DOCUMENT RESUME

ED 135 214 FL 008 358

AUTHOR Moisio, Risto; Valento, Eero
TITLE Testing Finnish Schoolchildren's Learning of English Consonants. Jyvaskyla Contrastive Studies, No. 3.
INSTITUTION Jyvaskyla Univ. (Finland).
PUB DATE 76
NOTE 109p.

EDRS PRICE MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS *Consonants; *Contrastive Linguistics; Distinctive Features; *English (Second Language); Finnish; *Interference (Language Learning); Language Instruction; Language Proficiency; Language Tests; *Learning Difficulties; Listening Tests; Phonemes; Phonetics; Pronunciation; Pronunciation Instruction; Secondary Education; *Second Language Learning; Testing
IDENTIFIERS Finland

ABSTRACT

The present study deals with two language skills, listening and speaking, and it is limited to single consonant phonemes only. The purpose is to try to find answers to the following problems: (1) Which Finnish consonants are given as substitutes for English consonants by Finnish pupils who have no previous (or practically no previous) knowledge of English? (2) Which English consonants are difficult for Finnish-speaking pupils to learn? (3) Are the areas of difficulty predictable on the basis of a contrastive analysis? (4) Is there a change in the amount and type of learning problems between second formers and fifth formers in secondary school? (5) Can success in the production test be predicted from the listening test results? (6) Are certain background variables related to pupils' ability to discriminate, identify and produce English consonants? The Finnish and English consonant systems are compared on the basis of physical, relational and distributional differences. Substitution, discrimination, sound analogy, written analogy, and production tests were constructed and administered to secondary school students; the total number of subjects involved varied from 48 for the production test to 329 for the discrimination test. The results are given by research question, and sample tests are included in the appendices. (CFM)

by
Risto

Reports from the Department of English
University of Jyväskylä

Jyväskylä Contrastive Studies
edited by
Kari Sajavaara and Jaakko Lehtonen

Reports from
the Department of English
University of Jyväskylä

No 3

Jyväskylä Contrastive Studies, 3

TESTING FINNISH SCHOOLCHILDREN'S LEARNING OF ENGLISH CONSONANTS

by
Risto Moisio and Eero Valento

Jyväskylä 1976

ISBN 951-677-663-9

Jyväskylän yliopiston monistuskeskus
Jyväskylä 1976

PREFACE

Nowadays pronunciation is gaining in importance as a special field of language teaching and learning. According to present practice, the teaching of pronunciation is based on mimicry after a model. To be able to mimic the student has to be taught to listen to particular features in the foreign language and to distinguish between foreign language sounds and the corresponding native language sounds. The criterion for correct pronunciation in secondary school is that the phonemes of the target language are kept distinct in such a way that a native speaker of that language is able to understand (Nykykielet 1971: 11, 29).

When learning a foreign language, we tend to transfer our native language habits into the target language (Lado 1957: 11). Therefore it seems feasible and logical to make a comparison between the native and foreign language systems. This is what contrastive analysis tries to do. Lado (1957: 12) holds that by comparing the two sound systems in contact it is possible to show where learning problems are likely to occur. However, he does not base his theory on any empirical data, which Brière (1966: 768, 769), for instance, considers necessary. This fact has given an impulse to the present study. Another factor justifying this study is that from the Finnish point of view English consonants have not been studied as thoroughly as English vowels.

The present work is based on the authors' master's thesis "On learning English consonants: an empirical study of learning problems met by Finnish-speaking pupils", which was prepared for a degree in English philology (under the supervision of Professor Esko Pennanen) and in Education at the University of Tampere. For practical reasons previous studies on background factors, the construction, administration, analysis and revision of the pretest versions of the tests and questionnaires, the final questionnaires and the data obtained from them are here only superficially touched upon. Those interested in them will find detailed information on them in the thesis.

We take this opportunity to present our sincere thanks to Juhani Miettinen, Lauri Myllykorpi, Kustaa Roine, Simo Tapio and Lauri Viljanmaa, headmasters of Tampereen yhteislyseo, Sammon yhteislyseo, Tampereen normaalilyseo, Harjun yhteiskoulu and Pirkanmaan yhteiskoulu, respectively,

for giving us their permission to administer the tests and questionnaires and thus making this study possible. We also express our gratitude to the teachers and pupils of these schools for their co-operation, especially to Marja Harkko, Rauni Kekoni, Toini Kuortti, Eila Rahkiola and Eila Rantanen for putting their lessons at our disposal and for transcribing their pupils' productions.

We are also greatly indebted to Professor Esko Pennanen and other members of the staff of the Department of English, University of Tampere, for critical comments and for assistance in the form of tapes, as well as to Timo Leino and Juhani Ikala from the Speech Department for arranging the recording of the tests.

Our heartfelt thanks are also due to Professor Kalevi Wiik of the University of Turku and Professor Jaakko Lehtonen of the University of Jyväskylä. Professor Wiik was kind enough to put his manuscript "Finnish and English Consonants" at our disposal and to offer critical and encouraging comments on the typescript of this study. Professor Lehtonen undertook the considerable task of reading the whole work in manuscript and gave freely of his time to advise us in the preparation of this paper.

Finally we owe a great debt of gratitude to James Crichton and Roy Parker, lecturers at the University of Tampere. Mr Crichton kindly offered to transcribe the productions of both the pretest and the final test subjects and made pertinent comments on the tests. Mr Parker did us an invaluable service in reading the pretest as well as the final test versions on tape and in transcribing the productions of the final test subjects. He also made valuable critical comments on the tests and the manuscript of our work and advised us in matters of English.

Tampere R.M. & E.V.

May, 1976

TABLE OF CONTENTS

INTRODUCTION

THE AIMS OF THIS STUDY

The purpose of this study is to try to find answers to the following problems: (1) Which Finnish consonants are given as substitutes for English consonants by Finnish pupils who have no previous (or practically no previous) knowledge of English? (2) Which English consonants are difficult for Finnish-speaking pupils to learn? (3) Are the areas of difficulty predictable on the basis of a contrastive analysis? (4) Is there a change in the amount and type of learning problems between second formers and fifth formers in secondary school? (5) Can success in the production test be predicted from the listening test results? (6) Are certain background variables related to pupils' ability to discriminate, identify and produce English consonants?

PREVIOUS STUDIES AND THEIR IMPLICATIONS FOR THIS STUDY

LINGUISTIC STUDIES. – As far as we know, empirical studies directly relevant to our study are not many. In Finland only those of Wiik (1965a, 1965b, 1966) and Hirvonen (1971) have dealt with problems similar to ours. Wiik used the substitution technique to find out which Finnish vowels/consonants the subjects tended to substitute for the English vowels/consonants they heard (Wiik 1965a; 37; Wiik 1966: 9; Wiik 1965b). By means of the substitution technique it is possible to pinpoint those English sounds that are confused by native speakers of Finnish with the similar Finnish sounds, i.e. where initial learning problems are likely to occur. This approach, however, does not reveal which English sounds Finns confuse with each other. These areas of difficulty can be explored by using other techniques, for instance the minimal pair technique. Tommola (1975) has explored the relationship between the discrimination and production of English sounds by Finnish secondary school pupils (third, fifth and seventh formers) and first-year university students of English. His discrimi-

nation test comprised both minimal pair contrasts and contrasts between two phonetic features, one typically English, the other a typical Finnish substitution feature. The production test was a repetition test in which the testees imitated unconnected sentences. Production performances were scored dichotomously: acceptable sound containing the idiomatic feature(s) was given 1 point, unacceptable sound with a substitution feature was marked 0. Tommola found that (1) on the whole it was statistically significantly more difficult to produce than to discriminate English sounds and (2) the correlations between receptive and productive skills were relatively low (all below .50). Thus, he considers discrimination and production separate and independent skills to such an extent that indirect measurement of production by means of discrimination does not seem feasible (see Tommola 1975: 14-15, 22, 25-26, 27). The minimal pair technique alone was used by Hirvonen in his sound discrimination test. In his pilot version of the test he found that only the items containing such sound contrasts that, according to the principles of contrastive analysis, are difficult for Finnish learners functioned well.

Hirvonen assumed that this gave support to contrastive analyses (Hirvonen 1971: 20). Tests based on minimal pairs presuppose ability to *discriminate* between two (or more) sounds, but they do not necessarily (possibly not at all) presuppose ability to *identify* and *categorize* the sounds in question. According to Lehtonen (1972a: 21), a normal hearer can, especially after some period of training, discriminate between different sounds much in the same way as he can make a distinction between various shades of colour.

But in language learning this is not enough. The learner must also be able to identify the sounds of the target language and to realize that certain new "differences in shade" can completely change the meaning of an utterance (e.g. a Finn may say either [*bi:sami*] or [*pi:sami*] and mean "mu rat" all the same, whereas in English the utterances "I got a *bike* yester day" and "I got a *pike* yesterday" do not mean the same thing). Thus, soun discrimination tests need not necessarily indicate that the testees maste the phonemic relationships between the sounds of the target language. The fore we used a combination of substitution and minimal pair techniques an in addition to them, devised altogether new types of tests which, we believe, also require identification and categorization of the sounds of th target language on the part of the learner.

There are also some empirical foreign studies based on contrastive principles, but their results must be interpreted with caution, as it is not feasible to assume that the sounds of one language would be universally difficult, whereas the sounds of another language would be universally easy to learn. For example, Finns and Swedes learning English may find quite different sounds difficult to learn (cf. Lehtonen 1972a: 25, 26). However, regardless of the languages being investigated, one interesting feature was common to all these studies: predictions based on "theoretical constructs of 'systems of distinctive versus redundant features', 'phonemic class memberships', and 'distribution of the phoneme classes' " between the native language and the target language were often considered inadequate and vague (see Brière 1966: 769; Nemser 1971: 95). Nemser even found that "different phonemic theories yield different predictions... and no theory... predicts or accounts for interference patterns as complex as those resulting from the contact of the Hungarian and English phoneme systems ..." "Only the Jakobsonian-based formulations yielded explicit predictions and they were generally erroneous" (Nemser 1971: 95; see also Hirvonen's opposite view, p. 2 above). Thus the faith put by many linguists on the predictive power of theoretical (not empirical) contrastive analyses is somewhat shaken by Brière's and Nemser's test results. Accordingly, it is dubious whether hierarchies of learning difficulty, arrived at in studies of given languages based on theoretical contrastive analyses are applicable to the present study. Still, it is interesting to make a brief review of the conclusions drawn by various linguists. According to Bloomfield (1935: 77, 79) and Trubetzkoy (1969: 51, 52), for instance, the speakers of a language learn to attend only to those features which are distinctive and to ignore those which are redundant. Lado and Fries also hold the same view. The implication of this view is that a person learning a foreign language "does not actually hear the foreign language sound units – phonemes. He hears his own" (Lado 1957: 11; see also Lehtonen 1972a: 27). Following the same line of thought many linguists assume that the higher the degree of similarity between the native and the target language phonological categories, the easier it is for the speaker to learn the target language phonological categories. For instance, if the sounds of a foreign language are physically similar to those of the native language, and also structure similarly to those of the native language, and are similarly distributed, they are believed to

be learnt by simple positive transfer without difficulty, while those sounds of a foreign language that are non existent or structure differently or are differently distributed in the native language, are learned more slowly (Lado 1957: 12). Weinreich, through his contrastive analysis of Romansch and Schwyzertütsch, came to the conclusion that "the greater the difference between the systems, the greater are the learning problems and the potential areas of interference" (Brière 1966: 768-769). These statements seem quite logical. But it is difficult to decide what kind of differences, physical, relational, or distributional or combinations thereof, will cause the greatest learning difficulties. Wolff (as quoted by Brière 1966: 768) feels that "it is easier for everyone to learn a completely new phoneme which does not exist in his native language ... than it is to learn a partially similar class in the target language that will involve negative transfer caused by the N system ..." In our study we shall make an attempt to answer this question of difficulty, not on the basis of ready-made hierarchies of difficulty or any single phonemic theory. The method we chose is to classify contrastive consonant pairs roughly into three main categories: (1) identical consonants occur in Finnish and English, e.g. /h/ and /m/ as in the pair *house - mouse*, (2) one of the two does not exist in Finnish, e.g. /v/ and /ð/ as in *van - than*, (3) both consonants occur only in English, e.g. /θ/ and /ð/ as in *loath - loathe*. Attention will also be paid to their distribution. Moulton (1962: 26) classifies teaching problems as phonemic, phonetic, allophonic and distributional. We shall take phonetic and allophonic differences into account only if they cause a phonemic error, e.g. if an initial /t/ uttered by a Finn is heard as /d/ by a native speaker of English, which may be due to the fact that Finns tend to produce their plosives without aspiration.

OTHER RELATED STUDIES. – So far we have dealt with linguistic factors, mainly phonological interference between the native language and the target language, that may affect the rate of learning English consonants. However, one might well assume that there are also factors other than linguistic ones that are related to pupils' ability to discriminate, identify and produce English consonants. The process of learning the consonants of a foreign language is such a specific problem that there are

virtually no studies on the subject. Only a few studies can be referred to.

In Takala's study the correlation between linguistic ability and the recognition of sounds was .30 (Takala 1968: 16). This correlation is fairly low. Linguistic ability explains only 9% of the variation in the ability [illegible].

[illegible] study concerning the structure of Engl[illegible] [illegible]pper secondary school Leino (1970: 7-8. [illegible]) [illegible] relevant to our study. It was the fa[illegible] pronunciation and comprehension of speech. It was made up of the following components (with respective correlations with the factor): (1) recognition of sounds (.66), measured by means of a minimal pair discrimination test, (2) production of sounds (.59), measured by a paper-and-pencil test, (3) production of stress (.55) and (4) listening comprehension (.50). In Leino's factor analysis the factors of general linguistic ability and of pronunciation and comprehension of speech emerged for both boys and girls.

Jorma Lehtovaara (1974) has studied the coherence of pronunciation as a skill when it is understood to contain both receptive and productive skills. His subjects were third formers at elementary schools in Tampere. By means of factor analysis he came to the conclusion that pronunciation consists of three factors: (1) mastery of sounds, (2) mastery of intonation and (3) fluency of speech. We are here interested in the components of the factor of mastery of sounds. They are (with respective correlations with the factor):

-free production of sounds through picture stimuli (.71)
-imitation of consonants (.71)
-discrimination of sounds through triplets based on minimal contrasts (.64)
-fluency of free production (.60)
-imitation of vowels (.58)
-sounds produced through reading aloud (.39)

It is to be noted that this factor comprised both receptive and productive skills. Lehtovaara points out, however, that discrimination of sounds was measured only by means of one test and thus it was not actually possible for a separate factor of sound discrimination to emerge in his study (see Jorma Lehtovaara 1974: 1, 34, 51-83, 96-97, 99-101).

Maija Lehtovaara (1974) has studied the relationships of certain pupil variables to the mastery of English sounds, which was measured by the following tests: free production of sounds through picture stimuli, imitation of consonants, imitation of vowels, and sound discrimination through triplets based on minimal contrasts. She found the following relationships with the mastery of English sounds:

-verbal ability (.65), measured by a conjoined variable of vocabulary, synonym and first letter test scores + the average of theoretical school subjects,
-the pupils' ability to concentrate (.53),
-the pleasantness of English (.37),
-social class (.27)

(see Maija Lehtovaara 1974: 34-36, 38,46-47, 50, 51, 53).

These studies are practically the only ones that deal with our specific topic to a noteworthy degree, whereas studies, both Finnish and foreign, concerning general school achievement and foreign language achievement are to be found in abundance. In these studies several factors have been found to correlate with success in foreign languages and with school achievement in general. The most important ones seem to be

(1) intelligence, especially verbal intelligence (see e.g. Ritvanen 1971, Leino 1972, and Konttinen 1970; Konttinen states (p. 1) the interesting fact that in Finnish studies the correlations of intelligence and of verbal ability in one's native language with foreign language achievement have been lower than in foreign studies. Leino (1972: 11) offers an interesting and plausible explanation: Finnish is not related to the foreign languages taught in our schools, whereas a majority of foreign studies deal with languages that are related),
(2) social and home background (see e.g. Jurama 1971, Ritvanen 1971, and Hämäläinen and Takala 1970),
(3) personality (see e.g. Leino 1972 and Ritvanen 1971),
(4) attitudes (see e.g. Smith 1971, Spolsky 1969, Leino 1972, Sysiharju 1970 and Heinonen 1968),
(5) motivation and goals (see e.g. Ritvanen 1971 and Jurama 1966),
(6) sex (see e.g. Jurama 1966, Heinonen 1964 and Takala 1968).

In addition to the variables referred to above, the following factors may also be related to the testees' ability to discriminate, identify and

produce English consonants: previous or concurrent experience with English outside school (e.g. private lessons, listening to English/American music, watching English/American TV-programmes), possible defects in hearing and speaking, the time spent on overt teaching of pronunciation at school, and the use of AV-aids.

Unfortunately we can here refer to only one previous study. Even that showed negative results. Brière (1967: 165, 168) found, when testing the perception and production of American English phonemes /d/ and /ð/ by Spanish-speaking pupils, that there was no significant correlation between the subjects' performances and the amount of time the subjects had been in the watched TV, listened to rock-and-roll records etc.

[illegible] variables will be taken into account in this study as shown in Diagram 1.

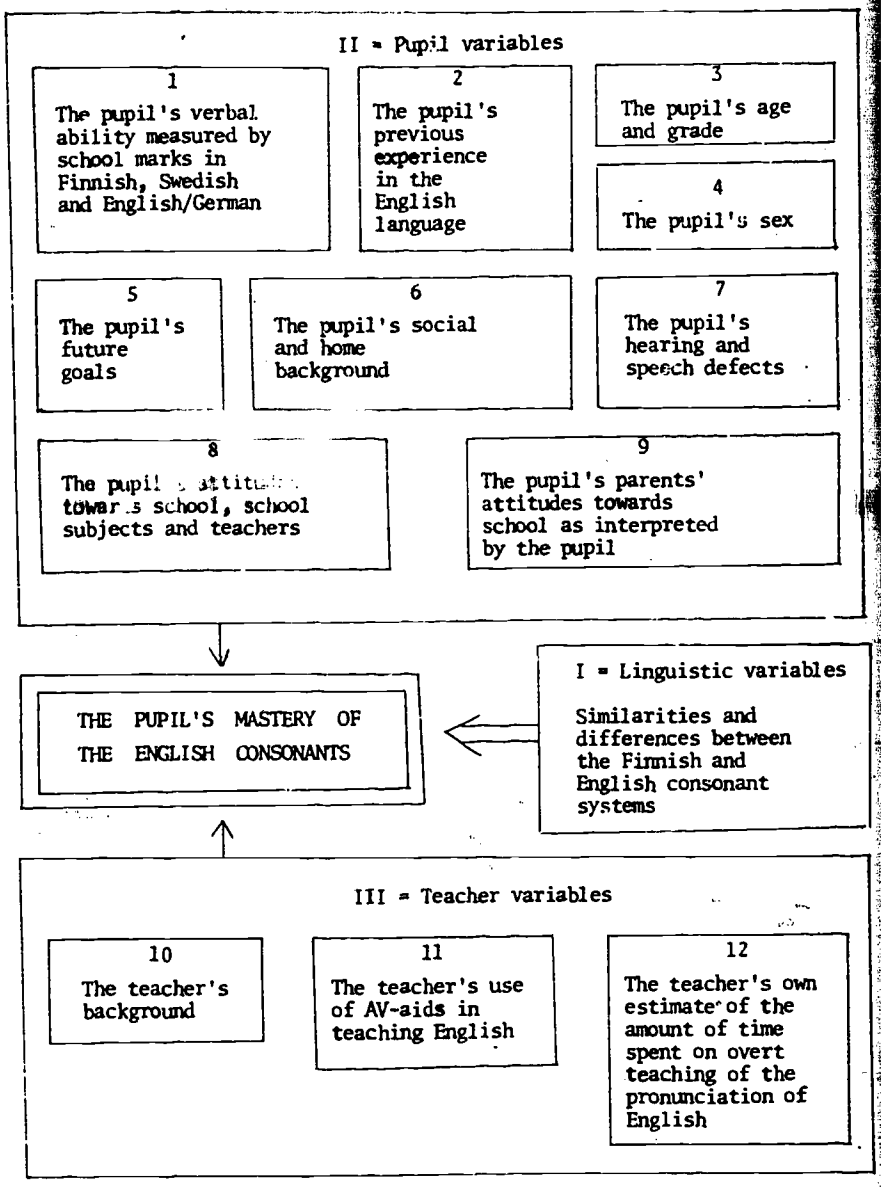


Diagram 1. The variable groups used in this study.

THE SCOPE AND THEORETICAL FRAMEWORK OF THIS STUDY

LANGUAGE SKILLS

The division of language skills into listening, speaking, reading and writing is generally accepted among present-day linguists (see e.g. Harris 1969: 100). Of these the present study deals with listening and speaking and it is limited to single consonant phonemes only. Consonant clusters or any higher-level phonological structures were not the objects of the present treatise.

THE CONCEPT OF THE CONSONANT

Some linguists, e.g. Branford (1967: 34, 35), Gimson (1962: 144, 153) and Gaeng (1971: 31), define consonants both phonetically and functionally, whereas some others, e.g. Jones (1967: 25) and Moulton (1962: 6), define consonants in purely phonetic terms. Wiik (1973: 69) prefers to define only vowels and say that all other sounds are consonants. In this study we do not define a consonant in any way; we merely state which concrete sounds we regard as consonants in Finnish and English. They are [p t k d h v s j r l m n ŋ b g f w z ʃ θ ð ʒ tʃ dʒ]. The first 13 of them are phonemes in both languages. They are here regarded as more or less physically equivalent; so also are [b g f w z ʃ], which are phonemes in English but allophones in Finnish. [θ ð ʒ tʃ dʒ] have no equivalents in Finnish.

THE CONCEPT OF THE PHONEME

A phoneme can be defined in several ways; no definition can however be considered complete and final (Branford 1967: 73). In principle there are two approaches to the definition of the phoneme: (1) according to the conventional approach, a phoneme is regarded as the smallest contrastive linguistic unit which can bring about a change of meaning (see e.g. Gimson 1962: 44), (2) the second approach regards a phoneme as a bundle of distinctive features without any reference to meaning (Harms 1968: 2; Jakobson 1962: 497, 498; Chomsky 1957: 94-100). The conventional approach is adopted here.

DISTINCTIVE FEATURES

As mentioned above (p. 3), different phonemic theories and distinctive feature categories, as such, were not sensitive enough to account for the complexity of learning problems resulting from the contact of the phonemic systems of given languages. This may be due to the fact that linguists strive after simplicity, economy and universality in constructing their theories. They do this by minimizing the number of distinctive features and for this purpose they are eager to regard as redundant all such features which do not serve a clearly distinctive function in the phonemic system. For example, aspiration is considered a redundant or non-functional feature in English, because the voiceless plosives [p t k] are classified as phonemes /p/, /t/ and /k/ whether they are aspirated (e.g. [t'] in *ten*) or not (e.g. [t] in *stick*). Further, the same opposition (e.g. voiced or voiceless) may in one case be distinctive (e.g. in the opposition /b/-/p/) and redundant in another (e.g. nasals, as nasality in English presupposes the occurrence of voice, or in the context of /s-/ the distinction between voiced and voiceless consonant is non-functional). Lyons (1971: 122-123) considers this an advantage because it enables linguists to state the restrictions upon the distribution of particular classes of phonemes more systematically and more economically. This may seem very promising, but what is a redundant feature to native speakers of English, for example, may not be redundant in communication with a learner of English. For example, Finns tend to pronounce the plosives [p t k] with little effort and without aspiration. If a Finn pronounces an initial [t], for instance, without due aspiration and force, there is a danger that a native speaker of English hears it as /d/. We can give here an autenthic example: In a Finnish TV-programme Danny Kaye asked the Finnish girl singers Tiina and Jaana to tell him their names. When he heard them he became very excited because he interpreted Tiina as /di:na/, which happened to be his daughter's name. This misinterpretation shows clearly that the concept of redundancy in phonological structure has to be kept apart from the actual cues of identification.

In our example, for instance, it is difficult to judge whether it was the lack of aspiration or the lack of force or both together which caused the misinterpretation. In fact, very little is known about what physical properties of speech really are redundant and insignificant regarding perception (Lehtonen 1972a: 35). It is possible, as Lehtonen puts it, that all the phonetic features that in the language of the hearer regularly belong

to the acoustic pattern of certain sounds are important to the hearer, no matter whether they are phonemically distinctive or automatically belong to a given distinction (Lehtonen 1972a: 35).

In this study we shall define distinctive features in articulatory terms. As it is not conclusively shown which features are distinctive and which are redundant in English, we prefer to take into account all the physical, articulatory differences between English consonant phonemes and regard them as separate distinctions.

PRINCIPLES OF CONTRASTIVE ANALYS

Because our study is concerned with single consonant phonemes of Finnish and English, i.e. with single segmental phonemes only, we shall restrict our discussion of the principles of contrastive analysis to those concerning the comparison of phonemic systems. According to Lado (1957: 13), the comparison of each phoneme "should include at least three checks: (1) Does the native language have a phonetically similar phoneme? (2) Are the variants of the phonemes similar in both languages? (3) Are the phonemes and the variants similarly distributed?" Thus the sound systems of the languages in question are juxtaposed to reveal similarities and differences. We do not pay so much attention to the similarities as to the differences in the phonemic systems for the following reasons: No serious learning problems should be involved if (1) a target language phoneme is in every respect fully identical to one in the native language (e.g. /n/ in Finnish and English), because it is obviously learnt by simple positive transfer, (2) a native language phoneme is sufficiently similar to that of the target language to be identified by a speaker of the target language as the phoneme intended, e.g. the Finnish phonemes /r/ and /l/ may sound un-English, but they are still recognizable as phonemes corresponding to English /r/ and /l/. Therefore from the point of view of communication it is not absolutely necessary for a Finn to learn phonetically correct variants of the English phonemes /r/ and /l/.

Here the main attention is focused on differences between the native language and the target language phonemic systems, because it is the differences that are more likely to cause learning problems. This is due to the fact that a learner of a foreign language cannot use native language phonemes as acceptable substitutes for phonemes in the target language.

For example, Finnish [s] cannot be used for English [ʃ], although the distinction [s]-[ʃ] need not be made in Finnish: you can equally well say [saka:li] or [ʃaka:li] and still be understood correctly. In English, [s] and [ʃ] belong to separate phonemes. In general, the differences between phonemic systems imply that entirely new phonemes or new uses of familiar sounds must be learnt. Differences between sound systems have been classified by Wiik (1965a: 15-16) into four major types:

(1) physical differences. A physical sound (or a group of sounds) occurs in one language but not in the other, e.g. the fricative [ʒ] occurs in English but not in Finnish.

(2) relational differences. Two physically similar sounds exist in both the native and the target language, but the sounds are grouped differently into phonemes, e.g. [w] is an allophone of /v/ in Finnish, but in English it is an independent phoneme /w/, which must be kept apart from /v/ as in the minimal pair *vent - went*.

(3) distributional differences. Similar sounds occur in both languages, but in different environments, e.g. /ŋ/ in Finnish occurs only word-medially and before /k/ or as a double consonant, e.g. *lanka* [laŋka] 'thread' - *langan* [laŋŋan] 'gen. of *lanka*', whereas in English /ŋ/ also occurs word-finally, e.g. *singer - wing* [siŋə-wiŋ]. In neither language does /ŋ/ occur word-initially.

(4) segmental differences Phonetically similar stretches of speech occur in both languages, but the stretches are differently divided into phonemic segments, e.g. Germans tend to treat the Englis affricate /tʃ/ as a sequence of /t/ + /ʃ/, whereas native speakers of English are apt to consider it a single phoneme.

Further, Wiik (1965a: 16-30) divides these major types into subclasses mainly by using free variation and complementary distribution as his criteria.

In this study physical, relational and distributional differences will be dealt with as follows.

(1) physical differences: Because the Finnish and English consonant systems are compared from the viewpoint of a Finnish learner only, there are only one-way physical differences to be dealt with, name ly those English consonants that do not occur in Finnish. Physical differences are assumed to cause both hearing and pronunciation difficulties.

(2) relational differences: They are assumed to cause maximum difficulty (both in hearing and pronunciation) in foreign language learning, because the allophones of the native language may be different phonemes in the target language. Psychologically it is, perhaps, more difficult to modify one's old habits, i.e. to use familiar sounds in a new way than to learn something entirely new, i.e. completely new sounds (cf. Wiik 1965a: 21 and Lado 1957: 14-15). For the learner it does not matter whether the allophones in the native language are in free variation (e.g. [s] and [ʃ] in Finnish) or in complementary distribution (e.g. [v] and [w] in Finnish). We shall illustrate this point. Finns may identify [ʃ] in English as [s], as they are not accustomed to keep them apart in their speech. Moreover, they can reproduce [ʃ] as [s]. In both cases they make a phonemic mistake. The same applies to [v] and [w], which in English belong to the phonemes /v/ and /w/. Thus relational differences must be taken into account in this study as well.

(3) distributional differences: To learn to use familiar sounds in unfamiliar environments may also prove difficult, especially where the distribution of a native language phoneme is more restricted than the distribution of the corresponding target language phoneme. For instance, /d/ occurs in Finnish only word-medially as in *madot* 'worms' v. *matot* 'carpets' (word-initially and word-finally only in loan words as *Daavid* 'David', *deodorantti* 'deodorant', *dieetti* 'diet'), whereas English /d/ occurs in all these positions (e.g. *day*, *ready* and *head*). We describe the distributions of consonant phonemes (possibly also allophones) in relation to words, not in relation to other phonemes or allophones for the simple reason that consonant clusters are not dealt with in this study. It is also extremely difficult to define a smaller phonological unit e.g. a syllable in English. Our test items are separate words. Thus we feel it is appropriate to use directly comparable units (i.e. words) as points of reference in describing distribution.

METHOD OF THIS STUDY

CONTRASTIVE ANALYSIS OF FINNISH AND ENGLISH CONSONANT SYSTEMS

THE FINNISH CONSONANT SYSTEM. – The number of Standard Finnish consonant phonemes varies from 13 to 18. This variation is due to a divergence of opinion whether [b g f ʃ ʔ] should be accepted as phonemes in Finnish or not. The generally accepted consonant phonemes are: / p t d k v s h j l r m n ŋ /. This view is also adopted here. [b g f ʃ] may occur in the speech of Finns with a knowledge of foreign languages, but in the first place < b g f sh š > are only letters which occur in loanwords in Finnish orthography. Our interpretation finds support in that such distinctions as /p/-/b/ or /s/-/ʃ/, for example, are not systematically maintained by native speakers of Finnish. Very often one hears people ask if a name is written with a "hard" or a "soft" < p >. Nowadays it is possible to write < s > instead of the old-fashioned < sh > or < š >, which was recommended earlier. The same non-functional status of [g] and [f] is reflected in everyday communication. These sounds are not kept apart from /k/ and /v/, because it is not necessary. The Finnish consonant system does not utilize a distinction between voiced and voiceless consonants. The opposition /t/ versus /d/ is the only exception. In our opinion the glottal plosive does not constitute a phoneme in Finnish, because it does not occur in any isolated word as do the other 13 consonant phonemes. Karlsson (1969: 357) also excludes [ʔ] from the consonant phoneme inventory of Finnish (see also Wiik 1965b).

THE ENGLISH CONSONANT SYSTEM. – The number of English consonant phonemes varies from a minimum of 22 to about 30 depending on the variety of English in question and on whether the affricates are treated as single phonemes or as phoneme sequences. The basic 22 consonant phonemes of Englis are: /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /θ/, /ð/, /h/, /s/, /z/, /ʃ/, /ʒ/, /l/, /r/, /m/, /n/, /ŋ/, /w/ and /j/. The status of the affricates is very problematic. The majority of linguists regard only [tʃ] and [dʒ] as

true affricates, while for instance Jones (1967: 163-167) discerns four more affricates: [ts], [dz], [tr] and [dr]. However, he excludes [ts] and [dz] from his chart of English consonants, because they occur only in loan words, e.g. *tsetse* and *Dzungaria*. On the other hand, he includes /tr/ and /dr/ in his consonant inventory, because they occur in native English words, e.g. *tree* and *dry*. Gimson (1962: 144) again places [tr] and [dr] in brackets in his consonant chart and thus does not attribute to them the status of independent phonemes. As there also exist the sequences /t/ + /s/ (e.g. *outside*), /d/ + /z/ (e.g. *heads*), /t/ + /r/ (e.g. *outrage*) and /d/ + /r/ (e.g. *blood-red*) the criterion of morpheme boundary has to be adopted to distinguish between the affricates and the corresponding sequences. In this treatise /ts/, /dz/, /tr/ and /dr/ will be treatedas sequences, because this interpretation results in a more economic phoneme inventory and most linguists tend to regard them as sequences. However, in accordance with most linguists (see Gleason 1969: 316-317) /tʃ/ and /dʒ/ will be taken, at least tentatively, as affricates in this study. The fact that /tʃ/ and /dʒ/ are generally felt to be single units among native speakers of English supports our view. Wiik (1965b) assumes that for Finns the problems of learning affricates are analogous to the problems of learning sequences of two consonants. The voiceless fricative [hw] can be thought of as a phoneme in, for example, the Scottish variety of English, where *witch* and *which* form a minimal pair. In Standard English (RP) it is an allophone of /w/. We regard [hw] as an allophone of /w/, too, because the Southern variety of English (RP) is used as the model for pronunciation in Finnish schools (see for instance POPS 1970: 122). The glottal plosive [ʔ] is not accepted as a phoneme in Standard English, either, and it will be excluded from our consonant inventory. Thus we have arrived at 24 consonant phonemes as the constituents of the Standard English consonant system.

A COMPARISON. – *Physical differences.* – The Finnish consonant system is characterized by a fairly resticted number of consonant phonemes (13), whereas the English system contains a large selection (24) of them. Plosives are frequent in both languages, 4 in Finnish and 6 in English. Thus both Finns and native speakers of English are at first sight accustomed to paying attention to the feature plosive. It would seem that the Finnish plosives / p d k / are fully acceptable as the corresponding English pho-

nemes and /t/ can be used as a substitute for the English /t/ in spite of a slight difference in the place of articulation. Thus learning to hear and produce English / p t d k / should not be too difficult for Finns. There are also two new plosives that must be learnt. They are /b/ and /g/. These may occur as sounds in loan-words in Finnish (e.g. *bussi* 'bus', *laboratorio* 'laboratory', *gallona* 'gallon', *agentti* 'agent', *Haag* 'the Hague'). Therefore one might assume that learning the English plosives is not difficult for a Finn. However, the picture is obscured by the fact that word-initially and at the beginning of a stressed syllable the fortis plosives / p t k / are aspirated in English, whereas in Finnish they are unaspirated. This difference should not cause any hearing problem, because Finns probably identify English / p t k / correctly whether they are aspirated or not. In production there may arise a difficulty, because Finns tend to pronounce their fortis plosives too laxly and without aspiration so that native speakers of English may hear them as / b d g /.

There is a marked difference in the number of spirants in the two languages, 1 in Finnish versus 5 in English. It could thus be assumed that Finns are not used to paying as much attention to the feature spirant as native speakers of English. The only Finnish spirant /h/ is quite acceptable as a substitute for its English counterpart. As to the other spirants, /θ/ and /ð/ are likely to present both hearing and pronunciation problems for Finns, because (1) there are no interdentals in Finnish and (2) they are kept apart only by the distinction fortis/lenis, whereas in Finnish no two consonants are separated by that distinction alone. In Finnish [f] occurs in loan-words and dialects (e.g. *fasaani* 'pheasant', *taifuuni* 'typhoon' and *fiini* 'fine'). Therefore [f] may be familiar to Finns and this may make it easier for Finns to hear and pronounce it than for instance to hear and pronounce /θ/ or /ð/. /v/ is a spirant in English but a semi-vowel in Finnish. However, both these phonemes are labio-dentals. The English /v/ should not cause any identification problems for Finns, as the nearest equivalent to it is the Finnish /v/. On the other hand if Finns use their own /v/ in speaking English native English speakers might identify it either correctly as /v/ or incorrectly as /w/, because the Finnish /v/ has the features belonging to the labiodental /v/ and semi-vowel /w/ in English.

Of semi-vowels, two in both languages, /j/ should not present any learning difficulties, as the Finnish /j/ is identical to the English /j/.

The learning of /w/ can be problematic. [w] is an allophone in Finnish, but as a phoneme it is new for Finns and for that reason alone it may cause problems. Finns may hear [w] as /v/, because the Finnish /v/, in spite of a difference in the place of articulation, shares the feature semi-vowel with the English /w/.

There is also a considerable difference in the number of sibilants between Finnish (one) and English (four). Even the only sibilant /s/ in Finnish is not fully identical to the /s/ of English. There is a difference in the place of articulation. The Finnish /s/ lies between the English /s/ and /ʃ/. Therefore native speakers of English may sometimes identify the Finnish /s/ as the English /s/, sometimes as /ʃ/. Thus Finns should learn to make a clear distinction between the English /s/ and /ʃ/ in their speech. On the other hand, /ʃ/ is, possibly, not a major hearing problem, because the letter < š > occurs in loan-words in Finnish (e.g. *šakki* 'chess') and so it may be familiar to Finns. There is also a danger that Finns confuse /s/ with the other sibilants /z/ and /ʒ/ in English. These are entirely new phonemes for Finns and in addition to that they are phonetically close to each other. Thus they are likely to cause both hearing and pronunciation problems for Finns.

In Finnish there are no affricates. Accordingly the two English affricates [tʃ] and [dʒ] are unfamiliar sounds to Finns. As /tʃ/ and /dʒ/ are separated by a fortis/lenis distinction and are articulated at a place where no Finnish consonants are articulated, it is obviously difficult for Finns to learn to distinguish them from each other and to pronounce them.

There is only one /r/ phoneme in both languages. The Finnish /r/ is a full tremulant. An identical [r] may occur in some varieties of English (e.g. in Scotland) and sometimes in RP, too. But the [r] commonly used in RP is a semi-tremulant or a glide and thus phonetically different from its Finnish counterpart. Yet from the point of view of communication the Finnish /r/ is interchangeable as a phoneme with the corresponding phoneme in English. The Finnish /r/ used as a substitute for the English one may sound un-English, but it does not in any way endanger communication, unless a native speaker of English feels so irritated at hearing it that he does not pay attention to the content of the message spoken. Therefore the English /r/ should not constitute a pronunciation problem for Finns. However, Finns may encounter some difficulty in identifying the English /r/. This is particularly true of a word-initial [r], which in English is often labialized,

i.e. very much like [b] or [w].

As to the laterals, there is one /l/ in both languages. The English /l/ is not likely to cause any serious pronunciation problems, because /l/ is produced in the same manner and place of articulation in both languages. Besides, from the point of view of communication it does not make any difference whether a clear [l] or a dark [l] is used, unless a native speaker of English is irritated by an inconsistent use of these two variants. On the other hand the English dark [l] presents identification problems for Finns due to its [u]-like and [o]-like formant positions as Wiik (1966: 25-26) has pointed out: Finns tend to hear dark [l] as [u ul o ol].

The physical properties of the nasals are exactly the same in the two languages. The three nasals / m n ŋ / are fully interchangeable in Finnish and English. Therfore Finns obviously learn the nasals of English without any difficulty.

Other systematic differences than those between separate consonant phonemes can also be found between the Finnish and English consonant systems. A really startling difference is the fact that there are no two consonants in Finnish that are kept apart from each other by fortis/lenis distinction alone (even in the case of /t/-/d/ a difference in the place of articulation accompanies that of voicing and duration), whereas eight such pairs are found in English: /p/-/b/, /t/-/d/, /k/-/g/, /f/-/v/, /θ/-/ð/, /s/-/z/, /ʃ/-/ʒ/, /tʃ/-/dʒ/. Here the difficulty is perhaps that the Finnish learner of English must learn to utilize a completely new criterion of distinguishing between speech sounds. It may cause both hearing and pronunciation problems. This usage of fortis/lenis opposition to alter the meaning of an utterance is perhaps comparable to the doubling of consonants which is typical of the Finnish consonant system. The length of consonants is functional in Finnish (e.g. *mato* 'worm' - *matto* 'carpet', *takana* 'behind' - *takkana* 'as a fireplace') but non-functional in English where the length of the consonant may vary freely. Thus a Finn may hear the English word *happy* as [hæpi] or [hæp:i] and he might consider them separate words in the beginning. This is a case of overdifferentiation, but no serious learning problems are involved: the learner soon learns to ignore the difference in length and it does not matter whether a Finn pronounces English consonants sometimes short or sometimes long.

Relational differences. –

allophones in Finnish, but phonemes in English	b	g	f w	ʃ z
the corresponding phoneme in Finnish	p	k	v	s

As mentioned earlier, relational differences between two sound systems may cause a maximum learning difficulty. The consonant sounds [b g f w ʃ z] are phonemes in English, whereas in Finnish they are regarded as allophones of / p k v s / respectively. There are several reasons for this interpretation. For example, the occurrence of [z] and [w] is fairly occasional in the speech of Finns. [z] may occur between two sonorants and in the speech of some educated Finns who have knowledge of foreign languages. [w] again can occur as an allophone of /v/ only between [u] and another back vowel, e.g. *vauva* [vauwa] 'baby', *hauva* [hauwa] 'doggie'. The sounds [b g f ʃ] seem to be on the way of acquiring the status of phonemes at least in the speech of educated Finns. However, it is dubious if even they make a consistent difference in their speech for example between [bussi] 'bus' and [pussi] 'bag' or between [ʃakki] 'chess' and [sakki] 'crowd, gang'. For the vast majority of Finns phonetic stretches like [liberaali] 'liberal', [gallona] 'gallon', [farmari] 'farmer' and [ʃakaali] 'jackal' are equal to [liperaali], [kallona], [varmari] and [sakaali], respectively.[1] This is also reflected in Finnish orthography. According to *Nykysuomen sanakirja* (1973: 468) it is equally correct to write *šaali* or *saali* 'shawl' and *šakaali* or *sakaali* 'jackal'. Pulkkinen (1966: 48) rightly notes that there is a trend in Finnish to replace < š > with < s >. This trend in orthography shows that the opposition /s/~/ʃ/ is felt to be foreign in Finnish.

In the case of the English phonemes / b g f w ʃ z / the learning problems may be much more complex than mere physical differences indicate,

[1] It is to be noted that there, in fact, is no real free variation between [p k s] and [b g z] , e.g. < bussi > can be pronounced either as [bussi] or as [pussi], but < pussi > always as [pussi].

as relational differences cause additional problems. It is perhaps difficult for Finns to distinguish /b/ from /p/, /g/ from /k/, /f/ from /v/ and possibly from /w/, and /s/ from /z/, because they need not make the fortis/lenis distinction in their own language. This hearing problem becomes more prominent, because the so-called voiced consonants in English are fully voiced only word-medially but partially de-voiced word-initially and word-finally. Thus the value of voicing as a clue for discriminating these sounds from each other is weakened.

Also the English /ʃ/ and /w/ may cause identification problems for Finns, as there is a danger that Finns hear and interpret them as /s/ and /v/ respectively.

In the case of these six allophones difficulties of pronunciation, too, are obvious. In all these cases Finns need not make any distinction between the allophones and the corresponding phonemes when speaking Finnish. They may well carry this habit over into English speech. In so doing they will be making a phoneme error.

Distributional differences. – Here we shall discuss the distribution of Finnish and English consonants in relation to words only. In order to be regarded as a genuine case of distribution, the phoneme in question has to fulfill the following conditions: the phoneme must occur (1) in isolated words, (2) in native words and (3) in words that are in no way marginal in the language. The distribution of Finnish and English consonants is shown in Chart 1. The cases that do not fulfill all the three conditions are inserted within brackets in the chart; as are also the Finnish words that may contain the sounds [w] and [z].

We do not accept /j/ and /r/ in [syöj+jo] and [hər aidiə] as word-final, because in isolated words *syö* and *her* /j/ and /r/ are never pronounced. According to condition (2), /d/ in *dia* in Finnish and /ǰ/ in *gigolo* in English, for example, cannot be considered to occur word-initially. Moreover, we regard interjections like *hep* and *huh* and onomatopoeic words like *vov-vov* and *pum* as marginal. Consequently, / p h v m / are not accepted as word-final in Finnish, neither is /k/ as it only appears in loanwords or onomatopoeic words like *sik-sak, tik-tak* 'tick-tock'.

From the point of view of a Finn learning English it is the differences in the distribution of the consonant phonemes occurring in both lan-

Chart 1. The distribution of consonant phonemes in Finnish and English.

	word-initial		word-medial		word-final	
	Finnish	English	Finnish	English	Finnish	English
/p/	poika 'boy'	poor	jopa 'even'	happy	(hep) an interjection	step
/t/	talo 'house'	tea	sota 'war'	water	neitsyt 'virgin'	fate
/d/	(dia) 'slide'	day	sade 'rain'	ready	(Daavid) 'David'	read
/k/	koti 'home'	key	joki 'river'	lucky	(sik-sak) 'zig-zag'	walk
/v/	vene 'boat'	valley	savi 'clay'	over	(vov-vov) 'bowwow'	save
/s/	sivu 'page'	say	kesä 'summer'	beside	mies 'man'	face
/h/	hieno 'fine'	home	lohi 'salmon'	behind	(huh) an interj.	-
/j/	juna 'train'	yet	vaja 'shed'	beyond	(syöj jo) 'eat at last'	-
/l/	lølu 'toy'	lap	melu 'noise'	alive	kyynel 'tear'	all
/r/	ruma 'ugly'	red	pari 'couple'	very	piennar 'edge'	(her idea)
/m/	meri 'sea'	milk	sama 'same'	limit	(pum) 'bang'	sum
/n/	nenä 'nose'	name	sana 'word'	many	pian 'soon'	can
/ŋ/	-	-	tanko 'bar'	singer	-	long
/b/	(basso) 'bass'	big	(tabu) 'taboo'	hobby	-	pub
/g/	(geeni) 'gene'	girl	(magia) 'magic'	figure	(Haag) 'the Hague'	leg
/f/	(firma) 'firm'	fire	(safari) 'safari'	suffer	(Joosef) 'Joseph'	wife
/w/	-	warm	(rouva) 'Mrs'	away	-	-
/ʃ/	(šakki) 'chess'	shoe	(tušši) 'drawing ink'	fashion	-	fish
/z/	-	zip	(hevosen) 'the gen. of horse'	busy	-	lose
/θ/	-	thin	-	author	-	tooth
/ð/	-	they	-	other	-	with
/ʒ/	-	(gigolo)	-	usual	-	rouge
/tʃ/	-	cheek	-	teacher	-	much
/dʒ/	-	just	-	wages	-	page

guages that need to be and will be discussed here in detail, Word-initially, the only difference is that /d/ occurs in English in that position, whereas in Finnish it may occur only in loan-words. All Finnish and English consonants may occur word-medially, whereas a number of differences are revealed in word-final position: / p d k v h j m ŋ / in Finnish versus / h j r / in English cannot be regarded as occurring word-finally. The distribution of / h j / is the same in both languages and thus learning problems should not appear. /r/ is a unique case in that its distribution is wider in Finnish than in English. Thus Finns must learn not to pronounce /r/ finally in isolated words (e.g. *star, there, dear, stare* etc.) but preserve it in the pronunciation when a vowel immediately follows (the so-called linking r). It is the spelling that may mislead Finns to pronounce final [r] sounds. Nevertheless, learning to use /r/ correctly in final position is perhaps not a serious problem, because from the point of view of communication the use of linking r is not absolutely necessary. On the other hand, it may be more difficult for a Finn to be able to hear and pronounce word-final / p d k v m ŋ / in English because of their restricted distribution in Finnish.

A REVIEW OF THE LEARNING PROBLEMS IMPLIED BY THE PRESENT CONTRASTIVE ANALYSIS. – At some points our contrastive analysis showed considerable differences between Finnish and English consonant systems. These differences may be postulated as a cause of learning problems. It seems then logical to make the following assumptions:

(1) It is more difficult for Finns to identify and produce such English consonant phonemes as do not occur in Finnish than those occurring in both languages. These consonant phonemes are: / b g f θ ð w ʃ z ʒ tʃ dʒ /.

(2) The fewer the distinctions between any two English consonant phonemes, the more difficult it is for Finns to keep them apart both in identification and pronunciation. This is to say that for example, /p/ in *pill* is more likely to be confused with /b/ in *bill* than than with /k/ in *kill*.

(3) It is difficult for Finns to identify and pronounce those English consonant phonemes that are distinguished from each other solely by the fortis/lenis opposition. These consonant pairs are: /p/-/b/, /t/-/d/, /k/-/g/, /f/-/v/, /θ/-/ð/, /s/-/z/, /ʃ/-/ʒ/ and /tʃ/-/dʒ/.

(4) It is more difficult for Finns to hear and produce word-final English consonants than word-initial or word-medial consonants. Thus /v/ in *Live* may be more difficult to learn than /v/ in *visit* or *heavy*.

(5) It is difficult for Finns to identify and produce English consonant phonemes which are allophones in Finnish. They are / b g f v ʃ z /.

In our opinion it is not safe to make any more detailed assumptions about the learning difficulties implied by our contrastive analysis. This is due to the fact that the problems in learning individual consonant phonemes are not likely to arise from one single difference between the given phonemic systems but from a complex of differences. For instance, when a Finn learns to distinguish /θ/ from /ð/ in *loath/loathe*, the following sources of difficulty are present: (1) they are both new sounds, (2) they differ only in one distinction, (3) the distinction is that of fortis/lenis and (4) they occur in word-final position. Now it is precarious to say whether these four factors are equally responsible for the learning problems or whether they form a hierarchy of difficulty. For that reason we are not able to arrange individual consonants in English into an order of difficulty on the basis of our contrastive analysis. We have to confine ourselves to the broader assumptions above. All these assumptions need to be verified empirically, which is the aim of this study.

CONSTRUCTION AND ADMINISTRATION OF TESTS

PRINCIPLES OF TEST CONSTRUCTION. – Tests are measuring instruments which are used to assign numerical values to the objects, events or properties being investigated. To be useful a test has to be, among other things,

(1) valid, i.e. it must measure what it is intended to measure,

(2) reliable, i.e. the results must be accurate, consistent and in no way dependent on chance,

(3) objective, i.e. the same scores are obteined regardless of the scorer,

(4) discriminatory, i.e. the objects of measurement can be arranged into an order of superiority, and

(5) practical in the sense that the test is easy and economical to administer and score.

Validity and reliability are commonly thought of as the most essential qualities of a good test (see for instance Downie 1967: 82,92; Harris 1969: 13; Heinonen 1961: 34; Kerlinger 1969: 429; Lado 1961: 30; Peltonen 1970: 15 and Valette 1967: 30). Therefore we shall discuss these concepts in more detail. According to Kerlinger (1969: 459), achieving reliability is mainly a technical matter while validity is much more than that: it involves philosophical considerations. For that reason, validity is more important than reliability (cf. Heinonen 1961: 34). The following types of validity are generally distinguished: (1) content validity, which means that a test covers the subject matter and objectives studied, (2) criterion related or predictive and concurrent validity, which means that test scores are correlated with some outside criterion, either future (predictive validity) or present (concurrent validity) criterion, (3) construct validity, which means "the degree to which certain exploratory concepts or constructs are responsible for performance in a certain test" (Downie 1967: 95), and (4) face validity, which merely means that a test seems to be valid for its purpose (Downie 1967: 93-96). It must be borne in mind that face validity is not validity in the technical sense and the validity of any test must be established in the other ways (1, 2 or 3) mentioned above. Nevertheless, face validity should not be overlooked. For example, if the content of a test looks irrelevant, silly or somehow inappropriate the examinees may lose their motivation or the test administrators will not want to use such a test (Harris 1969: 21).

There are four methods generally employed for assessing test reliability: (1) the test method, i.e. the same test is administered twice to the same examinees and the resulting scores are then correlated with each other, (2) the parallel forms method, i.e. two equivalent forms of the same test are administered to the same subjects and again the resulting two sets of scores are correlated, (3) the split-half method, i.e. the test is divided into two parts and the scores of the parts are then correlated, and (4) the Kuder-Richardson methods, i.e. special computation formulae that, like the split-half method, give a coefficient of the internal consistency of the test items. The Kuder-Richardson formulae can be regarded as an average of all possible splits.

In addition to the five requirements of a good test listed above, it

is desirable that a test efficiently answers the questions put forward in the study and that it can be repeated and its results can be statistically analysed and compared within different groups of subjects or with results arrived at in other tests.

The above principles apply to all tests. Therefore we tried to take them into account in constructing our listening and production tests. In our case special care had to be taken in the quality of the recording and the playback equipment to safeguard the reliability and validity of our tests. Lehtonen (1972b: 4, 11, 12, 18) states the minimum requirements for good recording and playback for research purposes:

(1) tape recorder 60 - 10,000 cps at $3^3/_4$ ips,
(2) signal-to-noise ratio >50 db
(3) tape speed $7^1/_2$ ips,
(4) microphone 40 - 15,000 cps,
(5) external loudspeaker 60 - 10,000 cps,
(6) sound-proof and echoless recording room, and
(7) in the case of minimal pair tests the test words should not contain any other clue than the one intended.

An attempt was made to meet these technical demands as fully as possible.

PRETEST VERSIONS. – All the tests and questionnaires were pretested at two secondary schools in Tampere (Pirkanmaan yhteiskoulu and Tampereen normaalilyseo). The number of the subjects was 110 second and fifth formers. Before the construction of the final versions the pretest data was thoroughly analysed (e.g. the tests were studied for reliability and an item analysis was performed to determine the discriminatory power of the test items). The tests and questionnaires were then revised.

FINAL TEST VERSIONS. – The final test battery included a substitution test, a discrimination test, a sound analogy test, a written analogy test and a production test. The necessary background data was gathered by means of a pupil questionnaire and a teacher questionnaire, which covered the variables in areas 1-12 in Diagram 1 on p. 8.

Substitution test. – The purpose of the substitution test (= S-test) was to find out which Finnish consonant phonemes Finns with no previous knowledge of English tend to substitute for the English consonants they hear. The subjects heard English words from the tape. Each word was uttered twice and the subjects were asked to write down on their answer sheets the words they heard using ordinary Finnish orthography. Since Finnish orthography is almost 100% phonemic, the subjects' answers should show sufficiently well how native speakers of Finnish identify English sounds in terms of Finnish phonemes. The test consisted of 70 items (35 test words), each consonant being 1 item, and of 15 practice items (4 words). The test is presented in Appendix 1.

Discrimination test. – By means of the discrimination test (= D-test) we wanted to find out how well our testees could differentiate between English consonant phonemes. Each item consisted of three English words, which the subjects heard on the tape, and the subjects were asked to mark on the answer sheets whether all the three words were (1) the same or (2) different, or which two words were the same, (3) the first two, (4) the last two or (5) the first and the third. Thus for example they heard from the tape and they should have marked on their answer sheets.

11. rum rum rum	11.	(X) (X) (X)
13. bet wet vet	13.	~~( ) ( ) ( )~~
69. bays bays beige	69.	(X) (X) ( )
29. strife strive strive	29.	( ) (X) (X)
9. cold gold cold	9.	(X) ( ) (X)

The order of the correct answer patterns was randomised to prevent any answer patterning. The use of five answer alternatives made the chance of successful guessing as low as 20%. Quadruplets (e.g. which - rich - which - which) instead of triplets would have further reduced the effect of guessing but such a test would have imposed a memory burden. In fact, Lado (1961: 54-55) is of the opinion that the triplet technique is "the most effective and satisfactory one to test aural perception that has been reported".

The actual test, preceded by 4 Finnish and 4 English practice items, comprised 75 items. It is given in Appendix 2.

Sound analogy test. – To measure how well foreign language sounds are identified, tests of the minimal pair type have so far solely been used, although there are some doubts (see for instance Lehtonen 1972b: 18) that they measure auditory discrimination rather than any mastery of the sound oppositions in a given language. A person with good hearing ability may well distinguish /θ/ from /ð/ without knowing any English. In such a case it cannot be maintained that the person has mastered the opposition /θ/-/ð/ in English, although one might easily be misled into drawing such a conclusion on the basis of discrimination tests of the minimal pair type.

Therefore we decided to devise tests that would measure the identification of English consonant phonemes without resorting to minimal pair techniques. Thus we arrived at the sound analogy (SA-test) and written analogy (WA-test) tests.

Every item in the sound analogy test consisted of three English words heard from the tape. The first word served as a stimulus and it was an unfamiliar word to the testees. They were instructed to listen carefully to its first sound. After a short pause they heard the other two words, which were absolutely familiar to them. Again they were instructed to listen to the first sound in the words. Then their task was to compare whether (1) both of the latter words, (2) neither of them, (3) the first of them or (4) the second of them began with the same sound as the stimulus. This is how word-initial consonant phoneme oppositions (25 items + 4 practice items) were tested. To test word-final consonant phoneme oppositions (20 items + 3 practice items), the same procedure was applied and the testees were asked to pay attention to the last sound in the words. We shall exemplify the four answer alternatives (the chance of successful guessing =25%) of our 45-item test:

	from the tape			on the answer sheet		
	stimulus	analogical words				
word-initial	3. faun	film	four	3.	(X)	(X)
	10. willow	very	boat	10.	~~( )~~	~~( )~~
	9. sear	say	shop	9.	(X)	( )
	11. turf	dark	today	11.	( )	(X)
word-final	27. tang	lying	ring	27.	(X)	(X)
	35. rude	lea	with	35.	~~( )~~	~~( )~~
	32. hawk	back	dog	32.	(X)	( )
	28. mash	miss	dish	28.	( )	(X)

In this test, too, the order of the correct answer patterns was rotated at random. The SA-test can be seen in Appendix 3.

The point of the sound analogy test was that the testees *had to identify*, from an unfamiliar stretch of sounds, a familiar consonant phoneme and to indicate what the consonant was by comparing it with the word-initial/word-final consonant in familiar analogical words.

Written analogy test. – The written analogy test (WA-test) was in principle similar to the sound analogy test, the only difference being that the familiar analogical words were not heard from the tape. Instead they were written on the answer sheet. Thus the subjects were asked to listen carefully to the initial/final sound of the unfamiliar stimulus word and to compare whether (1) both, (2) neither, (3) the first or (4) the second of the words on the answer sheet began with / ended in the same *sound* (not *letter*) as the stimulus. The testees were to mark their answers as follows:

word-initial		word-final	
from the tape	on the answer sheet	from the tape	on the answer sheet
2. fag	2. phone (X) five (X)	33. fuse	33. blouse (X) always (X)
16. locus	16. table (-) how (-)	36. tithe	36. tooth (-) give (-)
1. poke	1. pen (X) buy ( )	40. fade	40. bed (X) with ( )
25. sheer	25. seven ( ) show (X)	43. deem	43. strong ( ) home (X)

The test comprised 48 items: 26 items testing word-initial consonant oppositions (+ 4 practice items) and 22 items testing word-final consonant oppositions (+ 3 practice items). Here again the guessing rate was 25% and the order of the correct answer patterns was randomized. The WA-test is presented in Appendix 4.

Production test. – The production test (=P-test) was a reproduction test. The subjects were instructed to listen carefully to English words which they heard from the tape. Each word was heard twice and the testees' task was to reproduce the words. The test included 41 test words and every consonant in them formed an item. The number of items was either 103, if /tʃ/ and /dʒ/ were treated as consonant clusters, or 93, if they were considered unit phonemes. The P-test is to be seen in Appendix 5.

The tapes containing the testees' productions were so edited that the stimuli were erased. Thus the evaluators, JC and RP (native speakers of English), RM, EV and the subjects' English teachers (all native speakers of Finnish), heard only the subjects' responses and their task was to transcribe phonemically (using broad transcription) the responses on ready-made marking sheets. Consequently, in no phase did the evaluators have to decide whether the testees pronounced the phonemes correctly or incorrectly, they simply wrote down the phonemes they heard. RM and EV did the scoring afterwards on the basis of the transcriptions. The advantage of this method is that exact information can be obtained on what kind of mistakes were made. This information could not have been obtained, if, as often is the case, the responses had been directly marked right or wrong.

RECORDING AND ADMINISTRATION OF FINAL TESTS. – The tests were recorded in the studio of the Speech Department at the University of Tampere according to the criteria stated on p. 25. The test words were read on the tape by RP, a native speaker of English.

The listening tests (S-test, D-test, SA-test and WA-test) were administered in March 1973 to 329 secondary school pupils at 3 schools in Tampere: Harjun yhteiskoulu (HYK), Sammon yhteislyseo (SYL) and Tampereen yhteislyseo (TYL). The schools, the forms, the number of pupils in the forms, the number of testees, the failure rates and the tests administered are presented in Table 1.

Table 1. Testees of final tests.

school	form	number of pupils	number of testees in S-test	D-test	SA-test	WA-test	P-test	highest failure rate
HYK LE	II B	38		38	38	38	8	0
LE	V B	37		35	35	35	8	2
LG	V A	35	32	32				3
SYL LE	II A	38		37	37	37	8	1
LE	V C	41		40	40	40	8	1
LG	V A	38	38	38				0
TYL LE	II A	39		39	39	38	8	1
LE	V A	42		40	42	40	8	2
LG	V B	38	30	30				8
3	9	346	100	329	231	228	48	18 (5.4%)
LE= learners of English		235		229	231	228	48	7 (2.9%)
2nd formers		115		114	114	113	24	2 (1.7%)
5th formers		120		115	117	115	24	5 (4.1%)
LG= learners of German		111	100	100				11 (9.9%)

The production test was administered to 24 second formers and 24 fifth formers in the above schools. The testees were so selected that in each school 4 second and 4 fifth formers with the highest and 4 second and 4 fifth formers with the lowest sum total of D-test, SA-test and WA-test were chosen as subjects.

On the whole, the failure rates were low, the only exception being form V B of TYL with as many as 8 pupils (over 20%) who did not attend the two tests.

The selection of the learners of German to represent subjects with no previous knowledge of English was not an ideal solution, because their knowledge of Swedish and German obviously affected their interpretation

of English consonants. However, there was no subject group available in Tampere that would have fulfilled our original requirements, i.e. the testees should have been pupils (1) with absolutely no previous knowledge of English, (2) preferably with no knowledge of other languages than Finnish, (3) at the age to start learning English at school and (4) with sufficient writing ability.

Along with the substitution test we ran the discrimination test with the learners of German, because doubts have been expressed that tests of this type hardly measure the mastery of the sound oppositions in a given language. As the subjects had practically no knowledge of English (only 11 of the 100 testees had some knowledge of English), we could explore the construct validity of our discrimination test in the following way: if the learners of German obtain significantly lower scores than the learners of English it can be assumed that the D-test has construct validity, if they have equally high or even higher scores than the learners of English, the test obviously lacks construct validity, i.e. the D-test does not measure the mastery of the sound oppositions in a given language.

METHODS OF STATISTICAL ANALYSIS AND DATA PROCESSING

The following methods were used in the statistical analysis of the final tests and questionnaires and the data obtained from them:

(1) frequencies (f)
(2) percentages (%)
(3) means ($\bar{X}$)
(4) standard deviations (s)
(5) product-moment correlations (r)
(6) t-tests (t) for
    (a) the significance of the difference between means
    (b) the significance of correlation coefficients
(7) Kuder-Richardson Formula 20 to determine the reliability coefficients of the tests (KR_{20})
(8) regression analysis, free model.

In analysing the data we used the Statistical Data Processing System SURVO/71 developed at the University of Tampere. We ourselves made the necessary SURVO-programmes, which were run by a UNIVAC computer in Helsinki.

RESULTS

AN ATTEMPT TO ANSWER PROBLEM 1:

WHICH FINNISH CONSONANTS ARE GIVEN AS SUBSTITUTES FOR ENGLISH CONSONANTS BY PUPILS WHO HAVE NO PREVIOUS KNOWLEDGE OF ENGLISH

The substitution test was designed to gather the necessary information to enable us to find an answer to problem 1. The core of the results is presented in Tables 2-8, in which the first column shows the word containing the tested consonant phoneme (with the corresponding letter/letters underlined); the next columns enumerate the substitutes with their corresponding frequencies (only the substitutes with a frequency of 5 or more in some position in the word are reported); the column "others" gives the sum of frequencies of the rest of the substitutes; the column Ø reports the number of the cases where no substitute is given and the last column gives the total number of the different substitutes given for the English consonant in question. It is to be noticed here that in some cases there is doubt about what sounds the subjects mean by their substitutes. For instance it is uncertain what sound is meant by the substitute <z>, as <z> does not belong to the Finnish spelling convention. In our opinion the subjects may mean by <z> (1) the sound sequence [ts], as is the case in the Finnish product name *Fazer* [fatser] and in German (e.g. *Zahl* [tsa:l]) or (2) the voiced sibilant [z], as it occurs in German (e.g. *säugen* [zoigen]). The difficulty of interpreting the substitutes is primarily due to the fact that our subjects knew Swedish and German. It is possible that they interpreted the English consonant phonemes not only in terms of the Finnish consonant system, but also in terms of the consonant systems of Swedish and German. Therefore the results must be interpreted with caution.

As the number of the subjects was 100, the frequencies in the following tables are percentages at the same time. The sum totals make an exception: to get the percentages one must divide them by the number of times the consonant is tested. The sum total of the number of different substitutes is usually less than the sum of different substitutes in word-initial, word-medial and word-final positions, as there is overlapping: the same substitute can be given for the tested consonant in all those positions.

The results of the identification of the plosives is given in Table 2.

Table 2. Identification of plosives / p t k b d g / (N=100)

in item word	transcribed as						others	<ø>	No. of diff. subst.
/p/	<b>	<ph>	<h>	<bh>	<p>	<d>			
pack	36	22	15	14	12	0	1	0	6
appeal	35	8	0	2	48	1	6	0	9
zip	10	0	1	0	25	29	9	26	12
Σ	81	30	16	16	85	30	16	26	18
X̄%	27	10	5	5	28	10	5	9	
/t/	<t>	<th>	<d>	<tt>	<nd>				
teeth	76	11	9	0	0		4	0	5
better	48	22	1	20	0		9	0	13
fate	4	1	95	0	5		5	0	8
Σ	128	34	95	20	5		18	0	17
X̄%	43	11	32	7	2		6	0	
/k/	<k>	<kh>	<g>	<gh>	<nk>	<n>			
cab	51	23	16	6	0	0	4	0	5
viking	80	0	17	0	0	0	3	0	4
pack	42	2	10	1	15	5	23	2	23
Σ	173	25	43	7	15	5	30	2	25
X̄%	58	8	14	2	5	2	10	1	
/b/	<b>	<p>	<bh>	<d>	<v>	<w>	<m>	<n>			
better	89	5	5	1	0	0	0	0	0	0	4
beyond	97	3	0	0	0	0	0	0	0	0	2
rubber	35	2	0	11	33	5	1	0	13	0	15
cab	1	1	0	4	3	0	21	59	9	2	10
Σ	231	11	5	16	36	5	22	59	22	2	21
X̄%	56	3	1	4	9	1	5	14	5	0	
/d/	<d>	<g>	<t>	<n>	<l>	<v>	<nd>				
deserve	90	5	2	1	0	0	0		2	0	6
shady	55	1	0	11	12	8	1		10	2	14
thud	69	0	7	0	0	0	9		15	0	12
beyond	91	3	4	0	0	0	0		0	2	3
Σ	305	9	13	12	12	8	10		27	4	22
X̄%	76	2	3	3	3	2	3		7	1	
/g/	<g>	<k>	<n>								
garage	86	9	0						5	0	5
yoga	80	12	0						8	0	9
leg	55	27	7						11	0	13
Σ	221	48	7						24	0	20
X̄%	74	16	2						8	0	

Regardless of its position in the word the English /p/ is most frequently transcribed as <p> (85), almost as frequently as <b> (81). It is to be noticed, however, that word-initial [p] has been substituted with <ph> (22) and <bh> (14), where the h-element can obviously be regarded as representing the strong aspiration pertaining to word-initial English voiceless plosives. As can be seen [p] has been identified as an aspirated [b], although no such sound occurs in English. It would stand to reason to regard <p> and <ph> as representing the phoneme /p/; similarly <b> and <bh> as representing /b/. In 15 cases only [h] has been heard instead of word-initial [p]. Thus half (51) of our subjects have noted the aspiration. On the basis of the preceding one could perhaps sum up <p> and <ph> (85 + 30 = 115) and <b> and <bh> (81 + 16 = 97). This increases the proportion of the "correct" substitute, but still the sum of the phonetically nearest "incorrect" substitute remains surprisingly high. On the basis of the contrastive analysis one could have expected the English /p/ to have been transcribed more often as <p> or <ph>, at least word-initially and word-medially. Less unexpected seems the result that word-final [p] has been poorly identified. Still, the great number of different substitutes given (12), the low frequencies of the phonetically likely substitutes <p> (25) and <b> (10) and the high frequencies of the substitutes <d> (29) and the category "no substitute" (<Ø> =26) show that the subjects have had considerable difficulties in identifying [p] in word-final position. Something like this could be expected on the basis of the contrastive analysis (word-final consonants are rare in Finnish), but the number of <Ø> is startlingly high.

The plosive /t/ has been fairly often transcribed as <t> in word-initial (76) and word-medial (48) positions, but astonishingly seldom in word-final position (4). Again one can notice that a number of subjects have detected the aspiration. Further, [t] in medial has also been transcribed by rather many (20) as double (<tt>). There is reason to believe that the subjects have by their substitutes < t th tt > meant the phoneme /t/. Then the number of "correct" substitutes would be 87 word-initially and 90 word-medially. In these positions the substitute <d> is rare, whereas word-finally it is by far the most frequent (65). Even the substitute <nd> has been given as often as <t> and <th> together. We can find no feasible explanation for this, especially as /t/ occurs word-finally in Finnish as well (e.g. *pidot*, *halot*). Not even the fact that the Finnish /d/ and

English /t/ are articulated at the same place can explain this anomaly because [d] may occur in Finnish word-finally only in loan-words (e.g. *Daavid*).

A tendency similar to the identification of /p/ and /t/ is to be seen in the identification of /k/: word-initially and word-medially /k/ has been better identified than word-finally and the h-element after <k> and <g> in word-final position indicates that the aspiration has been registered (see above Table 2, p. 33). The substitutes <k> and <kh> apparently represent unaspirated and aspirated allophones of the English phoneme /k/: <g> and <gh> can similarly be thought of as representing the phoneme /g/. <k> and <kh> together have the frequency 74 word-initially, 80 word-medially and 44 word-finally; the corresponding figures for <g> + <gh> are 22, 17 and 11. Word-finally the number of different substitutes is strikingly high (23). This alone indicates how difficult it is for Finns to identify word-final consonants.

The word-initial [b] has been well identified: in *better* 89 subjects have given the "correct" <b> and 5 have marked <bh>; in *beyond* almost all the subjects (97%) have registered <b>. In word-medial, and especially in word-final, position a number of substitutes, mostly other than the most likely <b> and <p> have been given for the Enflish /b/. In *rubber* the substitute <v> (33) competes well with the "correct" <b> (35). In word-final position only one <b> was registered, the major categories of the substitutes being <m> (59) and <n> (21). It is difficult to find any explanation of why /b/ has been so well identified word-initially and so poorly identified word-medially and word-finally.

On the whole, /d/ does not seem to cause hearing problems: oddly enough even word-final [d] has been "correctly" identified in the majority of cases (the figures being 69 for *thud* and 91 for *beyond*), i.e. word-final [d] has been identified roughly as well as word-initial [d] (90). On the basis of the contrastive analysis it is somewhat surprising that word-medial [d] has the lowest "correct" answer percentage. It is only in word-medial position where /d/ occurs in native Finnish words, e.g. *sade*, *odotus*. Therefore one would have expected [d] to have been better identified.

As with the other plosives except /d/, /g/ has also received higher percentages of "correct" answers in word-initial (86) and word-medial (80)

positions than word-finally (55). The phonetically nearest consonant phoneme to /g/ is /k/ and therefore it is no wonder that <k> dominates among the "incorrect" substitutes, especially word-finally, where the English [g] is never fully voiced.

The column "others" contains rather many cases in which a consonant diagraph has been substituted for the tested single consonant. For instance, the diagraphs < ngh nd hr hk > have been given as substitutes for word-final [k] and < th ld nt dt lk nth > for word-final [d]. The cases in the category "others" are occasional in the sense that they have very low frequencies: usually the frequency is only one.

The identification of the spirants is given in Table 3. The spirant /f/ has been remarkably often transcribed as <f> in all positions (see Table 3, p. 37). This is obviously explained rather by the fact that /f/ was familiar to our subjects from Swedish and German than by the fact that [f] occurs in loan-words and dialects in Finnish. Learning to identify the English /f/ would thus not be a problem for our subjects. However, it remains an open question whether Finns without any knowledge of any other language than their own would have identified /f/ so well.

The English /v/ has been uniformly transcribed as <v>, the word-final /v/ being an exception. Rather many have written /v/ as <f>. This is particularly true of the interpretation of the word-final [v]. Apparently the fact that the English [v] is partly devoiced in this position at least to some extent accounts for the result. /v/ has also been marked as <w>. One cannot be sure whether sound [w] or [v] is meant. In Finnish and also in German the letter <w> stands for the consonant /v/ and therefore some of our subjects might have meant /v/ with their <w>. Some of the subjects may have had enough knowledge of English to indicate the consonant /w/ with their letter <w> (the background data revealed that 11 of our 100 subjects had studied English either in elementary school or privately).

It is to be noticed that there is a letter in the Finnish alphabet to represent each of the 8 consonants that have been dealt with so far. The Finnish alphabet lacks, however, the means of indicating the next two spirants in Table 3. Therefore it is no wonder that the phoneme /θ/ has been registered almost invariably as <f>, the word-medial [θ] also attracting other substitutes. How to interpret <th> is uncertain: it may denote (1) an aspirated [t] (cf. above p. 34), or (2) the sequence [t] + [h] as in *Saithan sen?* 'You got it didn't you?' or, less probably, (3) the

Table 3. Identification of spirants / f v θ ð h / (N=100)

in item word	transcribed as							others	<ø>	No. of diff. subst.
/f/	<f>	<pf>	<v>							
fate	90	5	1					4	0	6
foolish	98	2	0					0	0	2
surface	86	1	1					12	0	9
chief	91	0	5					4	0	5
Σ	365	8	7					20	0	14
X̄%	91	2	2					5	0	
/v/	<v>	<w>	<f>	<rv>						
viking	70	14	12	0				4	0	6
ever	81	4	2	2				11	0	15
deserve	53	0	17	9				15	6	16
Σ	204	18	31	11				30	6	30
X̄%	68	6	10	4				10	2	
/θ/	<f>	<ff>	<th>							
thud	96	0	0					3	0	3
author	64	7	14					15	0	13
teeth	97	0	0					3	0	4
Σ	257	7	14					21	1	16
X̄%	86	2	5					7	0	
/ð/	<v>	<w>	<f>	<d>	<t>	<th>	<lf>			
those	36	10	10	10	9	8	0	17	0	17
neither	43	4	30	9	5	4	0	5	0	10
with	3	0	43	3	3	12	11	25	0	25
Σ	82	14	83	22	17	24	11	47	0	38
X̄%	27	5	28	7	6	8	4	16		
/h/	<h>									
hanger	100							0	0	1
Σ	100							0	0	1
X̄%	100							0		

written manifestation of the phoneme /θ/ in English. Interpretation (3) is possible, because those 11 who had some knowledge of English might have known that the phoneme /θ/ is represented in writing by <th>. The results seem to suggest that /θ/ is a major learning problem for Finns, as they interpret both /f/ and /θ/ as <f>. Thus they are not likely to keep these phonemes apart from each other. Therefore special care must be taken to emphasize at a very early stage of learning English that /f/ and /θ/ are two separate consonants in English.

There is a lot of variation in the identification of the English consonant phoneme /ð/. The substitutes <v> and <f> attract the highest frequencies, but also <th>, <d> and <t> are fairly well represented among the substitutes (see Table 3). The substitutes with the highest frequencies, i.e. <v> and <f> are logical in the sense that they are phonetically the nearest possible consonants to replace /ð/. But the substitution of <v> and <f> for /ð/ results in a phonemic error and therefore /ð/ seems to constitute a serious learning problem. The high prcportion of different substitutes also implies that the phoneme /ð/ sounds very odd to the Finnish ear: substitutes like < 1th rf vd fn 1h vf 1d ds 1s > have been given for word-final [ð].

The identification of the spirant /h/ needs no comments: the frequency 100 for <h> speaks for itself.

Again, in the case of the spirants, the substitutes in the column "others" are individual cases and diagraphs are very common along them. Consonant diagraphs as substitutes for word-final [ð] have been exemplified above. As our examples show, consonant diagraphs are especially frequent word-finally. One further example: the subjects have found word-final [v] to be for instance < vf vs 1f 1h 1v ds 1d nt >.

As concerns the semi-vowels (Table 4), the majority of the subjects have marked /w/ as <v>. Quite many have also given <w> and it is not quite clear whether it denotes the phoneme /v/ as it does in Finnish and German or /w/ as it does in English. However, the frequencies of the substitute <w> (25 word-initially and 20 word-medially) are much higher than the number of those (11) who knew English would presuppose. Therefore one would be inclined to believe that /v/ rather than /w/ is meant by <w>. Possibly some subjects preferred <w> to <v> because of its more foreign appearance (the subjects knew that the test words were English). The vocalic nature of the semi-vowel /w/ is to be seen in the substitutes that contain a vowel

Table 4. Identification of semi-vowels / w j / (N=100)

in item word	transcribed as			others	<ø>	No. of diff. subst.
/w/	<v>	<w>	<uv>			
with	47	25	3	25	0	21
away	66	20	8	6	0	8
Σ	113	45	11	31	0	24
X%	57	23	6	16		
/j/	<j>					
yoga	97			3	0	4
beyond	16			0	84	1
Σ	113			3	84	4
X%	57			2	42	

<uv> is reported in Table 4 and substitutes like < vu ui vui uvi u gu bu bvu wu > are included in the category "others".

The semi-vowel /w/ causes hearing problems to native speakers of Finnish, as they tend to interpret both /v/ and /w/ as <v>. Therefore it is important to teach the pupils to make a difference between the two phonemes at the very outset of learning English.

The subjects did not meet any difficulties in identifying word-initial /j/. Only three subjects have offered other substitutes than <j>. They were < ij u y >. They give some evidence of the vowel-like quality of /j/. In word-medial position only 16 have marked <j>, while 84 have marked nothing to stand for [j]. This is obviously due to the environment: [j] is preceded by [i] in the test word *beyond*.

The test results for the sibilants are given in Table 5. The English /s/ has been most frequently transcribed as <s>. Word-medially the diagraph <ss> has a high frequency (53). Obviously the testees have recognized the loan-word *essee* in the test word *essay*. It seems legitimate to regard <s> and <ss> as representing the phoneme/s/. The substitutes <š> and <sh> most likely stand for the sound [ʃ], because <š> is the correct and <sh> the older way of indicating [ʃ] in Finnish orthography. This does not, however, exclude the possibility thas <sh> denotes the sequence

Table 5. Identification of sibilants / s z ʃ ʒ / (N=100)

in item word	transcribed as						others	<ø>	No. of diff. subst.
/s/	<s>	<š>	<ss>	<sh>					
surface	89	7	0	0			4	0	5
essay	45	1	53	0			1	0	4
surface	72	12	2	10			4	0	8
Σ	206	20	55	10			9	0	
X̄%	69	7	18	3			3	0	11
/z/	<s>	<š>							
zip	59	28					12	1	10
deserve	93	6					1	0	3
those	85	5					10	0	8
Σ	237	39					23	1	16
X̄%	79	13					8	0	
/ʃ/	<š>	<sh>	<s>	<šh>	<šš>	<ss>			
shady	58	20	13	7	0	0	2	0	6
Asia	47	5	18	2	11	11	5	1	10
foolish	54	10	27	5	0	0	4	0	6
Σ	159	35	58	14	11	11	11	1	12
X̄%	53	12	19	5	4	4	4	0	
/ʒ/	<š>	<s>	<šš>	<ls>					
azure	47	26	6	0			21	0	15
garage	49	26	0	6			17	2	15
Σ	96	52	6	6			38	2	23
X̄%	48	26	3	3			19	1	

[s] + [h] as in *Mieshän se oli?* 'It was a man, wasn't it?' Whichever interpretation is right, <š> and <sh> have been offered as substitutes for /s/ mainly in word-final position. For the majority of our subjects, /s/ presented no hearing problems, only some 10% confused it with [ʃ].

The English /z/ has been interpreted mainly as <s> (237) and to some extent as <š> (39). This is no wonder as in Finnish /s/ is the nearest equivalent to the English /z/ and Finnish orthography lacks the means of indicating the sound [z].

The great majority of the testees have interpreted the English /ʃ/ as some kind of wide sibilant as the substitutes < š sh šh šš > (159 + 35 + 14 + 11 = 219 = 73%) show. They all apparently denote /ʃ/, which the subjects knew from Swedish and German. In fact, the frequencies of < s ss> (58 + 11 = 69 = 23%) are unexpectedly low. Again, as in *better* and *essay*, double consonants have been given as substitutes for a word-medial English consonant: word-medial [ʃ] has been marked as <šš> (11) and as <ss> (11) by one-fifth of the testees. The preceding short syllable containing a lax vowel and, compared with Finnish, the longer duration of English fortis consonants may explain this tendency.

Roughly half of the subjects have marked /ʒ/ as <š>, a quarter of them as <s> and the rest have given various suggestions such as < sh šš ls sj ss nš sch z rsch >, the most frequent of them being <sh> (8), <šš> (6) and <ss> (4). The great number of different substitutes (23) reflects the difficulty of indicating /ʒ/ in terms of Finnish orthography. The subjects have tried to indicate /ʒ/ in a varying number of ways, mostly with consonant diagraphs as can be seen above.

Table 6 presents the data for the identification of the affricates. On the whole, /tʃ/ has been transcribed as <tš> (125), fairly often as <ts> (71) and the great number of individual cases (63) and of different substitutes (42) obviously reflects the subjects' difficulties in transcribing the affricate /tʃ/. The position of the affricate in the word seems to affect the interpretation. Word-initially the substitute <ts> is the most frequent, while <tš> is by far the most frequent word-finally. Also word-medially <tš> is the greatest category (35), but the category "others" is almost as great (33). The substitute <z> may denote [ts] or [z]. However, it seems probable that [ts] is meant (see above p. 32). Also <š> has gained some support (11). It is to be noted that most of the substitutes in the category "others" contain either a t-element or a š-element or both. Examples: < tj th dts tjs št tsj t tt ktš tz kts tx thj tts kš tsch tch tc ttš >. On the whole, the testees tended to consider [tʃ] to contain two segments.

The English /dʒ/ has been given a large selection of substitutes (34 altogether), reflecting the lack of an appropriate sign in Finnish orthography. <ts> and <tš> are the most frequent substitutes given (50%). If <z> is interpreted to denote [ts], it will further increase the proportion of <ts>-substitutes. In general, /dʒ/ as well as /tʃ/ has been

Table 6. Identification of affricates / tʃ dʒ / (N=100)

in item word	transcribed as					others	<ø>	No. of diff. subst.
/tʃ/	<ts>	<t͡s>	<s̆>	<z>	<d͡z>			
chief	38	26	10	11	0	15	0	16
etcher	19	35	1	12	0	33	0	25
itch	14	64	0	1	6	15	0	15
Σ	71	125	11	24	6	63	0	42
X̄%	24	42	4	8	2	21	0	

/d /	<ts>	<t͡s>	<tj>	<j>	<z>	<s>	<s>	<ds>	<tsj>			
gem	21	13	12	9	12	6	9	2	3	13	0	20
adjure	31	25	3	1	10	4	2	3	5	11	0	19
emerge	29	35	0	0	6	3	2	10	0	15	0	16
Σ	81	73	15	10	28	13	18	15	8	39	0	34
X̄%	27	24	5	3	9	4	6	5	3	13	0	

marked as some kind of consonant diagraph. In addition to those reported in Table 6, diagraphs like < d͡s̆ gs t͡sj sj s̆j zj st tz dj jh zd dts dds rds dz nz rs > were given as substitutes. Apparently the voicing contrast of the affricates presents learning problems for Finns, as they tend to interpret both of them as <t͡s̆> or <ts>. Therefore it seems essential to teach the Finns to make a distinction between [ts] and [tʃ], [ts] and [dʒ] and [tʃ] and [dʒ] at an early stage of learning English. Thus phonemic errors like *hats* for *hatch*, *bats* for *badge* and *cheap* for *jeep* could be avoided.

As to the identification of the English /r/ (see Table 7) most of the subjects identified it in both of the test words *rubber* and *garage*, but in word-initial position surprisingly many have given a diagraph containing <r>.
 in particular has a high frequency (49), almost twice as high as <r> (26). The fact that in English word-initial [r] is often strongly labialized, i.e. [b]- or [w]-like may explain this. Also the sound environment i.e. the following word-medial [b], may have confused the testees It is possible that they have wrongly segmented the word *rubber* and heard the word-medial [b] as word-initial. The word-medial [r] does not seem to cause hearing problems.

Table 7. Identification of /r/ and /l/ (N=100)

in item word	transcribed as				others	<ø>	No. of diff. subst.
/r/	 	<r>	<vr>	<wr>			
rubber	49	26	8	5	12	0	12
garage	0	86	0	0	13	1	6
Σ	49	112	8	5	25	1	16
X%	25	56	4	3	13	1	
/l/	<l>	<u>	<uu>				
leg	99	0	0	0	1	0	2
foolish	94	0	0	0	6	0	3
appeal	29	40	6	8	17	0	15
Σ	222	40	6	8	24	0	18
X%	74	13	2	3	8	0	

The English /l/ has been interpreted "correctly" as <l> word-initially and word-medially, the percentages being 99 and 94 respectively, while word-final [l] has been written as <l> in 29 cases only. The interesting thing about this is that the word-initial and word-medial l-sounds are clear allophones of /l/ and the word-final l-sound the dark allophone of /l/. The dark l seems to cause identification problems for Finns, who tend to hear it as a (back) vowel or as a sequence of a vowel and /l/: 40% of the subjects heard [u], 6% [uu], 3% [o], 8% [ul] and 4% [ol]. This may be explained by the fact that the Finnish /l/ is never as dark as its English counterpart. Further, in the context of front vowels the Finnish /l/ is nearer to the English clear l than to the dark variant. Compare e.g. *kyynel* 'tear' with the test word *appeal*. The dark l causes primarily a hearing problem (not being able to distinguish the dark l from a back vowel may make a word unintelligible), whereas if a Finn pronounces a clear l instead of a dark one when speaking English the mistake is not phonemic.

Table 8 indicates that the English nasals have been identified more or less correctly because of their close correspondence to their Finnish counterparts. Where they have been "incorrectly" identified, another nasal has usually been heard instead of the "correct" one. In the test word

Table 8. Identification of nasals / m n ŋ / (N=100)

in item word	transcribed as <m>	<n>	others	<ø>	No. of diff. subst.
/m/					
mountain	73	26	1	0	3
emerge	90	1	9	0	6
gem	69	23	8	0	9
Σ	232	50	18	0	12
X̄%	77	17	6	0	
/n/	<n>	<ng>			
neither	100	0	0	0	1
mountain	91	0	3	6	2
beyond	97	2	1	0	3
mountain	76	7	6	11	6
Σ	364	9	10	17	8
X̄%	91	2	3	4	
/ŋ/	<ng>	<m>	<n>		
hanger	85	6	4	5	0	8
viking	78	1	21	0	0	3
Σ	163	7	25	5	0	8
X̄%	82	4	13	3	0	

mountain the word-medial and word-final n-sounds may have affected the interpretation of the word-initial [m] retroactively, which explains the relatively high frequency of <n> (26). The substitute <ng> most likely stands for the phoneme /ŋ/, as it does in Finnish as well as in English orthography (e.g. *kangas* 'cloth', *hanger*). As one could have expected on the basis of the preceding results, the "correct" substitute percentages were the lowest in word-final position. Still, the nasals have been identified fairly well in that position, too. The results seem to suggest that Finns are not likely to encounter difficulties in identifying English nasals.

As one may have detected from the tables above, the number of different substitutes seems to be a rough estimate of the difficulty of the process of identifying consonants: the higher the number of different substi-

tutes, the more difficult the process of identification. As Finnish orthography lacks means of expressing "correct" substitutes for some English consonants (e.g. /θ/ and /ʒ/), it was not considered justified to compute the average correct answer percentages for consonants in word-initial, word-medial and word-final positions to find out whether the positions affected the identification. Instead, we computed the average number of different substitutes for word-initial, word-medial and word-final consonants. The average number of different substitutes for word-initial consonants was 7.45, for word-medial consonants 9.70 and for word-final consonants 11.45. This seems to suggest that the process of identifying word-final consonants is the most difficult.

It is interesting to compare our results with those of Wiik (1965b). On the whole, Wiik's subjects seem to have given similar substitutes to those of our subjects and in some cases even the frequencies of the substitutes are essentially the same. For instance, the word-final dark [l] has been transcribed as <l> in 41% of the cases, as <u> in 38%, as <ul> in 9%, as <o> in 3%, as <v> in 2% and as <ol> in 1% of the cases in Wiik's study. (For reference see our results in Table 7). Also the substitutes for word-initial [ð] and their frequencies in our study conform to those in Wiik's investigation: Wiik's percentages for <v> (46%), <t> (9%), <d> (8%) approximate to ours (36%, 9%, 10% and 10%, respectively). As to the identification of word-initial [θ], Wiik has a large selection of substitutes: <f> (51%), <s> (28%), <t> (3%), <th> (2%), <ts> (2%) and others (14%), whereas 96% of our subjects substituted <f> for [θ] word-initially. Thus [f] seemed to be the nearest equivalent to /θ/ in both studies. The results of the identification of word-initial <w> also coincide to a noteworthy degree: <v> 66%, <u> 10%, <vu> 8%, <uv> 2% and others 14% in Wiik's study versus <v> 47%, <w> 25%, <uv> 3% and others 25% in ours. Similar substitutes have been given for word-initial [r]: 38% of Wiik's subjects gave <r>, 8% <v>, 3% <u>, 15% <vr>, 3% <fr>, 2%
, 3% <ur> and 30% others. In our study 49% of the subjects gave
, 26% <r>, 8% <vr>, 5% <wr> and 12% others. The substitutes < r vr br > are the same in both studies, although their frequencies differ. Almost identical substitutes have been given in both studies for word-initial voiceless plosives. In Wiik's study the substitutes < p b ph bh > cover 92% of the total number of substitutes for /p/ versus 84% in ours; < t d th > make up 85% of the substitutes for /t/ in Wiik's investigation versus 96% in ours and the substitutes < k g kh gh > for /k/ comprise 92% of the total in Wiik's and 96% in our study.

Except for the substitutes for dark l, Wiik reported only utterance initial (in practice word-initial) substitutes for / p t k r θ ð w / in his pre-publication study, and thus any further comparisons between the results cannot be made. Generally speaking, the two studies yielded amazingly similar results in spite of the fact that the subjects differed considerably from each other: Wiik's subjects were junior secondary school first formers (11-12 years old) with no previous knowledge of English (and hardly knowing any other language than Finnish), whereas our subjects were fifth formers (15-17 years old) studying Swedish and German at school. The slight differences between Wiik's and our results may be due to the different populations.. For instance, in addition to <f> (51%) Wiik's subjects gave among others <s> (28%), <t> (3%), <th> (2%) and <ts> (2%) as substitutes for /θ/, whereas <f> occurred in all the substitutes given by our subjects: <f> (96%), <pf> (2%) and <fh> (1%). This seems to imply that /f/ belonged to the phoneme inventory of our subjects and [f] being phonetically nearest equivalent to /θ/ they did not have to resort to any other substitutes, while Wiik's subjects were apparently not so familiar with the sound [f]; hence the other substitutes.

AN ATTEMPT TO ANSWER PROBLEM 2: WHICH ENGLISH CONSONANTS ARE DIFFICULT FOR FINNISH-SPEAKING PUPILS TO LEARN ?

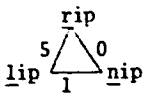
DISCRIMINATION AND IDENTIFICATION. – As the testees were to discriminate between different consonants (D-test) and to identify a certain consonant by comparing it with other consonants (SA-test and WA-test), it is more appropriate to speak in these tests of difficult consonant oppositions than of difficult consonants per se. We shall exemplify this standpoint with an extreme example. Let us suppose that researcher A has tested /p/ with such items as *pen - ten - pen, pit - pit - hit* and *mill - pill - pill*. The average correct answer percentage turns out to be 95% and researcher A concludes that /p/ is easy to discriminate. Researcher B has also tested /p/, but with items like *pan - pan - ban, weaver - weeper - weeper* and *rope - robe - rope*. He concludes that /p/ is fairly difficult to discriminate, because the average correct answer percentage was 58%.

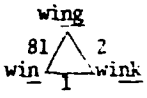
Why are the results so contradictory? The explanation is obvious: researcher A used decoys / t h m /, which are phonetically and acoustically quite dissimilar to /p/, while researcher B used phonetically and acoustically much closer decoys (/b/ and /v/). In addition, researcher A tested /p/ only word-initially, whereas B tested /p/ also word-medially and word-finally. Both researchers did not test /p/ per se, but /p/ in certain specific oppositions. It is clear from the above that we cannot answer problem 2 by making a list of difficult consonants per se, because the correct answer percentages for each consonant might be distorted due to varying oppositions. Therefore problem 2 must be restated as "Which English consonant oppositions are difficult for Finnish-speaking pupils to master?" Table 9 gives an answer to this problem. It is based on the correct answers of the 229 (in WA-test 228) learners of English. Table 9 shows (1) the tested oppositions, (2) the overall average correct answer percentages ($\bar{X}$%), (3) the number of times each opposition is tested and (4) the correct answer percentages testwise and itemwise. In items like *lip-rip-nip* where the tested consonants are different and thus more than one opposition is involved, the consonant that attracted least incorrect choices is inserted within brackets, because a careful error analysis reveals that in practice only one of the oppositions attracted the bulk of incorrect choices as shown below; The distribution of errors within three-member items is shown in form of triangles below. The figure given in the centre of the triangle indicates the number of those who have marked all the three consonants the same in SA-test and in WA-test. The error analysis shows that the mistakes (1) centre upon the opposition whose members are phonetically closest to each other and (2) are most frequent in items where all the three consonants are phonetically most closely related.

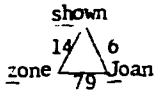
As can be seen from the table, the average correct answer percentages range from 100 to 18. The oppositions between the spirants (except /h/) and those between the affricates were the most difficult to discriminate and identify, while the oppositions between the consonant phonemes which occur in Finnish proved the easiest.

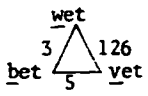
The spirants / f v θ ð / were extremely difficult for our subjects to distinguish from each other. Especially the oppositions /f/-/θ/ and /v/-/ð/ caused hearing problems. In addition, opposition 53, /θr/-/fr/-(/tr/), can be simplified to the opposition /f/-/θ/, because the error

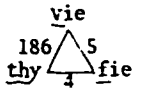
D-test (N=229)



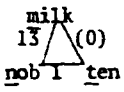


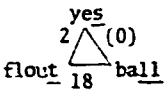


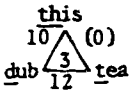


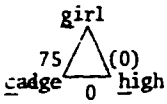


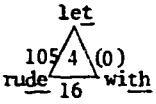
SA-test (N=229)

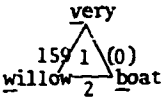


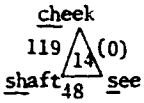


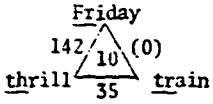


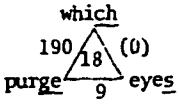




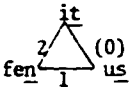


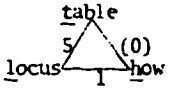


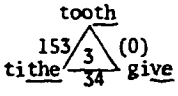


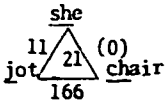


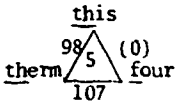
WA-test (N=228)

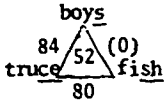












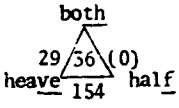


Table 9. Discrimination and identification (N=229) of English consonants.

rank number	oppos. tested	$\bar{X}$%	no. of times tested	correct answer percentages testwise	itemwise
1	h-m	100.0	1	D: 100	harsh-marsh-marsh
2	m-s	100.0	1	D: 100	mingle-mingle-single
3	n-t-(s)	98.0	1	WA: (98)	fen it (us)
4	b-w	96.0	1	D: 96	bill-will-bill
5	p-v	95.0	2	D: 97 93	weeper-weaver-weaver pallid-valid-pallid
6	s-k	94.0	1	D: 94	sing-sing-king
7	s-θ	93.0	2	D: 93 93	looser-Luther-Luther thick-sick-thick
8	v-r	93.0	1	D: 93	vain-rain-rain
9	ŋ-ŋk	92.0	1	D: 92	wink-wing-wing
10	no oppos.	92.0	1	D: 92	latches-latches-latches
11	l-w	92.0	2	D: 91 SA: 93	lean-wean-lean lumber wall long
12	j-Ø	91.5	2	D: 89 SA: 94	yeast-yeast-east yield-young easy
13	l-r-(n)	89.7	3	D: (97) 84 WA: 88	lip-rip-(nip) teller-terror-terror lax round learn
14	t-l-(s/h)	89.7	3	SA: (91) 81 WA: (97)	flout (yes) ball wail girl write locus table (how)
15	k-p	89.0	1	SA: 89	cot part count
16	no oppos.	89.0	1	D: 89	rum-rum-rum
17	s-t	87.0	1	SA: 87	sooth table summer
18	d-ð	86.0	2	WA: 90 82	fade bed with dote they desk
19	t-tʃ	85.0	1	D: 85	catty-catchy-catchy
20	dz-dʒ	83.5	2	D: 85 82	bards-bards-barge heads-hedge-heads
21	z-ʒ	83.0	1	D: 83	bays-bays-beige
22	tr-θr	82.0	2	D: 96 WA: 68	true-through-through thrush tree three
23	f-ð	81.5	2	D: 83 80	brief-breathe-breathe heifer-heather-heifer

rank number	oppos. tested	X̄	no. of times tested	correct answer percentages testwise	itemwise
24	h-h	80.0	1	SA: 80	hoist her home
25	s-ʃ	77.5	8	D: 94	shield-shield-sealed
				86	parcel-parcel-partial
				64	Paris-parish-parish
				SA: 86	mash miss dish
				79	ace house brush
				77	sear say shop
				WA: 84	sheer seven show
				50	sift shoe some
26	t-d-(ð)	76.9	11	D: 97	feed-feet-feed
				95	bleating-bleeding-bleating
				93	tub-dub-tub
				SA: (89)	dub (this) tea
				78	turf dark today
				(45)	rude let (with)
				41	helot cloud sit
				WA: 94	toil ten dark
				88	weird bread eat
				80	tilt door tall
				46	varlet read coat
27	s-l	76.0	1	WA: 76	omus face bell
28	ts-t	75.5	2	D: 79	pitch-pits-pitch
				WA: 72	perch hats watch
29	n-ŋ-(ŋk)	74.3	3	D: 74	singer-singer-sinner
				(63)	win-wing-wink
				WA: 86	fang in sing
30	n-m-(t)	74.2	5	D: 39	cunning-coming-coming
				SA: (93)	nob milk (ten)
				87	mole name man
				66	glean one room
				WA: 86	nil neck moon
31	w-r	73.5	2	D: 61	which-rich-rich
				WA: 86	rear run why
32	p-b	67.0	8	D: 82	ban-ban-pan
				79	lobe-lope-lobe
				76	pig-big-big
				46	staple-stable-staple
				SA: 55	booty pen bike
				55	pall book past
				WA: 82	bias put boy
				61	poke pen buy
33	v-b	66.0	1	D: 66	curve-curve-curb
34	n-n	66.0	1	WA: 66	nag know number
35	m-ŋ	64.3	3	D: 84	hanger-hammer-hanger
				63	ram-rang-ram
				WA: 46	deem strong home

rank number	oppos. tested	$\bar{X}$%	no. of times tested	correct answer percentages testwise		itemwise
36	k-g-(h)	64.0	11	D:	89	wick-wick-wig
					79	haggle-haggle-hackle
					78	cold-gold-cold
				SA:	(67)	cadge girl (high)
					54	hag break big
					48	guts good coffee
					39	hawk back dog
				WA:	71	gale come good
					70	wick bag back
					57	cane cat give
					52	fug bag work
37	f-f	61.3	3	SA:	87	faun film four
				WA:	49	reef enough wife
					48	fag phone five
38	ð-z	61.0	2	D:	91	seize-seize-seethe
					31	clothing-closing-clothing
39	f-v-(θ)	60.1	7	D:	89	surface-service-service
					87	fault-fault-vault
					72	strife-strive-strive
				SA:	56	vine four very
					20	thrive laugh eve
				WA:	93	foil very first
					(4)	heave (both) half
40	s-z-(ʃ)	59.8	8	D:	98	eyes-ice-eyes
					95	pace-pays-pace
					86	zip-zip-sip
					74	laser-lacer-laser
				SA:	51	dice plus boys
					41	hose days face
				WA:	28	booze ice days
					(5)	truce boys (fish)
41	z-dʒ-(ʃ)	56.0	1	D:	(56)	zone-(shown)-Joan
42	z-z	56.0	1	WA:	56	fuse blouse always
43	ʃ-tʃ-(s)	52.7	11	D:	83	lashes-latches-latches
					79	chair-share-chair
					51	cash-catch-catch
				SA:	78	chore child she
					66	leech wash much
					(21)	shaft cheek (see)
					16	leash fish teach
				WA:	71	chive shop cheek
					49	hutch British which
					35	sham child short
					31	trash dish much
44	dʒ-ʒ	52.0	1	D:	52	lesion-lesion-legion

rank number	oppos. tested	$\bar{X}$%	no. of times tested	correct answer percentages testwise	itemwise
45	v-w-(b)	51.0	7	D: 97	over-over-ower
				69	wary-vary-wary
				67	west-vest-west
				(40)	(bet)-wet-vet
				SA: (29)	willow very (boat)
				12	vigil wake very
				WA: 46	wail walk very
46	θ-ð-(v)	48.3	7	D: 92	ether-either-ether
				88	thy-thy-thigh
				63	teeth-teeth-teethe
				SA: 47	thee they thing
				15	soothe teeth with
				WA: (17)	tithe tooth (give)
				16	thine third there
47	ŋ-n	45.0	1	SA: 45	tang lying ring
48	tr-dr	40.0	1	D: 40	drain-train-drain
49	ʃ-ʒ	40.0	1	D: 40	mesher-mesher-measure
50	tʃ-dʒ-(z/ʃ)	32.8	11	D: 82	batch-badge-batch
				55	jaw-chore-jaw
				46	ledger-ledger-lecher
				SA: 32	gibe jump chair
				17	serge watch porridge
				6	chum June chalk
				(5)	purge which (eyes)
				WA: 44	parch Birch porridge
				36	gem church just
				25	badge page teach
				(13)	jot (she) chair
51	f-θ-(ʒ)	27.4	9	D: 19	thorn-thorn-faun
				14	deaf-death-death
				SA: 30	reef both knife
				28	heath mouth half
				9	thigh thing first
				WA: 58	thane think full
				50	serf teeth knife
				31	loath half mouth
				(8)	therm (this) four
52	v-ð-(f)	26.0	4	D: 19	clove-clothe-clothe
				14	than-van-than
				(14)	thy-vie-(fie)
				WA: 57	veil that very
53	θr-fr-(tr)	18.0	1	SA: (18)	thrill Friday (train)

	$\bar{X}$%	no. of times tested	D-test X%	SA-test X%	WA-test X%
all the 229 subjects	63.8	168	74.3	53.4	57.3
the 48 P-test subjects	63.8	168	73.0	54.7	57.9

analysis revealed that 75.5% of those who answered wrong marked /θr/ and /fr/ the same, 19% found /θr/ and /tr/ to be the same and the rest (5.5%) marked /θr/, /fr/ and /tr/ the same. It is interesting to notice that the fortis/lenis spirant oppositions (i.e. /f/-/v/ and /θ/-/ð/, 60.1% and 48.3%, respectively) were not confused with each other to the extent /f/-/θ/ (27.4%) and /v/-/ð/ (26%) were, although in all the four oppositions only one distinction keeps the members apart. When the number of distinctions is increased to two, as in opposition 23 (/f/-/ð/), the members of the opposition were much more easily distinguished from each other (81.5%). This is also reflected in oppositions 51 and 52 in such a way that /f/ and /ð/ were not mixed up. The error analysis showed that in opposition 52 (=*thy-vie-fie*) 94% of the subjects who answered wrongly marked /ð/ and /v/ the same, only 2.5% confused /v/ with /f/, even fewer (2%) mixed /f/ with /ð/ and the rest (1.5%) found all the tested consonants to be the same. In opposition 51 (=*therm-this four*) the wrong answer percentages were exceptionally evenly distributed: 51% for /θ/-/f/, 46.5% for /θ/-/ð/ and the rest (2.5%) made no distinction between all the tested consonants. These examples imply that the number of distinctions seems to play an important role in the discrimination and identification of consonants: the fewer the distinctions between the two consonants in opposition, the greater the probability of confusing them. We shall try to find out on pp. 70 ff. whether this conclusion holds true on a larger scale. On the whole, the spirants / ð θ f v / were mixed up only with each other. Of the other consonants merely /w/ and /b/ were confused with /v/, and /z/ with /ð/ to a notable degree, the average correct answer percentages being 51.4%, 66.0% and 61.0%, respectively. The identification of /f/ (no. 37 in Table 9) obviously needs to be commented on. In the SA-test /f/ has

been well identified (87%), whereas in the WA-test it has been considerably more difficult to identify (48% and 49%). This discrepancy has an obvious explanation: in the WA-test the orthography has probably misled the subjects: they heard *ʃag* from the tape and were asked to decide whether the words *phone* and *ʃive* began with the same sound as *ʃag*. It seems possible that more than 48% of the subjects identified the first sound in *ʃag* as /f/, but the orthography of *phone* misled them to choose only the alternative *ʃive*. Apparently the same applies to *teeʃ* (from the tape) - *enough* *witʃ* (on the answer sheet). Of the spirants, /h/ seemed to be easily distinguishable from the consonants with which it formed an opposition (see oppositions 1, 14, 24 and 36). In opposition 14 (=*locus* - *taste how*) only one of the six (3%) who answered wrongly confused /h/ with /ʃ/ and in opposition 36 (=*cadge* - *girl* *high*) all of the 33% who answered wrongly mixed /k/ with /g/; no-one chose the alternative /h/.

The affricates /tʃ/ and /dʒ/ are almost as difficult to keep apart from one another as the four spirants dealt with above. The average correct answer percentage was 32.8% (opposition 50). But the affricates are far less often confused with other consonants than with each other. This is clearly shown by items *purge* - *which* *eyes* and *jot* - *she* *chair*. In the former the vast majority of mistakes centered on the opposition /dʒ/ - /tʃ/ (88%), while in only 12% of the errors /z/ was marked as one or both of the affricates. In the latter the errors were distributed as follows: 51% marked the affricates as the same, and the remaining 16% marked /ʃ/ as one of the affricates. As can be seen from the two examples, the affricates are primarily confused with each other and in the second place with sibilants; the results seem to suggest that when sibilants are confused with affricates, it is most likely that lenis sibilants /z/ and /ʒ/ are confused with the lenis affricate /dʒ/ (e.g. oppositions 41 and 44) and fortis sibilants /s/ and /ʃ/ with the fortis affricate /tʃ/, e.g. opposition 43. Although sibilants and affricates are not confused with one another to the extent the affricates are, the average correct answer percentages (56%, 52% and 52.7% for oppositions 41, 44 and 43) are low enough to warrant attention. On the other hand, our subjects found it surprisingly easy to keep the clusters /ts/ and /dz/ apart from the affricates /tʃ/ and /dʒ/, respectively. The opposition /dz/-/dʒ/ was tested twice the average correct answer percentage was as high as 83.5; the opposition /ts/-/tʃ/ was also tested twice, the percentage being 75.5. Further, th

opposition /t/-/tʃ/ proved to be easy: the correct answer percentage was 85.

Oppositions 20, 21, 25 and 28 are of special interest as they seem to shed some light on the status of the affricates as perceived by Finnish pupils. The results would imply that our subjects tended the hear the affricates /tʃ/ and /dʒ/ rather as consonant clusters than as unit phonemes. If the affricate /dʒ/ is interpreted as a cluster /d/ + /ʒ/, then opposition 20 (/dz/-/dʒ/) would in fact be reduced to that of /z/-/ʒ/ (as in opposition 21), because /d/ is the common element in both /dz/ and /dʒ/. The almost identical correct answer percentages (82% and 85%, $\bar{X}$% = 83.5% for opposition 20 and 83% for opposition 21) seem to support this view. Similarly, if /tʃ/ is treated as a cluster /t/ + /ʃ/, opposition 28 (/ts/-/tʃ/) can be simplified to that of /s/ and /ʃ/, which is also tested in opposition 25. Again the average correct answer percentages are almost identical (77.5% in opposition 25 and 75.5% in opposition 28). Thus one would be inclined to draw the conclusion that the correspondence between the percentages of oppositions 20 and 21, and 28 and 25, respectively, are not due to mere chance but to the fact that they measure the same oppositions. However, there is no justification for making any far-reaching conclusions, as the number of items testing these oppositions is relatively small (3 for /z/-/ʒ/ and 10 for /s/-/ʃ/). Anyway, our results suggest that the status of the affricate deserves a more systematic empirical investigation than was possible in this study.

The purely sibilant oppositions seem to have been much easier to discriminate and identify than the affricate or spirant oppositions. In the discrimination and identification of the sibilants the number of distinctions again seems to play a crucial role: oppositions 49 (/ʃ/-/ʒ/) and 41 (/s/-/z/), where the members are distinguished from each other by the fortis/lenis distinctions alone, were far more difficult (40% and 59.8%) than oppositions 21 (/z/-/ʒ/) and 25 (/s/-/ʃ/) (83% and 77.5%, respectively), where there are three distinctions to keep the members apart. On the whole, sibilants were only confused with affricates or with each other. For instance, /s/ was well discriminated from /θ/ (93%), /k/ (94%), /t/ (87%) and /m/ (100%).

The majority of the plosive oppositions were of fortis/lenis type (i.e. /p/-/b/, /k/-/g/ and /t/-/d/) which on the basis of our contrastive analysis would appear to be more troublesome than other plosive oppositions

or oppositions where a plosive forms one of the two members. Our results seem to confirm this. As a rule, our subjects found the fortis/lenis plosive oppositions /k/-/g/ (X̄% = 6%), /p/-/b/ (67%) and /t/-/d/ (70.9%) to be more difficult to discriminate and identify than the only other plosive versus plosive opposition /k/-/p/ (89%) or the oppositions with a plosive as one member, e.g. oppositions 4, 5, 6, 14, 17, 18, 19, 22 and 33. It is interesting to notice that the opposition /t/-/d/ was more difficult in the cluster /tr/-/dr/ (opposition 48) than on the average.

Although the nasals are common to both Finnish and English, they turned out to be surprisingly difficult to discriminate and identify. In purely nasal oppositions X̄% ranged from 45% (/ŋ/-/n/) to 74.3% (/n/-/ŋ/). Nasals in opposition to other consonants did not cause any hearing problem (e.g. oppositions 1, 2, 3 and 13). In opposition 34 the written forms of the analogical words may have misled the subjects. The subjects were to decide whether *nay* began with the same sound as the analogical words *know* and *number*. The correct answer percentage is fairly low (66%) as compared with the 100% identification of /n/ in *neither* in the substitution test. Also the average correct answer percentages for oppositions 47 and 35 were surprisingly low (45% and 64.3%). This may be explained by the likelihood of the discrimination and identification of the nasals being affected by their position in the word: word-initial and word-medial nasals were easier than word-final nasals. For instance in opposition 47 (45%) /ŋ/ was word-final (*tang - lying - ring*). The same goes for /m/ and /ŋ/ in *deem - seeing - seem* and *ram - sang - ram*. The correct answer percentages were lower (46% and 63%) than that of *hanger - hammer - hanger* (94%), in which the opposition /m/-/ŋ/ occurs word-medially. This tendency was also noticed in the substitution test. There are, however, some exceptions to the rule. For example, *cunning - coming - coming* proved to be by far the most difficult item (39%) testing opposition 30 (/n/-/m/). But it was easier to keep /n/ and /m/ apart word-initially (*nob milk ten* 93%, *mole name man* 87%, *mid nack moon* 86%) than word-finally (*glean one room* 66%). The same tendency seems to present in the discrimination and identification of other consonants, too. The items testing the oppposition /f/-/v/-(/θ/) (no. 39) may serve as examples. The three word-initial items have the following percentages: 87%, 56% and 93%; the word-final items show considerably lower percentages: 72%, 20% and 4%. This will be systematically studied on pp. 70 ff.

The discrimination and identification of the phonemes /l/ and /r/ seems to be the least problematic. The oppositions /v/-/r/, /l/-/w/, /l/-/r/-/(n)/ and /t/-/l/-/(s/h)/ were all easy, and also the remaining two c sitions (/l/-/s/ and /w/-/r/) proved to be fairly easy as shown by the percentages 93%, 92%, 89.7%, 89.7%, 76% and 73.5%, respectively.

The opposition /j/-/Ø/ was included in our tests as Hirvonen (1972: 24) had found it to be problematic for upper secondary school pupils and thus included it in his trial version. In our discrimination and sound analogy tests the subjects (although junior secondary school pupils) found this opposition easy (91.5% on the average).

The average correct answer percentages for each test seem to suggest that the process of identification really requires more of the learner than mere discrimination does. The discrimination test has the highest mean percentage (74.3%), which is clearly higher than those of the sound analogy (53.4%) and written analogy (57.3%) tests, i.e. tests which we supposed to measure identification. The discrimination test contains a greater number of easy consonant oppositions (nos. 1-19) than the other tests. Therefore we may conclude that the difference in the correct answer percentages in favour of the discrimination test is due to this. To find out whether this was so we computed the average correct answer percentages for the oppositions common to all the three tests (i.e. for oppositions 25, 26, 30, 32, 36, 39, 40, 43, 45, 46, 50 and 51). The average correct answer percentages for the 12 oppositions in common were as follows: 69.7% in the D-test, 45.9% in the SA-test and 50.4% in the WA-test. The percentages show clearly that the difference remained essentially the same. Therefore it can be safely concluded that the process of discriminating consonants is easier than the process of identifying them.

PRODUCTION. – Unlike the listening tests, which were objective tests in the sense that the test scores were independent of the marker, the production test was subjective, because the testees' scores were dependent on what the transcriber heard them utter. Therefore more than one transcriber was needed. Table 10 below shows how severe and unanimous the different transcribers were in their interpretations of the subjects' productions. As the five teachers transcribed only their own pupils' productions, we shall treat them as if they were only one transcriber. The

Table 10. Means, standard deviations and intercorrelations of different transcriptions (N=48).

transcribed by	$\bar{X}$	s	Teachers	JC	RM	EV	RP
Teachers	87.98	9.55	1.00	.74	.69	.69	.53
JC	86.08	8.87		1.00	.71	.74	.64
RM	84.29	10.15			1.00	.90	.78
EV	76.98	11.88				1.00	.81
RP	72.98	10.50					1.00

acceptance level of the transcribers is shown by the mean of the subjects' total scores (the maximum score is here 103, as the affricates were treated as clusters at this stage; the transcriptions were scored as follows: the correct phoneme alone was given the value 1, all the other transcriptions were marked wrong (=0)). The intercorrelations of the subjects' total scores arrived at on the basis of the different sets of transcriptions reflect how unanimous the transcribers were. JC and RP are native speakers of English. RM and EV (the writers of this report) and the teachers are all native speakers of Finnish.

As to the level of acceptance, the means show that the transcribers fall into roughly two groups: (1) those whose means are far above 80 and (2) those whose means are clearly below 80. Group one comprises the teachers, JC (a university lecturer) and RM (one of the authors). Practically speaking, they have been equally severe: the teachers have been the least severe but JC's and RM's means are only slightly lower. RP, a trained phonetician, and EV, one of the writers, have been equally strict but markedly stricter than the transcribers in group one. This grouping is somewhat unexpected: one would have expected the native speakers of English, RP and JC, to form one group and the native speakers of Finnish, the teachers, RM and EV, the other. However, the native speakers of English diverged greatly in their level of acceptance. Nor did the native speakers of Finnish keep the same standard. This seems to suggest that the assessment of pronunciation is to a great extent subjective and independent of the transcriber's mother tongue. That RP was the most severe of the transcribers might be explained by the fact that he is a trained phonetician with many years' experience of assessing pronunciation. The highest mean, that of the teachers',

may be due to the fact that the teachers are accustomed to their pupils' pronunciation and thus some mistakes perhaps remained unnoticed.

All the transcribers' judgements seem to correlate positively with each other, but the intercorrelations between the transcriptions show great variation: they range from .53 to .90. The highest intercorrelation is between the transcribers RM and EV. This means that with an 81% precision RM and EV have managed to place the subjects in the same order of superiority. The lowest intercorrelation (between RP and the teachers) tells us that only a 28.09% agreement was reached on the order of the subjects. The fairly low correlation (.64=40.96% agreement) between the native speakers of English seems again to support the conclusion that the assessment of pronunciation is independent of the transcribers' mother tongue.

On the basis of the intercorrelations we can perhaps divide the transcribers into two groups: RP, EV and RM seem to form one group and JC and the teachers the other. The only difference in this grouping from that based on the means is that RM shifts his group: his level of acceptance was nearer to that of JC's and the teachers', while he is more in agreement with EV and RP on the order of the subjects. This group has the highest intercorrelations, RM - EV .90 (=81% agreement), RP - EV .81 (65.6% agreement) and RP - RM .78 (60.84% agreement), which might be explained by the fact that RM and EV have been RP's pupils. On the whole the intercorrelations point to the fact that it is extremely difficult to judge pronunciation consistently. However, in other studies, too, one has had to be content with intercorrelations of the same magnitude between different evaluators of pronunciation. For instance, Hirvonen (1974: 19, 93) seems to be quite happy with the average intercorrelations between his evaluators of the pronunciation test, although the intercorrelations are on the average about the same as in the present study. In Hirvonen's study the pupils' own teachers correlated .72 (=51.84% agreement) with the native speaker of English and .77 (=59.29% agreement) with the Finnish-speaking evaluator of the Matriculation Board and the correlation between the last two evaluators was .82 (=67.24% agreement).

It must be borne in mind, however, that the above intercorrelations (ours as well as Hirvonen's) strictly speaking tell us only how well different evaluators have been able to place the subjects in the same order of superiority. They do *not* indicate how unanimous the transcribers have

been about the mistakes made by the pupils. Let us illustrate this with a concrete example. EV and RP both found subject no. 171 to have made the same number of mistakes (20). Thus both EV and RP are in complete agreement on the total score achieved by the pupil. A further analysis reveals, however, that EV and RP disagree considerably as to which items the mistakes occurred in. They found a mistake in the same 12 items and in 9 cases they agreed on what the mistake was, while in 3 cases they disagreed: when EV heard the subject utter /θθ/,/p/ and /s/ instead of the correct phonemes /ð/, /b/ and /z/, respectively, RP heard /θ/, /d/ and an "in-between" phoneme /s~z/. A more noteworthy fact is, however, that EV marked 8 items wrong which RP accepted, and RP marked another 8 items wrong which EV found correct (e.g. in the test word *these* RP heard the subject say [bi:z], while EV heard [ði:s]. This shows clearly that it is not enough to compute the inter-marker correlations based on the subjects' total scores alone, as such correlations do not demonstrate the inter-marker agreement by items, inly subject by subject.

The statement above applies to the present study in particular, as we are interested in timing out which English consonant phonemes Finnish pupils find difficult to pronounce. Therefore we considered it appropriate to compute another inter-marker correlation, this time based on the number of correct answers in each item. The resulting correlation coefficient indicates the amount of agreement between the different markers on which consonants the subjects dound difficult/easy to produce. As this correlation could not be calculated by computer, it was computed between JC and RP alone; being native speakers of English they were the most relevant evaluators according to the foreign language teaching objectives in Finland (see Nykykielet 1971: 11, 29). The item correlation between RP and JC was .77 (=59.29% agreement). Thus RP and JC reached a considerably higher degree of unanimity about the difficulty of the English consonants than about the subjects' total scores (.64 = 40.96% agreement). The 59.29% agreement we felt to be sufficiently high and thus RP's and JC's transcriptions were used as the basis for the linguistic analysis of the production test data.

The answer to problem 2 is to be found in Table 11, where the tested consonants are presented in order of difficulty, beginning with the easiest, according to the average correct answer percentages for each consonant obtained from the conjoined transcriptions of JC and RP. For comparison the average correct answer percentages for each consonant in order of diffi-

Table 11. Production of English consonants (N=48)

the tested cons.	no. of times tested	$\bar{X}$%	RP	$\bar{X}$%	JC	$\bar{X}$%
1. /j	2	99.5	1. /j	100.0	1. /h	100.0
2. h	1	98.0	2. r	96.7	2. j	99.0
3. r	3	97.3	3. h	96.0	k	99.0
4. k	2	96.5	4. k	95.0	4. r	98.0
5. n	3	95.7	5. n	93.9	ŋ	98.0
6. l	4	93.6	6. l	89.8	6. n	97.4
7. ŋ	3	92.5	7. ŋ	86.9	7. l	97.0
8. b	7	87.3	8. b	85.6	8. f	93.4
9. ʃ	5	86.3	9. w	83.5	p	93.4
10. m	2	85.0	10. m	82.5	10. t	92.8
11. f	3	84.0	11. ʃ	80.6	11. ʃ	91.8
w	5	84.0	12. g	78.0	12. b	91.7
13. t	6	83.4	13. t	76.8	13. tʃ	90.1
14. p	4	80.1	14. f	74.4	14. m	87.5
15. g	3	80.0	15. d	73.2	15. v	85.4
16. tʃ	5	78.5	16. p	71.9	16. w	84.7
17. v	3	77.0	17. v	68.2	17. s	83.1
18. s	6	73.5	18. tʃ	66.3	18. g	82.0
19. d	7	72.7	19. s	63.8	19. ð	76.5
20. ð	3	63.3	20. ð	50.2	20. z	73.0
21. z	6	59.7	21. z	46.7	21. d	71.5
22. θ	3	54.3	22. θ	44.5	22. θ	63.9
23. ʒ	2	42.8	23. dʒ	34.2	23. ʒ	52.3
24. dʒ /	5	41.5	24. ʒ /	33.3	24. dʒ /	48.7
	93	$\bar{X}$%=77.9		$\bar{X}$%=75.4		$\bar{X}$%=84.8

culty are also reported separately for RP and JC. To allow comparison with the results of the listening tests the affricates are here treated as unit phonemes and thus the number of items is 93.

In accordance with the results of the listening tests the consonants occurring in both English and Finnish seem as a rule to be the

easiest to produce: seven of them top the list. At this point JC and RP agree admirably: they found the same seven consonants to be the easiest, only in a slightly different order. It is mainly the different level of acceptance alone (RP being more severe) that is reflected in the differing correct answer percentages.

RP and JC also reached considerable agreement on which of the consonants are the most difficult: both transcribers found / ʒ z θ ʃ dʒ / to be among the six most difficult consonants and in spite of the startling differences in the correct answer percentages (due to divergent levels of acceptance) they also placed them nearly in the same order, the only striking exception being the placing of /d/. RP noted it to be far easier in relation to the other consonants than JC did. It is worth noticing that these five consonants do not belong to the phoneme inventory of Finnish.

Thus our results seem to follow the lines suggested by our contrastive analysis: the subjects managed to produce well the consonant phonemes which occur in Finnish and they had difficulty in producing the consonant phonemes which do not exist in Finnish.

There is considerably more inter-marker fluctuation in the order of the consonants in the middle group (nos. 8-19) than in the top seven or the bottom five. The greatest variation is in the order of /w/, /g/, /f/, /d/ and /p/, their order being in JC's and RP's transcriptions as follows:

cons.	RP	JC	difference	RP $\bar{X}$%	JC $\bar{X}$%
/p	16.	8.	8	71.9	93.4
w	9.	16.	7	83.5	84.7
g	12.	18.	6	78.0	82.0
f	14.	8.	6	74.4	93.4
d/	15.	21.	6	73.2	71.5

The inter-marker differences can in our opinion be due to

(1) systematically different treatment of some consonants by RP and JC,

(2) JC's and RP's different levels of acceptance, and

(3) chance.

(1) *Systematic difference.* The difference in the evaluation of /p/ is mainly due to the fact that RP has obviously paid attention to aspiration, whereas JC seems to have primarily listened for voicing alone. Word-initial /p/ (which is strongly aspirated in English) has been heard by RP as /p/ 25, as /b/ 17, as /p~b/ 3 times and as miscellaneous 3 times

in *porridge* , and in *page* as /p/ 31 times, as /b/ 14 times, as /p~b/ twice and as non-recognizable phoneme once. Apparently the high proportion of /b/ speaks for the interpretation that quite a number of the subjects pronounced their word-initial /p/ without aspiration and therefore RP interpreted their /p/ as /b/. JC, on the other hand, must have paid more attention to voicing, because he has interpreted /p/ in *porridge* as /p/ 46 times and as /b/ only twice, and in *page* as /p/ 44 times, as /p~b/ 3 times and as /b/ only once. RP and JC do not differ much in their interpretation of word-final /p/ (not aspirated in English in this position): in *shop* and *zip* /p/ has been transcribed as /p/ 82 (out of 96) times by RP versus 89 times by JC. Thus the different placing of /p/ is for the most part due to the divergent interpretation of word-initial /p/. The same trend is noticed in RP's and JC's transcriptions of word-initial /t/: out of 96 cases, RP heard /t/ 40 times and /d/ 46 times, whereas JC heard /t/ 94 times and /d/ 0 times. Oddly enough, RP and JC transcribed word-initial /k/ similarly: 46 times as /k/ and twice as /g/ by RP and 47 times as /k/ and once as /g/ by JC.

The great difference in the order of /d/ between the two transcribers turned out to be due to their different treatment of word-final /d/: RP transcribed it nearly always either as /d/ (85 times out of the 144 possible) or as /t/ (53), whereas JC in addition to /d/ (69) and /t/ (31) marked a large number of cases as /t~d/ (43) versus only 2 in RP's transcription. Thus JC was notably uncertain whether /d/ or /t/ was pronounced in a number of cases. Such "in-between" phonemes as /d~t/ were scored wrong, because they leave the listener in doubt. For example, the listener may wonder whether a *dent* or *tent* is meant by "He's got a *dent* ~ *tent* in his car". JC's frequent use of /d~t/ has thus lowered his average correct answer percentage below that of RP's. JC's tendency to mark "in-between" phonemes seem to concern word-final consonants in particular. JC seems to have paid attention to voicing only and he had difficulty in deciding whether the subjects pronounced the consonants in question with enough voicing for them to be regarded as lenis consonants. RP, on the other hand, also seems to have taken the length of the preceding vowel into account, and thus if a subject uttered the lenis consonant devoiced and the preceding vowel long, RP presumably marked a lenis consonant; if, on the other hand, a subject pronounced the consonant devoiced but the preceding vowel short, RP transcribed a fortis consonant. The following

examples are cases in point:

	transcribed as	by RP	by JC	
(1) word-final /g/ in *dog*	/g/	23	26	times
	/k/	19	3	
	/k~g/	4	18	
	others	2	1	
(2) word-final /dʒ/ in *porridge*	/dʒ/	9	16	
	/tʃ/	26	9	
	/dʒ~tʃ/	0	13	
	/ts/	8	0	
	/dz~ts/	0	2	
	others	5	8	
(3) word-final /dʒ/ in *page*	/dʒ/	14	19	
	/tʃ/	27	12	
	/dʒ~tʃ/	0	12	
	/ts/	5	0	
	others	2	5	

The figures underlined show that in all of our examples RP has identified the majority of mistakes as clear fortis consonants (or the cluster /ts/), while JC has been in doubt about the voicing of the consonants in question and marked "in-between" phonemes. Thus the differences between RP and JC in the order of the above-mentioned consonants are for the most part due to a systematic difference in their treatment by the two transcribers.

(2) *Difference in the level of acceptance.* In other cases the difference between the correct answer percentages of RP and JC seem to result from a different level of acceptance alone (e.g. /f/ 74.4% versus 93.4%). This also applies to the correct answer percentages of the top seven and bottom five. As can be seen from Table 11, RP was stricter in his judgements than JC throughout the test, /j/ and /d/ being the only exceptions. Therefore it is not surprising that the difference in RP's and JC's means of the average correct answer percentages (75.4-84.8=-9.4%) is statisti-

cally highly significant (t=9.24, p<.001, df=92). This means that with 99.9% certainty the difference between RP's and JC's levels of acceptance is real, not caused by chance.

(3) *Chance.* It becomes evident from Table 11 that the order of the consonants in JC's list is on the whole determined by very minute differences, whereas in RP's list the "steps" between the consonants are longer. Thus JC's order of the consonants is statistically more susceptible to chance variation than RP's order. Let us take an example. In both RP's and JC's lists we find /ʃ/ in 11th place. If we suppose that one subject more had answered right/wrong every time /ʃ/ was tested, it would have meant a 2% increase/decrease in the percentage of /ʃ/ in both lists. In RP's list the 2% increase would have raised /ʃ/ one step higher (no. 10) and the corresponding decrease would not have affected its place in the list at all, while in JC's list the same 2% increase or decrease would have raised /ʃ/ three steps higher (to 8) or lowered it two steps (to 13). Thus one should not pay too much attention to minor differences in the order of the consonants in RP's and JC's lists: they may be real, but they may equally well be due to chance.

The status of the affricates /tʃ/ and /dʒ/ was also studied on the basis of the production test, because the results of the listening tests implied that some of our subjects tended to hear the affricates as consonant clusters. Thus we interpreted the affricates also as clusters of plosives and sibilants (i.e. as /t/ + /ʃ/ and /d/ + /ʒ/) and studied in which part of the cluster, in the plosive part /t/ or /d/ or in the sibilant part /ʃ/ or /ʒ/, the mistakes were mainly made. The distribution of mistakes is shown below:

	/t/ + /ʃ/ mistakes		/d/ + /ʒ/ mistakes	
RP	38	78	147	156
JC	6	21	107	117
total	44	99	254	273

In the case of /t/ + /ʃ/, considerably more mistakes were made in the sibilant part (99) than in the plosive part (44) of the cluster, while in the case of /d/ + /ʒ/ the distribution is almost even, although the same tendency is discerned: more mistakes were made in the sibilant part

(273) than in the plosive part (254). The results seem to support our earlier statement that Finns tend to hear the affricates /tʃ/ and /dʒ/ as clusters. This is interesting from the point of view of teaching English to Finns. Obviously /ʃ/ and /ʒ/ should be taught before /tʃ/ and /dʒ/, because teachers need not teach the affricates as new sounds, but as sequences of the familiar phonemes /t/ and /ʃ/ and /d/ and /ʒ/. Thus the learning of the affricates would be parallel to the learning of consonant clusters as is also claimed by Wiik (1965b).

It is interesting to compare our production test results with those of the D-test, SA-test, and WA-test. There seems to be considerable correspondence between the results. As a rule, the consonants which also occur in Finnish have been found easy to discriminate, identify and produce, whereas the consonants occurring only in English have been the most difficult. Some consonants seem to constitute a hearing problem primarily, some also a pronunciation problem. For instance, /f/ causes serious discrimination and identification problems when in opposition to /θ/ ($\bar{X}$% = 27.4), whereas our subjects have been fairly successful in producing /f/ (84%). The same seems to be true of /ʃ/ when in opposition either to /ʒ/ (40%) or /tʃ/ (52.7%), while it has been easy to produce /ʃ/ (86.3%). /θ/, /ʒ/, /dʒ/, /z/ and /ð/ constitute both hearing problems (especially when in opposition to /f/, /ʃ/, /tʃ/, /s/ and /v/, respectively) and pronunciation problems. The comparison between the percentages of the listening tests with those of the production test is complicated by the difference in their means of average correct answer percentages. The production test was much easier (77.9%) than the listening test battery (63.8%). The result is contrary to the general conception that pupils cannot be expected to pronounce the sounds of the target language correctly (especially such sounds as are phonetically close to each other) unless they are first able to hear them and to distinguish them from one another. This view is held for instance by the Finnish comprehensive school curriculum planning committee (POPS 1973: 14) and by Stratton (1970: vii). This unecpected result may simply be due to the following technical differences between the tests:

(1) In the production test the subjects heard the stimulus twice, whereas in the listening tests they heard the test words only once.

(2) All the stimulus words (except nos. 11, 20, 28 and 40, see Appendix 5) were familiar to the subjects, while all of the stimuli in the sound and written analogy tests were unfamiliar. In the discrimination test familiarity with the members of the triplets was not controlled; it contained a random number of triplets in which all members were unfamiliar (e.g. *thy - vie - fie*), one member was familiar (e.g. *teeth - teeth - teethe*) or all were familiar (e.g. *eyes - ice - eyes*).

(3) There were no distractors to mislead the subjects in the production test: they were asked simply to reproduce the word which they heard; in the listening tests the triplets or the analogical words contained distractors. The situation would have been more equal, if the subjects had been asked to produce for instance the different word in a triplet (e.g. *badge* in *batch - badge - batch*).

These technical differences alone may explain the subjects' better success in the production test. But Brière, too, has arrived at a similar result. He found that "production of sounds in isolation always preceded perception of sounds within the T system. Although this was especially noticeable in the case of perceptual confusion pairs, production in isolation preceded perception within the system for all sounds" (Brière 1966: 794). He found his result as unexpected as we do ours. He concludes that additional experimentation is needed to determine "the role of production as a possible mediator to perception" (Brière 1966: 795). We quite agree with him. But better success in production may not after all be as contradictory as it seems: it may well be that a learner is able to discriminate and identify foreign language sounds in the speech of others with ease only when he has learnt to make the appropriate distinctions in his own speech. The difference between hearing and production is perhaps analogous to the difference between theory and practice: a deeper understanding of theory grows from practice. Thus hearing distinctions in the speech of others remains "theory" until they are put into "practice" in the sense that the learner produces them himself. When he can control his own speech, he is better "equipped" to make the appropriate distinctions also in the speech of others.

It is not enough for a teacher to know that a mistake has been made; he must also know what the mistake was. Therefore, in addition to the cor-

rect answer percentages for each consonant, we shall report the major categories of incorrect responses in Table 12. The conjoined data of RP and JC is used. The column "wrong %" gives the proportion of incorrect answers. The symbol ∅ means that the transcribers have not heard any phoneme at all and the symbol ? indicates that they have not heard any recognizable English phoneme.

In general, the major categories of incorrect answers to each consonant conform to the results of the substitution test (see Tables 2-9): the nearest possible Finnish or English equivalent phonetically and acoustically was produced instead of the correct phoneme. The incorrect productions of / f w g d θ / may serve as examples (see Table 12). In some cases, other substitutes than the most probable (the nearest) were also given to a notable degree. Such substitutes are almost invariably due to the word-final position of the tested consonant. For instance, ?, ∅ and /nt/ instead of /n/ and ∅, /u/ and ? instead of /l/ are given word-finally. So is also ∅ instead of /ŋ/. In the case of /b/, too, the phonetically more unlikely mistakes ∅ and ʋ have nearly all been made in the word-final /b/ in *cab*: out of the 41 cases of ∅ and /v/ instead of /b/ 35 occurred word-finally. /b/ in *cab* proved problematic also in the substitution test, where the nasals /n/ and /m/ were the major substitutes (see above p. 33). On the whole those consonants that proved difficult to produce have been given a large number of different erroneous productions. For instance, /dʒ/ (33), /z/ (27) and /ð/ (22) are cases in point.

Table 12. Major categories of mistakes in the production test (N=48).

cons.	wrong %	distribution of mistakes in %							no. of different mistakes
/j	0.5	dj 100							1
h	2.0	Ø 100							1
r	2.7	br 63							4
k	3.5	g 50							3
n	4.3	? 25	Ø 17	nt 17	d 17				7
l	6.4	Ø 24	u 24	? 12					13
ŋ	7.5	Ø 41	n 27						7
b	12.7	Ø 26	v 25	p~b 14	p 12				14
ʃ	13.7	s 52	s~ʃ 31						11
m	15.0	n 69	m~n 17						6
f	16.0	v 36	v~f 18						13
w	16.0	v~w 39	v 26						14
t	16.6	d 57	t~d 10	Ø 10	k 10				14
p	19.9	b 63	p~b 21						8
g	20.0	k~g 47	k 45						6
tʃ	21.5	ts 35	dʒ 19	ts~tʃ 17	dʒ~tʃ 10				19
v	23.0	w 51	f 24	v~w 9	v~f 9				9
s	26.5	s~θ 19	ʃ 19	θ 18	s~ʃ 11	z 8	f 6	s~z 6	16
d	27.3	t 61	t~d 29						10
ð	36.7	v 20	θ 18	d 10	v~ð 9	θ~ð 8	f 8		22
z	40.3	s 32	θ 13	s~z 12	ʒ 7	ʃ 6	ð 5	ʒ~ʃ 4	27
θ	45.7	f 48	θ~f 33	tθ 5	ð 4				15
ʒ	57.2	z 44	ʒ~ʃ 12	s 11	ʃ 10				14
dʒ	58.5	tʃ 49	dʒ~tʃ 23	ts 7	dz 5				33

AN ATTEMPT TO ANSWER PROBLEM 3:

ARE THE AREAS OF DIFFICULTY PREDICTABLE ON THE BASIS OF CONTRASTIVE ANALYSIS ?

It would certainly be very helpful to teachers planning their teaching strategy if a contrastive analysis of the structures of the native language and the target language could reveal potential areas of difficulty in learning the target language. Appropriate material could be provided and appropriate methods could be used to overcome the most likely difficulties as soon as possible. Therefore we wanted to study whether it is possible on the basis of our contrastive analysis (see pp. 14 ff.) to point out the areas of difficulty in learning English consonants. Our contrastive analysis gave rise to five assumptions (see p. 22 above). If these assumptions could be verified empirically, i.e. if our test results confirmed them, then the contrastive analysis would fulfill the above aim: it would have enabled us to predict learning difficulties.

To test *Assumption 1* (it is more difficult for Finns to identify and produce such English consonants as do not occur in Finnish than those occurring in both languages), we divided the consonant oppositions in the D-test, SA-test and WA-test into three groups: (A) both members of the opposition are common to both Finnish and English (e.g. /k/-/p/), (B) one member of the opposition occurs in Finnish, the other only in English (e.g. /v/- /ð/) and (C) both members of the opposition occur only in English (e.g. /θ/-/ð/). Then we computed the average correct answer percentages for these groups of oppositions. In the case of the production test we could simply divide the test consonants into (A)those occurring in both languages and (C) those occurring in English alone. The average correct answer percentages were similarly computed. The results are presented in Table 13 below. The figures after the percentages indicate the number of items testing the opposition or consonant group in question.

The results seem to verify our assumption. In all tests the mean percentages are the highest in group A. They are notably higher than those in group C, the greatest difference being in the SA-test (44.4%) and the smallest in the P-test (13.3%). Although the values of t were not computed, the differences appear to be too high to be caused by mere chance. With reservations it may thus be concluded that it is more difficult for Finns to identify and produce English consonants which do not occur in Finnish than those that occur in both Finnish and English. Even the occurrence of only

Table 13. The average correct answer percentages of the consonant groups A, B and C.

test	whole test		group A		group B		group C	
	$\bar{X}$%	no. of items	$\bar{X}$%	no. of items	X%	no. of items	$\bar{X}$%	no. of items
D-test	74.3	75	86.1	20	72.7	32	67.1	23
SA-test	53.4	45	76.1	11	53.9	18	31.7	16
WA-test	57.3	48	79.3	14	58.0	16	45.3	18
listening test battery	63.8	168	81.2	45	64.3	66	49.6	57
P-test	77.9	93	84.6	46	-	-	71.3	47

One consonant, non-existent in Finnish, in an opposition (group B) seems to be enough to cause identification problems for Finns, as a comparison between the percentages in groups A and B shows.

To find out whether *Assumption 2* (the fewer the distinctions between any two English consonant phonemes, the more difficult it is for Finns to keep them apart both in identification and pronunciation) was true we divided the consonant oppositions in the listening tests into (1) those with 1 distinction, e.g. /v/-/ð/, (2) those with 2 distinctions, e.g. /b/-/w/, (3) those with 3 distinctions, e.g. /θ/-/s/ and (4) and those with 4 or more distinctions, e.g. /ʃ/-/tʃ/. We computed the average correct answer percentages for these groups of oppositions. In items like *dub - this tea* the number of distinctions is the same as the smallest number of distinctions between the three consonants in the item. This practice could be adopted, because the subjects usually confused the two nearest consonants in the item with each other. Thus the item *dub - this tea* was categorized as an opposition with 1 distinction (/d/-/t/). The results are presented in Table 14. The number of items testing the distinction in question is placed in brackets after the corresponding percentage.

On the whole, the higher the number of distinctions, the higher the correct answer percentage seems to be. This is in accordance with our assumption, but no definite conclusions can be drawn, because the differences between the adjacent groups are not particularly great. A closer look at the table reveals the following details:

Table 14. Average correct answer percentages of oppositions with 1, 2, 3 and 4 or more distinctions.

test	1 dist. $\bar{X}$%	2 dist. $\bar{X}$%	3 dist. $\bar{X}$%	4+ dist. $\bar{X}$%	whole test $\bar{X}$%
D-test	66.2 (36)	74.8 (9)	81.5 (16)	80.5 (12)	74.3 (73)
SA-test	40.3 (28)	23.3 (2)	80.7 (3)	86.8 (9)	53.4 (42)
WA-test	44.6 (30)	46.0 (1)	76.0 (7)	81.5 (6)	57.3 (44)
listening test battery	53.2 (94)	59.5 (12)	79.9 (26)	82.8 (27)	63.8 (159)

(1) In the D-test the differences between any of the groups are fairly small; even the greatest difference, that between 1 dist. and 3 dist., is only 15.3%. This may reflect the fact that the process of discrimination is so easy that subtle differences in the distinctions do not much affect the results: only the correct answer percentage of oppositions with 1 distinction remains below that of the whole test.

(2) In the SA-test and the WA-test the line of demarcation seems to go between the groups 2 dist. and 3 dist.:the oppositions with 1 distinction and 2 distinctions seem to be of roughly equal difficulty (clearly below the mean percentages of the tests) and the oppositions with 3 and 4 or more distinctions again have approximately the same average correct answer percentages (clearly above the mean percentages of the tests). This seems to imply that, in the process of identification, oppositions with 1 or 2 distinctions are difficult, while the leap from 2 to 3 distinctions is enough to make the opposition considerably easier. It is interesting to notice that in all the tests the average percentages are about the same in the groups 3 dist. and 4+ dist. as is also shown by the average correct answer percentages of the listening test battery (79.9% and 82.8%, respectively). Thus the difference in the average level of difficulty between the discrimination test and the sound/written analogy tests seems to result from the differences in the groups 1 dist. and 2 dist. alone. The percentages 66.2 and 74.8 in the D-test as against 40.3 and 23.3 in the SA-test and 44.6 and 46.0 in the WA-test seem to confirm, but also particularize, our statement that the process of discrimination is easier than the process of

identification: only when the consonants in opposition are phonetically close to each other (= distinguished from one another by 1 or 2 distinctions) is it more difficult to identify than to discriminate them.

This kind of "distinction analysis" could not be applied to the production test, because it tested the consonants as such, not in opposition to other consonants. The major categories of mistakes in the production test (see Table 12) seem, however, to suggest that distinctions play an important role in production in the sense that most frequently the nearest possible incorrect consonant is produced instead of the correct one.

Assumption 3 (it is difficult for Finns to identify and pronounce those English consonant phonemes that are distinguished from each other solely by the fortis/lenis opposition)was tested in the following way: (1) In the listening test battery the average correct answer percentage was computed separately for the fortis/lenis oppositions and for the remaining oppositions. As the fortis/lenis oppositions are special cases of oppositions with 1 distinction, we also computed the average correct answer percentage for oppositions with 1 distinction other than fortis/lenis. (2) In the production test the correct answer percentages were computed for the fortis consonants / p t k f θ s ʃ tʃ / and for their lenis counterparts / b d g v ð z ʒ dʒ / and for the rest of the consonants. The results are shown in Table 15. The number of items in each group is given after the corresponding percentage in the table.

(1) The average correct answer percentage of the fortis/lenis consonant oppositions is 9.6% lower than that of the rest of the oppositions and 5.9% lower than that of the whole battery. In this respect our assumption gains some support. It is interesting, however, to notice that other oppositions with one distinction have proved even more difficult than the fortis/lenis oppositions. This category comprised the oppositions /l/-/r/, /l/-/n/, /ʃ/-/tʃ/, /ʒ/-/dʒ/, /f/-/θ/ and /v/-/ð/. Among these, the last two in particular contributed to the low mean percentage. This result has an important implication for the teaching of English: special care should not only be taken to teach pupils to distinguish fortis consonants from their lenis counterparts as is frequently done (see e.g. POPS 1973: 20) but also to teach pupils to make a distinction between all consonants which form oppositions with one distinction alone (i.e. those in the oppositions /ʃ/-/tʃ/, /ʒ/-/dʒ/, and /f/-/θ/ and /v/-/ð/ in particular).

Table 15. Average correct answer percentages for fortis/lenis consonants and oppositions.

(1) the listening test battery			(2) the production test		
	$\bar{X}$%	no. of items		$\bar{X}$%	no. of items
battery	63.8	168	the whole test	77.9	93
fortis/lenis oppositions	57.9	65	fortis consonants	79.2	34
the rest of the oppos.	67.5	103	lenis consonants	67.6	36
oppos. with 1 dist. other than fortis/ lenis	43.8	29	the rest of the consonants	92.1	23

(2) Both the eight fortis and the eight lenis consonants seem to be more difficult to produce than the consonants incapable of forming oppositions with the fortis/lenis distinction as the only distinction. This seems to be in accordance with our assumption 3. The fact that the lenis consonants have, as a group, proved to be the most difficult to produce is by no means a surprise: out of the eight lenis consonants only two (/v/ and /d/) occur as phonemes in Finnish against four (/ p t k s /) of the eight fortis consonants. Nor is it surprising that the remaining eight consonant phonemes (in the category "the rest of the consonants") have been so easy (92.1%) to pronounce: seven of them occur also in Finnish, /w/ being the only exception.

Assumption 4 (it is more difficult for Finns to hear and produce word-final English consonants than word-initial or word-medial consonants) was empirically tested as follows: in the D-test, SA-test, WA-test and P-test, the average correct answer percentages were separately computed for word-initial, word-medial and word-final consonant phonemes.

The differences between the average correct answer percentages were tested for statistical significance. As the two transcribers (RP and JC) differed significantly in their treatment of word-initial, word-medial and word-final consonants in the production test, we found it legitimate to report the results in Table 16 separately for RP and JC.

Table 16. Average correct answer percentages of word-initial, word-medial and word-final consonants and the statistical significance of their differences.

test	word-initial cons. no.	$\bar{X}$%	word-medial cons. no.	$\bar{X}$%	word-final cons. no.	$\bar{X}$%	t	significant at % level [1]	df
listening tests:									
D-test	30	74.6	22	74.4	23	73.9	- [2]	-	228
SA-test	25	59.3	-	-	20	45.7	15.5	0.1	228
WA-test	26	63.2	-	-	22	50.8	14.8	0.1	227
production test:									
RP	39	73.0	20	82.7			-5.1	0.1	47
	39	73.0			34	70.4	1.1	-	47
			20	82.7	34	70.4	8.8	0.1	47
JC	39	88.4	20	90.3			-1.4	-	47
	39	88.4			34	75.8	9.3	0.1	47
			20	90.3	34	75.8	11.0	0.1	47

[1] The computational formula for t for testing the significance of the difference between two means for correlated samples was used, see formula 11.9 in Ferguson (1966: 169-170). Here, as well as elsewhere in this study, the differences are considered significant only if the risk is 5% or less.

[2] The differences between the mean percentages were minimal (all below 1%) and thus there was no point in testing their significance.

On the whole, our assumption seems to hold. In the SA-test and the WA-test the differences in the mean percentages are highly significant in favour of the word-initial consonants. Thus it can be concluded with 99.9% certainty that word-final consonants are more difficult for Finns to identify than word-initial consonants. The discrimination test, however, seems to be a case apart among the listening tests in this respect also: it seems to make no difference in the discrimination of consonants whether they occur word-initially, word-medially or word-finally. The drawback of the analogy tests is, of course, that word-medial consonants could not be tested.

In the production test both JC and RP seem to agree that word-medial

consonants have been the easiest and word-final consonants the most difficult to pronounce. But JC and RP differ in that the former has found the word-final consonants significantly more difficult than the word-initial or word-medial consonants and no statistical difference between the last two, whereas RP has found both word-final and word-initial consonants statistically equally difficult, but significantly more difficult than word-medial consonants. Thus JC's percentages are in complete accordance with our assumption, while the non-significant difference (2.6%) between word-initial and word-final consonants in RP's data does not directly support our assumption. Still, the difference is in favour of word-initial consonants and thus in conformity with our assumption.

As was reported earlier, JC's and RP's levels of acceptance differed significantly. The difference remained significant in all positions: word-initially (88.4 - 73.0 = 15.4) the difference JC - RP was significant at 0.1% risk ($t = 12.8$, $df = 47$), word-medially (90.3 - 82.7 = 7.6) it was also significant at 0.1% risk ($t = 4.8$, $df = 47$) and word-finally (75.8 - 70.4 = 5.4) it was significant at 2% risk ($t = 2.6$, $df = 47$).

The results imply that it is not enough to teach pupils to identify and pronounce English consonants per se: their position in the word should be taken into account in such a way that pupils get extra practice in identifying and producing word-final consonants.

Assumption 5 (it is difficult for Finns to identify and produce English consonant phonemes which are allophones in Finnish) was so tested that the average correct answer percentages were computed for (1) the "allophones" / b g w f ʃ z / and (2) for the rest of the consonants both in the listening tests and the production test. For comparison we also computed the corresponding percentages for (3) the consonants occurring in both Finnish and English and for (4) / ʒ tʃ dʒ θ ð /, which do not occur in Finnish at all, not even as allophones.

As consonant oppositions, not consonants per se, were tested in the listening tests, we divided the oppositions into the four groups as follows: in the D-test the consonant occurring twice in the triplet was considered the tested consonant, and if all the three consonants in opposition were different, the first was regarded as the tested consonant. The division into the groups was carried out according to the tested consonants; in the SA-test and the WA-test the consonants were divided into the four groups according to the consonants in the stimuli. In the production test the con-

sonants as such could be divided into these groups. The conjoined data of RP's and JC's transcriptions was used. The abbreviations "allo, "rest", "identical" and "only in English" are used for the sake of brevity to denote the above groups (1), (2), (3) and (4), respectively. The results are reported in Table 17.

The results do not support our assumption. The consonants of the "allo" group are roughly as difficult as the "rest" of the consonants both in the listening tests and in the production test. Not even the differences in favour of the "identical" group are greater than 9.4% in the listening tests and 4.9% in the production test. The differences could have been expected to be greater, as the consonants occurring in both Finnish and English were found to be by far the easiest (see Table 13). The most interesting and important result is that our subjects found the five consonant phonemes /ʒ tʃ dʒ θ ð/, which do not occur in Finnish at all, to be by far the most difficult both in hearing (47.6%) and production (5..7%). Thus the occurrence of [b g w f ʃ z] in Finnish seems to have facilitated rather than made the process of identification and production more difficult. This appears very surprising, as many linguists assume that it is easier to learn an entirely new phoneme of the target language than to learn a new usage of a familiar sound. They usually quote an example given by Lado. In Spanish there are two variants of the phoneme /d/. One resembles the English /d/ and the other the English /ð/. They are in complementary distribution, the first occurs word-initially and after /n/, the other between vowels and after /r/. Thus Spaniards are likely to say *father* pro *ladder* when speaking English (see Lado 1957: 14-15 and Lehtonen 1972a: 26) If linguists base their generalization on cases like this, our results are perhaps not so surprising after all. Of the allophones in Finnish, [b g f ʃ] occur in loan-words only, [w] and [z] are not such an integral part of the consonant system in Finnish as is [ð] in Spanish, where it is used every day by every speaker. In Finnish, /v/ is realized as [w] mainly in words like [rouwa] "Mrs", [vauva] "baby" (cf. Lehtonen 1972a: 27). In Finnish, /s/ tends to be voiced (approximating to English /z/) only in a fully voiced sound environment as in [hevozen], the genitive of 'horse', but it is not always realized as [z] in that position, whereas /d/ is always realized as [ð] between vowels and after /r/ in Spanish. Thus Finns are not accustomed to uttering any of the six allophones invariably in one position and another allophone of the phoneme in another as is the case in the use of the variants of

Table 17. Average correct answer percentages of English consonant phonemes occurring as allophones in Finnish.[1]

	"allo"		"rest"		"identical"		"only in English"	
	no.	$\bar{X}$%	no.	$\bar{X}$%	no.	$\bar{X}$%	no.	$\bar{X}$%
listening test battery $\bar{X}$% = 63.8	43	63.6	125	63.9	79	73.4	46	47.6
production test $\bar{X}$% = 77.9	29	79.7	64	77.1	46	84.6	18	57.7

[1] The groups "allo" and "only in English" in the above table correspond to group C, and the group "identical" to groups A and B together in Table 13. See also the footnote on p 19.

/d/ in Spanish. The native speakers of Spanish have thus grown into the habit of using the [ð] variant between vowels and after /r/ and therefore they transfer their habit into their English speech, while Finns have no such habit to be transferred. This may explain the relatively high average correct answer percentage of the allophones.

AN ATTEMPT TO ANSWER PROBLEM 4:

IS THERE A CHANGE IN THE AMOUNT AND TYPE OF LEARNING PROBLEMS BETWEEN SECOND FORMERS AND FIFTH FORMERS IN SECONDARY SCHOOL ?

Strictly speaking the answer to this problem would have presupposed a follow-up study of the second formers: we should have retested the same subjects in the fifth form. We could not wait for the necessary three years to pass. Therefore we decided to take two separate groups of subjects, (1) those pupils who were in the second form and (2) those who were in the fifth form during the spring term of 1975. As the two sets of subjects came from the same schools, one would not expect the groups to differ (as regards their background, talent and so on) from each other to such an extent that the results would be distorted.

To answer the first part of the problem (a change in the amount) we computed the mean scores in each test for the second and the fifth formers separately and tested the differences for statistical significance.

To answer the second part (a change in the type) we correlated the second formers' scores in each test item with those of the fifth formers. The resulting correlation coefficients indicate to what degree both the second and the fifth formers found the same items (i.e. the same consonants and consonant oppositions) difficult/easy. The higher the correlation coefficient the more the same types of learning problems occur in both groups. The results are reported in Table 18. It is to be noted here that in this case we treated the affricates as consonant clusters and thus the number of items in the production test is 103.

The fifth formers achieved significantly higher mean scores than the second formers in the listening tests and also in the production test according to RP's transcription. According to JC's transcription the difference is also in favour of the fifth formers, but it is not significant at the required 5% level, only at the 10% level. The evident conclusion from this is that there is a change in the amount of learning problems to the advantage of the fifth formers. The means and mean percentages do not, however, tell us whether the difference is primarily that of degree (the fifth formers have found the same consonants/consonant oppositions difficult/easy as the second formers, while they have achieved a somewhat better command of them) or that of number (the fifth formers have found fewer and thus different consonants/consonant oppositions difficult). The high correlation coefficients provide an answer to our question: to a very high degree the fifth formers have found the same consonants/consonant oppositions difficult/easy as the second formers. Thus the fifth formers face, only to a lesser degree, the same types of learning problems as the second formers do. In the case of the sound analogy test the correspondence is nearly complete (r=.97 = a 94% correspondence) and in the other two listening tests very high (r=.91 = an 83% correspondence). The fact that the production test was a subjective test naturally accounts for the somewhat lower correlation coefficients (RP r=.88 = a 77% correspondence and JC r=.86 = a 74% correspondence).

Thus the answer to problem 4 is that there is a change (towards a better command of the English consonants) in the amount of learning problems between the second formers and the fifth formers, but the same types

Table 18. The 2nd and the 5th formers' means and standard deviations in the tests and the significance of the differences between the means and the correlation of the item scores between the 2nd and the 5th formers.

test	5th formers $\bar{X}$	5th formers s	2nd formers $\bar{X}$	2nd formers s	t	significant at %	df	correlations r 2&5	correlations r as %
listening tests:									
D-test	58.6 =78.2%	5.0	52.9 =70.4%	7.1	7.12	0.1	227	.91	83
SA-test	25.9 =57.6%	5.2	22.2 =49.1%	4.2	6.03	0.1	227	.97	94
WA-test	30.0 =62.6%	4.5	24.7 =51.9%	6.3	7.42	0.1	226	.91	83
production test:									
RP	77.8 =75.6%	8.8	68.1 =66.1%	9.6	3.59	0.1	46	.88	77
JC	88.2 =85.6%	7.3	84.0 =81.5%	9.6	1.69	(10.0)	46	.86	74

of learning problems that occur in the second form still persist in the fifth form. However, the differences between the means and mean percentages in favour of the fifth formers, although statistically significant, are not as great as one would have expected. The fifth formers show on the average only an 8.9% superiority to the second formers in the listening tests. In the production test a comparison of the mean percentages is complicated by the fact that they, at least to some extent, depend on the evaluator. Therefore we shall report the mean percentages and their differences in both forms separately for each evaluator.

form	teachers	RM	JC	EV	RP
5th	85.6%	83.3%	85.6%	77.1%	75.6%
2nd	85.2%	80.3%	81.5%	72.4%	66.1%
difference	0.4%	3.0%	4.1%	4.7%	9.5%

The differences are very small except that based on RP's transcription. In fact, only RP found the fifth formers significantly better "producers" than the second formers. But even the 9.5% superiority does not mean that any great improvement in the production of English consonants had taken place. One factor which may have reduced the differences is that nearly all of the second formers (112 out of the 114) against only about one-fifth of the fifth formers (28 out of the 115) had studied English in elementary school.

The results suggest in any case that the fifth formers, too, need practice in discriminating, identifying and pronouncing English consonants. The most difficult English consonants are obviously so difficult for Finns that not even at the school leaving age have the pupils learnt to master them.

AN ATTEMPT TO ANSWER PROBLEM 5:

CAN SUCCESS IN THE PRODUCTION TEST BE PREDICTED BY THE LISTENING TEST RESULTS ?

The general belief that a correct pronunciation of the sounds of the target language cannot be expected before they are heard correctly, i.e. hearing precedes production, raised the question: Can we predict success in the production test by success in the listening tests? Therefore we selected the production test subjects in such a way that on the basis of the listening test battery the top 10% and the bottom 10% of the pupils in each of the six forms were taken as subjects. The underlying idea was that if those who did well/badly in the listening tests also did well/badly in the production test, then one could say that success in the production test is predictable on the basis of the listening test results. To find an answer to the problem, the correlation coefficients were computed between the production test scores (the criterion variable) and the listening test scores (the predictors) of the 48 subjects. In this case (as in connection with problem 6) the means of the five evaluators' scores were used as the criterion variable. The resulting correlations are reported in Table 19, where the correlation coefficients are presented above the dashes and the corresponding percentages showing the common variance as a mirror image below the dashes.

Table 19. Listening and production test means, standard deviations and correlations (N=48).

test	$\bar{X}$	s	predictors				criterion
			D-test	SA-test	WA-test	battery	P-test
D-test	54.8	9.7	-	.814	.811	.945	.796
SA-test	24.6	7.3	66.3%	-	.816	.927	.776
WA-test	27.8	8.4	65.8%	66.6%	-	.934	.833
battery	107.1	25.7	89.3%	85.9%	87.2%	-	.856
P-test	81.7	9.0	63.4%	60.2%	69.4%	73.3%	-

Table 19 shows that all the test correlate highly with each other. All the correlation coefficients are statistically significant at 1% risk. The listening test battery seems to be the best predictor of success in the production test (r=.856 = 73.3% prediction). Of the individual listening tests the WA-test is nearly as good a predictor (r=.833 = 69.4% prediction) as the battery. The D-test and the SA-test also correlate highly with the production test. That the test battery is only a slightly better predictor than the individual tests is due to the high intercorrelations between the three listening tests. The evident conclusion from the results is that in our case the listening tests yielded fairly accurate predictions (ranging from 60.2% to 73.3%) of success in the production test. It must be remembered, however, that our method of selecting high achievers and low achievers as our production test subjects enlarged the standard deviations and thus contributed to high predictions. It is obvious that such high predictions could only be obtained again if the subjects were similarly selected.

The fact that success in the production test could be predicted on the basis of the listening test results must not, however, be so interpreted that perception definitely precedes production. A correlation coefficient expresses only that two variables are mutually related; it does not indicate which is the cause and which the effect. Thus a high correlation coefficient between the listening test battery and the production test, for instance, tells us that knowing the subjects' performances in one, their performances in the other are predicatble, but one cannot say

that one causes the other. Which is the cause and which the effect must be logically determined. It is also possible that one variable (A) causes the other (B), which in turn brings about changes in the former (A). The last interpretation would appear to be the most likely one in our case. Obviously people with defective hearing cannot be expected to be able to produce foreign language sounds properly, but Brière's and, with reservations, our own results would seem to indicate that people with normal hearing ability gain mastery of perception through production (see pp. 66-67 above). Thus to be able to produce foreign language sounds seems to presuppose some skill in perceiving them, but to be able to perceive them accurately seems to presuppose practice in producing them. It has to be emphasized that we have not found conclusive evidence for this interpretation. In our opinion the implication of Brière's and our results for teaching would be that the teaching of foreign language sounds should *not* be divided into two separate sections, first training in perception, then training in production, as implied by the conviction that perception precedes production, but the training in perception and in production should alternate continuously.

AN ATTEMPT TO ANSWER PROBLEM 6:

ARE CERTAIN BACKGROUND VARIABLES RELATED TO PUPILS' ABILITY TO DISCRIMINATE, IDENTIFY AND PRODUCE ENGLISH CONSONANTS ?

The results indicated that the ability to discriminate, identify and produce English consonants is a specific skill that cannot be satisfactorily explained by means of the background variables used in this study. Of these only pupils' verbal ability (= school marks in languages), conceptions about the easiness of school subjects (of English particularly), home background, future educational goals and parents' favourable attitudes towards school seemed to be somewhat related to success in our tests. However, even the highest individual correlation with the listening tests, .507 (the easiness of English), explained only 25.7% of the fifth formers' performance in the sound analogy test. In most cases the significant correlation coefficients (at 5% significance level .195 or above) were low, usually between .20 and .30 and thus explaining only from 4% to 9% of the variance of the listening test scores. Obviously due to the selection of the production test subjects the seven significant (.288 or above) cor-

relations with the production test were considerably higher, ranging from .288 (grammar) to .638 (mark in English) and thus explaining from 8.29% to 40.7% of success in the production test.

Stepwise multiple regression analyses with the best individual background variables revealed that the chosen variables together did not explain more than 16.2% of the second formers' and 32.4% of the fifth formers' performance in the listening tests. In the production test the multiple correlation was as high as .753 (56.7%). Apparently the selection of the production test subjects largely contributed to this.

CHARACTERISTICS OF THE TESTS

Table 20 summarizes the properties of the final test versions for learners of English and learners of German.

The table shows that the S-test, SA-test and WA-test approximate to the ideal 50% difficulty, whereas the D-test and the P-test have proved rather easy. The means and standard deviations seem to indicate that the scores are normally distributed in the S-test, SA-test and WA-test while in the D-test and the P-test the distribution is negatively skewed. The forms of the distributions were graphically checked and the means and standard deviations were found to give a correct picture.

On the whole the tests were reliable, the KR_{20}-coefficients of the separate listening tests ranging from .59 to .79 and those of the P-test from .83 to .92 (depending on the transcriber). The battery (D-test + SA-test + WA-test) yielded reliability coefficients as high as .89 in the second form and .91 in the fifth form.

Of the four types of validity the criterion-related validity could not be determined as there were no valid outside criteria to correlate the test scores with. The content validity was secured by testing the English consonant phonemes in word-initial, word-medial and word-final positions. The construct validity of the tests had to be judged on the basis of logical inferences from the data. There seemed to be no doubt about the construct validity of the S-test, SA-test, WA-test and P-test, whereas the doubts that the D-test measures auditory discrimination rather than mastery of the sound oppositions gained support.

The learners of German achieved significantly (at 0.1% level, $t=3.32$, $df=212$) higher scores ($\bar{X}=55.9$) in the D-test than the second formers ($\bar{X}=52.9$). This clearly indicates that tests based on minimal pairs hardly measure the command of sound oppositions in a given language. It would illogical to think that the learners of German, practically without knowledge of English, have a better command of the English consonant phonemes than the second formers, the vast majority (112 out of 114) of whom had studied English already at elementary school. Not even the fact that the fifth formers proved significantly better than the learners of German (the difference between the means being $58.6 - 55.9 = 2.7$, $t = 3.62$, risk 0.1%, $df = 213$) reflutes our previous statement, because in

Table 20. Properties of the final test versions.

test	form	N	no. of items	$\bar{X}$	s	$\bar{X}$%	KR_{20}	time[1]
Learners of English								
D-test	2	114	75	52.9	7.1	70.4	.79	18 min.
	5	115	75	58.6	5.0	78.2	.64	18
SA-test	2	114	45	22.2	4.2	49.1	.59	18
	5	115	45	25.9	3.2	57.6	.72	18
WA-test	2	114	48	24.7	6.3	51.9	.77	16
	5	115	48	30.0	4.5	62.6	.63	16
battery	2	114	168	99.7	14.3	59.3	.89	52
	5	115	168	114.5	12.0	68.2	.91	52
P-test								
JC	2	24	103	84.0	9.6	81.5	.89	11
	5	24	103	88.2	7.3	85.6	.83	11
RP	2	24	103	68.1	9.6	66.1	.83	11
	5	24	103	77.8	8.8	75.6	.85	11
Teachers	2	24	103	87.8	8.5	85.2	.87	11
	5	24	103	88.2	10.3	85.6	.92	11
RM	2	24	103	82.8	10.6	80.3	.91	11
	5	24	103	85.8	9.2	83.3	.89	11
EV	2	24	103	74.5	11.8	72.4	.92	11
	5	24	103	79.4	11.2	77.1	.92	11
Learners of German								
S-test	5	100	70	38.3	4.5	54.8	.63	14
D-test	5	100	75	55.9	5.9	74.5	.75	18

[1] The time for administration includes instructions, practice items and the necessary pauses.

spite of the statistical significance the difference is only 3.7% in favour of the fifth formers. In fact when the learners of English are treated as one group, there is a slight difference in the average correct answer percentages in favour of the learners of German (74.5% against 74.3%). Therefore discrimination tests (based on minimal pair techniques) should be used to measure auditory discrimination alone.

BIBLIOGRAPHY

Bloomfield, L. 1935. *Language*. Revised edition. London: Allen & Unwin.

Branford, W. 1967. *The Elements of English. An Introduction to the Principles of the Study of Language*. London: Routledge & Kegan Paul.

Brière, E. 1966. An Investigation of Phonological Interference, *Language*, 42, 768-796.

Brière, E. 1967. Phonological Testing Reconsidered, *Language Learning*. 17, 163-171.

Chomsky, N. 1957. *Syntactic Structures*. Janua Linguarum, Series Minor, 4. The Hague: Mouton.

Downie, N.M. 1967. *Fundamentals of Measurement: Techniques and Practices*. Second edition. London: Oxford University Press.

Ferguson, G.A. 1966. *Statistical Analysis in Psychology and Education*. Ljubljana.

Gaeng, P.A. 1971. *Introduction to the Principles of Language*. Harper and Row.

Gimson, A.C. 1962. *An Introduction to the Pronunciation of English*. London: Edward Arnold.

Gleason, H.A. 1969. *An Introduction to Descriptive Linguistics*. Revised edition. London and Beccles: Holt, Rinehart & Winston.

Harms, R.T. 1968. *Introduction to Phonological Theory*. Englewood Cliffs: Prentice-Hall.

Harris, D.P. 1969. *Testing English as a Second Language*. New York and London: McGraw-Hill.

Heinonen, V. 1961. *Koulusaavutustestit*. Jyväskylä: Gummerus.

Heinonen, V. 1964. *Differentiaalipsykologia*. Jyväskylä: Työn Voima.

Heinonen, V. 1968. *Kansakoulun oppiaineitten arvioinnit. Oppilaitten ja opettajien käsitykset aineitten tärkeydestä, vaikeudesta ja mieluisuudesta*. Jyväskylä.

Hirvonen, P. 1971. *Englannin kielen taidon mittaaminen lukion päättyessä, 2: Kuuntelukoe*. Publications de l'Association Finlandaise de Linguistique Appliquée (AFinLA), 4. Turku.

Hirvonen, P. 1974. *Englannin kielen taidon mittaaminen lukion päättyessä, 5: Kielitaitokoe*. Publications de l'Association Finlandaise de Linguistique Appliquée (AFinLA), 9. Turku.

Table 15. Average correct answer percentages for fortis/lenis consonants and oppositions.

(1) the listening test battery			(2) the production test		
	$\bar{X}$%	no. of items		$\bar{X}$%	no. of items
battery	63.8	168	the whole test	77.9	93
fortis/lenis oppositions	57.9	65	fortis consonants	79.2	34
the rest of the oppos.	67.5	103	lenis consonants	67.6	36
oppos. with 1 dist. other than fortis/ lenis	43.8	29	the rest of the consonants	92.1	23

(2) Both the eight fortis and the eight lenis consonants seem to be more difficult to produce than the consonants incapable of forming oppositions with the fortis/lenis distinction as the only distinction. This seems to be in accordance with our assumption 3. The fact that the lenis consonants have, as a group, proved to be the most difficult to produce is by no means a surprise: out of the eight lenis consonants only two (/v/ and /d/) occur as phonemes in Finnish against four (/ p t k s /) of the eight fortis consonants. Nor is it surprising that the remaining eight consonant phonemes (in the category "the rest of the consonants") have been so easy (92.1%) to pronounce: seven of them occur also in Finnish, /w/ being the only exception.

Assumption 4 (it is more difficult for Finns to hear and produce word-final English consonants than word-initial or word-medial consonants) was empirically tested as follows: in the D-test, SA-test, WA-test and P-test, the average correct answer percentages were separately computed for word-initial, word-medial and word-final consonant phonemes.

The differences between the average correct answer percentages were tested for statistical significance. As the two transcribers (RP and JC) differed significantly in their treatment of word-initial, word-medial and word-final consonants in the production test, we found it legitimate to report the results in Table 16 separately for RP and JC.

Table 16. Average correct answer percentages of word-initial, word-medial and word-final consonants and the statistical significance of their differences.

test	word-initial cons. no.	$\bar{X}$%	word-medial cons. no.	$\bar{X}$%	word-final cons. no.	$\bar{X}$%	t	significant at % level [1]	df
listening tests:									
D-test	30	74.6	22	74.4	23	73.9	- [2]	-	228
SA-test	25	59.3	-	-	20	45.7	15.5	0.1	228
WA-test	26	63.2	-	-	22	50.8	14.8	0.1	227
production test:									
RP	39	73.0	20	82.7			-5.1	0.1	47
	39	73.0			34	70.4	1.1	-	47
			20	82.7	34	70.4	8.8	0.1	47
JC	39	88.4	20	90.3			-1.4	-	47
	39	88.4			34	75.8	9.3	0.1	47
			20	90.3	34	75.8	11.0	0.1	47

[1] The computational formula for t for testing the significance of the difference between two means for correlated samples was used, see formula 11.9 in Ferguson (1966: 169-170). Here, as well as elsewhere in this study, the differences are considered significant only if the risk is 5% or less.

[2] The differences between the mean percentages were minimal (all below 1%) and thus there was no point in testing their significance.

On the whole, our assumption seems to hold. In the SA-test and the WA-test the differences in the mean percentages are highly significant in favour of the word-initial consonants. Thus it can be concluded with 99.9% certainty that word-final consonants are more difficult for Finns to identify than word-initial consonants. The discrimination test, however, seems to be a case apart among the listening tests in this respect also: it seems to make no difference in the discrimination of consonants whether they occur word-initially, word-medially or word-finally. The drawback of the analogy tests is, of course, that word-medial consonants could not be tested.

In the production test both JC and RP seem to agree that word-medial

consonants have been the easiest and word-final consonants the most difficult to pronounce. But JC and RP differ in that the former has found the word-final consonants significantly more difficult than the word-initial or word-medial consonants and no statistical difference between the last two, whereas RP has found both word-final and word-initial consonants statistically equally difficult, but significantly more difficult than word-medial consonants. Thus JC's percentages are in complete accordance with our assumption, while the non-significant difference (2.6%) between word-initial and word-final consonants in RP's data does not directly support our assumption. Still, the difference is in favour of word-initial consonants and thus in conformity with our assumption.

As was reported earlier, JC's and RP's levels of acceptance differed significantly. The difference remained significant in all positions: word-initially (88.4 - 73.0 = 15.4) the difference JC - RP was significant at 0.1% risk ($t = 12.8$, $df = 47$), word-medially (90.3 - 82.7 = 7.6) it was also significant at 0.1% risk ($t = 4.8$, $df = 47$) and word-finally (75.8 - 70.4 = 5.4) it was significant at 2% risk ($t = 2.6$, $df = 47$).

The results imply that it is not enough to teach pupils to identify and pronounce English consonants per se: their position in the word should be taken into account in such a way that pupils get extra practice in identifying and producing word-final consonants.

Assumption 5 (it is difficult for Finns to identify and produce English consonant phonemes which are allophones in Finnish) was so tested that the average correct answer percentages were computed for (1) the "allophones" / b g w f ʃ z / and (2) for the rest of the consonants both in the listening tests and the production test. For comparison we also computed the corresponding percentages for (3) the consonants occurring in both Finnish and English and for (4) / ʒ tʃ dʒ θ ð /, which do not occur in Finnish at all, not even as allophones.

As consonant oppositions, not consonants per se, were tested in the listening tests, we divided the oppositions into the four groups as follows: in the D-test the consonant occurring twice in the triplet was considered the tested consonant, and if all the three consonants in opposition were different, the first was regarded as the tested consonant. The division into the groups was carried out according to the tested consonants; in the SA-test and the WA-test the consonants were divided into the four groups according to the consonants in the stimuli. In the production test the con-

sonants as such could be divided into these groups. The conjoined data of RP's and JC's transcriptions was used. The abbreviations "allo, "rest", "identical" and "only in English" are used for the sake of brevity to denote the above groups (1), (2), (3) and (4), respectively. The results are reported in Table 17.

The results do not support our assumption. The consonants of the "allo" group are roughly as difficult as the "rest" of the consonants both in the listening tests and in the production test. Not even the differences in favour of the "identical" group are greater than 9.4% in the listening tests and 4.9% in the production test. The differences could have been expected to be greater, as the consonants occurring in both Finnish and English were found to be by far the easiest (see Table 13). The most interesting and important result is that our subjects found the five consonant phonemes /ʒ tʃ dʒ θ ð/, which do not occur in Finnish at all, to be by far the most difficult both in hearing (47.6%) and production (5..7%). Thus the occurrence of [b g w f ʃ z] in Finnish seems to have facilitated rather than made the process of identification and production more difficult. This appears very surprising, as many linguists assume that it is easier to learn an entirely new phoneme of the target language than to learn a new usage of a familiar sound. They usually quote an example given by Lado. In Spanish there are two variants of the phoneme /d/. One resembles the English /d/ and the other the English /ð/. They are in complementary distribution, the first occurs word-initially and after /n/, the other between vowels and after /r/. Thus Spaniards are likely to say *father* pro *ladder* when speaking English (see Lado 1957: 14-15 and Lehtonen 1972a: 26) If linguists base their generalization on cases like this, our results are perhaps not so surprising after all. Of the allophones in Finnish, [b g f ʃ] occur in loan-words only, [w] and [z] are not such an integral part of the consonant system in Finnish as is [ð] in Spanish, where it is used every day by every speaker. In Finnish, /v/ is realized as [w] mainly in words like [rouwa] "Mrs", [vauwa] "baby" (cf. Lehtonen 1972a: 27). In Finnish, /s/ tends to be voiced (approximating to English /z/) only in a fully voiced sound environment as in [hevozen], the genitive of 'horse', but it is not always realized as [z] in that position, whereas /d/ is always realized as [ð] between vowels and after /r/ in Spanish. Thus Finns are not accustomed to uttering any of the six allophones invariably in one position and another allophone of the phoneme in another as is the case in the use of the variants of

Table 17. Average correct answer percentages of English consonant phonemes occurring as allophones in Finnish.[1]

	"allo"		"rest"		"identical"		"only in English"	
	no.	$\bar{X}$%	no.	$\bar{X}$%	no.	$\bar{X}$%	no.	$\bar{X}$%
listening test battery $\bar{X}$% = 65.8	43	63.6	125	63.9	79	73.4	46	47.6
production test $\bar{X}$% = 77.9	29	79.7	64	77.1	46	84.6	18	57.7

[1] The groups "allo" and "only in English" in the above table correspond to group C, and the group "identical" to groups A and B together in Table 13. See also the footnote on p 19.

/d/ in Spanish. The native speakers of Spanish have thus grown into the habit of using the [ð] variant between vowels and after /r/ and therefore they transfer their habit into their English speech, while Finns have no such habit to be transferred. This may explain the relatively high average correct answer percentage of the allophones.

AN ATTEMPT TO ANSWER PROBLEM 4:

IS THERE A CHANGE IN THE AMOUNT AND TYPE OF LEARNING PROBLEMS BETWEEN SECOND FORMERS AND FIFTH FORMERS IN SECONDARY SCHOOL ?

Strictly speaking the answer to this problem would have presupposed a follow-up study of the second formers: we should have retested the same subjects in the fifth form. We could not wait for the necessary three years to pass. Therefore we decided to take two separate groups of subjects, (1) those pupils who were in the second form and (2) those who were in the fifth form during the spring term of 1975. As the two sets of subjects came from the same schools, one would not expect the groups to differ (as regards their background, talent and so on) from each other to such an extent that the results would be distorted.

To answer the first part of the problem (a change in the amount) we computed the mean scores in each test for the second and the fifth formers separately and tested the differences for statistical significance.

To answer the second part (a change in the type) we correlated the second formers' scores in each test item with those of the fifth formers. The resulting correlation coefficients indicate to what degree both the second and the fifth formers found the same items (i.e. the same consonants and consonant oppositions) difficult/easy. The higher the correlation coefficient the more the same types of learning problems occur in both groups. The results are reported in Table 18. It is to be noted here that in this case we treated the affricates as consonant clusters and thus the number of items in the production test is 103.

The fifth formers achieved significantly higher mean scores than the second formers in the listening tests and also in the production test according to RP's transcription. According to JC's transcription the difference is also in favour of the fifth formers, but it is not significant at the required 5% level, only at the 10% level. The evident conclusion from this is that there is a change in the amount of learning problems to the advantage of the fifth formers. The means and mean percentages do not, however, tell us whether the difference is primarily that of degree (the fifth formers have found the same consonants/consonant oppositions difficult/easy as the second formers, while they have achieved a somewhat better command of them) or that of number (the fifth formers have found fewer and thus different consonants/consonant oppositions difficult). The high correlation coefficients provide an answer to our question: to a very high degree the fifth formers have found the same consonants/consonant oppositions difficult/easy as the second formers. Thus the fifth formers face, only to a lesser degree, the same types of learning problems as the second formers do. In the case of the sound analogy test the correspondence is nearly complete (r=.97 = a 94% correspondence) and in the other two listening tests very high (r=.91 = an 83% correspondence). The fact that the production test was a subjective test naturally accounts for the somewhat lower correlation coefficients (RP r=.88 = a 77% correspondence and JC r=.86 = a 74% correspondence).

Thus the answer to problem 4 is that there is a change (towards a better command of the English consonants) in the amount of learning problems between the second formers and the fifth formers, but the same types

Table 18. The 2nd and the 5th formers' means and standard deviations in the tests and the significance of the differences between the means and the correlation of the item scores between the 2nd and the 5th formers.

test	5th formers $\bar{X}$	s	2nd formers $\bar{X}$	s	t	significant at %	df	correlations r 2&5	r as %
listening tests:									
D-test	58.6 =78.2%	5.0	52.9 =70.4%	7.1	7.12	0.1	227	.91	83
SA-test	25.9 =57.6%	5.2	22.2 =49.1%	4.2	6.03	0.1	227	.97	94
WA-test	30.0 =62.6%	4.5	24.7 =51.9%	6.3	7.42	0.1	226	.91	83
production test:									
RP	77.8 =75.6%	8.8	68.1 =66.1%	9.6	3.59	0.1	46	.88	77
JC	88.2 =85.6%	7.3	84.0 =81.5%	9.6	1.69	(10.0)	46	.86	74

of learning problems that occur in the second form still persist in the fifth form. However, the differences between the means and mean percentages in favour of the fifth formers, although statistically significant, are not as great as one would have expected. The fifth formers show on the average only an 8.9% superiority to the second formers in the listening tests. In the production test a comparison of the mean percentages is complicated by the fact that they, at least to some extent, depend on the evaluator. Therefore we shall report the mean percentages and their differences in both forms separately for each evaluator.

form	teachers	RM	JC	EV	RP
5th	85.6%	83.3%	85.6%	77.1%	75.6%
2nd	85.2%	80.3%	81.5%	72.4%	66.1%
difference	0.4%	3.0%	4.1%	4.7%	9.5%

The differences are very small except that based on RP's transcription. In fact, only RP found the fifth formers significantly better "producers" than the second formers. But even the 9.5% superiority does not mean that any great improvement in the production of English consonants had taken place. One factor which may have reduced the differences is that nearly all of the second formers (112 out of the 114) against only about one-fifth of the fifth formers (28 out of the 115) had studied English in elementary school.

The results suggest in any case that the fifth formers, too, need practice in discriminating, identifying and pronouncing English consonants. The most difficult English consonants are obviously so difficult for Finns that not even at the school leaving age have the pupils learnt to master them.

AN ATTEMPT TO ANSWER PROBLEM 5:

CAN SUCCESS IN THE PRODUCTION TEST BE PREDICTED BY THE LISTENING TEST RESULTS ?

The general belief that a correct pronunciation of the sounds of the target language cannot be expected before they are heard correctly, i.e. hearing precedes production, raised the question: Can we predict success in the production test by success in the listening tests? Therefore we selected the production test subjects in such a way that on the basis of the listening test battery the top 10% and the bottom 10% of the pupils in each of the six forms were taken as subjects. The underlying idea was that if those who did well/badly in the listening tests also did well/badly in the production test, then one could say that success in the production test is predictable on the basis of the listening test results. To find an answer to the problem, the correlation coefficients were computed between the production test scores (the criterion variable) and the listening test scores (the predictors) of the 48 subjects. In this case (as in connection with problem 6) the means of the five evaluators' scores were used as the criterion variable. The resulting correlations are reported in Table 19, where the correlation coefficients are presented above the dashes and the corresponding percentages showing the common variance as a mirror image below the dashes.

Table 19. Listening and production test means, standard deviations and correlations (N=48).

test	$\bar{X}$	s	predictors D-test	SA-test	WA-test	battery	criterion P-test
D-test	54.8	9.7	-	.814	.811	.945	.796
SA-test	24.6	7.3	66.3%	-	.816	.927	.776
WA-test	27.8	8.4	65.8%	66.6%	-	.934	.833
battery	107.1	23.7	89.3%	85.9%	87.2%	-	.856
P-test	81.7	9.0	63.4%	60.2%	69.4%	73.3%	-

Table 19 shows that all the test correlate highly with each other. All the correlation coefficients are statistically significant at 1% risk. The listening test battery seems to be the best predictor of success in the production test (r=.856 = 73.3% prediction). Of the individual listening tests the WA-test is nearly as good a predictor (r=.833 = 69.4% prediction) as the battery. The D-test and the SA-test also correlate highly with the production test. That the test battery is only a slightly better predictor than the individual tests is due to the high intercorrelations between the three listening tests. The evident conclusion from the results is that in our case the listening tests yielded fairly accurate predictions (ranging from 60.2% to 73.3%) of success in the production test. It must be remembered, however, that our method of selecting high achievers and low achievers as our production test subjects enlarged the standard deviations and thus contributed to high predictions. It is obvious that such high predictions could only be obtained again if the subjects were similarly selected.

The fact that success in the production test could be predicted on the basis of the listening test results must not, however, be so interpreted that perception definitely precedes production. A correlation coefficient expresses only that two variables are mutually related; it does not indicate which is the cause and which the effect. Thus a high correlation coefficient between the listening test battery and the production test, for instance, tells us that knowing the subjects' performances in one, their performances in the other are predicatble, but one cannot say

that one causes the other. Which is the cause and which the effect must be logically determined. It is also possible that one variable (A) causes the other (B), which in turn brings about changes in the former (A). The last interpretation would appear to be the most likely one in our case. Obviously people with defective hearing cannot be expected to be able to produce foreign language sounds properly, but Brière's and, with reservations, our own results would seem to indicate that people with normal hearing ability gain mastery of perception through production (see pp. 66-67 above). Thus to be able to produce foreign language sounds seems to presuppose some skill in perceiving them, but to be able to perceive them accurately seems to presuppose practice in producing them. It has to be emphasized that we have not found conclusive evidence for this interpretation. In our opinion the implication of Brière's and our results for teaching would be that the teaching of foreign language sounds should *not* be divided into two separate sections, first training in perception, then training in production, as implied by the conviction that perception precedes production, but the training in perception and in production should alternate continuously.

AN ATTEMPT TO ANSWER PROBLEM 6:

ARE CERTAIN BACKGROUND VARIABLES RELATED TO PUPILS' ABILITY TO DISCRIMINATE, IDENTIFY AND PRODUCE ENGLISH CONSONANTS ?

The results indicated that the ability to discriminate, identify and produce English consonants is a specific skill that cannot be satisfactorily explained by means of the background variables used in this study. Of these only pupils' verbal ability (= school marks in languages), conceptions about the easiness of school subjects (of English particularly), home background, future educational goals and parents' favourable attitudes towards school seemed to be somewhat related to success in our tests. However, even the highest individual correlation with the listening tests, .507 (the easiness of English), explained only 25.7% of the fifth formers' performance in the sound analogy test. In most cases the significant correlation coefficients (at 5% significance level .195 or above) were low, usually between .20 and .30 and thus explaining only from 4% to 9% of the variance of the listening test scores. Obviously due to the selection of the production test subjects the seven significant (.288 or above) cor-

relations with the production test were considerably higher, ranging from .288 (grammar) to .638 (mark in English) and thus explaining from 8.29% to 40.7% of success in the production test.

Stepwise multiple regression analyses with the best individual background variables revealed that the chosen variables together did not explain more than 16.2% of the second formers' and 32.4% of the fifth formers' performance in the listening tests. In the production test the multiple correlation was as high as .753 (56.7%). Apparently the selection of the production test subjects largely contributed to this.

CHARACTERISTICS OF THE TESTS

Table 20 summarizes the properties of the final test versions for learners of English and learners of German.

The table shows that the S-test, SA-test and WA-test approximate to the ideal 50% difficulty, whereas the D-test and the P-test have proved rather easy. The means and standard deviations seem to indicate that the scores are normally distributed in the S-test, SA-test and WA-test while in the D-test and the P-test the distribution is negatively skewed. The forms of the distributions were graphically checked and the means and standard deviations were found to give a correct picture.

On the whole the tests were reliable, the KR_{20}-coefficients of the separate listening tests ranging from .59 to .79 and those of the P-test from .83 to .92 (depending on the transcriber). The battery (D-test + SA-test + WA-test) yielded reliability coefficients as high as .89 in the second form and .91 in the fifth form.

Of the four types of validity the criterion-related validity could not be determined as there were no valid outside criteria to correlate the test scores with. The content validity was secured by testing the English consonant phonemes in word-initial, word-medial and word-final positions. The construct validity of the tests had to be judged on the basis of logical inferences from the data. There seemed to be no doubt about the construct validity of the S-test, SA-test, WA-test and P-test, whereas the doubts that the D-test measures auditory discrimination rather than mastery of the sound oppositions gained support.

The learners of German achieved significantly (at 0.1% level, $t=3.32$, $df=212$) higher scores ($\bar{X}=55.9$) in the D-test than the second formers ($\bar{X}=52.9$). This clearly indicates that tests based on minimal pairs hardly measure the command of sound oppositions in a given language. It would illogical to think that the learners of German, practically without knowledge of English, have a better command of the English consonant phonemes than the second formers, the vast majority (112 out of 114) of whom had studied English already at elementary school. Not even the fact that the fifth formers proved significantly better than the learners of German (the difference between the means being $58.6 - 55.9 = 2.7$, $t = 3.62$, risk 0.1%, $df = 213$) reflutes our previous statement, because in

Table 20. Properties of the final test versions.

test	form	N	no. of items	$\bar{X}$	s	$\bar{X}$%	KR_{20}	time[1]
Learners of English								
D-test	2	114	75	52.9	7.1	70.4	.79	18 min.
	5	115	75	58.6	5.0	78.2	.64	18
SA-test	2	114	45	22.2	4.2	49.1	.59	18
	5	115	45	26.9	3.2	57.6	.72	18
WA-test	2	114	48	24.7	6.3	51.9	.77	16
	5	115	48	30.0	4.5	62.6	.63	16
Battery	2	114	168	99.7	14.3	59.3	.89	52
	5	115	168	114.5	12.0	68.2	.91	52
P-test								
JC	2	24	103	84.0	9.6	81.5	.89	11
	5	24	103	88.2	7.3	85.6	.83	11
RP	2	24	103	68.1	9.6	66.1	.83	11
	5	24	103	77.8	8.8	75.6	.83	11
Teachers	2	24	103	87.8	8.5	85.2	.87	11
	5	24	103	88.2	10.3	85.6	.92	11
RM	2	24	103	82.8	10.6	80.3	.91	11
	5	24	103	85.8	9.2	83.3	.89	11
EV	2	24	103	74.5	11.8	72.4	.92	11
	5	24	103	79.4	11.2	77.1	.92	11
Learners of German								
S-test	5	100	70	38.3	4.5	54.8	.63	14
D-test	5	100	75	55.9	5.9	74.5	.75	18

[1] The time for administration includes instructions, practice items and the necessary pauses.

spite of the statistical significance the difference is only 3.7% in favour of the fifth formers. In fact when the learners of English are treated as one group, there is a slight difference in the average correct answer percentages in favour of the learners of German (74.5% against 74.5%). Therefore discrimination tests (based on minimal pair techniques) should be used to measure auditory discrimination alone.

BIBLIOGRAPHY

Bloomfield, L. 1935. *Language*. Revised edition. London: Allen & Unwin.

Branford, W. 1967. *The Elements of English. An Introduction to the Principles of the Study of Language*. London: Routledge & Kegan Paul.

Brière, E. 1966. An Investigation of Phonological Interference, *Language*, 42, 768-796.

Brière, E. 1967. Phonological Testing Reconsidered, *Language Learning*. 17, 163-171.

Chomsky, N. 1957. *Syntactic Structures*. Janua Linguarum, Series Minor, 4. The Hague: Mouton.

Downie, N.M. 1967. *Fundamentals of Measurement: Techniques and Practices*. Second edition. London: Oxford University Press.

Ferguson, G.A. 1966. *Statistical Analysis in Psychology and Education*. Ljubljana.

Gaeng, P.A. 1971. *Introduction to the Principles of Language*. Harper and Row.

Gimson, A.C. 1962. *An Introduction to the Pronunciation of English*. London: Edward Arnold.

Gleason, H.A. 1969. *An Introduction to Descriptive Linguistics*. Revised edition. London and Beccles: Holt, Rinehart & Winston.

Harms, R.T. 1968. *Introduction to Phonological Theory*. Englewood Cliffs: Prentice-Hall.

Harris, D.P. 1969. *Testing English as a Second Language*. New York and London: McGraw-Hill.

Heinonen, V. 1961. *Koulusaavutustestit*. Jyväskylä: Gummerus.

Heinonen, V. 1964. *Differentiaalipsykologia*. Jyväskylä: Työn Voima.

Heinonen, V. 1968. *Kansakoulun oppiaineiden arvioinnit. Oppilaitten ja opettajien käsitykset aineitten tärkeydestä, vaikeudesta ja mieluisuudesta*. Jyväskylä.

Hirvonen, P. 1971. *Englannin kielen taidon mittaaminen lukion päättyessä, 2: Kuuntelukoe*. Publications de l'Association Finlandaise de Linguistique Appliquée (AFinLA), 4. Turku.

Hirvonen, P. 1974. *Englannin kielen taidon mittaaminen lukion päättyessä, 5: Kielitaitokoe*. Publications de l'Association Finlandaise de Linguistique Appliquée (AFinLA), 9. Turku.

Hämäläinen, S., & S. Takala 1970. *Kunnallisten kokeilukeskikoulujen ja oppikoulujen ruotsin ja englannin kielen koulusaavutukset.* Kouluhallituksen Julkaisuja, Kokeilu- ja tutkimusselosteita, 29. Helsinki: National Board of Schools.

Jakobson, R. 1962. *Selected Writings, 1. Phonological Studies.* The Hague: Mouton.

Jones, D. 1967. *An Outline of English Phonetics.* Cambridge: Heffer.

Jurama, V. 1966. Kansakoulun IV-luokkalaisten kaupunkilaisvanhempien lastensa koulunkäyntiin liittyvistä tavoitteista ja asenteista erilaisissa sosiaalisissa ryhmissä, *Kasvatus ja koulu,* 52, 120-125.

Jurama, V. 1971. *Kotitaustan yhteydestä oppilaiden koulunkäyntiin liittyviin muuttujiin.* Publications of the Institute for Educational Research, 88. Jyväskylä.

Karlsson, F. 1969. Suomen yleiskielen segmentaalifoneemien paradigma, *Virittäjä,* 73, 351-362.

Kerlinger, F. 1969. *Foundations of Behavioral Research, Educational and Psychological Inquiry.* Aylesbury (1972).

Konttinen, R. 1970. *Opiskelijoiden englannin kielen taitojen ja niiden oppimisen yhteydet verbaaliseen lahjakkuuteen ja persoonallisuuden piirteisiin.* Publications of the Institute for Educational Research, 70. Jyväskylä.

Lado, R. 1957. *Linguistics across Cultures.* Ann Arbor: University of Michigan.

Lado, R. 1961. *Language Testing: The Construction and Use of Foreign Language Tests.* Fifth impression. London: Lowe and Brydone (1967).

Lehtonen, J. 1972a. *Kielenopetuksen fonologiaa.* Publications of the Institute for Educational Research, 168. Jyväskylä.

Lehtonen, J. 1972b. *Äänitallenteet, tallennuslaitteet ja puheopetuksen apuvälineet.* Publications of the Institute for Educational Research, 155. Jyväskylä.

Lehtovaara, J. 1974. *Peruskoulun kolmasluokkalaisten ääntämistaidon rakenne.* Unpublished licentiate thesis in education. University of Tampere.

Lehtovaara, M. 1974. *Eräiden oppilaskohtaisten tekijöiden yhteyksistä englannin kielen ääntämistaidon osaamiseen peruskoulun kolmannella luokalla.* Unpublished licentiate thesis in education. University of Tampere.

Leino, A.-L. 1970. *Englannin kielen koulusaavutusten rakenteesta oppikoulun VII luokalla.* Helsingin yliopiston kasvatustieteen laitoksen tutkimuksia, 9. Helsinki.

Leino, A.-L. 1972. *English School Achievements and Some Student Characteristics 1 on the Relationships of Personality and Intelligence Variables to English School Achievements.* Research Bulletin, 33. Helsinki: Institute of Education.

Lyons, J. 1968. *Introduction to Theoretical Linguistics*. Cambridge: University Press.

Moulton, W.G. 1962. *The Sounds of English and German*. Chicago: The University of Chicago Press.

Nemser, W. 1971. The Predictability of Interference Phenomena in the English Speech of Native Speakers of Hungarian, in Nickel 1971, 89-96.

Nickel, G. (ed.) 1971. *Papers in Contrastive Linguistics*. Cambridge: University Press.

Nykykielet 1971 = *Valtion oppikoulujen opetussuunnitelmat: Nykykielet.* Helsinki: Valtion Painatuskeskus.

Nykysuomen Sanakirja. 1973. Porvoo: WSOY.

Peltonen, M. 1970. *Johdatus käyttäytymistieteiden tilastollisiin menetelmiin*. Porvoo: WSOY.

POPS 1970 = *Peruskoulun opetussuunnitelmakomitean mietintö, II. Oppiaineiden opetussuunnitelmat.* Komiteanmietintö 1970: A 5. Helsinki: Valtion Painatuskeskus.

POPS 1973 = POPS 1970, *Kielitaidon mittaaminen.* Opas 3 b. ed. S. Takala. Kouluhallitus, Kokeilu- ja tutkimustoimisto. Helsinki: Valtion Painatuskeskus.

Pulkkinen, P. 1966. *Asiasuomen opas*. Kolmas painos. Helsinki: Otava.

Ritvanen, R. 1971. *Oppilaan sosiaalisen taustan yhteys koulumenestykseen kokeiluperuskoulun VII ja VIII luokilla 1967-69*. Publications of the Institute for Educational Research, 85. Jyväskylä.

Smith, A.N. 1971. The Importance of Attitude in Foreign Language Learning. *The Modern Language Journal*, 55, 82-88.

Spolsky, B. 1969. Attitudinal Aspects of Second Language Learning. *Language Learning*, 19, 271-285.

Stratton, J. 1970. *The Sounds of English: a Practise Book of Contrastive Material*. Helsinki: Kirjayhtymä.

Sysiharju, A-L. 1970. *Mieluisa - yhdentekevä - vastenmielinen. Eri oppiaineiden oppilaissa herättämistä tunnereaktioista 1960-luvun keskikoulun päätösvaiheessa.* Helsingin yliopiston kasvatustieteen laitoksen tutkimuksia 8.

Takala, S. 1968. *Koko ikäluokan ruotsin kielen koulusaavutukset.* Kouluhallituksen julkaisuja, Kokeilu- ja tutkimusselosteita, 13. Helsinki.

Tommola, J. 1975. *On the Relationship between Discrimination and Production of English Sounds by Finnish Learners.* Publications in English Studies, University of Turku, 5. Turku: Department of English.

Trubetzkoy, N.S. 19[illegible]. *Principles of Phonology.* Berkeley and London: University of California Press.

Valette, R. 1967. *Modern Language Testing: a Handbook.* New York: Harcourt, Brace & World.

Wiik, K. 1965a. *Finnish and English Vowels: a Comparison with Special Reference to the Learning Problems Met by Native Speakers of Finnish Learning English.* Annales Universitatis Turkuensis, B 94. Turku: University of Turku.

Wiik, K. 1965b. *Finnish and English Consonants: Preliminary Version.* Unpublished.

Wiik, K. 1966. *Finnish and English Laterals: a Comparison with Special Reference to the Learning Problems Met by Native Speakers of Finnish Learning English.* Publications of the Phonetics Department of the University of Turku, 1. Turku: Department of Phonetics.

Wiik, K. 1975. *Taksonomista fonologiaa.* Publications of the Phonetics Department of the University of Turku, 11. Second edition. Turku: Department of Phonetics.

APPENDIX 1

TEST 1. SUBSTITUTION TEST

		TWICE FROM THE TAPE	CONSONANT TESTED		TWICE FROM THE TAPE	CONSONANT TESTED
Practice	1.	varstat	v r s t t	3.	beside	b s d
items	2.	katapultti	k t p l t	4.	mean	m n
Actual	1.	pack	p k	19.	emerge	m dʒ
test	2.	fate	f t	20.	yoga	j g
items	3.	those	ð z	21.	away	w
	4.	deserve	d z v	22.	garage	g r ʒ
	5.	shady	ʃ d	23.	author	θ
	6.	ever	v	24.	zip	z p
	7.	hanger	h ŋ	25.	Asia	ʃ
	8.	cab	k b	26.	viking	v k ŋ
	9.	better	b t	27.	itch	tʃ
	10.	gem	dʒ m	28.	neither	n ð
	11.	teeth	t θ	29.	foolish	f l ʃ
	12.	essay	s	30.	appeal	p l
	13.	with	w ð	31.	azure	ʒ
	14.	rubber	r b	32.	thud	θ d
	15.	chief	tʃ f	33.	adjure	dʒ
	16.	mountain	m n t n	34.	surface	s f s
	17.	etcher	tʃ	35.	beyond	b j n d
	18.	leg	l g			

APPENDIX 2

TEST 2. DISCRIMINATION TEST

		THE TRIPLETS FROM THE TAPE			OPPOSITION TESTED	Average correct answer percentage ($\bar{X}$%)
Practice items	1.	salo	palo	palo	s - p - p	
	2.	soma	soma	soma	no opposition	
	3.	multa	mutta	multa	l - t - l	
	4.	sana	sana	sama	n - n - m	
	5.	sound	sound	hound	s - s - h	
	6.	pen	ten	then	p - t - ð	
	7.	peal	peace	peace	l - s - s	
	8.	sum	sun	sum	m - n - m	
Actual test items	1.	harsh	marsh	marsh	h - m - m	100
	2.	mingle	mingle	single	m - m - s	100
	3.	tub	dub	tub	t - d - t	93
	4.	pig	big	big	p - b - b	76
	5.	feed	feet	feed	d - t - d	97
	6.	thorn	thorn	faun	θ - θ - f	19
	7.	chair	share	chair	tʃ - ʃ - tʃ	79
	8.	ether	either	ether	ɛ - ð - θ	92
	9.	cold	gold	cold	k - g - k	78
	10.	cash	catch	catch	ʃ - tʃ - tʃ	51
	11.	rum	rum	rum	no opposition	89
	12.	ledger	ledger	lecher	dʒ - dʒ - tʃ	46
	13.	bet	wet	vet	b - w - v	40
	14.	batch	badge	batch	tʃ - dʒ - tʃ	82
	15.	weeper	weaver	weaver	p - v - v	97
	16.	which	rich	rich	w - r - r	61
	17.	clothing	closing	clothing	ð - z - ð	31
	18.	ram	rang	ram	m - ŋ - m	63
	19.	over	over	ower	v - v - w	97
	20.	shield	shield	sealed	ʃ - ʃ - s	94

	THE TRIPLETS FROM THE TAPE			OPPOSITION TESTED	Average correct answer percentage ($\bar{X}$%)
Actual test items	21. wink	wing	wing	ŋk- ŋ - ŋ	92
	22. heads	hedge	heads	dz- dʒ- dz	82
	23. lip	rip	nip	l - r - n	97
	24. cunning	coming	coming	n - m - m	39
	25. Paris	parish	parish	s - ʃ - ʃ	64
	26. eyes	ice	eyes	z - s - z	98
	27. clove	clothe	clothe	v - ð - ð	19
	28. haggle	haggle	hackle	g - g - k	79
	29. strife	strive	strive	f - v - v	72
	30. yeast	yeast	east	ji:-ji:-i:	89
	31. lobe	lope	lobe	b - p - b	79
	32. parcel	parcel	partial	s - s - ʃ	86
	33. singer	singer	sinner	ŋ - ŋ - n	74
	34. latches	latches	latches	no opposition	92
	35. win	wing	wink	n - ŋ - ŋk	63
	36. teller	terror	terror	l - r - r	84
	37. fault	fault	vault	f - f - v	87
	38. teeth	teeth	teethe	θ - θ - ð	63
	39. lean	wean	lean	l - w - l	91
	40. lashes	latches	latches	ʃ - tʃ- tʃ	85
		Pause			
	41. pace	pays	pace	s - z - s	95
	42. lesion	lesion	legion	ʒ - ʒ - dʒ	52
	43. bleating	bleeding	bleating	t - d - t	95
	44. true	through	through	tr- θr- θr	96
	45. mesher	mesher	measure	ʃ - ʃ - ʒ	40
	46. than	van	than	ð - v - ð	14
	47. ban	ban	pan	b - b - p	82
	48. thy	vie	fie	ð - v - f	14
	49. brief	breathe	breathe	f - ð - ð	83
	50. pallid	valid	pallid	p - v - p	93

	THE TRIPLETS FROM THE TAPE			OPPOSITION TESTED	Average correct answer percentage ($\bar{X}$%)
Actual test items	51. wick	wick	wig	k - k - g	89
	52. curve	curve	curb	v - v - b	66
	53. sing	sing	king	s - s - k	94
	54. hanger	hammer	hanger	ŋ - m - ŋ	84
	55. west	vest	west	w - v - w	67
	56. zip	zip	sip	z - z - s	86
	57. pitch	pits	pitch	tʃ- ts- tʃ	79
	58. surface	service	service	f - v - v	89
	59. staple	stable	staple	p - b - p	46
	60. looser	Luther	Luther	s - θ - θ	93
	61. deaf	death	death	f - θ - θ	14
	62. drain	train	drain	dr- tr- dr	40
	63. bards	bards	barge	dz- dz - dʒ	85
	64. vain	rain	rain	v - r - r	93
	65. jaw	chore	jaw	dʒ- tʃ- dʒ	55
	66. zone	shown	Joan	z - ʃ - dʒ	56
	67. bill	will	bill	b - w - b	96
	68. heifer	heather	heifer	f - ð - f	80
	69. bays	bays	beige	z - z - ʒ	83
	70. catty	catchy	catchy	t - tʃ- tʃ	85
	71. laser	lacer	laser	z - s - z	74
	72. thy	thy	thigh	ð - ð - θ	88
	73. seize	seize	seethe	z - z - ð	91
	74. thick	sick	thick	θ - s - θ	93
	75. wary	vary	wary	w - v - w	69

APPENDIX 3

TEST 3. SOUND ANALOGY TEST

	FROM THE TAPE				Average correct answer percentage (X%)
	STIMULUS	ANALOGICAL WORDS		OPPOSITION TESTED	
Practice items	1. poika	palkka	voida	p - p - v	
	2. tila	peli	naru	t - p - n	
	3. doll	tea	day	d - t - d	
	4. she	shoe	short	ʃ - ʃ - ʃ	
Actual test items	1. cadge	girl	high	k - g - h	67
	2. booty	pen	bike	b - p - b	55
	3. faun	film	four	f - f - f	87
	4. chore	child	she	tʃ- tʃ- ʃ	78
	5. pall	book	past	p - b - p	55
	6. thigh	thing	first	θ - θ - f	9
	7. hoist	her	home	h - h - h	80
	8. cot	part	count	k - p - k	89
	9. sear	say	shop	s - s - ʃ	77
	10. willow	very	boat	w - v - b	29
	11. turf	dark	today	t - d - t	78
	12. gibe	jump	chair	dʒ- dʒ- tʃ	32
	13. lumber	wall	long	l - w - l	93
	14. vine	four	very	v - f - v	56
	15. nob	milk	ten	n - m - t	93
	16. shaft	cheek	see	ʃ - tʃ- s	21
	17. guts	good	coffee	g - g - k	48
	18. mole	name	man	m - n - m	87
	19. thee	they	thing	ð - ð - θ	47
	20. chum	June	chalk	tʃ- dʒ- tʃ	6
	21. dub	this	tea	d - ð - t	89
	22. yield	young	easy	j - j - i:	94
	23. sooth	table	summer	s - t - s	87

	FROM THE TAPE				Average correct answer percentage ($\bar{X}$%)
	STIMULUS	ANALOGICAL WORDS		OPPOSITION TESTED	
Actual test items	24. vigil	wake	very	v - w - v	12
	25. thrill	Friday	train	θr- fr - tr	18
		Pause			
Practice items	1. mies	sydän	pylväs	s - n - s	
	2. dog	big	speak	g - g - k	
	3. small	moon	sing	l - n - ŋ	
Actual test items	26. leash	fish	teach	ʃ - ʃ - tʃ	16
	27. tang	lying	ring	ŋ - ŋ - ŋ	45
	28. mash	miss	dish	ʃ - s - ʃ	86
	29. glean	one	room	n - n - m	66
	30. serge	watch	porridge	dʒ- tʃ- dʒ	17
	31. soothe	teeth	with	ð - θ - ð	15
	32. hawk	back	dog	k - k - g	39
	33. leech	wash	much	tʃ- ʃ - tʃ	66
	34. wail	girl	write	l - l - t	81
	35. rude	let	with	d - t - ð	45
	36. thrive	laugh	eve	v - f - v	20
	37. flout	yes	ball	t - s - l	91
	38. dice	plus	boys	s - s - z	51
	39. helot	cloud	sit	t - d - t	41
	40. purge	which	eyes	dʒ- tʃ- z	5
	41. hag	break	big	g - k - g	54
	42. heath	mouth	half	θ - θ - f	28
	43. ace	house	brush	s - s - ʃ	79
	44. hose	days	face	z - z - s	41
	45. reef	both	knife	f - θ - f	30

APPENDIX 4

TEST 4. WRITTEN ANALOGY TEST

		STIMULUS FROM THE TAPE	ANALOGICAL WORDS ON THE ANSWER SHEET		OPPOSITION TESTED	Average correct answer percentage (X%)
Practice items	1.	malli	nukkua	mitta	m - n - m	
	2.	heti	koti	talo	h - k - t	
	3.	car	coat	good	k - k - g	
	4.	river	write	rain	r - r - r	
Actual test items	1.	poke	pen	buy	p - p - b	61
	2.	fag	phone	five	f - f - f	48
	3.	dote	they	desk	d - ð - d	82
	4.	sham	child	short	ʃ - tʃ - ʃ	35
	5.	wail	walk	very	w - w - v	46
	6.	therm	this	four	θ - ð - f	8
	7.	lax	round	learn	l - r - l	88
	8.	cane	cat	give	k - k - g	57
	9.	thine	third	there	ð - θ - ð	16
	10.	tilt	door	tall	t - d - t	80
	11.	chive	shop	cheek	tʃ - ʃ - tʃ	71
	12.	nag	know	number	n - n - n	66
	13.	bias	put	boy	b - p - b	82
	14.	jot	she	chair	dʒ - ʃ - tʃ	13
	15.	rear	run	why	r - r - w	86
	16.	locus	table	how	l - t - h	97
	17.	gale	come	good	g - k - g	71
	18.	thrush	tree	three	θr - tr - θr	68
	19.	toil	ten	dark	t - t - d	94
	20.	sift	shoe	some	s - ʃ - s	50
	21.	gem	church	just	dʒ - tʃ - dʒ	36
	22.	nil	neck	moon	n - n - m	86
	23.	foil	very	first	f - v - f	93

	STIMULUS FROM THE TAPE	ANALOGICAL WORDS ON THE ANSWER SHEET		OPPOSITION TESTED	Average correct answer percentage ($\bar{X}$%)
Actual test items	24. veil	that	very	v - ð - v	57
	25. sheer	seven	show	ʃ - s - ʃ	84
	26. thane	think	full	θ - θ - f	58
		Pause			
Practice items	1. jalas	solan	keinäs	s - n - s	
	2. let	sit	bed	t - t - d	
	3. walk	five	big	k - v - g	
Actual test items	27. varlet	read	coat	t - d - t	46
	28. truce	boys	fish	s - z - ʃ	5
	29. loath	half	mouth	θ - f - θ	31
	30. fug	bag	work	g - g - k	52
	31. fang	in	sing	ŋ - n - ŋ	86
	32. serf	teeth	knife	f - θ - f	50
	33. fuse	blouse	always	z - z - z	56
	34. omis	face	bell	s - s - l	76
	35. wick	bag	back	k - g - k	70
	36. tithe	tooth	give	ð - θ - v	17
	37. perch	hats	watch	tʃ- ts -tʃ	72
	38. fen	it	us	n - t - s	98
	39. trash	dish	much	ʃ - ʃ - tʃ	31
	40. fade	bed	with	d - d - ð	90
	41. booze	ice	days	z - s - z	28
	42. reef	enough	wife	f - f - f	49
	43. deem	strong	home	m - ŋ - m	46
	44. badge	page	teach	dʒ- dʒ- tʃ	25
	45. heave	both	half	v - θ - f	4
	46. hutch	British	which	tʃ- ʃ - tʃ	49
	47. weird	bread	eat	d - d - t	88
	48. parch	Birch	porridge	tʃ- tʃ- dʒ	44

APPENDIX 5

TEST 5. PRODUCTION TEST

		TWICE FROM THE TAPE	CONSONANT TESTED		TWICE FROM THE TAPE	CONSONANT TESTED
Actual	1.	minute	m n t	22.	David	d v d
test	2.	house	h s	23.	porridge	p r dʒ
items	3.	sail	s l	24.	both	b θ
	4.	learn	l n	25.	sugar	ʃ g
	5.	ringing	r ŋ ŋ	26.	with	w ð
	6.	yards	j d z	27.	these	ð z
	7.	tooth	t θ	28.	cab	k b
	8.	zed	z d	29.	washing	w ʃ ŋ
	9.	fish	f ʃ	30.	sits	s t s
	10.	give	g v	31.	busy	b z
	11.	rouge	r ʒ	32.	face	f s
	12.	wife	w f	33.	church	tʃ tʃ
	13.	babies	b b z	34.	usually	j ʒ l
	14.	thirty	θ t	35.	dish	d ʃ
	15.	chalk	tʃ k	36.	vegetable	v dʒ t ð l
	16.	other	ð	37.	beside	b s d
	17.	which	w tʃ	38.	June	dʒ n
	18.	dog	d g	39.	page	p dʒ
	19.	shop	ʃ p	40.	zip	z p
	20.	aware	w	41.	Jim	dʒ m
	21.	teacher	t tʃ			