The Influence of Targeted Input on the Production of L2 Vowels

BA Thesis English Language and Culture, Utrecht University, November 2012

Marleen Berkhout
3711129

First reader: K.D.C.J. Sebregts, MA
Second reader: Dr. M.E.H. Schouten

Table of Contents

1. Introduction

The older second language (L2) learners are, the more trouble they seem to have with acquiring a new language (Baker et al. 318); consequently, native-like attainment of L2 pronunciation can pose difficulties for learners. For example, although the Dutch seem to think they speak good English (*NCRV's Altijd Wat* Par 1*),* they almost always produce their English with a distinct Dutch accent. It is interesting to observe the occurrence of this phenomenon regardless of plentiful input coming from, for example, available British and American television programmes, and English lessons in school. This study investigates the influence of native English input on Dutch native speakers, and what role perception and production play in this matter.

1.1 SLM

The *problem* with acquiring L2 sounds is that the learner is inclined to use already established first language (L1) phonemes for new L2 sounds (Polivanov in Escudero 3). For example, Dutch and English /ɪ/ are so similar, that Dutch speakers can and will use their Dutch /ɪ/ when speaking English (Flege 19). However, L2 sounds that have no truly corresponding L1 sound get the same treatment; when speaking English, Dutch speakers will probably not produce an aspirated /p/ in [p^hɛt] because the Dutch unaspirated /p/ is still quite similar, and should be recognised by native speakers of English as an allophone of the phoneme /p/, consequently not obstructing communication; the latter is known as equivalence classification (Flege 13).

No two sounds from two different languages are ever completely identical. However, to help categorise the different L2 sounds, they can be divided in three groups: identical, new, and similar sounds (Flege16). The "identical" sounds have corresponding versions in L1, and therefore need not be separately acquired, as the L1 versions can be used (as with Dutch and

English /ɪ/ (Flege 19)). Similar and new sounds have no truly corresponding versions in L1, but are replaced with L1 sounds that come close; similar sounds, such as English /i:/ for native speakers of Dutch (Flege 21)**,** only slightly differ from the replacing L1 sound; new sounds, such as for example English /æ/ for native Dutch speakers (Flege 20), differ significantly from their replacements. According to Flege's Speech Learning Model, only new sounds are still acquirable, even at a later stage in life, as, unlike the situation with similar sounds, there is still room for new phonetic categories in unoccupied phonetic space (12).

1.2 Perception versus production

It seems plausible that in order to be able to produce "new" sounds, speakers must be aware of them (Flege11); perception should precede production. However, several studies show disagreement on whether L2 perception is crucial for production or vice versa (Llisterri 92). For example, Flege shows that native speakers of French who were considered experienced speakers of English were able to perceive English aspirated /t/without producing it correctly (16). This implies that perception precedes production. Other research shares this view (Barry; Rochet). However, a study by van Heuven et al. showed that good production does not necessarily indicate proper perception; they found that Dutch speakers of good English, as "judged by phonetically trained native listeners", can still have "an imperfect perceptual representation of the target-language sound system" (115). Also, Sheldon and Strange found that Japanese speakers of English who produce "acceptable" English /l/s and /r/s do not always perceptive the contrast between /l/-/r/ perfectly (Sheldon and Strange in Escudero 3). Furthermore, Flege and Bohn have concluded that in early years of L2 acquisition, leaners' perception can precede production, but, when speakers gain more experience with the language, their production can actually outpace their perception, for example because of social pressure to perform (69). Regardless of this on-going debate, perception and production are inextricably related (Llisterri 98).

1.3 The current study

Irrespective of whether perception or production comes first, environmental influences, such as the length of residence in L2 countries, usage, and quality/quantity of input, can also be of importance (Best and Tyler; Flege and Bohn). If SLM is right, non-native speakers should be able to learn new L2 sounds if exposed to them sufficiently. The current Dutch accent situation therefore suggests a lack of input. With SLM in mind, the main research question of this study is: does targeted input from a native speaker of English help improve the production of new L2 sounds?

To answer the research question, ten native Dutch teenagers were recorded reading out an English text before and after targeted native input on the difference between English /ɛ/ and/æ/. For this study, English /ɛ/ is considered to be identical to Dutch /ɛ/; this means that this sound should pose no problems for the participants; /æ/ is considered a new sound for native speakers of Dutch that is initially usually replaced by Dutch /ɛ/ (van Heuven et al. 112). Flege's SLM hypothesises that targeted native input should cause the participants to improve both perception and production of /æ/. The drawback of this approach is that it will be hard to determine if contrast between /ɛ/ and /æ/ has been perceived if pronunciation does not improve. However, this approach will directly show the influence of this input on production if improvement of /æ/ is present.

2. Method

2.1 The current study

Ten native Dutch participants were each recorded while reading out a text consisting of several words with /ɛ/- and /æ/-sounds. Afterwards, they watched and listened to a short pre-recorded instructional PowerPoint presentation on the difference between /ɛ/ and /æ/. They were then recorded again while reading out the same text. This approach was reminiscent of the intuitive-imitation approach of pronunciation lessons; the participants only mimic L2 sounds without being explicitly told how to do so (Hashemian and Mahmood 969). Hashemian and Mahmood claimed that this approach is useful for vowel acquisition, albeit more for diphongs than monophongs (973). The recordings were analysed and compared to see if the subjects made a stronger contrast between /ɛ/ and /æ/ after receiving targeted native input. This part of the report will discuss the method used, starting with the setting. After this it will discuss the participants, and then describe the recorded text, the words selected for analysis, and the instructional presentation. Finally, it will describe how data were collected and analysed.

2.2. Setting

The experiment took place in a small, empty gymnasium. The recordings were made on a *Compactflash Audio Recorder* (Zoom H4) directly to the hard drive; they were recorded as WAV files at a sample frequency of 48 kHz, and a bit rate of 16 bits. The participants were placed approximately thirty centimetres from the microphone. Before reading, the participants were asked to read in a normal speaking tempo and to articulate clearly. They were also told that making mistakes would not matter, and that it was alright if any words were unfamiliar; pronunciation would probably not be affected by unfamiliarity, mainly because the spelling of the /ɛ/-/æ/ words used was consistent: twelve out of fifteen /ɛ/s were spelled with *e*, whereas all /æ/s were spelled with *a*. They read the text from a piece of paper they held in their hands. After this, they watched and listened to the presentation on a laptop placed on a table in front of them while wearing headphones. The second recording was done in the same manner as the first.

2.3 Participants

The ten participants were monolingually raised native speakers of Dutch between thirteen and sixteen years old. They were all members of Stichting Jeugdtheater Heiloo, a youth theatre group from Heiloo, North Holland. Three of them were male. They all attended regular Dutch secondary school. Before starting the experiment, every participant was asked to fill in a questionnaire in order to provide some general information, such as sex, age, their experience with the English language, etcetera. The results of this questionnaire are shown in table 2.1. The year they were in in secondary school ranged from second to fifth grade, and one participant was in her first year of vocational school. The number of years they were taught English in school varied from one to six. Five of them thought their teachers spoke RP, whereas one participant had an American teacher. The others did not know what accent they were taught. Fifty percent had never been to an English speaking country. The others had been at least once to either one or more of the following countries for a minimum of four days, though never longer than three weeks at a time: England, Scotland, Wales or the United States. Most of the participants reported that they watched a lot of English or American television, although four of them had no idea how many hours per week. The others, on average, watched 3.9 hours per week. Participants 1, 5, 7 and 8 reported dyslexia.

	Age	Sex	Grade	Dyslectic	Native language(s)	Years of English education so far/accent taught in school	Visits to English Speaking countries	English tv (hours/ week)
1	14	male	3rd	yes	Dutch	3, British	none	3 to 4
2	13	female	2nd	no	Dutch	1, don't know	none	2
3	15	male	4th	no	Dutch	6, British	none	don't know
4	13	female	2nd	no	Dutch	3, don't know	England 2 wks	don't know
5	14	male	2nd	yes	Dutch	4, British	England 2 wks, Scotland 3x1 wk, USA 3 wks	7
6	15	female	4th	no	Dutch	5, don't know	England 2 wks, USA 4 wks	3
7	16	female	4th	yes	Dutch	6, don't know	none	don't know
8	16	female	5th	yes	Dutch	6, American	England 4 days	don't know
9	13	female	2nd	no	Dutch	2, British	none	2
10	16	female	1st year MBO	no	Dutch	6, British	Scotland 1 wk, Wales 1 wk, England 4 days	don't know

Table 2.1 Results questionnaire.

2.4 Text

The recorded text was a short story (see Appendix A) to capture the vowels in connected speech at a normal speech rate. In total, thirty-two words from the text were analysed: fifteen with /ɛ/ and seventeen with /æ/ (see table 2.2). All but four words were monosyllabic; the four dissyllabic words had stress on the target vowel (/ɛ/ or /æ/).

As shown in table 2.2, the words include five minimal pairs, three near-minimal pairs because of an added bound morpheme to one of either words (e.g. *pecks-pack*), and sixteen words that could be compared to words with contrasting vowels because of corresponding codas or onsets (e.g. *met* could be compared to *man* and *cat*). *Impress* was the only word without an /æ/-partner with a corresponding coda or onset. The choice to use several (near) minimal pairs in the text originated from the idea that during analysis it would be easier and

/ɛ/	/æ/
Men	Man
Beg	Bag
Gem	Jam
Ben	Ban
Dead	Dad
Lend	Landed
Bed	Badly
Pecks	Pack
Said	Lads
Met	Cat
Bet	Patrick
Ten	Dan
	Can
	Plans
Friend	And
	Band
Mess	Mad
Impress	

Table 2.2 In-text words chosen for analysis.

clearer to compare similar words. Moreover, the minimal pairs created the opportunity to not only compare contrasts between vowels per person, per word, and per vowel, but also per minimal pair and corresponding coda/onset. Furthermore, minimal pairs were also used in the instructional presentation on the contrast between /ɛ/ and /æ/ because this was the plainest way for this relatively small experiment to make the distinction between the vowels clearest; hence it made sense to analyse the pronunciation of these words. The decision to not record and analyse every minimal pair from the PowerPoint was made on semantic grounds: by keeping the text semantically logical, the participants would recognize words more quickly, thus read more fluidly and confidently (Harley 202). Five in-text words were not part of the instructional PowerPoint: *friend*, *impress*, *plans*, *mad*, and *cat*. These words had a monitoring function: if increase in contrast occurred, it would be interesting to examine if the participants had only imitated what they had heard in the presentation, or if they had applied what they had learned about /ɛ/ and /æ/ to other words as well.

2.5 Instructional presentation

The instructional PowerPoint presentation on the contrast between /ɛ/ and /æ/ presented twenty minimal pairs (table 2.3), one pair per slide (figure 2.1):

/ɛ/	/æ/
Bed	Bad
Bet	Bat
Ben	Ban
Beg	Bag
Pen	Pan
Menx	Man
Mess	Mass
Ten	Tan
Dead	Dad
Den	Dan
Gem	Jam
Ken	Can
Send	Sand
Ent	Ant
Said	Sad
Less	Lass
Mat	Met
Pack	Peck
End	And
Lend	Land

Table 2.3 Minimal pairs presented in instructional PowerPoint presentation.

The pairs were depicted in image and their spelled form, and an audio file of a native adult male Briton saying the pairs played while the pairs were depicted. The British accent was chosen simply because it was available for recording. The PowerPoint started with a short instruction that told the participants to only watch and listen closely, and, again, that it was alright if a word was unfamiliar, this time not only because of spelling consistency, which again occurred, but also because it was assumed that the pictures would clarify any unknown words. The participants were able to go to the next slide by pressing space on the laptop. The audio file would play automatically when a slide appeared.

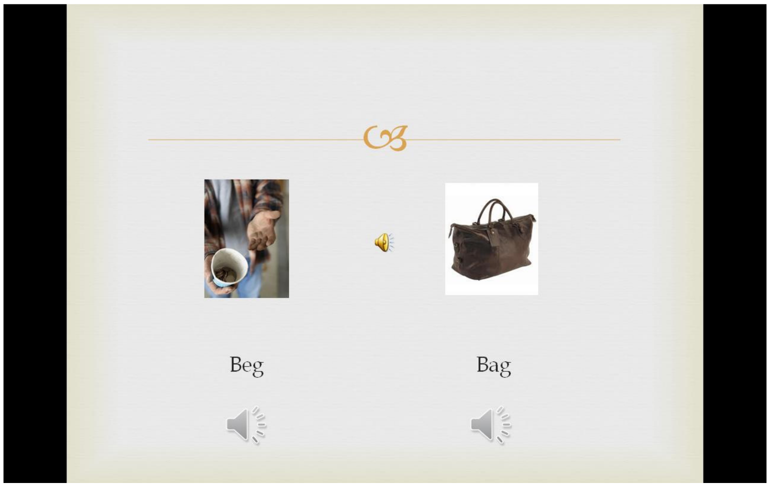

Figure 2.1 Example of PowerPoint slide. Source image left: http://blog.entrepreneurthearts.com/2009/12/25/avoiding-the-begging-cup/
Source image right: Barbour Leather Medium Travel Explorer Bag - Dark Brown; http://www.barbourbymail.co.uk/Barbour-Leather-Medium-Travel-Explorer.html

2.6 Data collection

The files were loaded into *Praat 5.3.23*, a speech analysis programme (Boersma and Wiersma 2012), after being transferred to a laptop without being resized. For each recording's thirty-two words, vowel length, and the first and second formant (F1 and F2) were measured for vowel contrast comparison. F1 and F2 stand for the high-low and the front-back dimension, respectively. Measurements in *Praat* were performed manually, and the formant settings used are displayed in figure 2.2.

To determine the vowel length, the vowels were isolated from the preceding and following consonants by auditory analysis; the sound waves depicted by *Praat* further assisted in determining these cut-off points. The point at which to measure formants was initially selected by putting the timeline in the middle of the vowel after isolation. The formants were then automatically extracted. Analysis showed some unlikely measurements, due to, for example, participants who spoke unclearly or too softly. In these cases the timeline would be

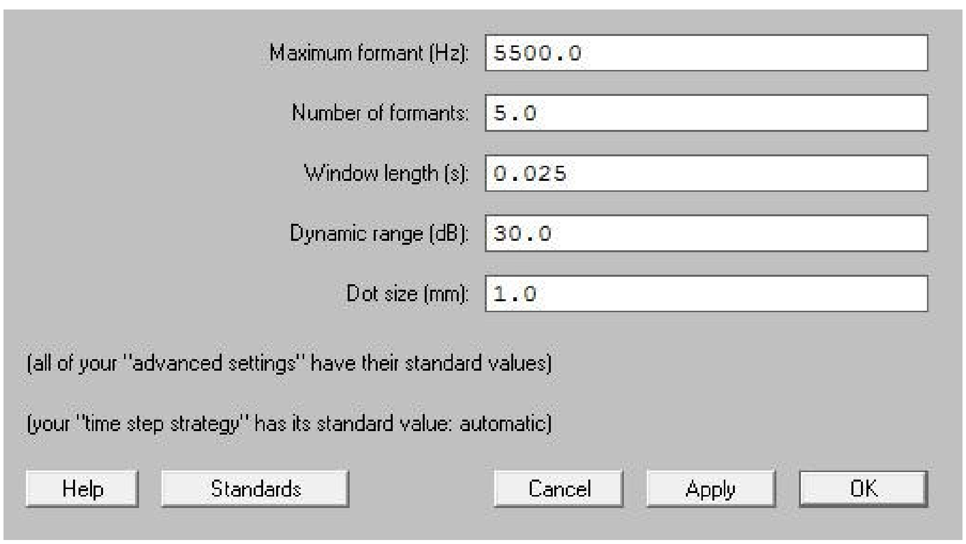

Figure 2.2 Formant settings.

manually placed at two clearer measuring points around the centre of the vowel, and the average of the two measurements was taken, or the vowel would be discarded. Participant 6's *mad* before input and participant 10's *cat* after input were discarded, as in both cases creaky voice made it impossible to measure formants. Also, *Praat* sometimes wrongly registered a formant in between the actual F1 and F2, which would be incorrectly listed as F2, or would present unlikely values. Again, in these cases either the timeline was manually placed at a better measuring point, or the number of formants shown was set to 4.0, which corrected these incorrect measurements. Any outliers discovered after all data was acquired were treated similarly to avoid incorrect deviations; participant 7's *band* and *badly* before input were discarded because their formants turned out to be unreadable, as were participant 9's *pecks* after input, and participant 10's *can* before input.

Some words were discarded before they were measured because of pronunciation errors: participant 1 said *just* instead of *dead* before input, and *planes* instead of *plans*, also before input; participant 2 said *laids* instead of *lads* after input, and *say* instead of *said* before input; participant 5 said *laids* instead of *lads* both before and after input; *and* underwent elision with participant 8 before input, and *pack* was realised as *pick*, also before input; participant 9 said *me* instead of *met* after input; participant 10 changed an entire sentence before input, in which she skipped *bet*.

Some mispronunciations did not affect the vowels, but merely the consonants that surrounded the vowels: *ten* was replaced by *fen* with participants 1 and 6 before input, and with participant 5 after input; *gem* was pronounced as /g/*em* with participants 5 and 7 before input; *beg* was realised as *beck* with participant 9 after input. During data processing, these results were taken out when the mean minimal pair values were computed. Furthermore, during comparison between corresponding onsets/codas, the mispronounced words were removed from the target group if the mispronounced consonant no longer corresponded with this target group, and, if possible, included in the group for which the target consonant was identical to mispronounced consonant (e.g. *gem* pronounced as /g/*em* was put in with the velar plosives).

The measured vowel lengths and formant readings were put in an excel file of which table 2.4 shows an example:

		friend before	friend after	man before	man after	etc.
1	length (ms)	84	85	88	105	
	formant 1 (Hz)	530.02	528.17	582.15	550.19	
	formant 2 (Hz)	1767.94	1794.54	2027.43	2034.26	
2	length (ms)	130	206	218	199	
	etc.					

Table 2.4 Example of collected data.

The numbers in the first column represent the participants: blue for males, pink for females. The first row displays the vowel length in milliseconds; the second and third row represents the first and second formant in Hertz respectively. Columns three, five, seven, etc. represent all the values for the vowels recorded before the instructional presentation; columns four, six, eight, etc. represent the values recorded after the presentation.

In order to observe if vowel contrast between /ɛ/ and /æ/ had or had not increased after native input, it was decided to make global comparisons along various dimensions: between speakers, the phonological context, and the minimal pairs. In order to do this, first the mean values per person were calculated. Secondly, overall mean values per vowel were computed. Thirdly, mean minimal pair values, and corresponding onset- and coda-mean values were calculated. The next section will present the results of these mean values.

3. Results

According to the English vowel triangle (figure 3.1), F1 for /æ/ with British adult male speakers approximately lies around 700 Hertz; in this study, a contrast of fifty Hertz or more was considered meaningful. The F2 values were normalised by deducting the first formant value from the second formant value (F2-F1) in order to neutralise the acoustic variation among individuals (Johnson 1). In figure 3.1, the F2-F1 value for /æ/ approximately lies around 1000 Hertz; a contrast of 100 Hertz or more was considered meaningful in the experiment. Both margins are rather conservative; this to avoid any precipitous conclusions. Although both /ɛ/ and /æ/ are considered short vowels (House 1175), contrast in vowel length could possibly be relevant, as /æ/ is often lengthened by speakers of British English in “certain common words” (Collins and Mees 94). Furthermore, speakers of American English also often lengthen /æ/ (Collins and Mees 137). In different syllable contexts vowel lengths are known to differ; for example, vowels in monosyllabic words act differently from stressed first syllable vowels in polysyllabic words (Umeda 434). To compare the difference in vowel length between the vowels, the absolute vowel lengths were normalised. To determine a normalising factor for each participant, the absolute length of the total text per participant was divided by 100. To then calculate the normalised vowel lengths, the absolute vowel lengths were divided by their factor. These normalised vowel lengths will be used throughout the rest of the results section.

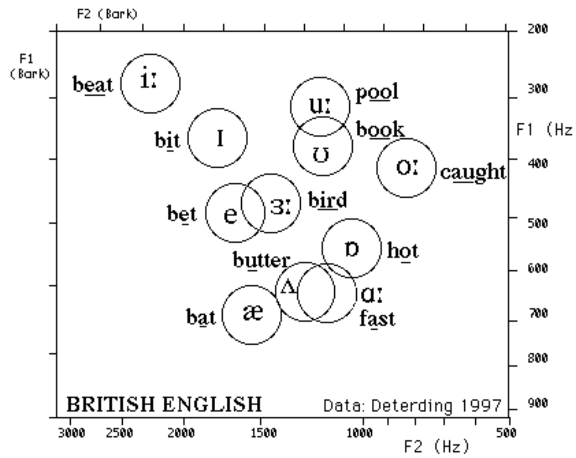

Figure 3.1 British Engligh vowel triangle for adult male speakers. Source: David Detering on University of Helsinki Institute of Behavioural Sciences’ website, par. 2.

In this study the vowel length contrast will be considered meaningful with a difference of 20 percent or more.

3.1 Overall mean values

Table 3.1 shows the mean values per vowel for all participants. L stands for mean vowel length in milliseconds (ms); F1 and F2-F1 represent the first formant and first formant subtracted from the second formant, respectively, in Hertz (Hz); negative figures indicate that the second value is smaller than the first.

	L before input (ms)	**L after input (ms)**	**diffe-rence**	**F1 b.i. (Hz)**	**F1 a.i. (Hz)**	**dif. (Hz)**	**F2-F1 b.i. (Hz)**	**F2-F1 a.i. (Hz)**	**dif. (Hz)**
mean /ɛ/	253	287	34	653	670	17	1277	1275	-2
mean /æ/	263	296	33	646	680	34	1303	1274	-29
contrast	10	9		-7	10		26	-1	

Table 3.1 Mean values for all vowels recorded.

As shown in table 3.1, before the presentation the difference in values between all /ɛ/s and /æ/s was unsubstantial for both formant values and the mean vowel length. This contrast did not change much after the presentation; the mean F1 for /æ/ became only slightly higher than the mean F1 for /ɛ/ value after input; the mean F2 for /æ/ approached the mean F2 for /ɛ/ further; the mean length contrast was only slightly smaller after input. Note that both mean F1 values went up, and both mean F2-F1 values went down. Also, both length values became slightly higher.

3.2 Individual Results

The results of the mean vowel values per person before and after native input are displayed in table 3.2. L here stands for absolute vowel length, in milliseconds (ms); F1 stands for the first formant in Hertz (Hz); F2-F1 is the normalised F2 value, where the first formant has been deducted from the second formant, in Hertz (Hz). Each participant’s colour represents his or her gender: blue for male, pink for female. Negatives indicate that the second value is smaller than the first.

speaker		mean /ɛ/ before input	mean /æ/ b.i.	difference	mean /ɛ/ after input	mean /æ/ a.i.	dif.
1	L (ms)	233	238	5	342	314	-28
	F1 (Hz)	549	543	-6	546	540	-6
	F2-F1	1294	1339	45	1334	1358	24
2	L (ms)	278	277	-1	356	281	-75
	F1 (Hz)	630	597	-33	645	623	-23
	F2-F1	1467	1539	72	1488	1561	73
3	L (ms)	228	294	66	231	296	65
	F1 (Hz)	570	577	8	585	601	16
	F2-F1	1061	1102	41	1066	1084	19
4	L (ms)	198	184	-14	242	250	8
	F1 (Hz)	620	629	8	617	680	64
	F2-F1	1206	1219	12	1219	1188	-32
5	L (ms)	275	255	-20	265	263	-2
	F1 (Hz)	564	552	-12	584	573	-11
	F2-F1	1316	1341	25	1303	1333	30
6	L (ms)	245	257	12	251	257	6
	F1 (Hz)	699	714	15	699	719	20
	F2-F1	1272	1203	-69	1307	1267	-40
7	L (ms)	311	320	9	373	381	8
	F1 (Hz)	728	730	1	731	744	13
	F2-F1	1282	1329	48	1261	1305	44
8	L (ms)	288	327	39	315	339	24
	F1 (Hz)	721	687	-34	735	691	-44
	F2-F1	1238	1302	64	1221	1241	20
9	L (ms)	252	218	-34	240	236	-4
	F1 (Hz)	725	681	-44	750	714	-36
	F2-F1	1416	1449	33	1370	1404	33
10	L (ms)	224	270	46	252	344	92
	F1 (Hz)	720	762	41	812	918	106
	F2-F1	1233	1203	-30	1191	1002	-188

Table1 3.2 Mean values per vowel per person.

It becomes clear from table 3.2 that before the instructional presentation none of the participants had considerable contrast for F1 or F2-F1 values between /ɛ/ and /æ/; both /ɛ/ and /æ/ were produced like /ɛ/. The biggest mean contrast in F1 values was only 44 Hz (participant 9), and her mean F1 for /ɛ/ was actually higher than that for /æ/.In fact, five participants' mean F1 for /æ/ frequencies were actually lower than their mean F1 for /ɛ/ frequencies, and eight participants' mean F2-F1 for /æ/ were higher than their mean F2-F1 for

/ɛ/ frequencies. The biggest F2-F1 contrast where /æ/ was lower than /ɛ/ was only 69 Hz (participant 6). As for mean vowel length, apart from participants 3 and 10, whose /æ/s were noticeably longer, no one had any notable mean vowel length contrast.

After input, eight participants had no notable difference in their mean F1 vowel contrast. Participants 4 and 10, who had higher meanF1 values for /æ/ than for /ɛ/ before input, showed an increase in contrast for their F1 vowel values. The mean F2-F1 contrast only increased considerably with participant 10. Although most mean /æ/'s became longer after input, so did most mean/ɛ/'s.[1] In fact, in seven cases the mean vowel length difference between /ɛ/ and /æ/ became smaller after input. Again, only participant 10's length contrast between /ɛ/ and /æ/ length increased considerably after input, while participant 3's contrast stayed the same.

3.3 Minimal pairs and syllable contexts

Table 3.3 compares the mean values for all participants per (close) minimal pair. L here stands for mean vowel length; F1 and F2-F1 represent the first formant and first formant subtracted from the second formant, respectively, in Hertz (Hz); negative figures indicate that the second value is smaller than the first.

Before input, only *peck-pack* and *gem-jam* showed real contrast between the F1 values. However, *pack* actually had a lower F1 value then *peck*. Furthermore, there was no contrast between the F2-F1 values, except between *peck* and *pack*. However, the F2-F1 value for *pack* was higher than the one for *peck*, which should have been reversed. Of the five minimal pairs with longer /æ/s, only three (*dad, bag and man*) were considerably longer; the /æ/ in pack was notably shorter than the /ɛ/ in peck.

[1] This would point to slower speech during the second recording; however, eight participants actually spoke faster after input, perhaps as a result of increased familiarity with the text. The fact that /ɛ/ and /æ/ actually were longer after input is therefore remarkable. It could indicate a task effect.

After input, the contrast between the F2-F1 values became notable for *beg-bag*. Of the three minimal pairs with longer /æ/s before input, only the contrast between *dead* and *dad* contrast increased. The contrast between *men* and *man* did not change, and the contrast between *beg* and *bag* decreased.

	L before input (ms)	**L after input (ms)**	**diffe-rence**	**F1 b.i. (Hz)**	**F1 a.i. (Hz)**	**dif. (Hz)**	**F2-F1 b.i. (Hz)**	**F2-F1 a.i. (Hz)**	**dif. (Hz)**
dead	232	257	25	649	654	5	1351	1271	-80
dad	342	428	86	637	685	48	1310	1281	-29
difference	110	171		-12	31		-41	10	
ben	240	293	53	666	700	34	1247	1267	20
ban	253	300	47	633	667	34	1307	1271	-36
dif.	13	7		-33	-33		60	4	
Beg	253	284	31	633	667	34	1298	1303	5
bag	340	321	-19	649	690	41	1280	1209	-71
dif.	87	37		16	23		-18	-94	
men	303	328	25	669	702	33	1327	1357	30
man	412	423	11	670	703	33	1333	1329	-4
dif.	109	95		1	1		6	-28	
gem	244	302	58	608	645	37	1313	1321	8
jam	270	305	35	653	695	42	1276	1243	-33
dif.	26	3		45	50		-37	-78	
peck(s)	227	269	42	705	695	-10	1219	1242	23
pack	154	222	68	633	694	61	1301	1200	-101
dif.	-73	-47		-72	-1		82	-42	

Table 3.3 Comparison minimal pairs

The length contrast between *bed* and *bad* increased; the contrast between *peck* and *pack* decreased.

Table 3.4 shows the mean values for all participants per onset context. The following contexts were compared: contrast between /ɛ/ and /æ/ after nasals; contrast between /ɛ/ and /æ/ after plosives; contrast between /ɛ/ and /æ/ after alveolar plosives; contrast between /ɛ/ and /æ/ after bilabial plosives. Table 3.5 shows the mean values for all participants per coda context. The following contexts were compared: contrast between /ɛ/ and /æ/ before nasals; contrast between/ɛ/ and /æ/ before plosives; contrast between /ɛ/ and /æ/ before alveolar

	L before input (ms)	L after input (ms)	diffe-rence	F1 b.i. (Hz)	F1a.i. (Hz)	dif.	F2-F1 b.i. (Hz)	F2-F1 a.i. (Hz)	dif.
/ɛ/ after nasals	255	289	34	653	676	23	1310	1306	-4
/æ/ after nasals	362	406	44	679	698	19	1299	1306	7
difference	107	117		25	21		-11	0	
/ɛ/ after plosives	253	291	38	664	681	17	1274	1268	-6
/æ/ after plos.	252	295	43	638	678	40	1298	1254	-44
dif.	-2	3		-26	-2		24	-14	
/ɛ/ after alveolar plos.	241	261	19	666	674	9	1311	1261	-50
/æ/ after alv. plos.	297	380	83	634	679	44	1338	1315	-23
dif.	56	119		-31	4		27	54	
/ɛ/ after bilabial plos.	257	303	46	650	670	20	1261	1271	10
/æ/ after bilab. plos.	236	266	30	678	639	-39	1283	1233	-50
dif.	-21	-37		29	-30		22	-37	

Table 3.4Mean vowel values per similar onset.

	L before input (ms)	L after input (ms)	diffe-rence	F1 b.i. (Hz)	F1a.i. (Hz)	dif.	F2-F1 b.i. (Hz)	F2-F1 a.i. (Hz)	dif.
/ɛ/ before nasals	259	284	26	657	677	19	1277	1283	6
/æ/ before nasals	248	279	31	641	677	36	1324	1296	-28
difference	-11	-6		-16	1		46	13	-11
/ɛ/ before plosives	257	304	47	649	669	20	1303	1288	-15
/æ/ before plos.	273	316	43	651	683	32	1280	1248	-31
dif.	16	12		2	14		-23	-39	
/ɛ/ before alveolar plos.	264	315	50	640	663	22	1322	1285	-37
/æ/ before alv. plos.	280	332	52	654	680	26	1276	1264	-13
dif.	16	17		14	17		-46	-22	
/ɛ/ before velar plos.	240	277	37	669	684	15	1258	1294	36
/æ/ before vel. plos.	252	271	20	641	692	51	1290	1205	-86
dif.	12	-6		-28	8		32	-89	
/ɛ/ before -nd	249	257	9	665	662	-4	1235	1250	15
/æ/ before -nd	220	254	34	647	679	33	1306	1268	-38
dif.	-28	-3		-18	18		71	18	

Table 3.5Mean vowel values per similar coda

plosives; contrast between /ɛ/ and /æ/ before velar plosives; contrast between /ɛ/ and /æ/ before /-nd/. L stands for mean vowel length in milliseconds (ms); F1 and F2-F1 again represent the first formant and first formant subtracted from the second formant respectively, in Hertz (Hz); negative figures indicate that the second value is smaller than the first.

As demonstrated by table 3.5, both before and after input, no considerable contrast between /ɛ/ and /æ/ was found in any factor between the corresponding codas. Conversely, although no contrast in either formant between similar onset contexts was found before or after input, table 3.4 shows that vowel length was considerably longer for /æ/ than for /ɛ/ after nasals and alveolar plosives before input. After input, this length difference stayed the same for nasals, but increased for the alveolar plosives.

4. Discussion

4.1 Lack of overall improvement

Table 3.1 showed that the overall production of /æ/ did not improve after targeted input. However, this lack of improvement does not necessarily disprove that targeted native input can cause improvement of both perception and production of /æ/.

First of all, the study was not extensive; the input was limited. As Flege and Bohn have argued, there should be enough input to make pronunciation improvement possible (2). More intensive input over a longer period of time might cause contrast increase between /ɛ/ and /æ/, and would perhaps eventually even lead to near-native production of /æ/. For example, the study by Hashemian and Fadaei showed that pupils who followed an intensive three-week English course that included vowel pronunciation lessons through the intuitive-imitative approach improved the pronunciation of these vowels (974). Although this study did not focus specifically on new versus similar vowel contrast, it did point out the importance of extensive input.

Secondly, the participants' perception of /æ/ might have improved, despite the participants' lack of improvement in production. As discussed earlier, there is still debate going on about what comes first: perception or production. The results of this study make it impossible to further comment on this discussion. However, if it is indeed the case that perception precedes production, the participants might have perceived the contrast between /ɛ/ and /æ/, but did not benefit enough from it to improve production. The only evidence that the participants perceived something was the fact that most participants' absolute text lengths were shorter after input, while their mean /ɛ/s and /æ/s were longer. This indicates that they were aware of both vowels during the second recording after having received input. However, this probably indicates a metalinguistic phenomenon caused by the task; not perceptual improvement of the contrast.

As little overall mean increase in contrast occurred, increase in contrast between minimal pairs and different syllable context was limited as well. However, some results require further interpretation. First of all, the minimal pair contrasts made in the PowerPoint seems to have had little influence on the participants, as none of the minimal pairs showed overall mean increase in contrast in all three contrast indicators; F2-F1 contrast with *beg-bag* has no clear explanation. Increase in length contrast only occurred with *dead-dad*. Mean length contrast after alveolar plosives did increase after input, which could account for this contrast. However, then *bed-bad* and *beg-bag* should have shown increase in length contrast as well. This could indicate that the contrast between *dead* and *dad* is a coincidence. Secondly, the contrast between /ɛ/ and /æ/ before nasals and alveolar plosives only increased in lengthwise after input, but, apart from with *dead-dad,* this is not reflected in the minimal pairs described above. It could be that the other words with nasal or alveolar plosive onsets had increase in length contrast that made the notable difference. Thirdly, the lack of increase of length contrast before nasal codas is peculiar, because length contrast between the vowels before nasals would not have been an unlikely result.

4.2 Individual cases

Although the overall mean contrast between /ɛ/ and /æ/ was not notable for any overall mean factor after input, it does seem that participant 10 improved the pronunciation of /æ/ somewhat after input. Furthermore, her contrast in length was already notable before input; this could indicate that she was already aware of the differences between /ɛ/ and /æ/ before input, as Flege and Bohn have claimed that vowel length is often a cue for inexperienced speakers (69). Moreover, some participants produced a considerable contrast difference after input with some words. Why this did not show in their individual mean values could be explained by either contrast decrease with other words, or inaccuracies in the measurements. It is difficult to say why particularly these words with these participants showed an increase in contrast, as in the overall minimal pair comparison, only *dead-dad* showed increase in contrast, and only with the mean vowel length. Furthermore, there also seems to be no relationship between these words and the corresponding codas/onsets, as the participants did not apply it to other words with corresponding codas and onsets. It therefore seems likely that all above discussed individual cases can be attributed to coincidence; the size of this study makes it hard to draw any firm conclusions. Nonetheless, it should be kept in mind that individual "language aptitude and awareness" can play a role L2 acquisition as well (van Boxtel 180), and that the individual differences in these factors are perhaps more important than differences between speakers before and after input, or between more or less input.

5. Conclusion

The overall mean results in table 3.1 indicate that this study's prediction based on SLM regarding production was incorrect. Although this lack of increase in contrast between /ɛ/ and /æ/ production does not necessarily exclude any perceptual improvement, in this context it is hard to test perceptual progress, and thus draw any conclusion on this matter. However, regardless of these results, previous studies have shown that L2 learners should be able to improve perception and production of "new" L2 sounds. It is therefore not possible to claim that, even though this study did not necessarily prove the hypothesis, targeted native input has no effect on L2 perception and production. Unfortunately, the current study was probably not comprehensive enough to obtain useful results. What can be deduced from the results is that the amount of input offered during the study was probably not enough to improve, at least, short term production. However, the task did cause some sort of metalinguistic awareness among most participants. Perhaps this awareness can indirectly contribute to improvement of the pronunciation of /æ/ on the long run, which would imply that any amount of targeted input is useful. Furthermore, participant 10's results could indicate that individual variation should also be taken into account; it seems to be the case that L2 acquisition success, unlike L1 acquisition success in most cases, also depends on the learner's abilities and ambitions. It is clear that more extensive research is necessary to understand the effects of targeted native input on L2 pronunciation, and to learn more about how to apply this input in schools and language courses to get the best results.

WORKS CITED

Baker, Wendy, et al. "Child–Adult Differences in Second-Language Phonological Learning: The Role of Cross-Language Similarity." *Language and Speech* 51.4 (2008): 317-42. Print.

Barry, William. "Perception and Production of English Vowels by German Learners: Instrumental-Phonetic Support in Language Teaching." *Phonetica* 46.4 (1989): 155-68. Print.

Best, Catherine, and Michael Tyler. "Non-Native and Second Language Speech Perception: Commonalities and Complementarities." *Second Language Speech Learning: The Role of Language Experience in Speech and Perception and Production.* Ed. M.J. Munro and Ocke-Schwen Bohn. Amsterdam: John Benjamins, 2007. 12-34. Print.

Boersma, Paul, and David Weenink. *Praat: Doing Phonetics by Computer* [Computer Programme]. Version 5.3.32.2012. http://www.praat.org. 2012. Web. 17 Sept. 2012.

Bohn, Ocke-Schwen, and James Flege. "Perception and Production of a New Vowel Category by Second-Language Learners."*Second-Language Speech: Structure and Process.* Ed. Allan James and Jonathan Leather Second. Berlin: Mouton de Gruyter, 1997. 53-74. Print.

Boxtel, Sonja van. "Can the Late Bird catch the Worm? Ultimate Attainment in L2 Syntax." Diss. U of Utrecht, 2005. Print.

Collins, Beverley, and Inger Mees. *Practical Phonetics and Phonology: A Resource Book for Students.* Ed. Peter Stockwell. Abingdon: Routledge, 2003. Print.

Detering, David. "The Formants of Monophthong Vowels in Standard Southern British English Pronunciation." *Journal of the International Phonetic Association* 27.1-2 (1997): 47-55.

Escudero, Paola. “Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization.” Diss. U of Utrecht, 2005. Print.

Flege, James. “English Vowel Production by Dutch Talkers: More Evidence for the “New” vs “Similar” Distinction.” *Second-Language Speech: Structure and Process.* Ed. Allan James and Jonathan Leather Second. Berlin: Mouton de Gruyter, 1997. 11-52. Print.

Harley, Trevor. *Talking the Talk: Language, Psychology and Science*. New York: Psychology Press, 2010. Print.

Hashemian, Mahmood, and Batool Fadaei. “A Comparative Study of Intuitive-imitative and Analytic-Linguistic Approaches towards Teaching English Vowels to L2 Learners.” *Journal of Language Teaching and Research.* 2.5 (2011): 969-79. Print.

Heuven, Vincent van, et al. “Perception of Checked Vowels by Early and Late Dutch/English Bilinguals.” *Sounds and Sounds: Studies Presented to M.E.H. (Bert) Schouten on the occasion of his 65th Birthday.*Ed. Wim Zonneveld, Hugo Quené and Willemijn Heeren. Utrecht: UiL OTS, 2011. 103-16. Print.

House, Arthur. “On Vowel Duration in English.” *The Journal of the Acoustical Society of America.* 33.9 (1961): 1174-8. Print.

Johnson, Keith. “Speaker Normalization in Speech Perception.” *The Handbook of Bilingualism.*Ed. David Pisoni and Robert Remez. Oxford: Blackwell Publishers, 2005. 363-89. Print.

Llisterri, Joachim. “Relationships between Speech Production and Speech Perception in a Second Language.”*ICPhS: Proceedings of the 13th International Congress of Phonetic Sciences Volume 4, August 13-19, 1995.* Ed. Kjell Elenius and Peter Branderud. Stockholm: Congress Organisers at KTH and Stockholm University, 1995. 92-9. Print.

Polivanov, Evgeni. “La Perception des Sons d’une Langue Étrangère.” *Travaux du Cercle Linguistique de Prague* 4 (1931): 79–96. Print.

“Reageer: Hoe Goed Spreekt de Nederlander Engels.” *Altijd Wat.* NCRV. 12 May 2011. Web. 1 Nov. 2012.

Rochet, Bernard. “Perception and Production of Second-Language Speech Sounds by Adults.” *Speech Perception and Linguistic Experience: Issues in Cross Language Speech Research*. Ed. Winifred Strange. Baltimore: New York Press, 1995. 379-410. Print.

Sheldon, Amy, and Winifred Strange. “The Acquisition of /r/ and /l/ by Japanese Learners of English: Evidence that Speech Production can Precede Speech Perception.” *Applied Psycholinguistics* 3.3 (1982): 243-261. Print.

Umeda, Noriko. “Vowel Duration in American English.” *The Journal of the Acoustical Society of America*. 58.2 (1975): 434-45. Print.

Vowel Charts. Institute of Behavioural Sciences University of Helsinki, 2011. Web. 20 Oct. 2012.

Appendix A - Recorded Text

“My friend has a new red dress. Man, I am dead jealous. She bought it to impress the ten men she has met at band camp. She plays the harp rather badly. However, Ben and Dan like her, especially when she is wearing her dress. It has landed her many dates and pecks on the cheek already. All the lads beg her dad: can I take her out? His little girl is a real gem, and I am glad she is my friend, because she said she will lend it to me for my plans tonight with Patrick. I hope I won’t mess up the dress with jam, because I bet that would make her really mad. I will pack it in a bag under my bed and ban the cat.”