser de Canarias	Es un carrito de la compra donde se ven unas judías verdes una tableta de chocolate unos garbanzos unas patatas y unos plátanos que deben ser de Canarias	Reconoce el uso de los diminutivos en <i>-ito</i> , así como la pronunciación tensa de la fricativa velar sorda /x/	
---	---	---	---	--

<b>Muestra</b>	19			
<b>Procedencia del hablante</b>	Madrileña (Natural de Madrid)			
<b>Información social</b>	Estudios medios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Confusión de <i>-d</i> por <i>-z</i> ( <i>Madrid</i> )	Me llamo Isabel y soy de aquí, de Madrid	Me llamo Isabel y soy de aquí de Madrid	Reconoce el yeísmo y la confusión de <i>-d</i> por <i>-z</i>	
Mantiene la <i>-s</i> Uso de acrónimos: <i>EGB</i> Pronunciación tensa de la fricativa velar sorda /x/	Hice los estudios de EGB	Hice los estudios de EGB	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la <i>-s</i> y <i>-d</i> Uso de acrónimos: <i>FP</i>	Hice una FP de grado medio en la rama de Administrativo	Hice una FP de grado medio en la rama de administrativo	Reconoce la totalidad del fragmento	
Mantiene la <i>-d</i> -intervocálica Pérdida de <i>-s</i> y <i>-d</i>	Nunca he vivido fuera de Madrid ni de España	Nunca he vivido fuera de Madrid ni de España	Reconoce la sinalefa y la pérdida de <i>-s</i> y <i>-d</i>	

Sinalefa				
Mantiene la -r y -s Pronunciación tensa de la fricativa velar sorda /x/ Variación léxica ( <i>pájaros</i> )	Veo un cazador, pues disparando contra unos pájaros	Veo un cazador pues disparando contra unos pájaros	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Pronunciación tensa de la fricativa velar sorda /x/ Pérdida de -s	La típica casa con su jardín y un día espléndido	La típica casa con su jardín y un día espléndido	Reconoce la totalidad del fragmento	
Mantiene la -s (en <i>productos, básicos, frutas, proteínas, verduras y legumbres</i> ) Pérdida de -s (en <i>los artículos</i> ) Variación léxica ( <i>carro</i> )	Veo un carro de compra con los productos básicos como son las frutas, proteínas, las verduras, las legumbres	Veo un carro de compra con los productos básicos como son las frutas proteínas las verduras la legumbres	Reconoce la pérdida de -s en <i>los artículos</i>	

<b>Muestra</b>	20			
<b>Procedencia del hablante</b>	Cántabra (Natural de Santander, Cantabria)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Distinción entre los fonemas /s/ y /θ/ Mantiene la -s, -d y -r	Me llamo Marta y soy de Santander, una ciudad de Cantabria, España	Me llamo Marta y soy de Santander en la ciudad de Cantabria España		Error fonético: «Una ciudad» entendido como «en la ciudad»
Sinalefa	A lo largo del tiempo, pues me	A largo del tiempo pues me he dado	Reconoce la sinalefa y la	

Pérdida de la <i>-d-</i> intervocálica Mantiene la <i>-s</i> y <i>-d</i>	he dado cuenta de que es una ciudad bastante pequeña	cuenta de que es una ciudad bastante pequeña	pérdida de la <i>-d-</i> intervocálica	
Distinción entre los fonemas /s/ y /θ/ Mantiene la <i>-s</i> y <i>-d</i>	Es una ciudad costera, tiene muchas playas	Es una ciudad costera tiene muchas playas	Reconoce la totalidad del fragmento	
Mantiene la <i>-s</i> y <i>-z</i> Pérdida de <i>-d</i> Preferencia por los diminutivos en <i>-ito</i>	Yo vivo en un barrio pues pequeñito, a 10 minutos a pie del centro de la ciudad	Yo vivo en un barrio pues pequeñito a 10 minutos a pie del centro de la ciudad	Reconoce la pérdida de <i>-d</i> y el uso de los diminutivos en <i>-ito</i>	
Mantiene la <i>-s</i> y <i>-d</i> Pronunciación tensa de la fricativa velar sorda /x/	Me fui un año de Erasmus a Mons, una ciudad de Bélgica	Me fui un año de Erasmus a Montse una ciudad de Bélgica		Error léxico: No reconoce el término «Mons»
Mantiene la <i>-s</i> y <i>-r</i> Pronunciación tensa de la fricativa velar sorda /x/	Se ve también pues vegetación así como típica de pantano alrededor	Se ve también pues vegetación así como típica de pantano alrededor	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la <i>-s</i> Pronunciación tensa de la fricativa velar sorda /x/ Yeísmo	Hay un sol dibujado, unos árboles frutales (...) también una valla	Hay un sol dibujado unos árboles frutales también una valla	Reconoce el yeísmo y la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la <i>-s</i> Pronunciación tensa de la fricativa velar sorda /x/ Variación léxica ( <i>patatas</i> , <i>garbanzos</i> , <i>judías</i> )	Dentro tiene una tableta de chocolate, unos plátanos, unas patatas, unos garbanzos y unas judías	Dentro tiene una tableta de chocolate unos plátanos unas patatas unos garbanzos y unas judías	Reconoce la totalidad del fragmento	

<b>Muestra</b>	21
----------------	----

<b>Procedencia del hablante</b>	Madrileña (Natural de Madrid)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Mantiene la -s y -d Yeísmo Pronunciación tensa de la fricativa velar sorda /x/	Me llamo Alejandra Gadea y soy de Madrid, España	Me llamo Alejandra Gadea y soy de Madrid España	Reconoce el yeísmo y la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la -s y -d Uso explícito del pronombre sujeto Yeísmo	Cuando era pequeña pues yo recuerdo la ciudad un poco más pequeña	Cuando era pequeña pues yo recuerdo la ciudad un poco más pequeño	Reconoce el yeísmo y el uso explícito del pronombre sujeto	Error fonético: «Pequeña» entendido como «pequeño»
Mantiene la -s Pronunciación tensa de la fricativa velar sorda /x/ Sinalefa	Eventualmente trabajo como azafata de eventos	Eventualmente trabajo como azafata de eventos	Reconoce la sinalefa y la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la -d- intervocálica Mantiene la -s Sinalefa	Nunca he vivido fuera de España	Nunca he vivido fuera de España		Error fonético: «He vivido» entendido como «vivido»
Mantiene la -d- intervocálica Preferencia por los diminutivos en -ito	Veo un día soleado en el que se ve una casita	Veo un día soleado en el que se ve una casita	Reconoce el uso de los diminutivos en -ito	
Distinción entre los fonemas /s/ y /θ/ Mantiene la -s	Veo un carrito de la compra en el que dentro vemos chocolate, plátanos, guisantes	Veo un carrito de la compra en el que dentro vemos chocolate plátanos guisantes	Reconoce la totalidad del fragmento	

Variación léxica ( <i>guisantes</i> )				
--	--	--	--	--

### III. Siri y el español de las comunidades autónomas bilingües

#### III.1 MUESTRAS RELACIONADAS CON EL ESPAÑOL DE CATALUÑA, ISLAS BALEARES Y COMUNIDAD VALENCIANA

<b>Muestra</b>	22			
<b>Procedencia del hablante</b>	Sant Cugat del Vallès, Barcelona, Cataluña (Nació en Barcelona, pero ha vivido los últimos diez años en Castilla-La Mancha)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
<b>Rasgos</b>	<b>Ejemplo en la grabación</b>	<b>Transcripción de Siri (aciertos / errores)</b>	<b>Tipos de aciertos</b>	<b>Tipos de errores</b>
Yeísmo Mantiene la -s y -t	Me llamo Meritxell, soy de Sant Cugat del Vallès, un pueblo de Barcelona	Me llamo Meritxell soy de San Cugat del Vallés un pueblo de Barcelona	Reconoce el yeísmo	
Pérdida de -d Mantiene la -s Pronunciación tensa de la fricativa velar sorda /x/	Ahora, en la actualidad estoy estudiando Magisterio	Ahora en la actualidad estoy estudiando magisterio	Reconoce la pérdida de -d y la pronunciación tensa de la fricativa velar sorda /x/	
Sinalefa Mantiene la -s	He estado seis años en Toledo	Estado seis años en Toledo		Error léxico-sintáctico: «He estado» entendido como «estado» por sinalefa
Mantiene la -s Variación léxica ( <i>pájaros</i> )	Podemos observar a un cazador que	Podemos observar a un cazador que está	Reconoce la totalidad del fragmento	

	está matando a unos pájaros	matando a unos pájaros		
Mantiene la -s Preferencia por los diminutivos en -ito	Podemos ver una casa con unas florecitas	Podemos ver una casa con unas florecitas	Reconoce el uso de los diminutivos en -ito	
Mantiene la -d- intervocálica	Es un día soleado	Es un día soleado	Reconoce la totalidad del fragmento	
Mantiene la -s Variación léxica (guisantes, patatas)	Contiene chocolate, garbanzos, plátanos, patatas y guisantes	Contiene chocolate garbanzos plátanos patatas y guisantes	Reconoce la totalidad del fragmento	

<b>Muestra</b>	23			
<b>Procedencia del hablante</b>	Sant Cugat del Vallès, Barcelona, Cataluña (Nació en Barcelona, pero ha vivido los últimos diez años en Castilla-La Mancha)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la -s	Me llamo Jordi, soy de Sant Cugat del Vallés	Me llamo Jordi soy de San Cugat del Vallés	Reconoce el yeísmo	
Uso explícito del pronombre sujeto Mantiene la -s Preferencia por los diminutivos en -ito	El pueblo donde yo vivía desde chiquitito	El pueblo donde yo vivía desde chiquitito	Reconoce el uso explícito del pronombre sujeto y el uso de los diminutivos en -ito	
Uso explícito del pronombre sujeto	Donde hay de todo y donde yo iba al colegio	Dónde hay de todo y donde yo iba al colegio	Reconoce la totalidad del fragmento	



Pronunciación tensa de la fricativa velar sorda /x/				
Mantiene la -s Pérdida de -d	Estoy estudiando en la Universidad de Castilla-La Mancha	Estoy estudiando en la universidad de Castilla-La Mancha	Reconoce la pérdida de -d	
Mantiene la -s y -r	Veo un cazador con el perro, que está apuntando a los patos	Veo un cazador con el perro que está apuntando a los patos	Reconoce la totalidad del fragmento	
Mantiene la -s y -r	Por lo que se puede apreciar es un día caluroso de verano	Por lo que se puede apreciar es un día caluroso de verano	Reconoce la totalidad del fragmento	
Mantiene la -s Pérdida de -d- intervocálica Variación léxica ( <i>guisantes</i> )	Tiene chocolate, plátanos, patatas, guisantes y un puñado de garbanzos	Tiene chocolate plátanos patatas guisantes y un puñado de garbanzos	Reconoce la pérdida de -d- intervocálica	

<b>Muestra</b>	24			
<b>Procedencia del hablante</b>	Balear (Natural de Santañí, Mallorca)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Distinción entre los fonemas palatales /k/ y /j/ Mantiene la -s	Me llamo Júlia y soy de Santanyí, un pueblo de Mallorca, en España	Me llamo Júlia y soy de Santañy un pueblo de Mallorca en España	Reconoce la distinción entre los fonemas palatales /k/ y /j/	

Mantiene la -s y -d	Estudio un doble grado en Traducción e Interpretación y Lenguas Aplicadas en la Universidad Pompeu Fabra	Estudio un doble grado de traducción e interpretación y lenguas aplicadas en la Universidad Pompeu Fabra	Reconoce la totalidad del fragmento	
Yeísmo Mantiene la -s y -l	Mi madre es de Brasil, con ella hablo portugués. Mi padre es de Mallorca, con él hablo mallorquín. Y mi hermana, con ella hablo castellano	Mi madre es de Brasil con ella hablo portugués mi padre es de Mallorca con el hablo mallorquín y mi hermana con ella hablo castellano	Reconoce el yeísmo	
Mantiene la -s y -r	El cazador está apuntando con un arma a cinco patos, que vuelan por el cielo	El cazador está apuntando con un arma a cinco patos vuelan por el cielo	Reconoce la totalidad del fragmento	
Mantiene la -s y -l Pronunciación tensa de la fricativa velar sorda /x/	Es verde y azul la imagen	Es verde y azul la imagen	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -s	La segunda imagen vemos un dibujo de una casa, con un sol, un arbusto, flores, una valla, dos árboles con manzanas...	La segunda imagen vemos un dibujo de una casa con un sol Augusto Flores una valla dos árboles con manzanas		Error léxico: «Arbusto» entendido como «Augusto»
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -s	En la tercera imagen, vemos un carrito de la compra que contiene chocolate, judías,	En la tercera imagen vemos un carrito de la compra que contiene chocolate judías patatas plátanos y garbanzos	Reconoce la totalidad del fragmento	

Variación léxica ( <i>patatas, judías, garbanzos</i> )	patatas, plátanos y garbanzos			
---	-------------------------------	--	--	--

<b>Muestra</b>	25			
<b>Procedencia del hablante</b>	Catalana (Natural de Terrassa, Barcelona)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Reducción de las vocales [e] y [a] átonas a una vocal central Ensondecimiento de la oclusiva alveolar sonora /d/ a final de palabra Mantiene la -d	Me llamo Marta y soy de Terrassa, una ciudad cerca de Barcelona, en España	Hola me llamo Marta soy de Terrassa una ciudad cerca de Barcelona en España	Reconoce el yeísmo, la reducción de las vocales [e] y [a] átonas a una vocal central y el ensondecimiento de la oclusiva alveolar sonora /d/ a final de palabra	
Uso del verbo <i>estar</i> en lugar del verbo <i>estudiar</i> Mantiene la -s Ensondecimiento de la oclusiva alveolar sonora /d/ a final de palabra	Estoy en la Universidad Pompeu Fabra, en Barcelona también	Estoy en la universidad Pompeu Fabra en Barcelona también	Reconoce el uso del verbo <i>estar</i> en lugar del verbo <i>estudiar</i> y el ensondecimiento de la oclusiva alveolar sonora /d/ a final de palabra	
Mantiene la -s Relajación de la /p/ en la articulación de la voz <i>septiembre</i>	Estuve viviendo durante un trimestre en Londres, de septiembre a diciembre del 2019	Estuve viviendo durante un trimestre en Londres de septiembre a diciembre en el 2019	Reconoce la relajación de la /p/ en la articulación de la voz <i>septiembre</i> y la velarización de la alveolar /l/	

Velarización de la alveolar /l/				
Velarización de la alveolar /l/ Yeísmo Mantiene la -l Pérdida de -s	Me comunicaba con mi familia a través de videollamadas y de mensajes en el móvil	Me comunicaba con mi familia a través de videollamada de Mensajes en el móvil	Reconoce el yeísmo	Error fonético: «Videollamadas» entendido como «videollamada» por pérdida de -s
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -s Variación léxica ( <i>escopeta, pájaros</i> )	En la primera imagen, veo un cazador que tiene en las manos una escopeta y está disparando a unos pájaros	En la primera imagen veo un cazador que tiene en las manos una escopeta y está disparando a unos pájaros	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Omisión del artículo femenino <i>la</i> Mantiene la -s Utiliza <i>vallo</i> en lugar de la voz <i>valla</i>	Delante de casa, hay unos arbustos con flores, una vallo y unos árboles que tienen manzanas	Delante de queso hay unos arbustos con flores un caballo y unos árboles que tienen manzanas		Error léxico: «Una vallo» entendido como «un caballo» por emplear <i>vallo</i> en lugar de <i>valla</i>
Preferencia por los diminutivos en <i>-ito</i> Mantenimiento de la -s Pronunciación tensa de la fricativa velar sorda /x/ Variación léxica ( <i>judías verdes, patatas, garbanzos</i> )	Dentro del carrito, hay una tableta de chocolate, plátanos, patatas, judías verdes y garbanzos	Dentro del carrito hay una tableta de chocolate plátanos patatas judías verdes y garbanzos	Reconoce el uso de los diminutivos en <i>-ito</i> , así como el resto del fragmento	

<b>Muestra</b>		26
<b>Procedencia del hablante</b>	Catalana (Natural de Santa Coloma de Farners, Girona)	

<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Reducción de las vocales [e] y [a] átonas a una vocal central Mantiene la -s	Me llamo Paula y soy de un pueblo que se llama Santa Coloma de Farners, un pueblo de Girona	Me llamo Paula y soy de un pueblo que se llama Santa Coloma de farners un pueblo de Girona	Reconoce la reducción de las vocales [e] y [a] átonas a una vocal central	
Mantiene la -s Variación léxica ( <i>hacer fiestas en lugar de celebrar fiestas</i> )	Se caracteriza pues por las fiestas tradicionales que hacemos, relacionadas con los productos gastronómicos típicos de la zona	Se caracteriza pues por las fiestas tradicionales que hacemos relacionadas con los productos gastronómicos típicos de la zona	Reconoce la totalidad del fragmento	
Mantiene la -s y la -d	No estudio en mi pueblo, sino en Barcelona, en la Universidad Pompeu Fabra	No estudio en mi pueblo sino en Barcelona en la Universidad Pompeu Fabra	Reconoce la totalidad del fragmento	
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -s y la -r Variación léxica ( <i>camuflaje</i> )	En la imagen 1, hay un cazador vestido de camuflaje	En la imagen uno hay un cazador vestido de camuflaje	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la -s Variación léxica ( <i>manzanos</i> )	Y a la derecha, hay dos árboles, dos manzanos y flores de color amarillo	Y a la derecha hay dos árboles dos manzanos y flores de color amarillo	Reconoce la totalidad del fragmento	

Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -s Preferencia por los diminutivos en <i>-ito</i>	La tercera imagen es un carrito de la compra	La tercera imagen es un carrito de la compra	Reconoce el uso de los diminutivos en <i>-ito</i> y la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la -s Variación léxica ( <i>patatas, judías verdes, garbanzos</i> )	Y dentro hay tres patatas, tres plátanos, una tableta de chocolate, tres judías verdes y [...] garbanzos	Y dentro hay tres patatas tres plátanos una tableta de chocolate tres judías verdes y (...) garbanzos	Reconoce la totalidad del fragmento	

<b>Muestra</b>	27			
<b>Procedencia del hablante</b>	Catalán (Natural de Vidreras, Girona)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la -s	Me llamo Pau y soy de Vidreras, un pueblo cerca de Girona, España	Me llamo Paul y soy de vidrieras un pueblo cerca de Girona España	Reconoce el yeísmo	Error léxico: «Vidreras» entendido como «vidrieras»
Mantiene la -s Ensondecimiento de la oclusiva alveolar sonora /d/	Estoy acabando la carrera de Informática en la Universidad Politécnica de Cataluña	Estoy acabando en la carrera de informática en la Universidad politécnica de Cataluña	Reconoce el ensondecimiento de la oclusiva alveolar sonora /d/	
Mantiene la -s Pronunciación tensa de la fricativa velar sorda /x/	Estoy trabajando en una empresa como ingeniero del <i>software</i>	Estoy trabajando en una empresa como ingeniero en el sofá	Reconoce la pronunciación tensa de la fricativa velar	Error léxico: «Del software» entendido como «en el sofá» por

Uso de anglicismos: <i>software</i>			sorda /x/	el uso de un anglicismo
Distinción entre los fonemas /s/ y /θ/ Reducción de las vocales [e] y [a] átonas a una vocal central	Ya que ir de Vidreras hasta Barcelona cada día era mucho trayecto	Ya que ir de Vidreras hasta Barcelona Cadera era mucho trayecto	Reconoce la reducción de las vocales [e] y [a] átonas a una vocal central	Error fonético: «Cada día» entendido como «cadera» por taquilalia
Mantiene la -r Reducción de las vocales [e] y [a] átonas a una vocal central	También hay un perro acompañando al cazador	También hay un perro acompañando al cazador	Reconoce la totalidad del fragmento	
Mantiene la -s Variación léxica ( <i>árboles que sacan manzanas</i> )	Sus árboles que sacan manzanas	Sus árboles que sacan manzanas	Reconoce la totalidad del fragmento	
Preferencia por el diminutivo en -ito Mantiene la -s Pronunciación tensa de la fricativa velar sorda /x/ Variación léxica ( <i>judías, garbanzos, patata</i> )	Un carrito de la compra, y dentro del carrito hay una patata, unas judías, unos garbanzos, un plátano y una tableta de chocolate de esas	Un carrito de la compra y entre el carrito hay una patata unas judías unos garbanzos un plátano y una tableta de chocolate de esas	Reconoce el uso de los diminutivos en -ito y la pronunciación tensa de la fricativa velar sorda /x/	Error fonético: «Dentro del» entendido como «entre el»

<b>Muestra</b>		28
<b>Procedencia del hablante</b>	Valenciana (Natural de Valencia)	
<b>Información social</b>	Estudios universitarios	
<b>Rasgos lingüísticos</b>		

Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Mantiene la <i>-d</i>	Me llamo Helena y soy de Valencia, de la Comunidad Valenciana	Me llamo Elena y soy de Valencia de la Comunidad Valenciana	Reconoce la totalidad del fragmento	
Mantiene la <i>-s</i> Mantiene la <i>-d-</i> intervocálica Pronunciación tensa de la fricativa velar sorda /x/	Es una zona bastante animada, en la que siempre hay gente joven	Es una zona bastante animada en la que siempre hay gente joven	Reconoce la totalidad del fragmento	
Mantiene la <i>-r</i> y la <i>-s</i> Yeísmo Uso de catalanismos: <i>universitat</i>	He ido a la Universitat Jaume primer, que está en Castellón	He ido a la universidad Jaume primer que está en Castellón	Reconoce el yeísmo	Error léxico: «Universitat» entendido como «universidad» por el uso de un catalanismo
Mantiene la <i>-s</i> y la <i>-r</i> Sinalefa Variación léxica ( <i>sabueso</i> )	Hay un cazador que es un hombre con un perro, que será un sabueso	Hay un cazador que es un hombre con un perro que será un sabueso	Reconoce la sinalefa	
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la <i>-s</i> y la <i>-r</i> Variación léxica ( <i>atuendo, estampado</i> )	El hombre lleva un atuendo, con estampado militar de camuflaje	El hombre lleva un atuendo con estampado militar de camuflaje	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Uso de <i>como</i> como muletilla Mantiene la <i>-s</i> Variación léxica ( <i>cacahuets, vaines de guisantes, patatas</i> )	Hay como unos cacahuets, chocolate, una especie de vaines con guisantes, patatas y plátanos	Hay como unos cacahuets chocolate una especie de vaines con guisantes patatas y plátanos	Reconoce la totalidad del fragmento	



### III.2 MUESTRAS RELACIONADAS CON EL ESPAÑOL DEL PAÍS VASCO Y

#### EL NORTE DE NAVARRA

<b>Muestra</b>	29			
<b>Procedencia del hablante</b>	Navarro (Procedente de Mutilva, Navarra)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
<b>Rasgos</b>	<b>Ejemplo en la grabación</b>	<b>Transcripción de Siri (aciertos / errores)</b>	<b>Tipos de aciertos</b>	<b>Tipos de errores</b>
Elisión de la <i>-d-</i> intervocálica Preferencia por los diminutivos en <i>-ito</i>	Hola, me llamo Asier y soy de Mutilva, un pueblito de... de Navarra, al <i>lao</i> de Pamplona	Hola me llamo Siri soy de Mutilva un polito de de Navarra la de Pamplona	Reconoce el diminutivo en <i>-ito</i>	Error léxico: No reconoce el nombre Error fonético: «Al lado de» entendido como «la de» por elisión de la <i>-d-</i> intervocálica
Mantiene la <i>-s</i>	Llevo viviendo aquí desde que tengo 5 años	Yo viviendo aquí desde que tengo cinco años		Error fonético: «Llevo» entendido como «yo» por taquilalia (habla rápida)
Mantiene la <i>-s</i> Pronunciación tensa de la fricativa velar sorda /x/	La parte vieja del pueblo que era pues casas, pisos altos, y casas un poquito mas viejas	La parte vieja del pueblo que era pues casas pisos altos y casas un poquito más viejas	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	

Uso preferente del pretérito perfecto simple	Viví fuera de España en 4º de la ESO	Viví fuera de España en cuarto de la eso	Reconoce el uso del pretérito perfecto simple	
Mantiene la -s y -r	Un par de manzanos, algún que otro arbusto y una valla	Un par de manzanos algún que otro arbusto y una valla		

<b>Muestra</b>	30			
<b>Procedencia del hablante</b>	Vasco (Natural de Vitoria-Gasteiz, Álava)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la -s, -d y -z	Me llamo Miguel González, soy de Vitoria-Gasteiz, una ciudad de la provincia de Álava, en el País Vasco	Me llamo Miguel González soy de Vitoria-Gasteiz una ciudad de la provincia de Álava en el País Vasco	Reconoce el yeísmo y el mantenimiento de -s, -d y -z	
Uso preferente del pretérito perfecto simple	Estudiando en la universidad me trasladé a vivir a Pamplona	Estudiando en la universidad me trasladé ir a vivir a Pamplona	Reconoce el uso del pretérito perfecto simple	
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -r y -s	En la primera imagen se ve un cazador cazando patos	La primera imagen se ve un cazador cazando patos	Reconoce la pronunciación tensa de la fricativa velar sorda /x/ y el mantenimiento de -r y -s	

Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -d- intervocálica Variación léxica ( <i>niño pequeño</i> )	Parece dibujado por un niño pequeño	Parece dibujado por un niño pequeño	Reconoce la pronunciación tensa de la fricativa velar sorda /x/ y la -d- intervocálica	
Mantiene la -s	Con algunos alimentos como son chocolate, plátanos, patatas, garbanzos y guisantes	Con algunos alimentos como son chocolate plátanos patatas garbanzos y guisantes	Reconoce la -s	

<b>Muestra</b>	31			
<b>Procedencia del hablante</b>	Navarro (Natural de Fitero, Navarra)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la -s	Me llamo Adrián y soy de Fitero, un pueblo de la provincia de Navarra, España	Me llamo Adrian y soy de Fitero un pueblo de la provincia de Navarra España	Reconoce el yeísmo y la -s	
Mantiene la -s y -d	Era un pueblo bastante pequeño, la verdad	Era un vuelo bastante pequeño la verdad	Reconoce la -s y -d	
Preferencia por los diminutivos en -ito	Un pueblo pequeñito de la zona de la ribera	Un pueblo pequeñito de la zona de la Ribera	Reconoce el diminutivo en -ito	
Mantiene la -s	Con mi familia pues me comunico	Con mi familia pues me comunico sobre	Reconoce el yeísmo, el leísmo	

Leísmo Yeísmo	sobre todo por teléfono, hablo con ellos por WhatsApp o les llamo	todo por teléfono hablo con ellos por WhatsApp o les llamo	y la -s	
Pronunciación tensa de la fricativa velar sorda /x/ Uso explícito del pronombre sujeto	En la primera imagen, yo veo a un hombre cazando	En la primera imagen yo veo a un hombre cazando	Reconoce la pronunciación tensa de la fricativa velar sorda /x/ y el uso explícito del pronombre sujeto	
Mantiene la -s Variación léxica ( <i>aves</i> )	Está disparando a unas aves	Está disparando o no sabes		Error fonético: «A unas aves» entendido como «o no sabes»
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -d- intervocálica	Un tejado naranja, el sol al lado	Un tejado naranja el sol al lado	Reconoce la pronunciación tensa de la fricativa velar sorda /x/ y la -d- intervocálica	

<b>Muestra</b>	32			
<b>Procedencia del hablante</b>	Navarra (Natural de Mutilva, Navarra)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo	Me llamo Beatriz Montes y soy de Mutilva Alta	Me llamo Beatriz Montes y soy de Mutilva Alta	Reconoce el yeísmo	
Yeísmo	Íbamos con los mayores	Íbamos con los mayores	Reconoce el yeísmo	

Elisión de la dental sonora, acompañada por el cierre de /o/ en /u/, en las terminaciones en - <i>ado</i>	Estaba rodeado, como todavía no estaba muy desarrollado, pues estaba rodeado de un montón de trigales	Estaba rodeado como todavía no estaba muy desarrollado pues estaba rodeado de un montón de trigales	Reconoce la elisión de la dental sonora, acompañada por el cierre de /o/ en /u/, en las terminaciones en - <i>ado</i>	
Distinción entre los fonemas /s/ y /θ/ Pronunciación tensa de la fricativa velar sorda /x/	En la primera imagen, veo como una especie de lago y un terreno de hierba	En la primera imagen veo como una especie de lago y un terreno de hierba	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Sinalefa Elisión de la dental sonora, acompañada por el cierre de /o/ en /u/, en las terminaciones en - <i>ado</i>	He estudiado	Estudiado		Error fonético-sintáctico: «He estudiado» entendido como «estudiado» por sinalefa y elisión de la - <i>d</i> -intervocálica
Mantiene la -s Variación léxica ( <i>carro, patatas, alubias</i> )	Parece... pues el típico carro de supermercado, pues con plátanos, chocolate, patatas, alubias y garbanzos en el centro	Parece pues el típico carro de supermercado pues con plátanos chocolate patatas alubias y garbanzos en el centro	Reconoce la -s	

<b>Muestra</b>		33
<b>Procedencia del hablante</b>	Navarro (Natural de Barañáin, Navarra)	
<b>Información social</b>	Estudios universitarios	
<b>Rasgos lingüísticos</b>		

Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Distinción entre los fonemas /s/ y /θ/	Me llamo Juanjo, soy de Pamplona, nací en Pamplona	Me llamo Juanjo soy de Pamplona nací en Pamplona	Reconoce la totalidad del fragmento	
Preferencias por los diminutivos en <i>-ito</i>	Era un pueblecito colindante, justo al lado de Pamplona, o sea, dentro de nada será Pamplona y ya estaba bastante industrializado cuando yo nací	Era un parecido colindante justo al lado de Pamplona o sea dentro de nada será Pamplona y yo estaba bastante industrializado cuando yo nací		Error fonético: «Pueblecito» entendido como «parecido» por no reconocer el uso del diminutivo en <i>-ito</i>
Sinalefa Mantiene la <i>-d-</i> intervocálica Pérdida de la <i>-d</i>	He estudiado la carrera de Medicina en la Universidad de Navarra y, en cuanto a otros estudios, no tengo otros estudios más que la ESO y Bachiller	Estudiado la carrera de medicina de la Universidad de Navarra y en cuanto a otros estudios no tengo otros estudios más que la eso y bachiller	Reconoce la <i>-d-</i> intervocálica y la pérdida de <i>-d</i>	Error fonético: «He estudiado» entendido como «estudiado» por sinalefa
Variación léxica ( <i>gansos, lago</i> )	Veo a un cazador junto con un perro intentando cazar a unos gansos y en un lago	Veo un cazador junto con un perro intentando cazar a unos gansos y en un lago	Reconoce la totalidad del fragmento	
Yeísmo Variación léxica ( <i>arbustos, valla</i> )	Veo una casa con sus arbustos, su valla, el sol, un entorno idílico	Veo una casa con sus arbustos su vaya el sol un internet de Lico		Error fonético: «Valla» entendido como «vaya» por yeísmo Error léxico: No identifica la expresión «entorno idílico»

<b>Muestra</b>	34			
<b>Procedencia del hablante</b>	Navarro (Natural de Mutilva, Navarra)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Preferencia por los diminutivos en <i>-ito</i> Variación léxica ( <i>pueblito</i> )	Me llamo Javi y soy de Mutilva, un pueblito cerca de Pamplona	Me llamo Javi soy de Mutilva un polito cerca de Pamplona	Reconoce el yeísmo	Error léxico: «Pueblito» entendido como «polito» por confusión léxica y falta de identificación del diminutivo
Mantiene la <i>-s</i> Preferencia por los diminutivos en <i>-ito</i>	Antes eran solo casitas de campo y unas pocas urbanizaciones pequeñas	Antes eran solo casitas de campo y unas pocas organizaciones pequeñas	Reconoce la <i>-s</i> y el uso del diminutivo en <i>-ito</i>	
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la <i>-r</i>	Se podía jugar y correr	Se podía jugar y correr	Reconoce la <i>-r</i> y la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la <i>-s</i> , <i>-r</i> y <i>-c</i> Pérdida de <i>-d</i>	No tengo proyecto de estudiar nada más, la verdad	No tengo proyecto de estudiar nada más la verdad	Reconoce la <i>-s</i> , <i>-r</i> y <i>-c</i> , así como la pérdida de <i>-d</i>	
Mantiene la <i>-s</i>	El máximo, dos meses que estuve en Francia	El máximo dos meses que estuve en Francia	Reconoce la <i>-s</i>	
Mantiene la <i>-s</i> Yeísmo	Unos arbustos, unas flores, unos	Unos arbustos unas flores unas manzanas y una galleta		Error léxico: «Galleta» entendido como

Variación léxica ( <i>manzanos</i> )	manzanos y una vallita			«vallita» por yeísmo y no identificación del diminutivo en <i>-ito</i>
Mantiene la <i>-s</i> Preferencia por los diminutivos en <i>-ito</i> Variación léxica ( <i>alubias verdes</i> )	Un carrito de la compra en el que hay unas alubias verdes	Un carrito de la compra en el que hay unas alubias verdes	Reconoce la totalidad del fragmento	

<b>Muestra</b>	35			
<b>Procedencia del hablante</b>	Vasca (Natural de San Sebastián)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la <i>-s</i>	Me llamo Laura y soy de San Sebastián	Me llamo Laura y soy de San Sebastián	Reconoce el yeísmo y la <i>-s</i>	
Mantiene la <i>-s</i>	Es muy acogedora y está rodeada de montañas	Es muy acogedora y está rodeada de montañas	Reconoce la totalidad del fragmento	
Sinalefa Mantiene la <i>-s</i> y <i>-d</i>	He estudiado en la Universidad de Navarra	Estudiado en la Universidad de Navarra	Reconoce la <i>-s</i> y <i>-d</i>	Error fonético-sintáctico: «He estudiado» entendido como «estudiado» por sinalefa
Uso preferente del pretérito perfecto simple Mantiene la <i>-s</i>	En Alemania, donde viví tres meses	En Alemania donde viví tres meses	Reconoce el uso del pretérito perfecto simple y la <i>-s</i>	



Mantiene la -s y -r Variación léxica (aves)	Un cazador con un perro que ha ido a cazar aves	Un cazador con un perro que ha ido a cazar aves	Reconoce la totalidad del fragmento	
Mantiene la -s Preferencia por los diminutivos en -ito Variación léxica (verja)	Una casa en el centro, con una verja y unos arbolitos	Una casa en el centro con una verja y unos arbolitos	Reconoce la -s y el uso del diminutivo en -ito	
Mantiene la -s Variación léxica (vainas)	Un carro de la compra con plátanos, garbanzos, vainas y chocolate	Un carro de la compra con plátanos garbanzos vainas y chocolate	Reconoce la totalidad del fragmento	

### III.3 MUESTRAS RELACIONADAS CON EL ESPAÑOL DE GALICIA

<b>Muestra</b>	36			
<b>Procedencia del hablante</b>	Gallega (Natural de La Coruña)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la -s	Me llamo Blanca y soy de La Coruña, España	Me llamo Blanca y soy de La Coruña España	Reconoce el yeísmo	
Mantiene la -d	Es una ciudad mediana	Es una ciudad mediana	Reconoce la totalidad del fragmento	
Mantiene la -s y -r Pronunciación tensa de la fricativa velar sorda /x/	Tiene mucho verde, muchos jardines y es muy	Tiene mucho verde muchos jardines y es muy agradable pasear por allí	Reconoce la pronunciación tensa de la fricativa velar	

	agradable pasear por allí		sorda /x/	
Variación léxica ( <i>aves</i> )	Es un cazador que está apuntando pues a unas aves con su escopeta	Es un cazador que está apuntando pues <b>ahora</b> <b>sabes</b> con su escopeta		Error léxico: «Unas aves» entendido como «ahora sabes»
Mantiene la -s Mantiene la -d- intervocálica Pronunciación tensa de la fricativa velar sorda /x/	Es un día pues con el cielo bastante despejado	Es un día pues con el cielo bastante despejado	Reconoce la totalidad del fragmento	
Yeísmo Mantiene la -s	Tiene una pequeña valla y unos manzanos	Tiene una pequeña valla y unos manzanos	Reconoce el yeísmo	
Variación léxica ( <i>vainas de guisantes, patatas</i> ) Mantiene la -s	Que hay como unas vainas de guisantes, garbanzos, una tableta de chocolate , unos plátanos y unas patatas	Que hay como unas vainas de guisantes garbanzos una tableta de chocolate unos plátanos y unas patatas	Reconoce la totalidad del fragmento	

<b>Muestra</b>	37			
<b>Procedencia del hablante</b>	Gallega (Natural de Vigo, Pontevedra)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la -d	Me llamo Iria y soy de Vigo, una ciudad de Galicia	Me llamo Iria y soy de Vigo una ciudad de Galicia	Reconoce el yeísmo	

Mantiene la <i>-s</i> , <i>-r</i> y <i>-d</i> Uso de galleguismos: <i>Rías Baixas</i>	Es una ciudad costera, que está en las <i>Rías Baixas</i> en el sur de Galicia	Es una ciudad costera que está en la rías Baixas en el sur de Galicia	Reconoce el uso de galleguismos	
Mantiene la <i>-s</i> Omisión del artículo determinado <i>el</i>	Siempre nos llamábamos por móvil o WhatsApp	Siempre nos llamábamos por móvil o WhatsApp	Reconoce la omisión del artículo determinado <i>el</i>	
Mantiene la <i>-s</i> y <i>-r</i> Variación léxica ( <i>pájaros</i> ) Pronunciación tensa de la fricativa velar sorda /x/ Sinalefa	Un cazador que está apuntando a unos pájaros	Un cazador que está apuntando a unos pájaros	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	Error léxico-sintáctico: «Apuntando a unos» entendido como «apuntando unos» por sinalefa
Mantiene la <i>-s</i> Pronunciación tensa de la fricativa velar sorda /x/	Y se ve una casa, con dos ventanas, una puerta, un tejado rojo	Y sobre una casa con dos ventanas una puerta un tejado rojo	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	Error fonético: «Se ve» entendido como «sobre»
Mantiene la <i>-s</i> Preferencia por los diminutivos en <i>-ito</i>	Es un carrito de la compra	Es un carrito de la compra	Reconoce el uso de los diminutivos en <i>-ito</i>	
Sinalefa Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la <i>-s</i> Variación léxica ( <i>judías</i> , <i>edamame</i> )	Lo otro no sé si son judías o edamame, o algo así	El otro no sé si son judíos o de Mami o algo así	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	Error léxico-sintáctico: «Lo otro» entendido como «el otro» por sinalefa Error léxico: «Edamame» entendido como «de mami»

<b>Muestra</b>	38
----------------	----

<b>Procedencia del hablante</b>	Gallego (Natural de Oleiros, La Coruña)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la <i>-r</i> y <i>-s</i> Sinalefa	Me llamo Rober y soy de un municipio de Coruña, llamado Oleiros	Me llamo Rober y soy de un municipio de Coruña llamado Oleiros	Reconoce el yeísmo y la sinalefa	
Mantiene la <i>-s</i> y <i>-r</i> Mantiene la <i>-d-</i> intervocálica Pronunciación tensa de la fricativa velar sorda /x/	Tenemos mucha vegetación y el mar justo al lado	Tenemos mucha vegetación y el mar justo al lado	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la <i>-d-</i> intervocálica Mantiene la <i>-s</i> Pérdida de la <i>-d</i>	Nunca he vivido fuera de mi casa, de Oleiros. ¡Una pena, la verdad!	Nunca he vivido fuera de mi casa de Oleiros una pena la verdad	Reconoce la pérdida de la <i>-d</i>	
Sinalefa Mantiene la <i>-s</i> y <i>-r</i> Pronunciación tensa de la fricativa velar sorda /x/ Preferencia por los diminutivos en <i>-ito</i>	Que está matando o intentando matar deben ser pues unos pajaritos	Qué está matando intentando matar deben ser pues unos pajaritos	Reconoce la sinalefa, la pronunciación tensa de la fricativa velar sorda /x/ y el uso de los diminutivos en <i>-ito</i>	
Uso de los diminutivos en <i>-ete</i> Variación léxica ( <i>manzano</i> )	Una casa con un solete, la chimenea echando humo, un manzano	Una casa con un solete la chimenea echando humo un manzano	Reconoce el uso de los diminutivos en <i>-ete</i>	

Mantiene la -s y -r Variación léxica ( <i>patatas, habas, guisantes</i> )	Hay patatas, plátanos, chocolate, habas o guisantes debe ser	Hay patatas plátanos chocolate habas o guisantes debe ser	Reconoce la totalidad del fragmento	
---	--	---	-------------------------------------	--

<b>Muestra</b>	39			
<b>Procedencia del hablante</b>	Gallega (Natural de Sada, La Coruña)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Sinalefa Yeísmo Mantiene la -s	Me llamo Paloma y soy de Sada, un pueblo de La Coruña, en España	Me llamo Paloma y soy desada un pueblo de La Coruña en España	Reconoce el yeísmo	Error léxico: «De Sada» entendido como «desada»
Mantiene la -s Uso agramatical de la tercera persona del plural con un sujeto singular Mantiene la -d- intervocálica Pronunciación tensa de la fricativa velar sorda /x/ Sinalefa	A lo largo de los años han venido mucha gente de diferentes partes de España	A lo largo de los años han venido mucha gente de diferentes partes de España	Reconoce el uso agramatical de la tercera persona del plural con un sujeto singular y la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la -s y -r Variación léxica ( <i>venideros</i> )	Tenemos muchos venideros como se le suele decir aquí	Tenemos muchos venideros como se le suele decir aquí	Reconoce la totalidad del fragmento	
Leísmo Yeísmo	Un pueblo marinero, le llaman a perla das mariñas	Un pueblo marinero le llaman a perlas marinas	Reconoce el leísmo y el yeísmo	Error léxico: «A perla das mariñas» entendido como

Uso de galleguismos: <i>a perla das mariñas</i> Mantiene la -s				«perlas marinas» por ser un galleguismo
Mantiene la -s Sinalefa Preferencia por los diminutivos en <i>-ito</i> Variación léxica ( <i>lago, fauna, flora</i> )	Hay un lago por detrás y muchos arbolitos, y mucha fauna por ahí y flora	Hay un lago por detrás y muchos arbolitos y mucha fauna por ahí y flora	Reconoce el uso de los diminutivos en <i>-ito</i>	Error fonético-sintáctico: «Ahí y flora» entendido como «ahí flora» por sinalefa
Uso repetitivo de los diminutivos en <i>-ito</i>	La casita con una chimenea echando humito, con dos ventanitas	La casita con una chimenea echando un mito con dos ventanitas		Error léxico: «Humito» entendido por «un mito» por la errónea identificación de un diminutivo en <i>-ito</i>
Mantiene la -s Uso de los diminutivos en <i>-illo</i> Variación léxica ( <i>manzanos</i> )	Hay unos manzanos con unas florecillas	Hay unos manzanos con unas florecillas	Reconoce el uso de los diminutivos en <i>-illo</i>	
Variación léxica ( <i>patatas, garbanzos, hebras de los guisantes</i> )	Dentro han echado chocolate, (...) plátano, patatas, garbanzos y (...) las hebras de los guisantes	Dentro han echado chocolate plátano patatas garbanzos y las cebras de los guisantes	Reconoce la totalidad del fragmento	

<b>Muestra</b>		40
<b>Procedencia del hablante</b>	Gallega (Natural de Monforte de Lemos, Lugo)	
<b>Información social</b>	Estudios universitarios	

Rasgos lingüísticos				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la -s Sinalefa	Me llamo Marta y soy de Monforte de Lemos, que es un pueblo en Galicia	Me llamo Marta y soy de Monforte de Lemos que es un pueblo en Galicia	Reconoce el yeísmo y la sinalefa	
Mantiene la -d- intervocálica Sinalefa	Mi pueblo no ha cambiado mucho de cuando yo era pequeña a ahora	Mi pueblo no ha cambiado mucho de cuando yo era pequeña a ahora		Error fonético-sintáctico: «Pequeña a ahora» entendido como «pequeña ahora» por sinalefa
Uso del pretérito perfecto simple vinculado al presente Pérdida de -d Mantiene la -s	Así que me fui a la Universidad de Alcalá de Henares	Así que me fui a la Universidad de Alcalá de Henares	Reconoce el uso del pretérito perfecto simple vinculado al presente y la pérdida de -d	
Mantiene la -r Pronunciación tensa de la fricativa velar sorda /x/ Uso explícito del pronombre sujeto	En la primera imagen yo puedo ver a un cazador	En la primera imagen yo puedo ver a un cazador	Reconoce el uso explícito del pronombre sujeto, así como la pronunciación tensa de la fricativa velar sorda /x/	
Mantiene la -s y -r Variación léxica (patos, faisanes)	El cazador pues está cazando pues patos o faisanes	El cazador pues está cazando pues patos o faisanes	Reconoce la totalidad del fragmento	
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -s	Tiene un jardín con una valla y unas flores, arbustos y árboles frutales	Tiene un jardín con una valla y unas flores arbustos y árboles frutales	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	

Mantiene la -s Preferencia por los diminutivos en <i>-ito</i> Pronunciación tensa de la fricativa velar sorda /x/ Variación léxica ( <i>judías, patatas</i> ) Error léxico por influencia del inglés ( <i>barra de chocolate = chocolate bar</i> )	Es un carrito de la compra, en el que hay judías, garbanzos, patatas, unos plátanos y una barra de chocolate	Es un carrito de la compra en el que hay judías garbanzos patatas unos plátanos y una barra de chocolate	Reconoce el uso de los diminutivos en <i>-ito</i> , la pronunciación tensa de la fricativa velar sorda /x/ y algún error léxico por influencia del inglés	
--	--	--	---	--

<b>Muestra</b>	41			
<b>Procedencia del hablante</b>	Gallega (Natural de República Dominicana. Desde los seis años, vive en Monforte de Lemos, Lugo)			
<b>Información social</b>	Estudios medios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la -s	Me llamo Claudia y soy de Monforte de Lemos, Lugo	Me llamo Claudia y soy de Monforte de Lemos Lugo	Reconoce el yeísmo	
Mantiene la -s Uso etimológico de los pronombres clíticos Pronunciación tensa de la fricativa velar sorda /x/ Variación léxica ( <i>jovial</i> )	Antes lo recordaba como más jovial, juvenil	Antes lo recordaba como más jovial juvenil	Reconoce el uso etimológico de los pronombres clíticos y la pronunciación tensa de la fricativa velar sorda /x/	



Mantiene la -s	Cuando era más pequeña y estaba en él	Cuando era más pequeña y estaba en el	Reconoce la totalidad del fragmento	
Uso del pretérito perfecto simple vinculado al presente	Lo que viene siendo estéticamente no cambió mucho	Lo que viene siendo estéticamente no cambió mucho	Reconoce el uso del pretérito perfecto simple vinculado al presente	
Yeísmo Mantiene la -s Uso de la construcción <i>llevar + participio</i>	Llevo ya tres años trabajados en una productora audiovisual	Llevo ya tres años trabajados en una productora audiovisual	Reconoce el yeísmo y el uso de la construcción <i>llevar + participio</i> ( <i>llevo trabajados</i> )	
Mantiene la -s y -r Pronunciación tensa de la fricativa velar sorda /x/	Por temas de trabajo, tengo que viajar bastante	Temas de trabajo tengo que viajar bastante	Reconoce la totalidad del fragmento	
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -s Variación léxica ( <i>aves, escopeta, camuflaje</i> )	Una persona con su traje de camuflaje, su escopeta y apuntando a unas aves	Una persona con su traje de camuflaje su escopeta y apunta que no sabes		Error léxico: «Apuntando a unas ave» entendido como «apunta que no sabes»
Pronunciación tensa de la fricativa velar sorda /x/ Uso del pretérito perfecto simple vinculado al presente	El dibujo que siempre se hizo	El dibujo que siempre se hizo	Reconoce la pronunciación tensa de la fricativa velar sorda /x/ y el uso del pretérito perfecto simple vinculado al presente	
Mantiene la -s Preferencia por los diminutivos en -ito	Es un carrito de la compra con comida dentro	Es un carrito de la compra con comida dentro	Reconoce el uso de los diminutivos en -ito	

<b>Muestra</b>	42			
<b>Procedencia del hablante</b>	Gallega (Natural de Monforte de Lemos, Lugo)			
<b>Información social</b>	Estudios universitarios			
<b>Rasgos lingüísticos</b>				
Rasgos	Ejemplo en la grabación	Transcripción de Siri (aciertos / errores)	Tipos de aciertos	Tipos de errores
Yeísmo Mantiene la -s y -z	Me llamo Sabela Fernández Fernández, soy de Monforte de Lemos, Lugo, Galicia, España	Me llamo Sabela Fernández Fernández soy de Monforte de Lemos Lugo Galicia España	Reconoce la totalidad del fragmento	
Mantiene la -s Leísmo	Para mí Monforte de Lemos... le tengo muy buen recuerdo	Para mí Monforte de Lemos lee tengo muy buen recuerdo		Error sintáctico: «Le tengo» entendido como «lee tengo» por leísmo
Uso del pretérito perfecto simple vinculado al presente Mantiene la -s	Nunca me dio como una sensación de claustrofobia	Nunca me dio como una sensación de claustrofobia	Reconoce el uso del pretérito perfecto simple vinculado al presente	
Mantiene la -s, -r y -d Mantiene la -d- intervocálica	Al vivir en Coruña, he compartido piso con personas de mi ciudad de nacimiento	Al vivir en Coruña he compartido piso con personas de mi ciudad de nacimiento	Reconoce la totalidad del fragmento	
Pronunciación tensa de la fricativa velar sorda /x/ Mantiene la -s Variación léxica (lago)	La primera imagen la describiría como un escenario de caza, cerca de un lago	La primera imagen la describiría como un escenario de caza cerca de un lago	Reconoce la pronunciación tensa de la fricativa velar sorda /x/	

Mantiene la -s Omisión de la preposición <i>de</i>	La describiría como la casa ensueño	La describiría como la casa en sueño		Error léxico-sintáctico: «Casa ensueño» entendido como «casa en sueño» por omisión de la preposición <i>de</i>
Mantiene la -s Yeísmo Acortamiento (y <i>na</i> )	Es una fotografía con trazos infantiles, colores muy llamativos y... y nada	Es una fotografía con trazos infantiles colores muy llamativos y y nada	Reconoce el yeísmo y el acortamiento de y <i>nada</i> por y <i>na</i>	
Preferencia por los diminutivos en <i>-ito</i>	Un carrito de compra	Un carrito de compra	Reconoce el uso de los diminutivos en <i>-ito</i>	


#### IV. Protocolo de encuesta para el Trabajo Fin de Grado

##### IV.1 INSTRUCCIONES SOBRE EL DESARROLLO DE LA ENCUESTA

Al no poder desplazarnos, se les pedirá a los interesados que quieran realizar la encuesta que se graben con sus teléfonos móviles (con la grabadora o a través de las notas de voz de la aplicación WhatsApp) y que nos envíen los archivos de audio a través de WhatsApp a alguno de nuestros números de teléfono o por correo electrónico a esta dirección que hemos creado expresamente con este propósito ([tfgvariedades21@gmail.com](mailto:tfgvariedades21@gmail.com)).

##### IV.2 PREGUNTAS

PREGUNTA	OBJETIVO PERSEGUIDO CON LA PREGUNTA
Empiece la grabación diciendo su nombre y lugar de nacimiento (ciudad/pueblo y país). Es importante que diga « <i>Me llamo...</i> » para que todas las encuestas empiecen igual. Por ejemplo, « <i>Me llamo Ana y soy de Pozoblanco, un pueblo de Córdoba, España</i> ».	Pronunciación de <i>ll</i> - yeísmo (Me llamo) Pronunciación de <i>s</i> - seseo, ceceo, heheo Otros rasgos de los diferentes dialectos del español

<p>¿Podría decirnos cómo era su ciudad/pueblo cuando era pequeño/a?</p>	<p>Pluralización del verbo <i>haber</i> impersonal (<i>Habían parques...</i>).</p> <p>Rasgos del andaluz como lenición: fricativización del sonido <i>ch</i> (<i>Cuando era chico...</i>)</p>
<p>¿Estudia o trabaja en la actualidad? Si trabaja, ¿cuál es su profesión? Si estudia, ¿qué estudia o ha estudiado? Si estudia o ha estudiado en la universidad, indique cuál, por favor.</p>	<p>Uso de pretérito indefinido o pretérito perfecto compuesto</p>
<p>¿Ha vivido alguna vez fuera de su lugar de nacimiento? ¿Dónde? ¿Cuánto tiempo? ¿Cómo se comunicaba con su familia desde allí?</p>	<p>Importante por razones sociolingüísticas: debemos saber si ha estado en contacto con alguna otra variedad del español.</p> <p>También para analizar el uso de los pronombres <i>le/les/lo/los/la/las</i> referidos a personas («<i>A mi madre la llamaba por teléfono</i>» o «<i>A mi madre le llamaba por teléfono</i>»)</p>
<p>Describe la siguiente imagen:</p>  <p>The illustration shows a man in a green jacket and cap holding a long rifle, standing in a grassy field. A brown dog is on the ground next to him. In the background, there are trees, a blue sky with birds flying, and a body of water.</p>	<p><b>Nivel fónico:</b></p> <p>Seseo, ceceo o heheo (a través de la pronunciación de las palabras <i>cazador</i>, <i>cazando</i>, etc.)</p> <p>Pérdida de <i>-d-</i> intervocálica (<i>al lado</i> o <i>al lao</i>)</p> <p>Conservación de consonantes finales (<i>pájaros</i>, <i>árboles</i>)</p> <p>Pérdida de consonantes finales</p> <p>Mantenimiento de <i>-s</i> final de sílaba (<i>escopeta</i>)</p> <p>Pérdida o aspiración de <i>-s</i></p> <p>Aspiración de la fricativa velar sorda /x/ (<i>pájaro</i>)</p> <p><b>Nivel morfosintáctico:</b></p> <p>Queísmo</p> <p>Dequeísmo</p> <p>Leísmo</p> <p>Loísmo</p> <p>Laísmo</p>
<p>Describe la siguiente imagen:</p>	<p><b>Nivel fónico:</b></p> <p>Seseo, ceceo o heheo (a través de la pronunciación de las palabras <i>sol</i>, <i>casa</i>, <i>manzana</i>, etc.)</p> <p>Debilitamiento u omisión de <i>-d-</i> (<i>al lado</i> o <i>al lao</i>)</p> <p>Conservación de consonantes finales (<i>flores</i>, <i>árboles</i>, etc.)</p> <p>Pérdida de consonantes finales (<i>árbo</i> - <i>árbol</i>)</p>



Mantenimiento de -s  
 Debilitamiento o aspiración de -s

**Nivel morfosintáctico:**

Queísmo  
 Dequeísmo  
 Leísmo  
 Loísmo  
 Laísmo

Describe la siguiente imagen:



**Nivel fónico:**

Seseo, ceceo o heceo (a través de la pronunciación de la palabra *garbanzos*)

Debilitamiento u emisión de d (*al lado* o *al lao*)

Conservación de consonantes finales

Pérdida de consonantes finales

Mantenimiento de -s

Pérdida o aspiración de -s final de sílaba

Lenición: fricativización de *ch* (chocolate)

Aspiración de la fricativa velar sorda /x/ (judía)

**Nivel morfosintáctico:**

Queísmo  
 Dequeísmo  
 Leísmo  
 Loísmo  
 Laísmo

**Nivel léxico:**

Patatas/papas

Plátanos/bananas

Judías

Habichuelas

Garbanzos

Chícharos

Guisantes

## **V. Informe favorable del Comité de Ética**

En la siguiente página se encuentra disponible el informe favorable a esta investigación que emitió el Comité de Ética en Investigación de la Universidad de Granada.



## COMITE DE ETICA EN INVESTIGACION DE LA UNIVERSIDAD DE GRANADA

La Comisión de Ética en Investigación de la Universidad de Granada, visto el informe preceptivo emitido por la Presidenta del Comité en Investigación Humana, tras la valoración colegiada del Comité en sesión plenaria, en el que se hace constar que la investigación propuesta respeta los principios establecidos en la legislación internacional y nacional en el ámbito de la biomedicina, la biotecnología y la bioética, así como los derechos derivados de la protección de datos de carácter personal,

Emite un Informe Favorable en relación a la investigación titulada: 'SESGOS EN EL SOFTWARE DE PROCESAMIENTO DEL LENGUAJE NATURAL DE LA APLICACIÓN SIRI EN DISTINTAS VARIEDADES DEL ESPAÑOL DE ESPAÑA' que dirige D./Dña. BEGOÑA FERNÁNDEZ MARTÍNEZ, con NIF 31.019.668-M, quedando registrada con el nº: 2170/CEIH/2021.

Granada, a 13 de Mayo de 2021.

EL PRESIDENTE  
Fdo: Enrique Herrera Viedma

EL SECRETARIO  
Fdo: Francisco Javier O'Valle Ravassa