

**GLASÚN 1.0: A NEW DATA MANAGEMENT  
AND PROMPTING SYSTEM FOR RESEARCH  
IN ACOUSTIC PHONETICS**

**GLASÚN 1.0: UN NUEVO SISTEMA  
DE GESTIÓN DE DATOS PARA LA INVESTIGACIÓN  
EN FONÉTICA ACÚSTICA**

MARK GIBSON  
*Universidad de Navarra*  
mgibson@unav.es

IGNACIO DE LORENZO RODRÍGUEZ  
*Universidad de Navarra*  
idelorenzo@unav.es

*Artículo recibido el día: 31/10/2013*  
*Artículo aceptado definitivamente el día: 30/06/2014*  
*Estudios de Fonética Experimental, ISSN 1575-5533, XXIII, 2014, pp. 247-264*

---

### ABSTRACT

This methodological study presents the results of an initiative to create a new data management and prompting system which can be interfaced with acoustic data for research in acoustic phonetics. The objectives of this initiative were to provide an economical way to record, organize and manipulate acoustic data and to construct a system which allows the use of external recording equipment for production experiments. For the project, a program was designed with a built-in visual prompting system which simultaneously synchronizes the text in accordance with defined temporal regions in the long file for easy extraction for use in perception experiments. Results of external assessment were considerably positive due to the user-friendly platform which does not require scripting, meaning the organization and manipulation of data files is less time-consuming and more efficient.

Keywords: *data management system, acoustic phonetics research, visual prompting.*

### RESUMEN

Este estudio metodológico presenta los resultados de una iniciativa para crear un nuevo sistema de gestión de datos que se puede interconectar con un archivo acústico para la investigación en fonética acústica. Los objetivos de esta iniciativa eran crear una forma eficiente de registrar, organizar y manipular los datos acústicos y de diseñar un sistema que permita el uso de un equipo de grabación externo para los experimentos de producción. Para el proyecto, se diseñó un programa que presenta el corpus al sujeto de una manera visual y organiza de forma simultánea los datos de acuerdo con las regiones temporales definidas en el archivo largo para una fácil extracción para su uso en experimentos de percepción. Los resultados de la evaluación externa fueron en general positivos debido a la plataforma que no requiere shell scripting, lo cual significa que la organización y la manipulación de los archivos de datos eran menos laboriosas y más eficientes.

Palabras clave: *sistema de gestión de datos, investigación en fonética acústica, presentación visual del corpus.*

## 1. INTRODUCTION

The following initiative arose out of the need to create a data management system for acoustic data collection which permitted the use of an external recording system during the data collection phase of a Master's thesis entitled *L'aspiration du /s/ en position coda dans l'espagnol d'Andalousie occidentale : analyse acoustico-perceptive autour de l'ambigüité du message* (Ivent, 2013) at the *Université Sorbonne Nouvelle – Paris 3*. This acoustic-perceptual study presented a variety of challenges for conventional acoustic analysis. The first of these challenges was related to the topic under study, aspiration of coda /s/ in the Western Andalusian dialect of Southern Spain. In this dialect, /s/ is aspirated or eliminated in a gradient way in a variety of prosodic, phonological, morphological, syntactic and pragmatic contexts.

As has been well established in the literature, the aspiration of /s/ in Spanish involves the substitution of the oral features of /s/ by a glottal [h], characterized by a gradient current of air with no frication quality. Articulatorily, aspiration is characterized by a parameter setting specifying a wide glottis which is superimposed on the form of the oral cavity (for a partial review of /s/ aspiration and deletion in Spanish, as well as their residual effects, please see Alonso et al 1950; Salvador 1957; Alarcos Llorach 1958; Mondéjar 1970; Contreras Jurado 1975; Salvador 1977; Terrell 1979; Lipski 1983, 1984, 1986; Penny 2000; Gerfen 2002; O'Neill, 2005, among many others). Ladefoged and Maddieson (1996:50) suggest that for breathy voice (most akin to the process of aspiration in Spanish) the rate of airflow through the glottis for an average adult male registers between 500-250 ml/s.

Aspiration and deletion manifest in a number of ways in the spectrogram. Features of the segment may appear in the temporal position of the gesture they replace (/s/), but the effects of aspiration also appear as residual effects over the following consonant (and emerge as a post aspiration for the consonant at the right margin of the segment). In many cases, the energy concentrates around the higher frequencies, making it difficult to discern clear cases of aspiration from ambient noise. The extent to which this phonetic cue is perceptibly salient, or at least audible, depends on many factors such as the prosodic position and the following consonant. Often, the timing of aspiration does not coincide with that of /s/ and the only phonetic cues available for analysis reside in the residual effects on the following or preceding segment. In the event that the following segment is consonantal, aspiration often centers as noise around F4 and F5 (figure 1):

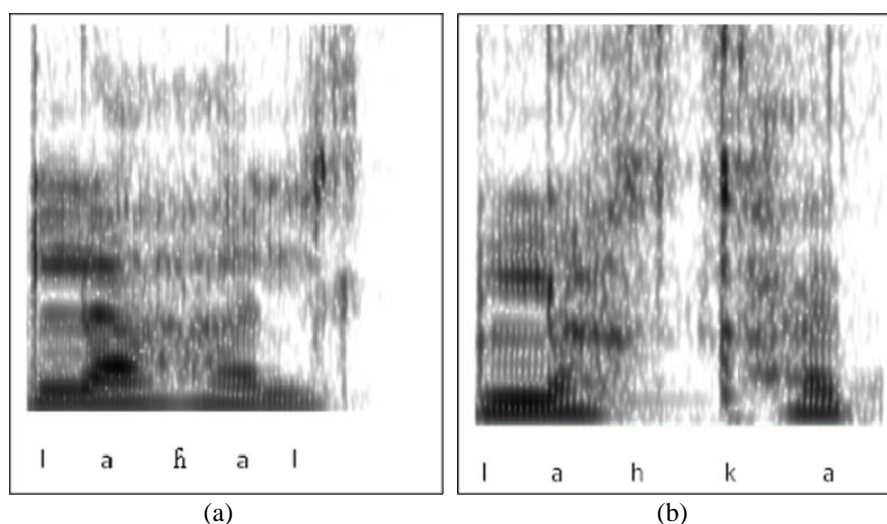


Figure 1. *The first image (a) shows aspiration in the intervocalic context (las alas ‘the wings’). The second image (b) is an example of aspiration in coda position preceding a stop across word boundaries. As opposed to standard segments in the Spanish phonological inventory, there is relatively low stability with regard to phonetic realization.*

The variant nature of the segment in the phonetic environment and the fact that the effects of aspiration are often times not discernible from ambient noise motivated the need to collect the data for the study in an unconventional way. The fact that the acoustic data would be used for perception testing following the post-data processing required that there be minimal ambiguity in the auditory signal which may distort the subjects’ responses.

The second challenge which emerged in the planning stages of the study was related to the nature of the corpus. For the study, we wanted to collect samples of aspiration, deletion and full realization of /s/ in a variety of contexts. One of the weaknesses we discovered in past studies treating the aspiration and deletion of /s/ was that the corpus was restricted to very few phonological, prosodic and morphological contexts. This means that statistics may be skewed with regard to the emergence of aspiration and deletion, and the environment which triggers these processes.

The corpus for the present study contained 328 total phrases, of which 151 contained the target structure (20 tokens of lexical /s/ and 131 tokens of morpheme /s/). The remaining 177 tokens contained the target structure in different prosodic contexts. Three repetitions of the entire corpus were solicited by two native speakers of Western Andalusian Spanish, rendering a total of 1968 phrases for post-data processing. Given the robust quantity of tokens, an economical way to organize, fragment and annotate the tokens was crucial.

These two challenges were resolved by designing a program which permitted the use of a powerful recording design and facilitated the fragmentation of the acoustic data. The program was designed for use when laboratory instrumentation and environmental controls are either insufficient or unavailable. For the present study, the program was designed for use in the 98.3 Radio recording studio of the University of Navarra because of the superior quality of the recording equipment and the double-insulated sound-attenuating booth.

## 2. ACOUSTIC DESIGN

The data for the acoustic experiments were to serve two purposes. First, a production analysis was performed in order to categorize the acoustic parameters of aspiration and deletion in syllable-final position. Afterward, the data were to be extracted from their carrier phrases and used as stimuli for a series of perception tests. To collect the data, a recording design was devised which used three Sennheiser 441 professional microphones connected to an AEQ BC2500 audio mixer which directed the acoustic signal to a personal computer outfitted with an M-audio fast-track pro sound card (permitting a resolution of 16 bytes and signal compression of 256 KB/second (figure 2):

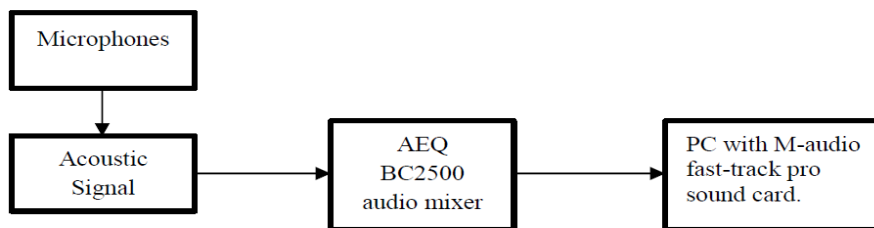


Figure 2. *Design scheme to extract the acoustic signal.*

---

The acoustic signal was channeled through a super-cardioid microphone for use in speech recording. The particular model used for data collection rejects auditory feedback and allowed for accurate signal response and low distortion at high sound pressure levels. A compensation coil prevented ambient noise from registering in the sound file.

The signal was directed to the audio mixer where it was pre-processed by a radiofrequency filter of 100 (Hz) before being directed and registered to the PC sound card. This process permitted the adjustment of signal saturation by an LED (Light Emitting Diode) which alerts to maximum levels. A potentiometer regulated channel output signal balance. The result was a registered signal with minimum noise and natural saturation levels.

An external sound card was chosen to register the signal in order to avoid EMI/RF (electromagnetic interference/radio frequency) and PSU (power supply unit) noise. The signal was received by the external sound card and registered digitally before being transferred to the PC for analysis.

The subjects for the production study repeated the tokens contained in the corpus and the acoustic system registered the repetitions in one long file. The data management system was superimposed over the acoustic file, and the temporally defined acoustic regions were synchronized to the data text automatically. Thus, two types of files were generated for the two distinct purposes of the study: (1) long file which permitted signal analysis using Matlab, and (2) automatically fragmented acoustic files to allow easy extraction for perception stimuli. In the following section, a description of the data management system will be outlined.

### **3. DATA MANAGEMENT SYSTEM: GLASÚN 1.0**

The acoustic data were collected for both production analysis and use in perception tests. For the former, long files are desirable in order to easily extract duration and frequency values by way of a script. However, very long acoustic files with various repetitions per token render extraction for perception tests time-consuming.

The data management system interfaces two readily available programs: Microsoft Access, which was amplified and modified with Visual Basic (VB), and Sound Forge. For simple functions such as window design and display, Access was employed. The more complex functions such as the exportation of temporal

regions and recording control were structured with VB codes (but may be done in Java). The information in the data base was organized according to the scheme in figure 3:

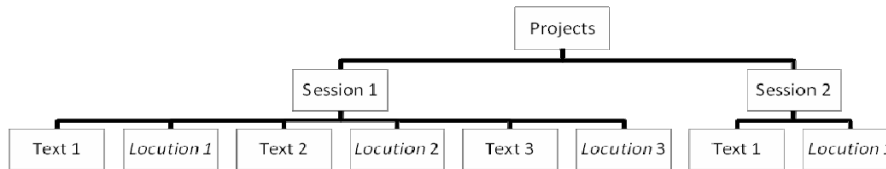


Figure 3. *Conceptual scheme of GLASÚN 1.0.*

The different elements of the data base include:

1. *Projects* (*proyectos* in the original Spanish version): can be composed of one or various sessions. This permits the transferal of data pertaining to different projects within the same system.
2. *Sessions* (*sesiones*): within the same project corresponds to diverse recordings of various texts.
3. *Texts* (*textos*): correspond to the tokens which will be recorded in each session.
4. *Locutions* (*locuciones*): are the result of the distinct locution sessions. These locutions are non-existent until recording is realized.

The texts are shared between the distinct projects and sessions in such a way that within the same projects, the sessions can be varied to include determined texts, but not others.

This organization is expressed in the data base in the following way:

1. *Projects* (fields: Id, name of project, author, starting date of the project).

2. *Sessions* (fields: Id, Locutor, Id of the project, date of recording, starting and end times of the recording sessions).
3. *Text* (fields: Id of interlocutor and text to be recorded).
4. *Locutions* (fields: Id of interlocutor, text to be read, start and end times for recordings).
5. *Unisesiontexto*: This is a table which permits association between texts and sessions, as well as their organization. The fields are (1) identification of subject, (2) text, (3) session and (4) order. This table interfaces the identification fields of the subjects of the individual sessions as well as the text.

These data are displayed afterward in a window which facilitates data management and manipulation (figures 4 & 5):



Figure 4. Main window of the application.



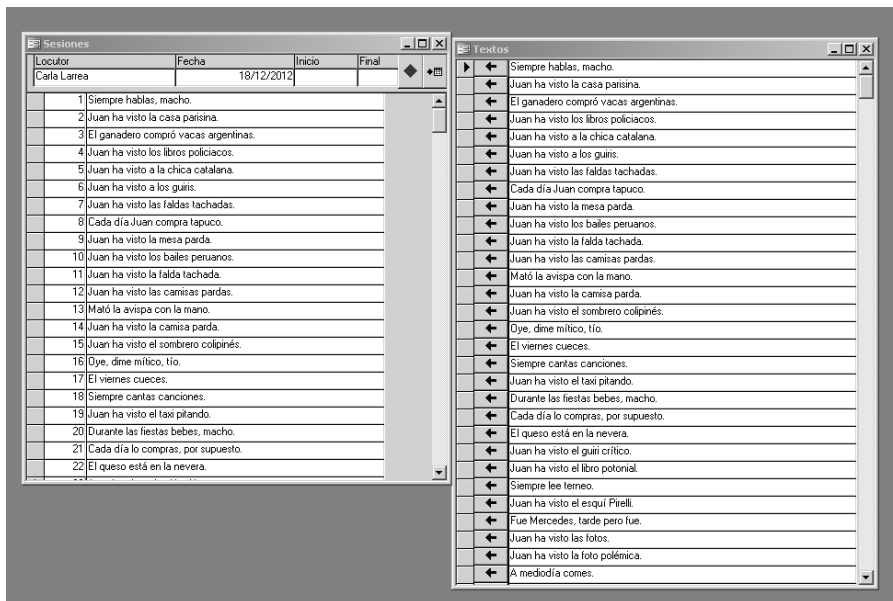


Figure 5. Window of a program session. To the right we can see the randomized text which the author enters into the system. To the left, the text is ordered for recording. The system also permits a randomization function for data sets.

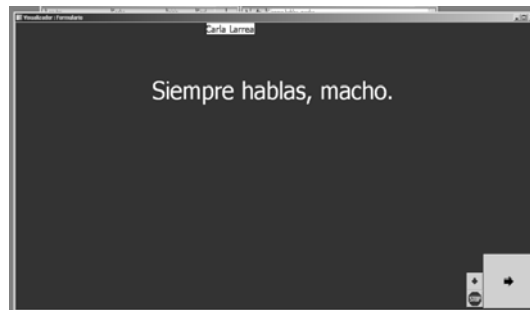


Figure 6. Prompt window. The button which appears at the bottom right-hand corner is used by the controller to advance to the next prompt.

The prompt window which presents the text to the subject appears in the figure 6. This window is interfaced with an internal chronometer which registers the start and end times of the locutions associated with that particular window.

This window synchronizes the location of the prompt with the recording using Sound Forge 9.0. Effectively, the saved text is superimposed over the sound file, becoming synchronized to permit easy access and organization for the perception tests.

Once the recording is finished, the program exports the duration data of all the locutions to a text file in txt. and Windows Media Script. When the temporal parameters of the locutions are introduced in Sound Forge for post-data editing, the long file can be fragmented according to the temporal regions defined for each locution and saved as individual files. This proves to be especially beneficial for large corpuses in which non-target tokens are randomized with target structures (figure 7):

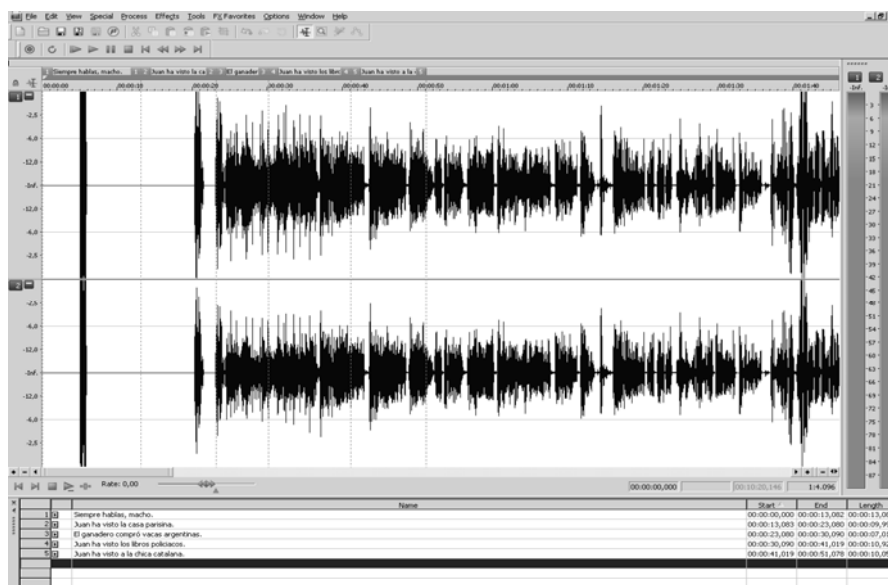


Figure 7. Sound Forge 9.0 Window. In the bottom portion of the screen we can see the created regions which the author chooses to separate from the long file.

Afterward, the text sentences and their synchronized sound file can be reorganized in any number of ways to facilitate isolation of a single group of tokens (figure 8).

Id	Texto
309	A Marta le encantan las obras de Picasso
242	A mediodía come perritos.
344	A mediodía come, claro.
82	A mediodía comes, claro.
41	A mediodía comes.
194	A menudo Lola va cansada.
196	A menudo Lola va, sí.
243	A menudo Maribel va caluda.
336	A menudo vas cansada.
214	A menudo vas, sí.
235	A menudo vas.
236	Ahora alquila casas.
251	Ahora alquilas casas.
213	Ahora con el trabajo, paso menos.
241	Ahora las alquila, chiquillo.
153	Ahora las alquilas, chiquillo.
107	Ahora las alquilas.
160	Ahora las alquilas.
239	Ahora las alquilas.
42	Ahora, Luis alquila cats.
328	Atento lo aprende, sí.
209	Atento lo aprendes, sí.
323	Atento, aprende teorías.
335	Atento, aprendes teorías.
149	Atento, Juan aprende telomanas.
326	Atento, lo aprende.
159	Atento, lo aprendes.
199	Cada día compra tabaco.
276	Cada día compra tabaco.
19	Cada día Juan compra tabaco.
304	Cada día lo aprende, por cuarenta.

Figure 8. The corpus is arranged in this window alphabetically (although other options exist) in order for similar tokens to appear together for easy access. The original position in the corpus and long file appears to the left of the carrier sentence.

The corpus for the study was randomized and saved as a simple unformatted text file which was uploaded to the system without shell-scripting. The corpus is shown as individual text sentences which appear as a prompt for the subject to read.

#### 4. EXTERNAL ASSESSMENT OF THE ACOUSTIC DESIGN AND DATA MANAGEMENT SYSTEM

To evaluate the efficacy of the present initiative both qualitative and quantitative evaluations were solicited. For both the data design scheme and quality of the acoustic files, the data were evaluated both by students and trained phoneticians at universities in three countries: Spain, France and the United States (four participants per country, rendering twelve evaluators in total (six per group for the following tasks)). With regard to the quality of the acoustic files, the evaluators were asked to segment two files for the same speaker using Praat and give their opinions based on a five point scale; 5 being the best quality, 1 being the worst.

---

Evaluators were also asked to compare the acoustic files recorded using the system outlined in section 2, with an identical file which was recorded using an internal sound card (recorded under studio conditions). Qualitative comments were also solicited at the end of the evaluation.

In general the data recorded using the professional-level external microphone system described in section 2 were rated consistently higher on all assessment points: (1) absence of ambient noise, (2) clarity of auditory signal, (3) voice quality, and (4) identification of target feature (aspiration). The elimination of ambient noise and clarity of the auditory signal were reported to be the most positive aspects (absence of ambient noise: mean score of 4.3 for the system outlined in section 2 versus a mean score of 3.4 for the same corpus recorded using an internal sound card; signal clarity: mean score of 4.6 for our system versus 3.8 for recording with internal microphone). Those evaluators familiar with aspiration reported quantitative differences in duration (up to 20 ms.) when comparing the two types of files, due to the fact that the acoustic properties of aspiration were more clearly discernible in the files registered by the external sound equipment (presumably due to the clarity of the signal upon diminishing the ambient noise).

With regard to the data management system, two groups of evaluators were chosen (from the twelve evaluators mentioned above) to assess the efficacy of the system and two different tasks were outlined. The first task involved fragmenting the long sound file into individual sentence files for one speaker (984 total sentences). As this is done automatically by GLASÚN 1.0 as the sound file is being recorded, there was no need to measure results for this system. However, many systems currently used for acoustic data collection and analysis either require manual fragmentation, the use of a script to match manually segmented intervals to text, or a combination thereof (Praat, Sound Forge, Vegas Pro, MatLab, Esps Xwaves). For this task, Group 1 was given a text file with the individual sentences and asked to mark the sentence's temporal intervals on a Praat Text Grid. They were then provided a script (see Annex I) which reads the text file and inserts the text to the intervals marked on the TextGrid. Group 2 was asked to perform the procedure by hand from the long sound file, copying and pasting to make individual files for each prompt. Both Groups were asked to record the time spent on this procedure.

The second task involved extracting individual tokens from the long file, a necessary yet time-consuming procedure for the perception tests. For this task, the first group of evaluators was asked to extract a short list of tokens (50) from the long file and record the time they spent in doing so. They were given a copy of the corpus to aid in locating the tokens. The second group was instructed how to organize the corpus using the functions in GLASÚN 1.0. They were given the

same list of tokens as Group 1, in addition to a copy of the original corpus, and instructed to extract the words and record their time as well.

For the first task, Group 1 encountered many problems related to the script due to runtime errors (reasons for errors were not disclosed). The average time for those evaluators who finished the task was roughly one hour and twenty minutes. Group 2 had relatively few problems given the archaic nature of the task, but spent more time to complete the task (two hours and fifteen minutes average, yet one reported to have stopped before finishing). For the second task, the average time to locate and extract the tokens for Group 1 was twenty-seven minutes and thirty-four seconds, while that time was greatly reduced for Group 2; fifteen minutes and twenty-two seconds.

For both tasks, the efficacy of the data management system presented here rendered superior results. The data files were easier and less time-consuming to organize and manipulate.

## 5. CONCLUSIONS

This study presented the results of a new data management system for research in acoustic phonetics. Assessment of the system by external reviewers provides positive feedback related to the usefulness of this system for an array of different purposes in phonetic research. Aside from the benefits reported by external review, the system is also robustly adequate for data collection for fieldwork settings in which large corpuses are recorded in a short time and under non-laboratory conditions. Although for the pilot study the corpus was defined previously to data collection and interfaced directly with the acoustic recording, the system also permits post facto interfacing between the corpus file and acoustic file in order to collect and organize natural speech data.

Currently, the system is undergoing revision to integrate a data analysis system (spectrographic software) into the main program. In addition, the program is being developed using freeware programs to support mass use and diffusion of the system in academic settings. Finally, the possibility of developing the system for use with articulatory instrumentation such as aerodynamics is being explored as a source for future enhancements.

---

## 6. BIBLIOGRAPHIC REFERENCES

- ALARCOS LLORACH, E. (1958): «Fonología y fonética: a propósito de las vocales andaluzas», *Archivum*, VIII, pp. 193-205.
- ALONSO, D. (1950): «Vocales andaluzas. Contribución al estudio de la fonología peninsular», *Nueva Revista de Filología Hispánica*, 4, 3, pp. 209-230.
- CONTRERAS JURADO, A. (1975): «Vocales abiertas del plural en andaluz oriental ¿fonemas o prosodemas?», *Yelmo*, 26, pp. 23-25.
- GERFEN, C. (2002): «Andalusian codas», *Probus*, 14, 2, pp. 303-333.
- IVENT, F. (2013): *L'aspiration du /s/ en position coda dans l'espagnol d'Andalousie occidentale: analyse acoustico-perceptive autour de l'ambigüité du message*, M.A. thesis (unpublished), Université Sorbonne Nouvelle-Paris 3.
- LADEFOGED, P. & I. MADDIESON (1996): *The Sounds of the World's Languages*, Oxford, Blackwell.
- LIPSKI, J. (1983): «Reducción de /s/ en el español de Honduras», *Nueva Revista de Filología Hispánica*, 32, pp. 272-88.
- LIPSKI, J. (1984): «/s/ in the Spanish of Nicaragua», *Orbis*, 33, pp. 171-181.
- LIPSKI, J. (1986): «Instability and reduction of /s/ in the Spanish of Honduras», *Revista Canadiense de Estudios Hispánicos*, 11, pp. 27-47.
- O'NEILL, P. (2005): «Utterance final /S/ in Andalusian Spanish. The phonetic neutralization of a phonological contrast», *Language Design*, 7, pp. 151-166.
- MONDÉJAR, J. (1970): «*El verbo andaluz. Formas y estructuras*», Madrid, C.S.I.C. (annex X of RFE).
- PENNY, R. (2000): *Variation and change in Spanish*, Cambridge, Cambridge University Press.
- SALVADOR, G. (1957): «El habla de Cúllar-Baza (Contribución al estudio de la frontera del andaluz)», *Revista de Filología Española*, XLI, pp. 161-249.

SALVADOR, G. (1977): «Unidades fonológicas vocálicas en andaluz oriental», *Revista de la Sociedad Española de Lingüística*, 7, pp. 1-23.

TERRELL, T. (1979): «Final /s/ in Cuban Spanish», *Hispania*, 62, 4, pp. 599-612.

---

**ANNEX I. PRAAT SCRIPT USED BY ASSESSORS TO MATCH TEXT TO TIME INTERVALS.**

(Downloaded at: [http://www.helsinki.fi/~lennes/praat-scripts/public/label\\_from\\_text\\_file.praat](http://www.helsinki.fi/~lennes/praat-scripts/public/label_from_text_file.praat).

Author: Mietta Lennes)

```
# This script reads lines from a text file (called labels.txt
and saved in the home directory)
# and adds them line by line as labels for intervals in a
selected TextTier in the selected TextGrid object.
#
# You should check that the boundaries are correct before
running the script.
# The script will jump over intervals labeled as "xxx". Use
this marking if there are intervals that
# you will remove later.
# Hint: This tool is useful if you use the mark_pauses script
before it! It is easy to check that the
# pause boundaries are in correct places, if you know what
kind of content should be in the sound
# segments - e.g., read sentences.
#
#
soundname$ = selected$ ("TextGrid", 1)
select TextGrid 'soundname$'
stringlength = 0
filelength = 0
firstnewline = 0
oldlabel$ = ""
newlabel$ = ""
form Label intervals in an IntervalTier from text file
comment Give the path of the text file containing the label
lines:
sentence Filename /home/lennes/labels.txt
comment Which IntervalTier in the selected TextGrid do you
want to label?
integer Tier_(index) 1 (= the first IntervalTier)
comment Which interval do we start labeling from?
integer Starting_interval_(index) 1 (= the first interval)
comment Do you want to overwrite old labels?
comment (Intervals previously marked with 'xxx' will be
skipped despite this!)
boolean Overwrite 1
endform
```



---

```

if fileReadable (filename$)
numberOfIntervals = Get number of intervals... tier
if starting_interval > numberOfIntervals
exit There are not that many intervals in the IntervalTier!
endif
leftoverlength = 0
# Read the text file and put it to the string file$
file$ < 'filename$'
if file$ = ""
    exit The text file is empty.
endif
filelength = length (file$)
leftover$ = file$
# Loop through intervals from the selected interval on:
for interval from starting_interval to numberOfIntervals
oldlabel$ = Get label of interval... tier interval
if oldlabel$ <> "xxx"
# Here we read a line from the text file and put it to
newlabel$:
firstnewline = index (leftover$, newline$)
newlabel$ = left$ (leftover$, (firstnewline - 1))
leftoverlength = length (leftover$)
leftover$ = right$ (leftover$, (leftoverlength -
firstnewline))
# Then we check if the interval label is empty. If it is or
if we decided to overwrite,
# we add the new label we collected from the text file:
    if overwrite = 1
        Set interval text... tier interval
        'newlabel$'
    elseif oldlabel$ = ""
        Set interval text... tier interval
        'newlabel$'
    else
        exit Stopped labeling, will not overwrite old labels!
endif
endif
endif
endfor
else
    exit The label text file 'filename$' does not exist
where it should!
endif

```