Language Mode Influences Language-Specific Categorization

Haily Merritt
Faculty Mentor: Dr. Isabelle Darcy, Department of Second Language Studies, Indiana University Bloomington

ABSTRACT

The present study aims to fill a gap at the intersection of the phenomena of language mode—the state of activation of the bilingual’s languages and language processing mechanisms—and the subset problem—issues learners face when the second language has fewer of some kind of contrast than the first language. In studying advanced learners of Spanish (which has fewer vowels than English, thus giving rise to the subset problem) and considering language mode, we investigate whether learners create separate categories for Spanish vowels—as opposed to simply adapting their English categories—and, if so, whether the use of such categories depends on the language being used. With this, we ask: “Does language mode influence language-specific categorization?” To investigate this question, we had native English-speaking, advanced Spanish learners perform an AX task in both English and Spanish, in which they identified whether two aurally presented vowel stimuli were the same or different. If language mode does influence language-specific categorization, then participants should perform better on tasks that include a single language than on tasks that include dual languages (both English and Spanish). According to our data, there was no strong effect of language mode across conditions (single versus dual language), but we found that reaction times were significantly slower and that error rates were higher in dual language tasks. Thus, we conclude that, when multiple languages are activated, it is more difficult to process a given language.

KEYWORDS: second language, phonology acquisition, language mode, subset problem, psycholinguistics

INTRODUCTION

Second language learners are faced with an enormous task and many opportunities to make mistakes. Just what kind of mistakes depends, to some degree, on the learner’s native language and target language. To date, much of the research on second language (L2) acquisition of phonology has focused on the acquisition of novel phonemes and how/whether L2 learners can learn to perceive and properly categorize them. Escudero and Boersma (2002) were among the first to ask the opposite question: What happens when a target language has fewer contrasts than a native language? For learners in such situations, several possible issues could arise. The vowel space that a single vowel in the L2 occupies might be shared by more than one vowel in the first language (L1). This might create spurious contrasts, such that learners then map the single L2 vowel onto multiple L1 vowels, which may then be realized erroneously in both perception and production.

Upon becoming aware of the single vowel category in the L2, how does the learner handle it? Learners might merge their L1 categories such that the L2 vowel occupies the combined space of the L1 vowels. Similarly, learners might delete one of their L1 vowel categories for their L2, mapping the L2 vowel directly onto either one of the corresponding L1 vowels. Less commonly, learners might create a new L2 vowel category that, while occupying the same vowel space as the L1 vowels, is separate from the L1 vowels. Escudero and Boersma called this issue the subset problem and stated that it gives rise to multiple-category assimilation, which they define as “perceiving a binary contrast in a second language as more than two categories in the first language” (Escudero & Boersma, 2002, p. 209).

To investigate the psychological reality of multiple-category assimilation, Escudero and Boersma tested Dutch-speaking learners of Spanish (which has fewer vowels than Dutch) on their perception of front versus back vowels using embedding consonants chosen to sound ambiguously Dutch/Spanish. The consonant-vowel combinations were included within a carrier phrase in either Dutch or Spanish and were ultimately presented to participants in a forced-choice labeling task. Escudero and Boersma found that, when learners heard the Dutch carrier phrase, they used three Dutch categories, whereas Spanish has two corresponding categories. Morrison (2003) and Gordon (2011) found similar results for English-speaking learners of Spanish. Morrison had participants listen to a vowel sound and indicate the Spanish vowel (identified with key words) that was closest to the one they heard. Gordon had participants listen to vowels in a CVVC context under different conditions to elicit different perceptual manners. That is, the contexts informed the participants of which language they should use to make judgments about the stimuli. She asked participants to classify the vowels according to vowel categories but gave them no indication from what language vowels were. Using methods similar to Escudero and Boersma, both researchers found that participants classified the vowels differently depending on which language they thought they were hearing. Furthermore, Escudero and Boersma (2002) found that participants employed the extraneous categories less often when they thought they were listening to just Spanish vowels as opposed to just Dutch vowels. Escudero and Boersma attributed this difference to what they called perception mode (Escudero & Boersma, 2002, p. 216). By altering their perception mode, participants could limit the degree of influence of multiple-category assimilation. Grosjean (1989) investigated this very concept of a perception-altering mode. Language mode, as Grosjean calls it, is the state of activation of a bilingual’s language processing mechanisms at a given point in time. Bearing this phenomenon in mind, it is no surprise that Escudero and Boersma’s participants performed differently depending on which language they thought they were hearing.

Language mode, according to Grosjean, can be mediated by multitudinous factors such as environment, context, speakers, and purpose of the conversation (Grosjean, 1998), and it can be affected at all levels of language and in all modalities (Grosjean, 1989). The state of activation of a bilingual’s language processing mechanisms could range from one language being fully activated—the perpetual state of the monolingual—to both languages being highly activated—as is the case with the code-switching phenomenon or in balanced bilinguals. Because of this range, Grosjean (1989) warns against direct comparison between bilingual competency and monolingual competency. He maintains that, instead of being the sum of two monolinguals with separate language faculties, the bilingual is, in fact, an integrated whole. Thus, to evaluate a bilingual’s linguistic competency in comparison with a monolingual would fail to take into account the combined system unique to the bilingual.
Furthermore, Grosjean (1999) argues that—whether it be in the base language or in the comparative level of activation—any change in language mode is largely unconscious and effortless. There is, however, variability among bilinguals in how “far” they travel along the language mode continuum. Simultaneous interpreters, for example, spend a large part of their time with both languages highly activated. Still, for these individuals Grosjean suggests that it is unlikely that the two languages are equally highly activated. On the other end of the continuum, it is also unlikely that a bilingual individual is ever entirely in monolingual mode (Grosjean, 1999).

Outside of research, the phenomenon of language mode is more common than one might think. For example, a native French speaker who is also fluent in English might be chatting with a colleague that speaks English as their native language. A second colleague that also happens to be fluent in French greets the native French speaker in French. Because the native French speaker was conversing in English, she does not immediately understand what her French-speaking colleague said, despite French being her native language. This is because, although both languages are activated, her English is much more active than her French in that specific situation. On a smaller scale, the struggle to remember a word in one’s native language when its translation in another language is accessible is another example of the prevalence of language mode.

Although many have cited Grosjean’s and Escudero and Boersma’s work separately, few have investigated the topics of language mode and the subset problem concomitantly. Bullock, Toribio, González, and Dalola (2006), however, kept the above researchers in mind when they studied how language mode affects voice onset time (VOT) production in code-switching tasks. They found that, contrary to much of Grosjean’s work, the switch between bilingual phonetic systems in production may not actually be free of cost. Bullock and colleagues observed in L2 Spanish speakers a temporary lowering of VOT both in anticipation of and directly after a code-switch to Spanish from their native English. While Bullock and colleagues’ work is illuminating, it cannot comment on the perception of bilinguals as it relates to language mode and the subset problem.

In the present study, we investigated how perception and categorization of vowels change depending on language mode when the languages spoken by the individual give rise to the subset problem. Although a few researchers have investigated these phenomena in tandem (Bullock, et al., 2006), very little attention has been directed toward the influences of language mode on perception using the subset problem to guide the choice in languages being studied. For this reason, we were interested in learning how language mode influences perception in the form of language-specific vowel categorization. Spanish and English serve as convenient languages for this kind of comparison because Spanish has fewer vowels than English. With these phenomena and languages in mind, we ask the question: “Does language mode influence language-specific categorization?” We examined whether participants struggle to categorize extraneous vowels when operating in a language with fewer vowel categories. We tested this using a forced-choice AX task and manipulated the language mode they were in within the same experimental session. Possible effects may be realized in terms of processing speed, error rate, and interference by the less activated language.

**METHODS**

**Participants**

Eight undergraduate students (6 women) at Indiana University participated for either candy or extra credit in a Spanish class. Participants’ ages ranged from 20 to 22 (M = 21.13). All participants were enrolled in an advanced Spanish class and had been studying Spanish for 6 to 15 years (M = 7.88). All but one participant had spent at least two weeks abroad in a Spanish-speaking country to study Spanish. Participants were told that the experiment would take place in both English and Spanish and that they should come prepared to think in both English and Spanish.

**Materials**

A female early bilingual of English and Spanish produced nine English vowels (/oʊ/, /ɑ/, /eɪ/, /æ/, /i/, /I/, /eɪ/, /aɪ/, and /o/) and five Spanish vowels (/a/, /e/, /l/, /o/, /u/) in a noise-isolated recording booth. Stimuli were edited using Praat version 5.4.01 (Boersma & Weenink, 2011), such that each token consisted of the consonant sound /h/ (which exists in both English and Spanish) and one vowel sound (the “target” vowel for that item) and that all tokens were the same length in milliseconds. For more information on the stimuli and their organization across conditions, refer to the Design and Predictions section or email the first author. Five native English speakers judged the identity of the English vowels, and three native Spanish speakers judged the quality of the Spanish vowels. Each judge heard three instances of each vowel (27 vowels for English judges and 15 vowels for Spanish judges). Judges saw carrier words for each vowel as orthographic exemplars. Judges indicated which vowel they heard by writing the carrier word they think contained the vowel they heard. All vowels were correctly identified at least 80 percent of the time. It is important to note that not all English vowels were tested. Schwa was excluded from the study because it was not correctly identified at least 80 percent of the time. ‘Open o’ was excluded because it generally does not exist in southcentral Indiana dialect (the region where this study took place). That is, its location in the English vowel space is usually occupied by /a/.

**Design and Predictions**

This study has a 2 (Language mode: English vs. Spanish) x 2 (Language Duality: Single vs. Dual) x 2 (Sameness: same vowel vs. different vowel) within-subjects design. The stimuli were paired with each other in order to create an AX task. In each trial, participants heard two stimuli (separated by 150 ms of silence) and decided as quickly as possible whether they thought the two stimuli contained the same vowel or not. Stimuli were either paired with themselves (“same” pairs) or with a different vowel (“different” pairs). In addition, stimuli could be paired with a vowel from the same language (“Single Language”) or from the other language (“Dual Language”).

There were four blocks of trials, each one differing in the kind of pairing of stimuli. Each block was counterbalanced such that it included the same number (or as close to the same number as possible) of “same” and “different” pairs. Because there were more “different” pairs than “same” pairs, the “same” pairs were presented with a higher frequency to avoid a bias of responding “different.” It should also be noted that stimuli in the “same” pairs were acoustically the same token, in order to eliminate the chance of responding to minor acoustic differences. *Table 1* presents an overview of the block structure of the task.

The first block, Spanish Mode Single Language (SMSL), acted as a baseline for participants’ ability to distinguish Spanish vowels from each other, ensuring their perceived Spanish vowel space was accurate enough to have developed a “Spanish Mode.” Spanish was presented first in accordance with some of the methodological suggestions given by Grosjean (1998). The second block, Spanish Mode Dual Language (SMDL), served as an indicator for how language mode might affect the categorization, or even perception, of vowels. The third block, English Mode Single Language (EMSL),
acted as a baseline for participants’ categorization of English vowels. Finally, Block 4 (English Mode Dual Language, or EMDL) served as an indicator for how language mode might affect the categorization, or even perception, of vowels, particularly when compared with the SMDL block. Participants were given no indication that any blocks would include both English and Spanish and were furthermore instructed to “think just in English” or “think just in Spanish,” similar to Gordon (2011), so that they would be in a more monolingual language mode. For each item, the first vowel participants heard aligned with the language mode they were in (in accordance with some of the methodological suggestions provided by Grosjean (1998). The second vowel either did or did not match the participant’s language mode, as indicated in Table 1. An item with stimuli from a single language should see no negative effects on response time or accuracy, whereas participants should perform worse on an item with stimuli from dual languages.

We hypothesize that, because the participants are English-speaking learners of Spanish, they will “group” together vowels that are phonemic in English but not in Spanish (e.g. /i/ and /Ir/). This difference in phonological status might influence a categorical decision, such as in the AX task. More precisely, we expect that English-speaking learners of Spanish will hesitate about whether the vowel they hear is the same or different when they hear /i/ - /Ir/ but only if they are in Spanish mode, and only if language mode indeed influences categorization.

Thus, we have two specific predictions. First, when in English Mode, we expect participants to have high accuracy and quick response times in the Single Language condition for easy items, which are baseline items. We also expect the same for Spanish Mode in the Single Language condition, which includes baseline items.

Second, for Dual Language conditions, we might see a difference between English and Spanish Modes: If language mode influences categorization, participants will hesitate about categorizing difficult items (e.g. an [i] – [Ir] pair) as “different.” Therefore, they will be slower and make more errors on such items in Spanish Mode. This will not be the case in English Mode, however, since English does have a phonemic distinction between these two vowels.

### Procedure

In order to become accustomed to the layout of the experiment, participants completed a practice test consisting of five items, all of which were present in the experiment proper. The practice test resembled the experiment itself in that it was an AX Task. For each item, participants heard two stimuli with a 500-ms interval between them. They then decided whether the second vowel they heard was exactly the same as the first. Participants indicated their responses by pressing one of two buttons on a keyboard. Participants had the same instructions for all tasks: Indicate whether the second vowel is exactly the same as the first by pressing the LEFT ALT key for “same” and the RIGHT ALT key for “different”.

Immediately before beginning the experiment and the first Spanish block (SMDL), participants read a passage in Spanish to “induce” Spanish mode. The passage was an excerpt from a fictional story originally written in Spanish. While Escudero and Boersma (2002) used only a carrier phrase, Grosjean (1999) suggests that much more may be necessary to truly induce a language mode. Participants then completed the four blocks of AX tasks. The SMDL block was conducted immediately after SMDL so that participants were still operating in Spanish Mode. SMDL, however, contained tokens that are both Spanish and English. We gave participants no indication that they would hear non-Spanish vowels. After completing SMDL, participants completed a language background questionnaire that assessed their linguistic history. The questionnaire was entirely in English to “induce” English mode. Grosjean suggests that sequential bilinguals require very little influence to induce the mode of their native language (Grosjean, 1999). Participants then completed the EMDL block, which contained only English vowels. Finally, participants completed the EMSL block. Testing lasted approximately 1.5 hours for each participant. The independent variables were Language Mode (E vs. S), Language Duality (Single vs. Dual), and Item Difficulty (easy vs. difficult); the measured dependent variables included error rate/accuracy and response time (measured in milliseconds). Participants were tested individually in a research lab. Auditory stimuli were presented through Sennheiser HD515 headphones. The experimental presentation was controlled by the DMDX software (Forster & Forster, 2003).

### RESULTS AND DISCUSSION

The accuracy and response time data were log-transformed to create a normal distribution and analyzed using a Linear Mixed Effects Model. Overall, there was no difference in accuracy between English (M = 87) and Spanish Modes (M = 86) (p > 0.1). This indicates that the participants had attained sufficient fluency to make proper judgments. As expected, language duality influenced accuracy (see Figure 1). When in English Mode, participants were more accurate in Single Language trials (M = 0.95) than in Dual Language trials (M = 0.83), $F(1, 5938) = 93.06, p < .001$. Likewise, when in Spanish Mode, participants were more accurate in Single Language trials (M = 0.94) than in Dual Language trials (M = 0.82), $F(1, 5933) = 84.12, p < .001$. These data suggest that language mode does, in fact, influence language-specific perception, since participants struggled to categorize vowels from an unexpected language.

---

### Table 1.

Overview of Block Structure and Item Pairings, Showing Mode, Language Duality, and Item Difficulty

<table>
<thead>
<tr>
<th>Language Mode</th>
<th>Language Duality</th>
<th>“Same” pairs</th>
<th>“Different” pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>Block 1: SINGLE (Spanish only)</td>
<td>Easy items (N = 45)</td>
<td>Easy items (N = 45)</td>
</tr>
<tr>
<td></td>
<td>SMDL</td>
<td>(e.g. /e/ - /e/)</td>
<td>(e.g. /e/ - /a/)</td>
</tr>
<tr>
<td>Spanish</td>
<td>Block 2: DUAL (Spanish + English)</td>
<td>Easy items (N = 115)</td>
<td>Easy items (N = 66)</td>
</tr>
</tbody>
</table>
|               | SMDL             | (e.g. /e/ - /e/) | (e.g. /e/ - /a/)
|               |                  | Difficult items (N = 48) |                  |
|               |                  | (e.g. /e/ - /e/) |                  |
| English       | Block 3: SINGLE (English only) | Easy items (N = 72) | Easy items (N = 72) |
|               | EMSL             | (e.g. /e/ - /e/) | (e.g. /e/ - /e/) |
| English       | Block 4: DUAL (Spanish + English) | Easy items (N = 171) | Easy items (N = 117) |
|               | EMDL             | (e.g. /e/ - /e/) | (e.g. /e/ - /e/) |
|               |                  | Difficult items (N = 54) |                  |
|               |                  | (e.g. /e/ - /e/) |                  |

For clarity, subscripts on the vowels indicate which language they are in (E = English, S = Spanish).
As hypothesized, language duality influenced reaction time (Figure 2). When in English Mode, participants had faster reaction times in single language trials ($M = 2.96$) than in Dual Language trials ($M = 3.00$), $F(1, 5383) = 65.39, p < .001$. Contrary to expectations, when in Spanish Mode, participants had slower reaction times in Single Language trials ($M = 3.02$) than in Dual Language trials ($M = 2.99$), $F(1, 5382) = 18.39, p < .001$. The English Mode reaction time data support our hypothesis that participants will be more hesitant in categorizing vowels from an unexpected language. The absence of such support from the Spanish Mode reaction time data could be due to a task effect, as the SMSL task was the first task participants performed, and thus might have been initially a bit slower. Further research is needed to examine this possibility.

As we predicted, reaction times were faster for easy stimuli ($M = 3.00$) than for difficult stimuli ($M = 3.06$, $F(1, 5933) = 129.8$, $p < .001$), meaning participants needed less time to process the similarity or lack thereof between the stimuli. This supports our hypothesis that participants will not hesitate as much with stimuli of the same language or with stimuli that occupy very different locations in the vowel space. Similarly, there was a difference in accuracy between easy stimuli ($M = 94$) and difficult stimuli ($M = 72$), $F(1, 5933) = 448.4, p < .001$. There was also an interaction between mode and language duality for different vowel responses (see Figure 3), in terms of the reaction times ($F(1, 2985) = 30.1, p < .001$). This interaction may be due to the fact that response times are slower in the SMSL block. While it may mirror a real effect, it could also be due to this block being the first task participants completed; thus, we cannot exclude this possibility entirely.

**CONCLUSIONS**

We conducted this study to investigate the possible effect of language mode on language-specific vowel categorization using languages for which the *subset problem* is relevant. We expected performance to be negatively affected in dual language conditions due to increased difficulty in differentiating the stimuli. Because error rates were higher and response times were slower in dual language conditions for both English and Spanish, we argue that language mode does, in fact, influence language-specific perception and categorization. In other words, the category to which an individual maps a sound they perceive depends on the language mode they are in and the vowel categories provided by that language. An individual’s expectations regarding the language they are listening to also influence how they perceive input in that language. For example, if an individual is in Spanish mode and he/she hears a mid-front unrounded vowel, they might perceive it as /e/. The same individual could hear the same vowel while in English mode, but in this instance they perceive it as /ɛ/. This kind of ambiguity is realized in our study through the increased response times and error rates in the dual language conditions. Because participants were presented with
stimuli that did not exactly adhere to the vowel categories of the language participants were thinking in, participants struggled to map the vowels onto a sensible category. Since the stimuli in the single language conditions did not pose this problem, error rates were lower and response times were faster. It warrants mention that there were no difficult items in the single language conditions. This provides an area for future investigation in that easy items in the single language condition versus easy items in the dual language condition might be compared.

At the same time, the increased error rates and response times in dual language conditions in our participants do not align with Grosjean’s claim that changes in language mode are free of cost. Although we cannot comment on participants’ conscious knowledge of the mixed-language nature of the stimuli in the dual language conditions, the data suggest that this affected participants’ processing of the vowels. Investigating participants’ awareness of the language of the stimuli might be of interest to those conducting future research in this area. It could be the case that introducing an unexpected language into a communicative event negatively affects event processing and perception, but this may only be true for bilinguals who do not use both languages with comparable frequency. This is supported by Grosjean’s claims (1998). It should be noted, however, that the bilinguals who participated in Grosjean’s studies and on whom Grosjean’s models are based were often highly proficient bilinguals who were functionally balanced. The participants in our study were all English-dominant.

For those with interest in studying language mode, we should point out that our study included confounds in inducing language mode. Grosjean (1999) suggests extensive caution when doing research related to bilinguals and language mode, as mode is quite susceptible to influence. He went so far as to propose—among other things—that researchers test the two languages at different times of day, on different days, in different locations, and with different researchers (whose linguistic capabilities the participant is not aware of). The present study was limited in terms of time and personnel and was not able to hold such factors constant. We therefore acknowledge that controlling for such factors could have produced more robust results. Our intention, however, was not to induce entirely different modes, but rather it was to see if there is a difference in perception when only a small shift in mode is made. The small shift we were able to make in language mode still produced differences. Still, we do not have a quantitative measure of what participants’ modes were at any given point in time, as the presentation of a stimulus from a language that is not presently activated could change the activation levels and present a misleading reading to the researcher.

Additionally, the sample size is another shortcoming of the present study. Any trends we found could have been the result of the small sample size. We recommend a replication with a larger sample size. We hope that this study serves as a starting point for future investigation considering language mode and the subset problem, as well as a friendly reminder that our bilingual colleagues may use language mode as an excuse upon forgetting words in their native language.

ACKNOWLEDGEMENTS
I would like to extend a profound thank you to Dr. Isabelle Darcy for her guidance and support throughout this project. It is because of her direction and advice that this project is what it is. I am very appreciative of Dr. Kenneth de Jong and his comments on a previous draft of this paper. I would like to express my gratitude to the Second Language Psycholinguistics Laboratory as a whole for their suggestions and constructive criticisms. I would also like to thank Binyan Li from the Indiana University Statistical Consulting Center for her assistance in the statistical analysis.

AUTHOR INFORMATION
All correspondence should be sent to the first author: merritthaily@gmail.com

REFERENCES


