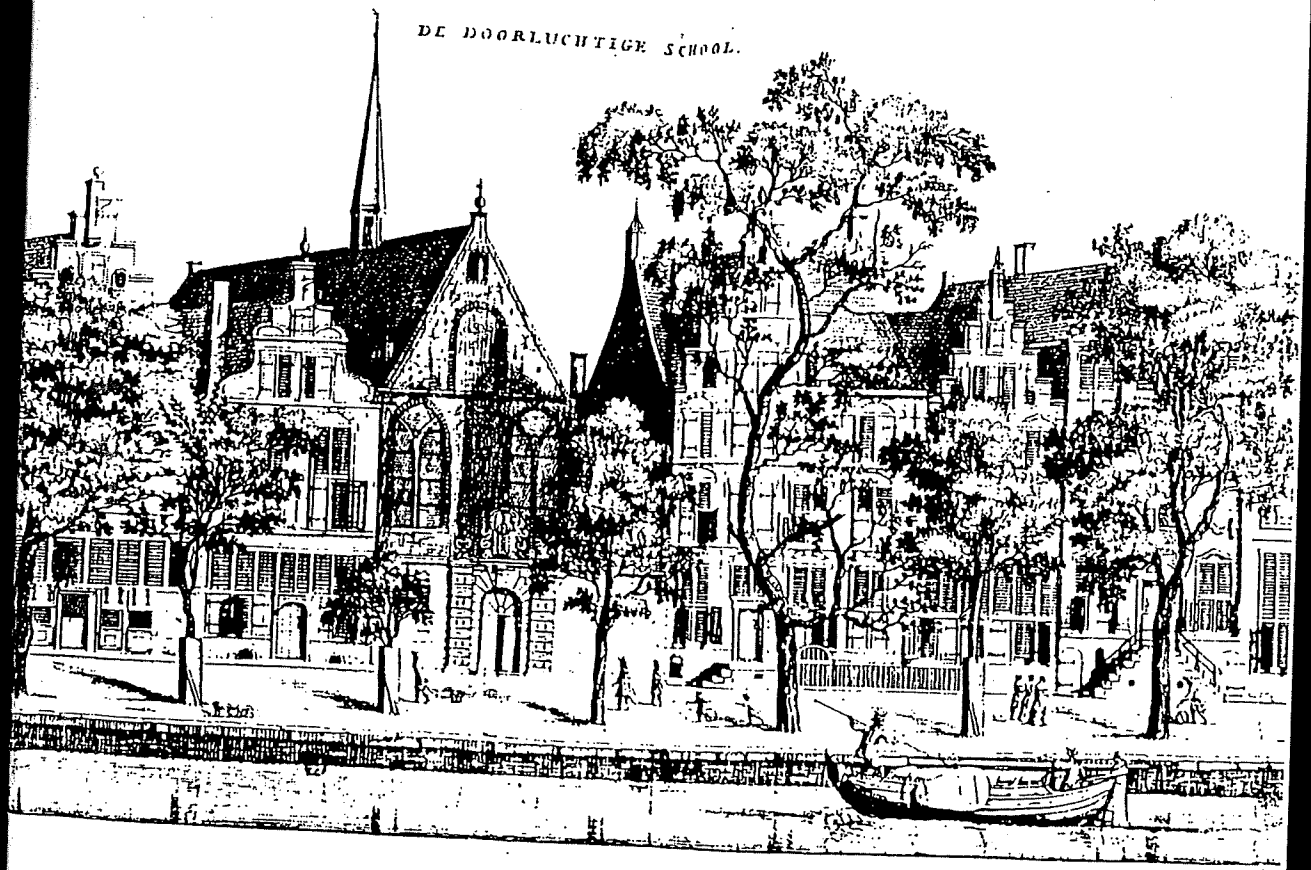
new sounds 92



ADULT ABILITIES IN L2 SPEECH: EVIDENCE FROM STRESS

John Archibald
University of Calgary

The question of adult ability in second language acquisition is, of course, well-covered territory. Traditionally, this issue was investigated under the heading of the Critical Period Hypothesis. More recently, a variety of concerns have been discussed under the heading of Age-Dependent Effects (Flynn & Manuel 1991). This term groups together such diverse elements as ultimate attainment in a second language, early advantage in L2 learning, and adult access to UG.

Studies now are looking very closely at *what* is acquired in order to assess the abilities of adults. Studies have ranged from Voice Onset Time to Subjacency. In spite of the breadth of coverage this question has received, the conclusions to be drawn are still controversial as Figure #1 illustrates:

Adults Can't Do "It"	Adults Can Do "It"	Arguable
Lenneberg (1967)	Olson & Samuels (1973)	Curtiss (1977)
Asher & Garcia (1969)	Snow & Hoefnagel-Höhle (1977)	McLaughlin (1984)
Seliger, Krashen &Ladefoged (1975)	Snow & Hoefnagel -Höhle (1978)	Singleton (1989)
Oyama (1976)	Neufeld (1979)	
Oyama (1978)	Snow (1983)	
Patkowski (1980)	Obler (1984, etc.)	
Newport (1984)	Ellis (1985)	
Newport & Supalla (1987)	Flege (1987, etc.)	
Coppieters (1987)	Major (1987, etc.)	
Scovel (1988)	Snow (1987)	
Johnson & Newport (1989)	Genesee (1988)	
Patkowski (1990)	Birdsong (1991)	
Long (1990)	Flynn & Manuel (1991)	
Hyltenstam (1990)	Singleton (1992)	
Hurford (1991)		
Johnson & Newport (1991)		

Figure #1. Previous studies of age-dependent effects in second language learning.

In order to make sense of these conflicting results and interpretations, Long (1990) distinguishes between the *Whether* question and the *Where* question. The *whether* question is concerned with such issues as whether adults have an initial advantage over children in early learning and *whether* children's ultimate attainment outstrips adults'. The *where* question is concerned with which aspects of grammar may be affected by this critical period.

As Birdsong (1991 & personal communication) points out, the Coppieters (1987) study is at the centre of both the *whether* and *where* questions. Birdsong has replicated Coppieters study (with methodological refinements) and claims that several of the NNS's

are not performing significantly differently than the NS's. This is a response to Long's call for a demonstration that a learner who had begun learning after the proposed sensitive period could attain native-like proficiency.

In his survey, Long suggests that a decline in phonological abilities occurs around the age of six. He claims that the results "all suggest that SL phonological attainment is strongly conditioned by learner age. Specifically, a native-like accent is impossible unless first exposure is quite early, probably before 6 in many individuals and by about 12 in the remainder" (266). Scovel argues that there is also a sensitive period for accent *recognition* in non-native speakers. His adult subjects did not detect accents as accurately as children. Later in this paper we will see how adult abilities in stress perception are very good.

One thing that I think needs to be said was said by Cook (1992). He points out that many researchers appear to be looking down their noses at a bilingual's use of the L2. We refer to it as an interlanguage in which the L2 grammar has not yet reached the stage of proficiency of a native speaker. We think of the bilingual's competence as being somehow deficient. Would it not make more sense, he said, to think of monolingual's competence as deficient for being monolingual? According to Cook (1991), multi-competence is the norm, and we should think of it that way.

My purpose here is not to address the broader question as to the existence of the critical period, but rather to consider some data from one domain that may help us to define the characteristics of the critical period.

Even in the studies which argue *against* the existence of a critical or sensitive period in second language learning, I think it is fair to say that many people still hold the view that phonology may be the heart and soul of the critical period hypothesis. Adults do tend to retain L1 accents, afterall. Scovel's (1988) book took this stance very clearly, in assuming that there was a critical period for *speech* but not for *language*.

What I would like to consider in this paper is just *what* aspects of L2 speech have been looked at under the heading of *phonology*. I will argue that most of the studies until recently have been concerned with segmental phonology and not with suprasegmental phenomena such as tone and stress (the work of Leather and James and very few others excepted). These areas have the potential of informing SLA research greatly.

Stress

In this paper I describe the results and implications of an empirical investigation of one aspect of the developing grammars of second language learners: stress assignment. The studies are designed to investigate the acquisition of English stress patterns by adult, nonnative speakers of English. Specifically, I look at the behaviour of learners from two different language types (with respect to stress): Polish and Hungarian. These

languages have different stress-assignment properties that influence the acquisition of English stress. The learners' acquisition of stress is examined in light of the metrical parameters proposed by Dresher & Kaye (1990).

Relevance of the Project

The project is relevant to the current research programme in learnability in 3 ways:

(1) it investigates one aspect of the final-state grammar and examines learners attempting to acquire that particular system of knowledge.

(2) the issue of how the knowledge underlying correct stress placement is acquired has been virtually ignored in both first and second language acquisition studies. Those who have looked at it have often assumed that stress is a single thing to be acquired (i.e. a single rule of stress placement). Second language researchers who have avoided this topic may have done so for similar reasons, assuming that the results would be uninteresting ("so Poles tend to stress the penult in English, so what?"). These studies show that stress is *not* a single thing to be acquired, and that a careful investigation of this phenomenon can help to explain *why* the learners behave in the way that they do. It can help us to come to a better understanding of the first language (L1) structures which will influence the second language (L2) grammar.

(3) the study will allow us to collect some empirical evidence relevant to a principles and parameters model of language acquisition. Within current linguistic research, it is generally accepted that theories of language acquisition which posit a mainly inductive learner (i.e. a hypothesis tester) are problematic (cf. Gold 1967, Lightfoot 1982). More recent work has been conducted within a deductive framework known as the principles and parameters model (cf. Chomsky 1981, Roeper & Williams 1987). Most of the L2 studies that have been done have been done in the area of syntax (Flynn 1987, 1989; White 1988, 1989, etc.). There has been very little done in the area of phonology (Dresher & Kaye 1990; Broselow & Finer 1991; Singh 1991; James 1990; Leather 1990). The phonological phenomenon of stress seems to be particularly well-suited to this type of study.

Research Design

The subjects of these studies were 23 Poles and 20 Hungarians. There was no significant difference between the two groups in terms of (1) number of errors made, (2) proficiency, and (3) vocabulary knowledge. Differences in their performance, then, can be attributed to their first language. The subjects had to perform both perception and production tasks related to stress. First of all, they had to read a list of words (see Appendix A) and a list of sentences (see Appendix B) out loud onto a tape recorder (this is the production test). Then, after a training

session, they had to listen to the same words and sentences read to them and they had to mark where they perceived the stress to be on the target words (this is the perception test). Later the production utterances were transcribed.

The experimental items were chosen to illustrate particular characteristics of English stress (quantity-sensitivity, extrametricality, stress retraction, differences between Noun and Verb stress, etc.)

The Study

The metrical parameters in question are drawn from Dresher & Kaye (1990) and are shown in Figure #2:

P1: The word-tree is strong on the [Left/Right].
P2: Feet are [Binary/Unbounded].
P3: Feet are built from the [Left/Right].
P4: Feet are strong on the [Left/Right].
P5: Feet are quantity-sensitive (QS) [Yes/No].
P6: Feet are QS to the [Rime/Nucleus].
P8A: There is an extrametrical syllable [No/Yes].
P8: It is extrametrical on the [Left/Right].

Figure #2. Universal metrical parameters.

Polish Stress

Polish is a language of essentially fixed stress. In words of more than one syllable, main stress falls on the penult. If at least two syllables precede the penult, then a secondary stress falls on the initial syllable. The parameter settings are shown in Figure #3 (taken in part from Hayes & Puppel,1984):

P1: The word tree is strong on the [Right].
P2: Feet are [Binary].
P3: Feet are built from the [Right].
P4: Feet are strong on the [Left].
P5: Feet are quantity-sensitive (QS) [No].
P6: Feet are QS to the [N/A].
P8A: There is an extrametrical syllable [No].
P8: It is extrametrical on the [N/A].

Figure #3. Polish metrical parameter settings.

Hungarian Stress

Hungarian also has fixed stress. Main stress is assigned to the initial syllable of every word. A secondary stress is assigned to alternating syllables after that. Hungarian is also quantity-sensitive (like English) but it is quantity-sensitive to the nucleus not the rime. The parameter settings are shown in Figure #4 (taken in part from Kerek (1971)):

P1: The word tree is strong on the [Left].
P2: Feet are [Binary].
P3: Feet are built from the [Left].
P4: Feet are strong on the [Left].
P5: Feet are quantity-sensitive (QS) [Yes].
P6: Feet are QS to the [Nucleus].
P8A: There is an extrametrical syllable [No].

Figure #4. Hungarian metrical parameter settings.

The bulk of the evidence from these studies provides support for the claim that adult L2 learners are transferring their L1 parameter settings. I will not go into those details here as I have presented them elsewhere (Archibald 1989a, 1989b, 1989c, 1990, 1991a, 1991b, 1992, forthcoming, to appear). Many of the errors of the Polish subjects resulted from the transfer of their settings of [P5] Quantity-Insensitive; [P8] No Extrametricality.

The Hungarian errors were often the result of transferring the setting of [P1] Word tree strong on the left; and [P6] Quantity-sensitive to the Nucleus. Some of the explanation for this behaviour can be made clear by examining Figure #5 which compares the various settings given:

	Default	Polish	Hungarian	English
P1 (Word tree)	left	right	left	right
P2 (Feet)	unbounded	binary	binary	binary
P3 (Direction)	left	right	left	right
P4 (Head)	left	left	left	left
P5 (QS/QI)	QI	QI	QS	QS
P6 (QS domain)	rime	NA	nucleus	rime
P8 (Extrametrical)	no	no	no	yes

Figure #5. Comparison of relevant parameter settings.

These data show that the learners are not reverting to a default setting. Elsewhere (Archibald (forthcoming)) I argue that Spanish speakers learning English do not revert to an unmarked [-extrametrical] setting; they transfer their L1 [+extrametrical] markings.

In general, though, it must be remembered that the kinds of subset relationships evident in many of the proposed syntactic parameters are absent with regard to these phonological parameters. This makes it difficult to say whether or not the notion of default setting is useful at all in these cases.

Findings of the Project

The following is a summary of some of the general points which have emerged from this project:

1. The subjects' L1 metrical parameter settings transfer into English.

2. Adult L2 learners do seem to be able to reset their parameters.

3. Adult L2 learners do not appear to violate proposed universals of metrical phonology.

4. Subjects' perception of stress is more accurate than their production of stress.

5. General grammatical ability does not correlate with accurate stress placement.

6. Not all phonological transfer has a muscular cause.

Adult Abilities in L2 Stress

But what I want to focus on in this paper is the accuracy of the subjects. They did very well at acquiring something as insanely complex as English stress. Figure #6 shows the profiles of the subjects in terms of the numbers of errors they made, their L1, their score on the Michigan test, a vocabulary test, and the order in which the tests were given to them. Figure #7 shows their accuracy rates more clearly.

The vast majority of the errors can be accounted for by obvious L1 transfer (Poles stressing the penult and Hungarians stressing the initial syllable). The forms which were not explicable by transfer were less than 5% for every class, and could be attributed to random variation.

As you can see, the adults performed very well on the perception task (significantly better than on the production task in terms of number of errors made. It could perhaps be argued that this is in some ways parallel to a grammaticality judgement task in syntax. But there is an obvious difference. In this task the subjects are being asked to make judgements about something for which there is a physical correlate. They can perceive things like pitch, duration, and loudness directly. In a way that they cannot perceive subjacency violations directly. There is the chance that they could be performing this task with non-linguistic faculties. But if this were the case, we would not expect to see L1 influence on these perception tasks, and we do. This argues that the tasks *are* tapping into linguistic competence.

Several pilot tests were conducted prior to this study which are relevant to this question. Tests were given to determine whether the phrasing of the instructions affected the outcome of the testing (e.g circle the syllable you perceive to be stressed; place a mark over the letters you perceive to be stressed, etc.) and whether native speakers performed reliably on the same test from time 1 to time 2. Regarding the first question, the method of marking stress did not significantly affect the results. Regarding the second question, native speakers did not perform significantly differently from one trial to the next. But the native speakers were only given perception tasks. I assumed that

		#Errors						
Subj #	W-Prod	S-Prod	W-Perc	S-Perc	Michigan Score (/40)	Order	L1	Vocab. Score (/35)
3	18	15	8	5	32	2	P	32
10	3	0	0	0	36	2	P	33
100	20	14	11	13	16	2	P	0
103	11	13	0	0	20	2	P	33
104	7	6	0	0	23	2	P	31
105	3	3	0	0	16	2	P	29
106	16	14	9	11	18	2	P	30
108	4	10	5	5	18	2	P	28
109	10	12	0	0	25	2	P	20
113	7	7	2	4	25	2	P	34
200	15	11	6	7	21	1	P	32
201	11	9	9	11	0	1	P	0
202	9	6	2	5	0	1	P	0
203	6	5	3	2	0	1	P	0
204	9	10	9	5	0	1	P	0
205	11	7	6	4	12	2	P	0
206	12	3	1	1	16	1	P	31
207	10	12	9	8	31	1	P	33
208	7	7	7	5	22	1	P	33
209	9	10	2	1	23	1	P	31
210	9	8	10	5	26	1	P	32
211	18	8	3	1	20	1	P	34
20	7	4	9	14	22	2	P	30
1	12	16	1	0	4	1	H	0
2	11	11	4	2	14	1	H	10
4	11	9	0	0	20	1	H	28
6	18	16	17	16	10	2	H	5
7	13	15	4	9	5	2	H	8
8	10	17	15	15	13	2	H	12
9	17	20	5	5	0	2	H	•
11	8	7	0	1	11	1	H	15
12	15	7	1	1	18	2	H	24
13	4	6	1	0	21	2	H	28
14	11	12	9	9	22	1	H	27
15	12	16	0	0	7	1	H	0
16	14	8	2	2	19	2	H	24
17	11	9	1	0	19	2	H	20
18	14	8	12	9	15	2	H	13
1800	18	16	6	6	15	1	H	19
319	11	11	4	2	14	1	H	10
320	11	9	0	0	20	1	H	20
321	4	6	1	0	21	2	H	28
322	11	12	9	9	22	1	H	27

Figure #6. Profile of subject performance.

1 = Word-production test (i.e. word list)
2 = Sentence-production test (i.e. sentence list)
3 = Word-perception test (i.e. word list)
4 = Sentence-perception test (i.e. sentence list)

Accuracy Rates

	Polish Subjects	Hungarian Subjects
% Correct - production	71.9	65.7
% Correct - perception	84.9	84
% Simple transfer error - production	17.3	24.9
% Simple transfer error - perception	7.6	8.3
Correct + Simple Transfer (prod)	89.2	90.6
Correct + Simple Transfer (perc)	92.5	92.3

Figure #7. Accuracy rates of subjects.

their production tasks would be 100% accurate (though, I now wish I had gathered that comparison data). And while they did very well on the perception tasks (thus validating the test instrument, they did not perform perfectly. Figure #8 shows how well they performed. In this chart a score of every word correct would receive a score of 10 (1 = correct; 2 = incorrect).

Test 1	(mark the syllable): n=6	Mean=11	Std. Dev.=1.673
Test 2	(mark the letters): n=9	Mean=10.833	Std. Dev.=.983
Test 3	(circle the syllable): n=7	Mean=10	Std. Dev.=0
Test 4	(circle the letters): n=7	Mean=10.5	Std. Dev.=.837

Figure #8. Performance of native speakers on pilot tests.

These results show that native speakers are able to do this type of task.

Let us now look at the native speakers' performance on these tests and on the tests given to check the reliability of the instrument. Thirty-six native speakers, in all, were given 10 words to mark the stress on. Thus, there are 360 items in the database. Of these 360 words, 36 were marked incorrectly: 10%. This is very similar to the non-native speakers (remember though, that the tasks were not the same).

Hungarian and Polish Age Effects

I did not find any significant age effects in these subjects and the number of errors they made as Figure #9 shows:

	W-prod	S-prod	W-perc	S-perc	Age	Years
Age	-.048	-.123	-.126	-.174	1	
Years	-.031	.052	-.056	-.084	-.181	1

Figure #9. Age effects

For the Polish subjects the mean Age was 34.9, and the mean Years of Study was 1.6. For the Hungarian subjects the mean Age was

32.3, and the mean Years of Study was 1.1. The difference in Years Of Study between the two groups was not significant (p=.4386).

Of course, not all subjects performed in the same fashion. There was a considerable amount of individual variation, as Figure #10 shows.

Subject #	#Errors			
	W-prod	S-prod	W-perc	S-perc
100	20	14	11	13
10	3	0	0	0
103	11	13	0	0
203	6	5	3	2
20	7	4	9	14

Figure #10. Individual variation in performance.

Notice that subject 100 had high numbers of both production and perception errors; subject 10 had very few of both production and perception errors; subject 103 had a high number of production errors and no perception errors; subject 203 had a few of both production and perception errors; subject 20 had a few production errors and more perception errors. The only pattern that didn't emerge was a subject who had very few production errors but lots of perception errors. This supports Menn and Matthei's (1990) two-lexicon model in which input and output forms have independent representations.

Subject 10 is a Polish speaker age 34 who has been studying English for five years; the subject began studying English at age 29- well past the age usually posited as the close of some kind of sensitive period. And yet, this subject performs extraordinarily well on these tasks.

Why would stress be different?

One question that still has to answered is, of course, why stress would be different than other elements of phonology. These are, for the most part, subjects who haven't had a lot of exposure to English, and they're doing very well. If their abilities have fallen off because of their age, then they haven't fallen very far. Why would this be so?
I'd like to propose a few factors that I think might be relevant here.

(1) stress implementation is not muscular in the same way that segmental phonology is. In this respect, it is more like syntax in that it is governed by abstract principles. The learners in this study were not violating any of the proposed universals of metrical systems (Halle & Vergnaud 1987).

(2) infants have been shown to have the ability to distinguish which syllable is being stressed (Spring & Dale (1977)). This seemingly innate ability may be quite robust further evidence that the critical window has not come down.

(3) stress is a categorial not a continuous phenomenon. A syllable is judged as being stressed or unstressed even though pitch, loudness, and duration are continuous. Perhaps the non-native speakers retain accents in their phonetic implementation of stress in terms of things like pitch, loudness, and duration (I have not checked this instrumentally) but they *were* able to alter their L1 settings enough to be perceived as changing categories (from unstressed to stressed). Native speakers would have judged someone as native-like if they had altered the continuous phenomena enough to cross a category boundary. This would also account for the fact that acquiring new vowel distinctions, for example tense/lax, is generally problematic for L2 learners. They cannot alter the continuous phenomena (formant structure?) enough to cross a category boundary. The model in Figure #11 may illustrate what I'm thinking:

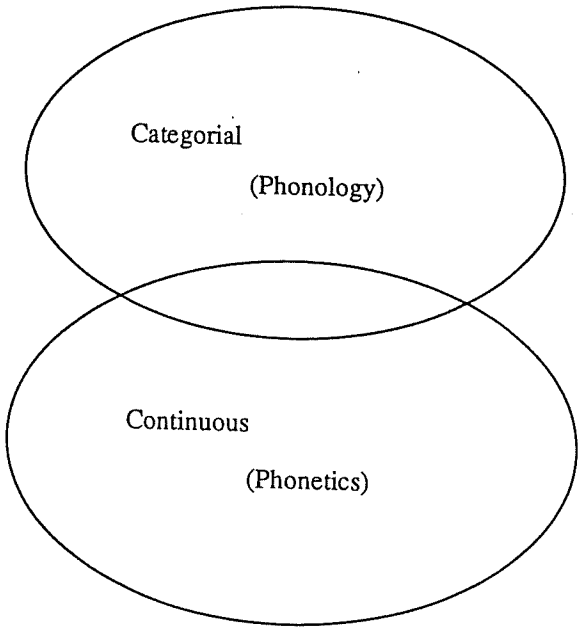


Figure #11. Categorical versus continuous phenomena.

Perhaps our ability to deal with categorial phenomena (e.g. consonants, tone, and stress) in second language learning is better than our ability to deal with continuous phenomena (e.g. vowels). Maybe age-dependent effects are more pronounced for phonetics than phonology. The overlap area on the diagram would seem to deal with things like voice onset time which can be altered but perhaps not in such a way to match the L2 category structure.

Conclusions

I believe that the project has been fruitful in that the empirical results seem to be largely what the theoretical model would predict. Many studies have taken the general approach of saying something like. "We've come up with a certain structural model of a particular linguistic phenomenon. On the basis of this, we'd predict that people learning this system would behave in a certain way. For example, if Italian speakers can move WH words in ways that English speakers can't, we might expect them to produce English sentences which show their L1 patterns". Largely as a result of working with problematic parameters, the empirical studies have not provided clear-cut support for the structural descriptions. In contrast, the empirical study discussed here provides support for the principles and parameters model of description.

It also shows that age-dependent effects in SLA may not extend to suprasegmental phonology in the same way they do to segmental phonology.

Jonathan Leather and Allan James (eds)

new sounds 92

Proceedings of the 1992 Amsterdam Symposium on the Acquisition of Second-language Speech (University of Amsterdam, 13 - 16 April 1992)

University of Amsterdam, 1992