THE MAPPING OF PHONETIC INFORMATION TO LEXICAL REPRESENTATIONS IN SPANISH: EVIDENCE FROM EYE MOVEMENTS

Andrea Weber[1], Alissa Melinger[1, 2], & Lourdes Lara Tapia[1]

[1] Saarland University, [2] University of Dundee
aweber@coli.uni-sb.de, a.melinger@dundee.ac.uk, lara@alsancak.de

ABSTRACT

In a visual-world study, we examined spoken-word recognition in Spanish. Spanish listeners followed spoken instructions to click on pictures while their eye movements were monitored. When instructed to click on the picture of a door (*puerta*), they experienced interference from the picture of a pig (*puerco*). The same interference from phonologically related items was observed when the displays contained printed names or a combination of pictures with their names printed underneath, although the effect was strongest for displays with printed names. Implications of the finding that the interference effect can be induced with standard pictorial displays as well as with orthographic displays are discussed.

Keywords: spoken-word recognition, eye-tracking, Spanish, pictures, printed words.

1. INTRODUCTION

It is generally assumed that the recognition of spoken words entails the competitive evaluation of multiple lexical hypotheses [for a recent overview see 1]. All word candidates that match the incoming speech signal are activated and compete with each other until they no longer agree with the acoustic input. For example, the onset of the English word *steak* will activate initially not only *steak*, but also words with similar sounds like *stay*, *stain*, and *stale*.

The visual-world paradigm has been found to be particularly suited for the investigation of phonological competitor activation [2]. The paradigm makes use of the fact that participants make saccadic eye movements to pictures of objects on a computer screen as the names of the objects are being heard. Locations and latencies of eye movements on pictures are recorded using a camera mounted on a headband. While participants hear the name of a target picture, they look more often to pictures with names that are similar in onset with the target name than to pictures with phonologically unrelated names. Fixation proportions to pictures are assumed to reflect activation levels of word candidates. The interesting aspect of eye fixations are that they provide a window into the listeners' processing when alternative word candidates still compete for recognition.

The main goal of the present study was to expand the methodology. Visual-world studies usually present listeners with pictures. This restricts the domain of investigation to depictable objects. It has often proved difficult to find enough suitable items for a particular question under investigation Recent evidence from Dutch [3, 4] suggests that written displays can also induce lexical competition effects in eye movements. We want to establish the viability of written displays by directly comparing eye movements to pictorial and written displays in one study. An extension to written displays would open up the visual-world paradigm to the investigation of a variety of research questions. The influence of presentation mode was investigated in a Spanish visual-world study.

2. EXPERIMENT

2.1. Method

2.1.1. Participants

Twenty-four native speakers of Mexican Spanish with a mean age of 25 years, took part in the experiment. They had normal or corrected-to-normal vision and normal hearing. The experiment was conducted in Germany at Saarland University. On average, participants had lived in Germany for one and a half years at the time of testing.

2.1.2. Materials

Twenty-two Spanish nouns referring to picturable objects were chosen as targets. Each target was

paired with a competitor. The onset of the competitor overlapped phonemically with the onset of the target in Spanish (e.g., target *puerta* /pu̯érta/ 'door' was paired with competitor *puerco* /pu̯érko/ 'pig'). On average, target and competitor overlapped by three phonemes at an average word length of 6 letters.

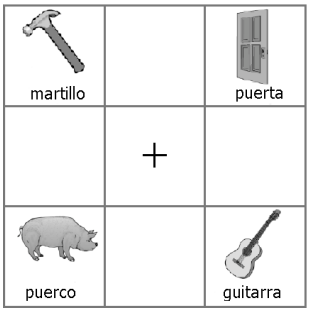
Two phonologically unrelated distractors were added for each target (e.g., *guitarra* /gitára/ 'guitar' and *martillo* /martíʝo/ 'hammer'). A target item, its competitor, and two unrelated distractors were displayed together in one trial set. The target words were actively named in the spoken instructions, whereas the competitor and the unrelated distractors were not named.

Lexical frequencies for Spanish lemmas were computed with the Corpus del Español [5]. The overall lexical frequencies for targets and competitors did not differ significantly.

To prevent participants from developing expectations that items with phonologically similar names were likely targets, 45 additional filler trials with no phonemic overlap between items were constructed. Three more representative trials were constructed as practice trials.

All items in a trial were presented to participants either as pictures, printed words, or pictures with words printed underneath (see Fig. 1). Pictures were colored line drawings, taken from the IMSI MasterClips Image Collection [6]. Printed words were written in the AvantGarde font (font size 30).

Fig. 1: Example of display with pictures and printed names.



In Spanish, definite articles mark the gender of nouns. To avoid gender cues, we by embedded targets in the article-free carrier *Selecciona:...* 'click on:…'. Spoken instructions were recorded onto minidisk by a female native speaker of Mexican Spanish, sampling at 44.1 kHz (later down-sampled to 22 kHz). Target noun onsets were measured by visual and auditory checking. The average duration of putative overlap between target and competitor (e.g., /pu̯ér/ in *puerta*) was 335 ms.

Three lists with all experimental and filler trials in pseudo-random order were constructed, such that before each experimental trial there was at least one filler trial. Experimental trials appeared once in a given list, and presentation mode (pictorial, written, pictorial+written) was counterbalanced across lists. Each list contained an equal number of filler trials in the three presentation modes.

2.1.3. Procedure

The whole experiment was conducted in Spanish. Participants were tested individually. At the beginning of a session, they were told that on each trial they should click on the object or word on the screen that was mentioned in the instructions. Instructions were presented auditorily over headphones.

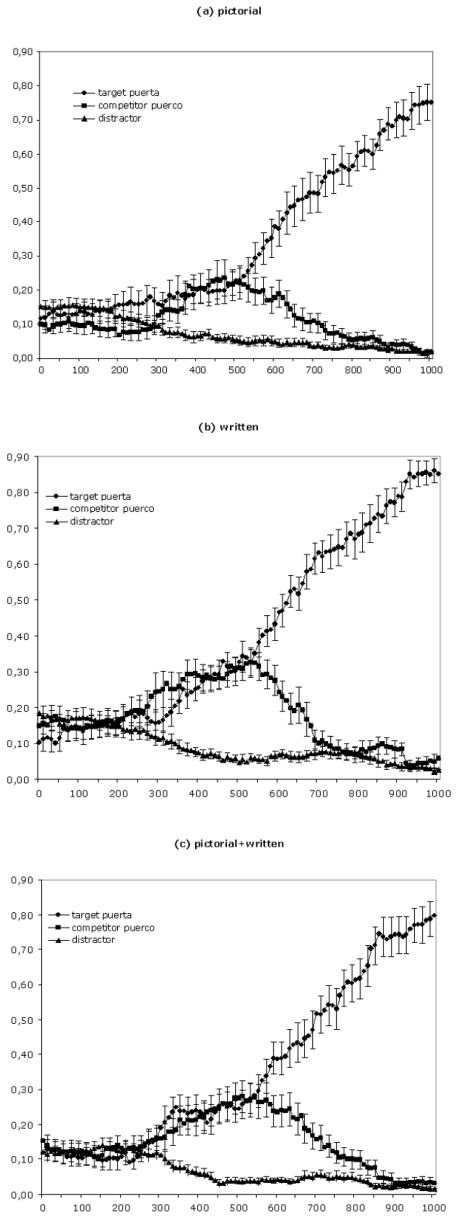
While they were listening, participants' eye movements were monitored using an SMI EyeLink head-mounted eye-tracker. Onset and offset times and the spatial coordinates of the participants' fixations were recorded (250 Hz sampling rate). All pictures were presented in color on a 3 x 3 gray grid; each cell measured 7.5 x 7.5 cm. Positions of target and competitor objects on the grid were randomized across trials.

For the analysis, graphical software was used to display the locations of the participants' fixations as dots superimposed on the grid. Fixations were coded as pertaining to the cell of the target, the competitor, one of the two unrelated distractors, or to the background. Saccade times were not added to fixation times.

2.2. Results

Two items had to be removed from the analysis due to scripting errors. For the remaining 20 items, proportions of fixations were calculated for each item type (target, competitor, disctractor), aggregated over participants or items, during successive 10 ms time frames. Fig. 2, presents the averaged proportions of fixations after target noun onset for trials with a pictorial display (Fig. 2a), a written display (Fig. 2b), and a combination of pictorial and written display (Fig. 2c) respectively. To simplify the figures, fixation proportions for the two unrelated distractors were averaged.

Fig. 2: Fixation proportions over time from target noun onset for pictorial trials (a), written trials (b), and trials with a combination of pictorial and written display (c).



For our analyses, we compared fixation for the competitor and the distractors. If all words that matched the unfolding acoustic input were activated, then target (e.g., *puerta*) and competitor (e.g., *puerco*) should compete for the listeners' visual attention. Thus, the competitor should be fixated more often than the distractors. Furthermore, we were interested in whether the presentation mode interacted with the competition effect.

Note that it typically takes about 200 ms before a programmed eye movement is launched [7]. Thus fixations that are triggered by the first 50 ms of acoustic target information are observable around 250 ms after target onset.

In all three presentation modes, Spanish participants fixated competitor pictures more often than distractor pictures (see Fig. 2). Between 250 ms and 700 ms, the proportion of fixations in pictorial trials was on average 18% for the competitor and 6% for the distractors; in trials with written displays, 26% for the competitor and 8% for the distractors; in trials with a combination of pictures and their printed names, 22% for the competitor and 6 % for the distractors.

We conducted two-factor ANOVAs on the mean proportion of fixations between 250 and 700 ms with picture type (competitor and distractors) and presentation mode (picture, written word, and combined presentations) as the within-participants and within-items factors. The analyses revealed that, across presentation modes, the competitor was fixated significantly more than the average of the unrelated distractors ($F_1[1, 23] = 49.46$, $p < .001$; $F_2[1, 19] = 42.24$, $p < .001$); there was a marginally significant difference between presentation modes ($F_1[2, 46] = 3.17$, $p > .05$; $F_2[2, 38] = 2.67$, $p > .08$), and no interaction between presentation mode and picture type ($F_1[2, 46] = 1.35$, $p > .2$; $F_2 < 1$).

In order to further investigate the difference between presentation modes, we analyzed fixation patterns in pairwise comparisons. In all three comparisons, we found a main effect of picture type and no interaction with presentation mode in the 250 to 700 time window. The only significant effect of presentation mode was found between written trials and pictorial trials ($F_1[1, 23] = 8.16$, $p < .01$; $F_2[1, 19] = 4.30$, $p > .05$), such that in written trials the competition effect was stronger. Furthermore, in written trials the competition effect started to emerge already between 250 and 350 ms after target noun onset ($F_1[1, 23] = 4.48$, $p < .05$; $F_2[1, 19] = 3.45$, $p > .05$), whereas in pictorial trials or in trials with a combination of pictures and printed names the effect started to emerged 50 ms later between 300 and 400 ms.

3. DISCUSSION

The present study is the first to show the time course of the mapping of phonetic information to lexical representations in Spanish using the visual-world paradigm. Paralleling similar findings in other languages [e.g., 8, 9, 10], we found that during the recognition of spoken words, Spanish

listeners consider multiple lexical candidates that match the unfolding speech signal. This was illustrated by the listeners' tendency to fixate pictures and words in a display more when they were consistent with the incoming speech signal than they were inconsistent with it.

The present study is also the first to compare phonological competitor activation for three different presentation modes. We found competitor activation not only in trials with pictorial displays, but also in trials with printed names, and in trials with a combination of pictures and names. The finding of more looks to orthographic forms of phonologically-related competitors in Spanish is in line with the results of [3, 4] for Dutch.

Interestingly, written displays in the present study produced stronger competition effects than pictorial displays. The letters that constitute a printed name are typically much less variable than the lines that represent the picture of an object, since objects can be depicted in a variety of forms. Identifying one object in a display might therefore be more complicated and require systematic inspections of all displayed objects; this would reduce the number of looks to a particular object, i.e. the competitor. Also listeners started to look at the competitor earlier with written displays, which might be additional evidence for letters being easier to recognize than objects. The advantage of printed names might further have been enhanced by the transparent orthography of Spanish [11], a language with a high degree of correspondence between phonemes and graphemes. Whether printed names also produce stronger competition effects in other languages with less transparent orthography needs to be established in further research.

Whereas we found a stronger competition effect for written displays, Huettig and McQueen [3] found a longer competition effect with a different setup. In their study, only competitor items were displayed (never targets), and listeners fixated phonological competitors more than unrelated distractors, both when they were shown as pictures and as printed names. For printed names the effect was, however, persistent: listeners kept fixating the competitor long after the acoustic offset of the target word. In the present study, the target was always shown, and listeners obviously fixated the target once they could acoustically distinguish it from the competitor.

The standard use of pictures in visual-world studies constrains the domain of investigation to picturable objects. The present study confirms written stimuli (and the combination of pictures plus written words) as a valid alternative. This demonstration opens the door to new investigations into spoken lexical access processes with items that are not necessarily depictable. For example, investigations comparing the processing of different word classes will benefit from this method. But also debates about the influence of distributional cues of particular phonetic characteristics to word class membership might be settled [12].

4. REFERENCES

[1] McQueen, J. in press. Eight questions about spoken-word recognition. In M. Gaskell (Ed.), *The Oxford handbook of psycholinguistics.* Oxford: OUP.

[2] Tanenhaus, M., Magnuson, J., Dahan, D. Chambers, C. 2000. Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *JPR*, 29, 55-580.

[3] Huettig, F., McQueen, J. in press. The tug of war between phonological, semantic and shape information in language-mediated visual search. *JML*.

[4] McQueen, J., Viebahn, M. 2007. Tracking recognition of spoken words by tracking looks to printed words. *QJEP*, 60, 661-671.

[5] Davies, M. 2001. Corpus del Español. (http://www.corpusdelespanol.org).

[6] INSI Master Clips 1990. Premium image collection 303,000. (http://www.imsisoft.com).

[7] Matin, E., Shao, K., Buff, K. 1993. Saccadic overhead: information processing time with and without saccades. *Perception & Psychophysics*, 53, 372-380.

[8] Spivey, M., Marian, V. 1999. Crosstalk between native and second languages: Partial activation of an irrelevant lexicon. *Psycholoical Science*, 10, 281-284.

[9] Weber, A., Paris, G. 2004. The origin of the linguistic gender effect in spoken-word recognition: Evidence from non-native listening. In *Proc. of CogSci* (pp. 1446-1451).

[10] Weber, A., Cutler, A. 2004. Lexical competition in non-native spoken-word recognition. *JML*, 50, 1-25.

[11] Justicia, F., Defior, S., Pelegrina, S., Martos, F. 1999. The sources of error in Spanish writing. *Journal of Research in Reading*, 22, 198-202.

[12] Monaghan, P., Chater, N., Christiansen M. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*, 96, 143-182.