Implementing Escudero's model for the SUBSET problem

Klara Weiand
kweiand@science.uva.nl
Master Studies in Artificial Intelligence, University of Amsterdam*

May 13, 2007

Contents

*I would like to thank Paola Escudero for bringing the problem to my attention and I am grateful to her and Paul Boersma for their comments on an earlier draft. All remaining errors are my own.

Abstract

This paper reports on the results of a series of experiments that put Escudero's theoretical model[4] to a practical test. This model was developed to explain how Dutch learners of Spanish vowels solve the SUBSET problem. Escuderos model uses Stochastic Optimality Theory and, in this paper, we use this same framework with the Gradual Learning Algorithm to simulate learning on the basis of Escuderos proposal.

Escudero's model is based on empirical data which shows that Dutch learners initially use three native categories to classify the two Spanish vowels /i/ and /e/. The model proposes that this initial situation leads to non-optimal perception as well as non-optimal recognition because three words instead of two will be stored in this learners L2 lexicon. The empirical data also shows that advanced learners can attain optimal L2 perception. To explain this L2 development, the model proposes that both perceptual learning and lexical feedback are involved in the reduction of one of the three native categories to acquire optimal L2 sound perception and word recognition.

The proposal for the initial state in Dutch learners of Spanish vowels relies on the existence of minimal pairs in Spanish which are assumed to lead to the storage of lexical entries different in form but identical in meaning and entries identical in form but different in meaning. Escudero proposes that the existence of two entries with the same form but with different meanings leads to what she calls semantic errors. The proposal also says that, when the learner notices these errors through contextual cues, perceptual learning will occur, just like in the lexicon-driven learning proposed in Boersma[1] for L1 acquisition and in Escudero and Boersma[2] for L2 acquisition.

In order to computationally test Escudero's proposal, several different parameter settings and correct candidate selection mechanisms were compared. It was found that the model as proposed in principle is viable but is very sensitive to noise. A different selection mechanism and alternate rankings remedy this problem, leading to complete category elimination and boundary shifts to the native positions.

1 Introduction

When a learner acquires a new language, the number of vowel categories in his L2 can differ from that in his native language in two ways, posing two different learning problems: The new language can have more vowel categories than his L1 –this is called a NEW scenario– or it can have fewer categories, leading to a SUBSET learning scenario.

Escudero[4], extending previous work by Boersma and Escudero[2] to the learning of lexical items and the interaction between perceptual and lexical learning, proposes a learning model for the SUBSET scenario using Stochastic Optimality Theory and the Gradual Learning Algorithm[1] and the interaction between perceptual and recognition learning.

Stochastic OT differs from classical OT in that constraint rankings are not discrete and ordinal but rather arranged on a continuous scale, which means that the distance between constraints is not fixed. A noise component is added to the ranking values during evaluation which makes the selection of the optimal candidate non-deterministic.

This is because the order of constraints close (relative to the amount of noise) in ranking value can temporarily switch ranking orders as the result of the added jitter, yielding a different total ranking used for the computation of the optimal candidate.

During learning with the Gradual Learning Algorithm, a Stochastic OT grammar is fed input tokens from the training data and in the case of a mismatch between the respective training output and the actual output of the Grammar, the ranking values of the constraints violated by either output are adjusted.

The goal of this work is to simulate the described proposal that so far is only of theoretic nature, and to additionally investigate the influence of different parameters on its performance.

Escudero uses the situation of a Dutch learner of Spanish to exemplify her proposal: Whereas Dutch has three vowels, /i/, /ɪ/ and /ɛ/, Spanish only has two vowels, namely /i/ and /e/, that cover the same part of the F1 continuum for front vowels (see figure 1).

Assuming that the L1 categories are used as a starting point in acquiring the vowel system of an L2, Spanish /i/ and /e/ tokens are initially recognized as /i/ and /ɪ/ and /ɛ/ and /ɪ/ respectively by the Dutch learner since the Spanish /i/-/e/ boundary at 407 Hertz lies between the Dutch /i/-/ɪ/ and /ɪ/-/ɛ/ boundaries. The beginning learner then faces two kinds of problems,

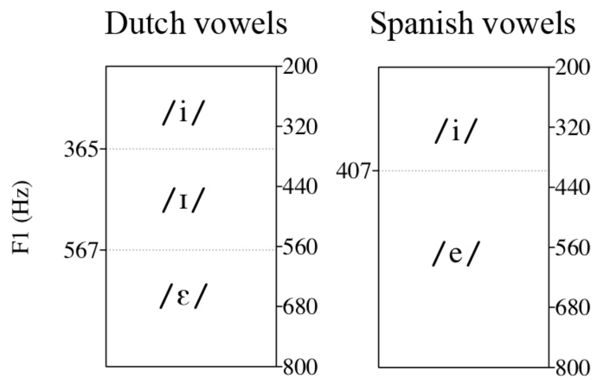

Figure 1: Optimal category boundaries for Dutch and Spanish

she perceives more categories than is optimal and, as a consequence, she also creates unnecessary lexical entries.

For example, when hearing the Spanish word |tʃika|, "girl", she will perceive

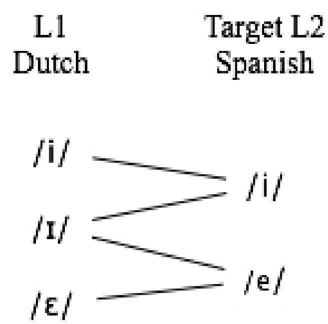

Figure 2: Perception of Spanish /i/ and /e/ in Dutch

the |i| as /i/ some of the time and as /ɪ/ the rest of the time (see Figure 2), resulting in two different lexical entries, |tʃika| and |tʃɪka|, that are associated with the same meaning, "girl".

Similarly, the Spanish word |tʃeka|, "Czech" results in two lexical entries |tʃɛka| and |tʃɪka|. The minimal pair |tʃika|-|tʃeka| thus leads to four different lexical entries where one pair of words has the same form and two words have the same meaning.

Escudero proposes that the learner reduces his categories and lexical entries by means of an interaction between perceptual and recognition learning, or, on a more practical level, the interaction between GLA learning in a perception and a recognition grammar.

The recognition grammar contains lexical constraints for all candidates, in our example the four forms mentioned above, and faithfulness constraints which penalize a mismatch in vowel quality between the perceptual input and the candidates.

In the recognition of the words |tʃika| and |tʃeka| two different kinds of errors are possible: Semantic errors, when e.g. the perceptual input /tʃika/ with the intended meaning "girl" is recognized as "|tʃɪka| 'Czech'", and phonological errors where the form of the input does not match that of the correct candidate.

While the former leads to recognition learning, perceptual errors trigger perceptual learning, a change in the ranking of the cue constraints in the perception grammar, which in turn has an influence on the recognition grammar and its performance, and in the ranking of the faithfulness constraints.

This interaction results in a shift of category boundaries, reducing the perception of /ɪ/ and finally converging on the reduction of the number of categories to two (see Figure 3).

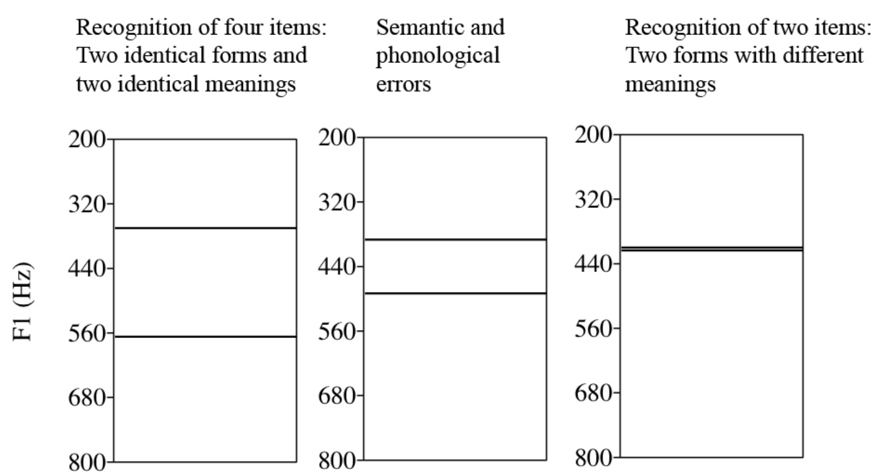

Figure 3: The learning process

This works aims at implementing Escudero's proposal to test and verify her predictions about the course and outcome of the learning given the described constraints and learning mechanisms.

2 Escudero's model

As described in the previous section, the models makes use of two different grammars, a perception grammar containing cue constraints and a recognition grammar with faithfulness and lexical constraints.

The input to the perception grammar is a Hertz value for which the optimal candidate, the perceived vowel, is selected[1] through cue constraints of the form "`number` Hz is not /`vowel`/", which exist for every Hertz-vowel pair. The optimal candidate then is that whose relevant cue constraint given the input is ranked the lowest during evaluation (see figure 4).

[F1 = 550 Hz]	[550 Hz] not /i/	[550 Hz] not /ɛ/	[550 Hz] not /ɪ/
/tʃika/	*!		
/tʃɛka/		*!	
☞/tʃɪka/			*

Figure 4: Classification in the perception grammar

The perceived form is then used as the input to the recognition grammar which selects the optimal candidate, a pair of underlying form and meaning (e.g. "|tʃika| 'girl'") for the given input.

In the recognition grammar (see figure 5 for an example), the faithfulness constraints (e.g. "FAITH */ɪ/ → |i|") penalize a vowel change between the input, that is, the perceived surface form, and the underlying form in the candidate, while the lexical constraints (e.g. "*LEX |tʃɪka| 'girl'") work against observing specific combinations of underlying form and meaning.

It is important to note that only faithfulness constraints that are ranked higher than at least some lexical constraints have an influence on which candidate is selected, since every candidate violates exactly one lexical constraint and no two candidates violate the same lexical constraint.

Consequently, initially all faithfulness constraints are ranked higher than all lexical constraints. The faithfulness constraints are ranked by the per-

[1]Note that Escudero's proposal shows not simple vowels but vowels embedded into words (e.g. /chica/ and /chɛca/) as the output of the perception grammar, while in this implementation the perception grammar selects vowels (e.g. /i/ and /ɛ/ which are only then embedded into the word template. This difference does not have an effect of the workings of the model.

/tʃɪka/ Intended meaning: ’girl’	FAITH */ɪ/ → |i|	FAITH */ɪ/ → | ɛ|	*LEX |tʃɛka| ’Czech’	*LEX |tʃɪka| ’girl’	*LEX |tʃika| ’girl’	*LEX |tʃɪka| ’Czech’
✓|tʃika| ’girl’	*!				*	
|tʃɛka| ’Czech’		*!	*			
✓|tʃɪka| ’girl’				*!		
☞|tʃɪka| ’Czech’						*

Figure 5: A semantic error in the recognition grammar.

ceptual distance between categories and the lexical constraints are ranked by perceived frequency. The perceived frequency describes how many Spanish /i/ tokens are perceived as /i/ and how many are perceived as /ɪ/ and what percentage of /e/ tokens are perceived as /ɛ/ or /ɪ/. Escudero shows that 79.5 % of /i/ tokens are perceived as /i/, while 20.5 % are perceived as /ɪ/. For /e/, the numbers found are 13.7 % for /ɛ/ and 87.3 % for /ɪ/. The initial ranking of lexical constraints is thus *LEX |ɛ| ‘Czech’ >> *LEX |ɪ| ‘girl’ >> *LEX |i| ‘girl’ >> *LEX |ɪ| ‘Czech’.

Overall, three types of different constraints are used in the two grammars: Cue constraints where every candidate violates one cue constraint and where which cue constraint is violated depends on the combination of candidate and input, faithfulness constraints whose violation also depends on the candidate and input but which are only violated for certain combinations, and lexical constraints which are violated once by every candidate and whose violation only depends on the candidate.

Escudero describes two kinds of learning following the recognition of an underlying form-meaning pair, namely semantic and phonetic learning. When the learner detects that the meaning she recognized and that intended by the speaker mismatch, this semantic error triggers learning in the recognition grammar. Since there are two candidates with the correct meaning, one of them is chosen at random, thus determining the correct underlying form. For example, in the situation displayed in figure 5, the first candidate, “|tʃika| ‘girl’”, is chosen half of the time, and the third candidate “|tʃɪka| ‘girl’” the other half of the time.

Semantic learning then means that the ranking values of the lexical constraints violated by the wrongly selected optimal candidate are increased by the amount specified as the learning rate, while all constraints violated by the correct form are demoted by the same amount.

Additionally, if there is a mismatch between the vowel contained in the correct form and that in the input to the recognition grammar, the ranking value of the violated faithfulness constraint is decreased. The error is propagated back to the perception grammar: The constraint pertaining to the original Hertz value and the vowel found in the correct underlying form is demoted, while the ranking value for that for the wrongly winning vowel and the input value is increased. Escudero calls this *phonological learning*.

The model relies on the existence of minimal pairs for /i/ and /e/ in Spanish and the assumption that multiple, partially identical lexical entries are created as a consequence of the existence of one more category than is optimal. Entries, that is, candidates, with identical form and different meanings then give rise to semantic errors.

All learning is thus based on the learner's realization that the recognized meaning differs from the intended meaning, which triggers semantic learning in the recognition grammar and the selection of a correct candidate, which in turn can give rise to phonological learning in the perception and recognition grammar.

Escudero predicts that the entries that share their form but differ in semantics will lead to semantic errors. Consequently, the *FAITH constraints that prohibit the recognition of an underlying form containing |i| when an /ɪ/ was originally perceived will be demoted to a position below the lexical constraints.

*FAITH constraints apply to all inputs to the recognition grammar but *LEX constraints do not. For example the input /tʃɪka/ can lead to a violation of *LEX |tʃɪka| 'girl'" in the correct candidate but never to one of *LEX |mɪsa| 'mass'", while both the inputs /tʃɪka/ and /mɪsa/ can lead to a violation and demotion of FAITH */ɪ/ → |i|.

The same is predicted to occur with the *FAITH constraint penalizing the recognition of an underlying form containing |ɛ| where there is a /ɪ/ in the perceived form.

Since additionally *FAITH constraints only ever decrease their ranking value, the *FAITH constraint is consequently expected to become the lowest-ranked constraint relatively quickly.

As described above, this means that for an input to the recognition grammar containing /ɪ/, the selection of the optimal candidate between the underlying forms containing |ɪ| and those containing |i| will depend only on the rankings of the lexical constraints.

Since /i/ and /ɛ/ inputs are uniquely associated with one meaning, semantic errors are only expected to happen for inputs containing /ɪ/.

The lexicon-driven learning caused by semantic errors then gives rise to phonological learning when the correct candidate that violates the demoted *FAITH constraint is selected. This is the case for example when the meaning intended by the speaker is 'girl' and the form is perceived as /ʃɪka/ and recognized to mean 'Czech' and the correct candidate consequently picked is "|tʃika| 'girl'".

The phonological learning then raises the ranking values of the cue constraints pertaining to /ɪ/, causing /ɪ/ to be perceived less often and thus shifting the vowel boundaries, decreasing the range of and finally eliminating the /ɪ/ category.

3 Implementation

The model was implemented in the PRAAT[3] scripting language. The initial perception grammar representing a native Dutch speaker was created by training a cue constraint grammar containing 243 constraints –one constraint for every combination of vowel and Hertz values between 200 Hz and 800 Hz in steps of 10 Hz– on vowel-value pairs randomly drawn from Gaussian distributions. Median values F1 values of 733, 438 and 305 Hertz for /ɛ/, /ɪ/ and /i/ respectively, and a standard deviation of 0.16 octaves were used to generate the distributions. All constraints were initially set to the same ranking value, 100.

There were 180,000 training repetitions with the initial learning rate of 1 decreasing by a factor of 0.7 every 10,000 steps[2].

The responses of the resulting grammar with the evaluation noise set to 2.0, collected by presenting each input value to the grammar 1000 times and recording the outputs, are displayed in figure 6. As can be seen from the figure, the equal likelihood points in the trained grammar match those reported by Escudero and Boersma[2] and Escudero[4], namely 365 Hertz for /i/ and /ɪ/ and 567 Hertz for /ɪ/ and /ɛ/. Accordingly, the fraction correct without evaluation noise matches the maximum likelihood fraction correct of the training data at 0.984.

The F1 values for the simulation of the learner's resolution of the SUBSET problem in turn were drawn from Gaussian distributions with median values 340 and 500 for Spanish /i/ and /e/ and a standard deviation of 0.16 octaves.

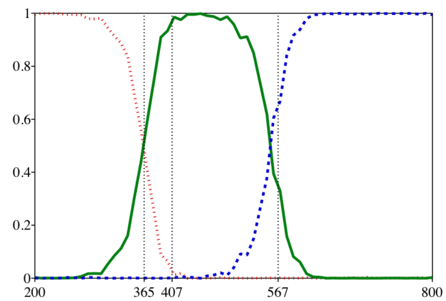

Figure 6: The output of the native Dutch perception grammar. The x-axis displays Hertz values and the y-axis the probability. The red dotted line stand for /i/, the green solid line for /ɪ/ and the dashed blue line for /ɛ/.

In order to select the speaker's intended meaning and the word template needed, word-meaning forms where the word contained the selected vowel were then drawn from distributions over word frequency. The frequencies of the different words were determined using Google[2]. The list of minimal pairs was created through manual selection from an automatically generated list of minimal pairs extracted from a list of Spanish words provided online by the Carnegie Mellon School of Computer Science[3]

The vowel, word and meaning were taken to represent the speaker's intentions, while the Hertz value was used as an input to the perception grammar.

Learning in the implementation proceeds as described in the last section with two additional optional variations: The number and nature of the lexical constraints and the selection of the correct candidate in the case of a semantic mismatch.

Escudero only describes the creation of four lexical entries for one minimal pair, in the case of /tʃika/-/tʃeka/ these are "|tʃika| 'girl'", "|tʃeka| 'Czech'","|tʃɪka| 'girl'" and "|tʃɪka| 'girl'".

As the distributions for /i/ and /ɛ/ overlap to a (very small) degree, there are cases where an intended /i/ is perceived as an /ɛ/ and vice versa. Since there are thus rare instances where for example /tʃika/ with the meaning 'Czech' is perceived, this might be represented in the lexical inventory. Due to their low frequency, these lexical constraints are initially ranked higher

[2]http://www.google.com

[3]http://www.cs.cmu.edu/afs/cs.cmu.edu/project/ai-repository/ai/areas/nlp/corpora/dicts/spanish/

than the other lexical constraints. As the random correct candidate selection assigns all candidates with the found correct meaning the same probability, these rare cases might be over-represented during learning which could lead to problems with learning in the recognition grammar. This will be investigated in the next section.

Also, instead of selecting the correct candidate randomly from those candidates with the correct meaning, a variation where the most harmonic candidate was chosen was implemented. This means that the candidate with the correct meaning that incurs the least serious violations is selected as the correct candidate. That is, given the new knowledge that the optimal candidate has the wrong meaning, the candidates bearing that meaning are discarded and out of the candidates left, the one that is optimal is selected. For example, in figure 5, the third candidate is always selected when using this scheme for correct candidate selection.

Another variation of correct candidate selection tested selects the most harmonic candidate with regard to only the lexical constraints, ignoring the candidates' violations in the faithfulness constraints. In the case shown in 5, the first candidate is the correct candidate according to this selection mechanism.

4 Experiments and Discussion

This section describes the evaluation and results of the implemented model.

As mentioned in the previous sections, there are several parameters that can be varied when running the model: The number of minimal pairs in the lexicon, the evaluation noise, the initial ranking values in the recognition grammar, the learning rate and its changes over time, the standard deviation of the Gaussian distributions, the lexical entries assumed and the mechanism according to which the correct candidate is chosen.

The learning rate and evaluation noise were kept constant across all experiments described in the following and were set to 0.1 and 2.0 respectively.

The parameters in the first experiment were identical to those described by Escudero with a standard deviation of 0.16 octaves for the Gaussian distributions, a random selection of correct candidates, an initial ranking of lexical constraints inverse to perceptual frequency for the lexical constraints and rarely encountered observations (e.g. /tʃika/ - 'Czech') not being represented in the recognition grammar.

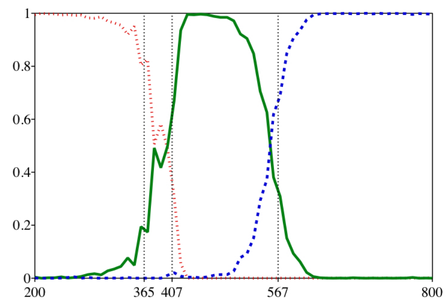

Figure 7: The output of the perception grammar after 10,000 trials.

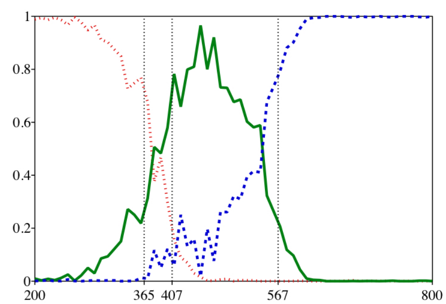

Figure 8: The output of the perception grammar after 60,000 trials.

Figure 7 shows the perception grammar after 10,000 training steps with a dataset of 13 Spanish minimal pairs. As can be seen when comparing this plot to figure 6, the boundary between /i/ and /ɪ/ has shifted by about 25 Hertz, increasing the area of the /i/ category and shifting the boundary almost to the position of the /i/-/e/ boundary in Spanish, 407 Hertz. The boundary between /ɪ/ and /ɛ/ on the other hand remains constant. This is reflected in the recognition grammar where the faithfulness constraint that prohibits changing an incoming /ɪ/ to an underlying |i| is only 20th highest ranked out of the 58 constraints –six faithfulness and 52 lexical constraints–, while the equivalent constraint for /ɪ/ and |ɛ| is still ranked higher than all lexical constraints.

Note that not every repetition leads to learning, but only those where a semantic error occurs, that is, where the learner notices she has made a mistake. Phonological learning in turn only happens when it is preceded by a

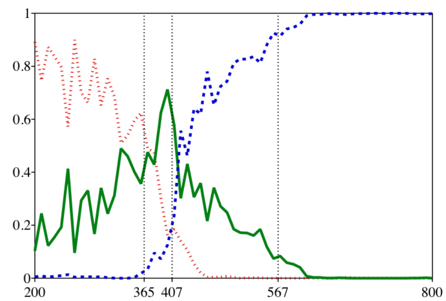

Figure 9: The output of the perception grammar after 80,000 trials.

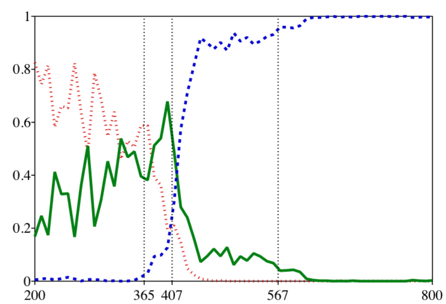

Figure 10: The output of the perception grammar after 100,000 trials.

step of semantic learning. In the present case, 606 cases of semantic learning and 351 cases of phonological learning were observed, meaning that in almost 94 percent of the trials no learning takes place. This rate remained roughly stable in the consequent experiments with about one in ten trials leading to semantic learning and one in two cases of semantic learning leading to phonological learning.

The fact that the /i/ category starts to widen first is due to the initial ranking of the lexical constraints: The constraint matching |ɪ| with the meaning belonging to the word containing Spanish |e| (e.g. * LEX |ɪ| 'Czech') is ranked lower than that matching it to the Spanish |i| word (e.g. *LEX |ɪ| 'Girl'). This means that initially, incoming forms containing /ɪ/ are always matched to the meaning belonging to the |e| word, only leading to a semantic error –and thus semantic and potentially phonological learning and a consequent shift of category boundaries in the perception grammar– when the

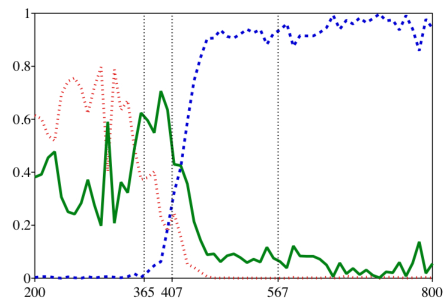

Figure 11: The output of the perception grammar after 200,000 trials.

/tʃɪka/ Intended meaning: ’Czech’	FAITH */ɪ/ → |i|	FAITH */ɪ/ → | ɛ|	*LEX |tʃɛka| ’Czech’	*LEX |tʃɪka| ’girl’	*LEX |tʃika| ’girl’	*LEX |tʃɪka| ’Czech’
|tʃika| ’girl’	*!				*	
|tʃɛka| ’Czech’		*!	*			
|tʃɪka| ’girl’				*!		
☞|tʃɪka| ’Czech’						*

Figure 12: Correct recognition

speaker had really intended |i| (see figures 5 and 12).

Since the overlap of the |ɪ| category with Spanish |e| is bigger than the overlap with Spanish |i| (see figure 1), initially the latter situation will only occur in the minority of trials.

As learning progresses, the two lexical constraints mentioned eventually switch their order and FAITH */ɪ/ → |i| becomes the lowest ranked constraint in the recognition grammar, leading to a further decrease of the /ɪ/ category (see figure 8).

After around 100,000 trials (see figure 10), the /ɪ/ category has been significantly reduced and almost eliminated, covering an area of only about 10 Hertz. Figure 13 displays the equal likelihood category boundaries at different positions during learning. As can be seen, the boundaries and their development are nearly identical to the prediction shown in figure 3.

Overall, it can thus be concluded that the implementation confirms the predictions made by Escudero and succeeds in reducing the /ɪ/ category through the interaction of semantic and phonological learning.

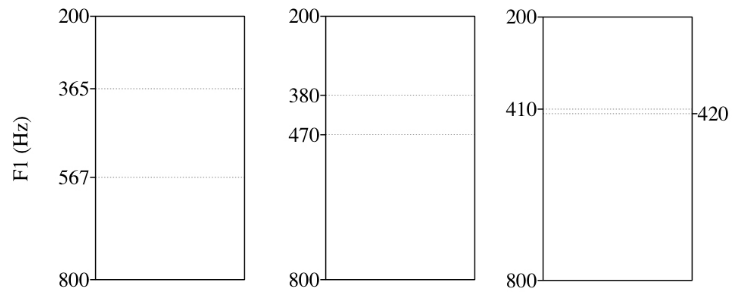

Figure 13: Category boundaries for the initial state and after 70,000 and 100,000 steps

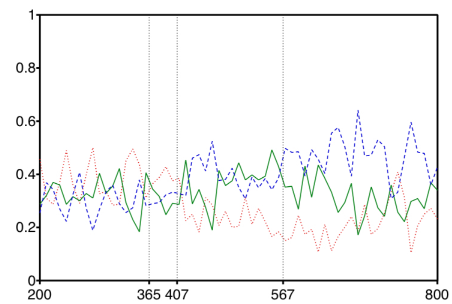

Figure 14: The output of the perception grammar after 100,000 trials (six candidates).

However, looking at the plots as the amount of learning increases, it becomes apparent that, while the /ɪ/ category is decreased, the amount of noise overall increases, actually leading to a higher probability for the perception of /ɪ/ for low Hertz values after 100,000 trials than after 80,000 trials. As all learning in the perception grammar is preceded by semantic errors in the recognition grammar, the cause of the noise must lie within the recognition grammar. It seems that only input to the grammar that contains /ɪ/ can cause semantic errors since /i/ and /ɛ/ input is unambiguously mapped to only one meaning each and Faithfulness constraints prohibit a change of vowels.

But there is another source of semantic errors that happen when an intended |e| is perceived as an /i/ or an intended |i| as an /ɛ/. Each of these cases leads to a semantic error and phonological learning since |ɛ| and |i| are

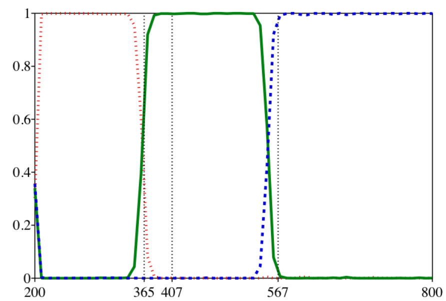

Figure 15: The output of the initial perception grammar for a decreased standard deviation.

uniquely associated with different meanings.

The chance of this happening is initially very small, for example, the probabilities assigned to 300 Hertz in the Spanish data is 99.6 % for /i/ and 0.4 % for /e/, that is, according to the Gaussian distributions, a stimulus of 300 Hertz is intended to be an |e| in only about one in 200 cases. If an intended /e/ with a Hertz value that is more typical for an /i/ is indeed classified as an /i/ in the perception grammar, the correct candidate chosen after the occurrence of the semantic error can either contain |ɛ| or |ɪ|. If it contains |ɛ|, FAITH */i/ → |ɛ| decreases its ranking value and the cue constraints are changed so that it becomes more likely to perceive /ɛ/ the next time a stimulus with an F1 of 300 Hertz is encountered, thus slightly diminishing the /ɛ/-/i/ contrast in both grammars.

In the analogous case when the chosen correct candidate contains |ɪ|, FAITH */i/ → |ɪ| is decreased and the chance to perceive /ɪ/ next time is increased, thus working against the reduction of the /ɪ/ category.

The noise thus can not only undo the reduction of the /ɪ/ category, but is also self-reinforcing, which means, that in the long run it prohibits convergence and after a big number of learning steps significantly diminishes category boundaries. Analogous experiments with two and 32 minimal pairs showed the same effect and proved that the number of minimal pairs is not critical to this problem.

Noise and its consequences are also the reason why the addition of lexical constraints for "|tʃɛka| 'Girl'" and "|tʃɪka| 'Czech'" as discussed in section 3 proves unsuccessful (figure 14), since under this condition the chance to

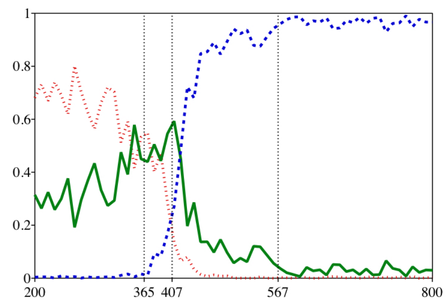

Figure 16: The output of the perception grammar after 200,000 trials (decreased standard deviation).

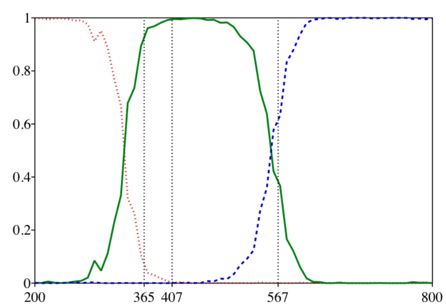

Figure 17: The output of the perception grammar after 100,000 trials (harmonic selection).

encounter seemingly contradictory evidence is increased all categories have been unlearned after only 100,000 trials.

One approach to solving this noise problem is to reduce the standard deviations of the Gaussians from which the training data for the L1 and L2 is drawn. If the distributions do not overlap (or overlap less), the number of atypical /i/ and /e/ tokens is reduced, thus slowing down or eliminating further noise increase.

Figure 15 shows the initial perception grammar of a virtual Dutch speaker trained on data drawn from Gaussian distributions with identical medians but a standard deviation of 0.08, that is, half of what was proposed by Escudero and used in the previous experiment.

Figure 16 displays the output of the same grammar after 200,000 cycles

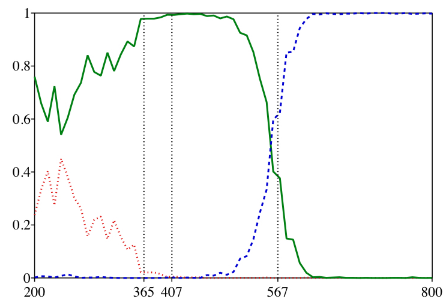

Figure 18: The output of the perception grammar after 160,000 trials (harmonic selection).

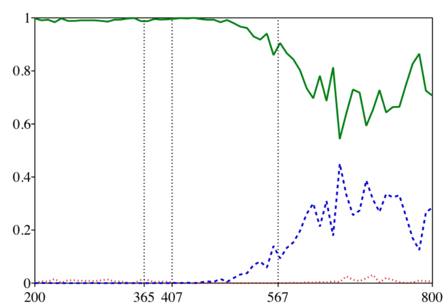

Figure 19: The output of the perception grammar after 400,000 trials (harmonic selection).

of training, the |ɪ| category having been successfully reduced. As can be observed in comparison with figure 11 which shows the output of the perception grammar for a standard deviation of 0.16 after the same number of repetitions, reducing the standard deviation successfully decreases the noise, thus prohibiting an "unlearning" of the categories. However, some noise can still be observed.

In a further series of experiments, an alternative correct candidate selection mechanism, choosing the most harmonic candidate (see section 3), was evaluated. Figures 17 to 19 show the resulting perception grammar after 100,000, 160,000 and 400,000 trials respectively.

Noticeably, learning with this alternate correct candidate selection not only progresses more slowly than in the random selection, but also results in

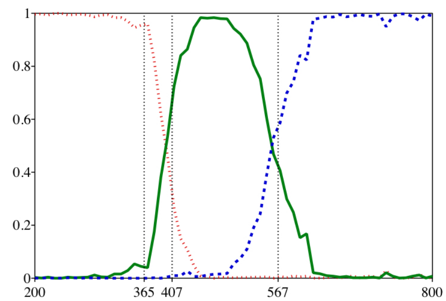

Figure 20: The output of the perception grammar after 80,000 trials (alternative harmonic selection).

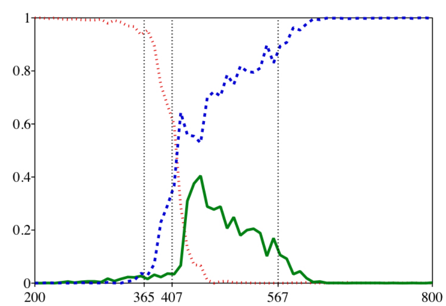

Figure 21: The output of the perception grammar after 60,000 trials (alternative harmonic selection with different constraint order).

the opposite of the targeted goal: All categories but |ɪ| are unlearned.

The reason for both becomes apparent when the four possible types of input to the initial recognition grammar and their respective consequences are considered; In the case that the input to the recognition grammar contains a |i| (e.g. "/tʃika/) or |ɛ| (e.g. "/tʃɛka/) that was intended as such by the speaker, the optimal candidate selected is also the correct candidate since the respective form is uniquely associated with one meaning - no learning takes place.

If the input contains |ɪ| and the intended meaning is the one belonging to the corresponding Spanish word containing |e|, the meaning of the optimal candidate will also be that which is intended (see figure 12) and, again, no learning takes place.

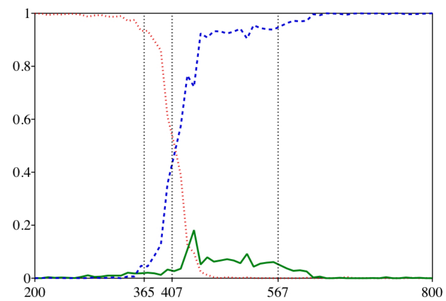

Figure 22: The output of the perception grammar after 80,000 trials (alternative harmonic selection with different constraint order).

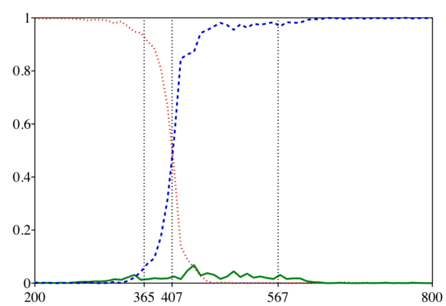

Figure 23: The output of the perception grammar after 100,000 trials (alternative harmonic selection with different constraint order).

Finally, when the input contains |ɪ| and the intended meaning is that belonging to the |i|-word, the most harmonic candidate will be that which maintains the input form, thus resulting only in semantic learning.

But if no phonological learning takes place, how are the |i| and |ɛ| categories eliminated? The reason for this lies in the occasional misperception of |i| as /ɛ/ and |e| as /i/ as discussed above. If for example the speaker intends to convey "Czech" and thus the word "|tʃeka|", but the virtual learner perceives the vowels as an /i/, resulting in "/tʃika/" as the input to the recognition grammar, the meaning of the optimal candidate in the recognition grammar will be "girl". Following this semantic error, the candidate "|tʃɪka| - 'Czech'" is selected as the correct form since it is more harmonic than "|tʃɛka| - 'Czech'". Consequently, there is a mismatch between the perceived vowel –/i/– and the vowel in the correct form, /ɪ/ and the cue constraint against the perception of /ɪ/ at the Hertz value in question is lowered while the ranking value of the constraint permitting the perception of /i/ is increased. This explains not only why after many trials only the |ɪ| category remains, but also why learning proceeds only slowly; as explained above, this "noisy" perception caused by an overlap of Gaussians is relatively rare and only about one in 195 trials leads to learning in the perception grammar.

In a final experiment, an adaptation of the harmonic candidate selection mechanism where only the violations in the lexical constraints were considered, was tested.

Figure 20 shows the output of the perception grammar after 80,000 trials. The boundary between /i/ and /ɪ/ has shifted to the position of the Spanish /i/ and /e/ boundary, but the /ɪ/ and /ɛ/ boundary has remained unchanged. Further trials showed that this situation does not not change when the number of trials is increased.

The reason for this can be found in the interplay between the harmonic selection mechanism and the initial ranking in the recognition grammar. The selection mechanism always choses the candidate with the intended meaning that incurs the least serious violation in the lexical constraints in the recognition grammar. That means that out of the two candidates with the correct meaning, the one that violates the lexical constraint that is lower ranked is picked as the winner. As the ranking of the lexical constraint is consequently decreased, the low position of the constraint is further reinforced. This in turn means that the most harmonic candidate for a given meaning is constant across trials and determined by the initial ranking order.

For the minimal pair "/tʃika/-/tʃɛka/", the ranking order is *LEX |ɛ| 'Czech' >> *LEX |ɪ| 'girl' >> *LEX |i| 'girl' >> *LEX |ɪ| 'Czech'. Consequently, if an input intended as 'girl' is recognized as 'Czech', the most harmonic candidate is "|tʃika| 'girl'". On the other hand, for the opposite case, the most harmonic candidate with the meaning Czech is |tʃɪka| 'Czech', meaning that the correct Spanish lexical item is never given as feedback, thus inhibiting the adjustment of the /ɪ/-/ɛ/ boundary.

While the harmonic selection strategy fails to yield good learning results given the initial rankings so far, the dependence of the mechanism on the ranking order means that a different order should yield better learning results.

To this end, the experiment was repeated with lexical constraints ranked according to the following scheme:
*LEX |ɪ| 'girl' >> *LEX |i| 'girl' >> *LEX |ɪ| 'Czech'>>*LEX |ɛ| 'Czech' .

The output of the resulting perception grammar (see figures 21 to 23) shows that this strategy leads to very good results, namely successful acquisition of the Spanish /i/-/e/ category boundary without notable interference from noise and thus stable results. However, the motivation for the alternative ranking is purely practical and it is not clear to what extent it lends itself to a theoretical explanation.

5 Conclusion

I have shown that Escudero's model for the SUBSET problem leads to the predicted result, the reduction of the /ɪ/ category for a virtual Dutch learner of Spanish.

The result however is highly dependent on the parameters used and noise and consequent lack of convergence poses a problem for the model.

Reducing the standard deviation of the Gaussian distributions from which the training data are drawn can solve this problem to a certain degree.

Using a harmonic selection mechanism for the winner gives results that are worse with both the |i| and |ɛ| categories being unlearned.

An adaptation of the harmonic selection mechanism where only lexical constraints are considered for finding the most harmonic candidate and were a different initial ranking in the recognition grammar is used yields the best results, but is practically and not theoretically motivated.

Another approach to improving performance might be to employ a decreasing learning rate.

A Appendix

The following minimal pairs were used in the experiments:

checa - chica *Czech - girl*	mesa - misa *table - Mass*
checo - chico *Czech - boy*	memo - mimo *fool - mime*
fecha - ficha *date - token*	reto - rito *dare - rite*
gres - gris *stoneware - gray*	rezo - rizo *prayer - curl*
lega - liga *layman - league*	veda - vida *prohibition - life*
lema - lima *motto - file*	peso - piso *weight - floor*
meca - mica *Mecca - mica*	

References

[1] Paul Boersma. *Functional Phonology. Formalizing the interactions between articulatory and perceptual drives.* PhD thesis, Universiteit van Amsterdam, 1998.

[2] Paul Boersma and Paola Escudero. Learning to perceive a smaller L2 vowel inventory: An optimality theory account. Rutgers Optimality Archive 684, 2004.

[3] Paul Boersma and David Weenink. Praat, a system for doing phonetics by computer. Technical report, Institute of Phonetic Sciences of the University of Amsterdam, 1999.

[4] Paola Escudero. *Linguistic Perception and Second Language Acquisition. Explaining the attainment of optimal phonological categorization.* LOT Dissertation Series 113, 2005.